BMJ Open

Effect of smoking on physical and cognitive capability in later life: a multicohort study using observational and genetic approaches

Teri-Louise North,[1] Tom M Palmer,[1,2,3] Sarah J Lewis,[1] Rachel Cooper,[4] Chris Power,[5] Alison Pattie,[6] John M Starr,[7] Ian J Deary,[6,7] Richard M Martin,[1,8] Avan Aihie Sayer,[9] Meena Kumari,[10,11] Cyrus Cooper,[9,12,13] Mika Kivimaki,[11] Diana Kuh,[4] Yoav Ben-Shlomo,[1] Ian N M Day[1]

To cite: North T-L, Palmer TM, Lewis SJ, *et al*. Effect of smoking on physical and cognitive capability in later life: a multicohort study using observational and genetic approaches. *BMJ Open* 2015;**5**:e008393. doi:10.1136/bmjopen-2015-008393

▸ Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2015-008393).

Received 2 April 2015
Revised 26 October 2015
Accepted 29 October 2015

For numbered affiliations see end of article.

Correspondence to
Teri-Louise North;
teri.north.2005@my.bristol.ac.uk

ABSTRACT

Objectives: The observed associations between smoking and functional measures at older ages are vulnerable to bias and confounding. Mendelian randomisation (MR) uses genotype as an instrumental variable to estimate unconfounded causal associations. We conducted a meta-analysis of the observational associations and implemented an MR approach using the smoking-related single nucleotide polymorphism rs16969968 to explore their causal nature.

Setting: 9 British cohorts belonging to the HALCyon collaboration.

Participants: Individual participant data on N=26 692 individuals of European ancestry (N from earliest phase analysed per study) of mean ages 50–79 years were available for inclusion in observational meta-analyses of the primary outcomes.

Primary outcomes: Physical capability, cognitive capability and cognitive decline. The smoking exposures were cigarettes per day, current versus ex-smoker, current versus never smoker and ever versus never smoker.

Results: In observational analyses current and ever smoking were generally associated with poorer physical and cognitive capability. For example, current smokers had a general fluid cognition score which was 0.17 z-score units (95% CI −0.221 to −0.124) lower than ex-smokers in cross-sectional analyses. Current smokers had a walk speed which was 0.25 z-score units lower than never smokers (95% CI −0.338 to −0.170). An MR instrumental variable approach for current versus ex-smoker and number of cigarettes smoked per day produced CIs which neither confirmed nor refuted the observational estimates. The number of genetic associations stratified by smoking status were consistent with type I error.

Conclusions: Our observational analysis supports the hypothesis that smoking is detrimental to physical and cognitive capability. Further studies are needed for a suitably powered MR approach.

Strengths and limitations of this study

- Using individual participant data from nine UK cohorts of ageing individuals, this study meta-analyses the associations between smoking and physical and cognitive capability in later life.
- We consider cognitive and physical capability, in addition to cognitive capability decline.
- We derive a score for general fluid cognition and include this in cross-sectional analyses.
- We use the rs16969968 single nucleotide polymorphism, which associates with nicotine dependence, in a Mendelian randomisation (MR) to explore the causality of the observational associations.
- While our study has demonstrated the feasibility of using an MR approach to understand the association of smoking with ageing outcomes, it has demonstrated that a larger sample size is required for a suitably powered analysis.

INTRODUCTION

Epidemiological studies have been conducted to explore the associations between smoking and physical and cognitive capability in mid to later life, generally concluding that smoking is associated with worse capability outcomes.[1–8] Physical and cognitive capability, otherwise known as the physical and intellectual tasks of daily living, are often used as markers of ageing having been consistently shown to be associated with survival and onset of disease and disability. For example, a recent meta-analysis of over 30 studies showed that poorer objective physical capability is associated with higher mortality rates.[9] Smoking is a modifiable behaviour and understanding the extent to which it influences biological

ageing is crucial given the burden of morbidity in today's ageing populations.

The associations between smoking and markers of ageing are likely to be confounded (or mediated) by factors such as socioeconomic position (SEP), body mass index, health status and prior IQ. Although studies have adjusted for these factors, residual confounding and bias may have affected the interpretation of results. Mendelian randomisation (MR) uses genotype as an instrumental variable (IV) to estimate the causal effect of an exposure on an outcome free of confounding and reverse causation bias.[10 11] If genotype is associated with the exposure under consideration and is not associated with the confounders of the observational association, nor directly with the outcome of interest, it may be used to conduct an IV analysis to generate a causal estimate of the observational association. This can be implemented, for example, using a two-stage approach where the predicted exposure based on genotype is used to measure the association with the outcome. Associations of genotype with the outcome in different strata of the exposure can also contest or support the causality of the observational association. If a genetic association is observed in exposed individuals and not in unexposed individuals, this supports a causal observational association. The minor allele of the rs16969968 single nucleotide polymorphism (SNP) in the *CHRNA5* gene has been consistently associated with increased nicotine dependence,[12] therefore providing a potential instrument for MR analyses of the effects of smoking.

Harnessing the availability of data across cohorts belonging to the HALCyon collaboration,[13] we have conducted a meta-analysis of smoking and physical and cognitive capability using individual participant data (IPD) from middle-aged to older individuals in the UK. Given the known associations between rs16969968 and nicotine dependence, we explored the associations between this SNP and our ageing outcomes in different smoking classes to supplement the observational associations. We considered that if associations are observed in current or ever smokers but not in never smokers, this provides evidence to support causality. We implemented instrumental variable regression to generate IV estimates for the true causal associations of smoking with continuous physical and cognitive capability measures. Our aim was to examine whether the observational associations are consistent with estimates obtained using an MR approach.

METHOD

IPD was meta-analysed across nine cohorts belonging to the HALCyon collaboration. These included the Boyd Orr Cohort (BO), the Caerphilly Prospective Study (CaPS), the English Longitudinal Study of Ageing (ELSA), the Hertfordshire Ageing Study (HAS), the Hertfordshire Cohort Study (HCS), the Lothian Birth Cohort 1921 (LBC1921), the National Child Development Study (NCDS), the MRC National Survey of Health and Development (NSHD) and the Whitehall II Study (WHII). Further information about the cohorts is provided in previous HALCyon publications.[14]

Outcome measures

Physical capability was assessed using the objective measures grip strength, chair rise speed, inability to balance on one leg for 5 s with eyes open, walking speed and timed get up and go (TUG) speed. Cognitive capability measures included tests of crystallised intelligence (National Adult Reading Test (NART) and Mill Hill) and fluid intelligence (Alice-Heim 4-I test (AH4), semantic fluency, phonemic fluency, search speed, word recall, four choice reaction time (FCRT), logical memory and Raven's Progressive Matrices). Crystallised intelligence measures knowledge accumulated across the life course like vocabulary and captures premorbid IQ, while fluid intelligence measures problem-solving skills.[15] Cognitive decline was calculated by taking the percentage change in continuous cognitive measure between the baseline wave and the last available wave. This was then converted to the percentage change per year using the age difference between waves. Individuals were categorised into a binary variable of the top 25% of decliners versus others in each cohort. This approach is similar to previous studies.[16 17] We used factor analysis to derive a general fluid (Gf) cognitive ability score and included this in the cross-sectional analyses.

Smoking behaviour

Participants were classified at the time of baseline capability assessment as current, ex or never smokers. A 'smoker' was defined as an individual who smoked pipes, cigars, manufactured cigarettes or hand-rolled cigarettes, if this information was available. For current smokers, we additionally analysed number of cigarettes smoked per day (CPD). Where possible this was restricted to manufactured cigarettes to maintain consistency in tobacco quantities. Individuals who were occasional smokers (smokes less than 1 CPD or does not smoke daily but does smoke occasionally) were re-coded as smoking 0 CPD. In Boyd Orr, we estimated CPD by taking the median of intervals of cigarettes smoked per day.

Genetic analyses

We genotyped rs16969968 across cohorts where this was not previously available. Rs1051730 was substituted into analyses where available and when rs16969968 was unavailable (see online supplementary table S1). These two SNPs are highly correlated and thus interchangeable.[12]

Covariates

We selected a range of covariates a priori. These were sex, age (continuous), socioeconomic position obtained from the earliest wave of outcome assessment (the Registrar-General's Social Class, RGSC), body mass

index (BMI), height and disease status (history of diabetes, stroke or heart disease: see online supplementary table S2). BMI, height and disease status were derived using data from the same wave as each outcome measure if available unless the outcome was a decline measure in which case the covariates were taken from the baseline wave. Age, sex and SEP were included as potential confounders of the observational associations of smoking and physical and cognitive capability. Disease history was included to examine the extent of mediation of smoking effects via diabetes, stroke and heart disease. BMI may confound and mediate the association of smoking and physical capability, while height is strongly correlated with physical capability outcomes.

Further information on the genotyping and the derivation and harmonisation of the exposures, covariates and outcomes is provided in online supplementary material.

Statistical analysis

All analyses were conducted using Stata 13.1.[18] Known ethnic outliers were excluded from the analysis (self-reported non-European ancestry or previously detected from genome-wide data), as were related individuals. For harmonisation continuous outcome measures were standardised within cohorts using all data available. Logistic regression was implemented for analyses of binary outcomes and linear regression for continuous outcomes.

Four choice reaction time was inverse transformed and search speed was natural log transformed to improve normality. A score for Gf was derived in several cohorts using the -factor- command in Stata[18] using the pcf option and imposing one factor, to supplement cross-sectional analyses.

Observational analyses assessed the associations between CPD or smoking status (current vs ex, current vs never or ever vs never smoker) and each of the physical and cognitive capability measures. Three models were run for physical capability adjusted for (1) age, SEP and sex (2) age, SEP, sex and disease status (3) age, SEP, sex, disease status, height and BMI. Models (1) and (2) were run for the cognitive outcomes.

The rs16969968 genotype was coded additively as 0, 1 or 2 minor alleles. The associations between genotype and smoking behaviour were calculated using linear or logistic regression. The associations between genotype and physical and cognitive capability were calculated in current, never and ever smokers to test for both pleiotropy and whether an association was observed in current and ever but not never smokers. All genetic associations were adjusted for age and sex.

For the smoking exposures which correlated with genotype (cigarettes per day and current vs ex smoking) and for the continuous outcomes which associated with these smoking exposures, we performed IV estimation using the two-stage least squares (2SLS) estimator and compared the observational estimates with the IV estimates. The IV assumptions, that can be checked using observational data, were checked by testing the unadjusted associations between genotype and the covariates.

All associations were analysed within cohorts and the effect estimates meta-analysed using a random effects two-stage approach.[19 20] The observational analyses included all individuals with data available on the exposures and outcomes of interest. Secondary models were restricted to individuals who had relevant covariates available. The genetic and IV analyses on the associations between genotype and cognitive and physical capability were conducted on a subset of the observational sample with genotype data. The associations between genotype and smoking behaviour or covariates were examined in all individuals with data available.

RESULTS

The observational and genetic samples are characterised in table 1 and online supplementary tables S3–5. The mean age in the observational analysis was 50–79 years and the majority of studies had similar numbers of men and women. The total sample size taking the earliest wave of outcome assessment was 26 692. As shown in table 2, each of the physical capability outcomes was associated with at least one comparison of smoking status. In particular, current compared to never smoker status resulted in a decrease in walking or TUG speed of between 0.23 and 0.29 z-scores ($p<0.0001$). Across outcomes and models, the effect estimates generally suggest that being a current or an ever smoker was associated with worse physical capability compared to never smokers. Adjustment for BMI and height often resulted in an increased magnitude of effect on physical capability for the smoking exposures. Heterogeneity statistics are provided in the Supplement.

Current compared to ex or never smokers performed more poorly across all cross-sectional cognitive outcomes tested for both models with the exception of phonemic fluency and the single cohort analyses (table 3, see online supplementary table S6). An extra cigarette per

Table 1 Age and sex by cohort study

Cohort	Mean age in years (SE)	Percentage female	Total sample size, n
BO	69.64 (0.25)	55.56	279
CaPS	61.77 (0.10)	0	1831
ELSA	66.01 (0.13)	54.56	5425
HAS	67.39 (0.09)	35.85	636
HCS	66.13 (0.05)	47.84	2803
LBC	79.06 (0.02)	57.38	542
NCDS	50 (NA)	50.63	7652
NSHD	53.45 (0.00)	50.9	2949
WHII	55.39 (0.09)	26.56	4575
TOTAL			26 692

Numbers based on all individuals with age, sex, smoking status, socioeconomic position (SEP) and at least one outcome measure at the earliest phase analysed. Numbers based on grip strength analysis for HCS.

Table 2 Observational estimates for the associations between smoking and physical capabilities

Outcome	Model†	Cigarette per day‡	Current vs ex-smoker	Current vs never smoker	Ever vs never smoker
Regression coefficient (95% CI)					
Grip strength	M1	0.000 (−0.003 to 0.003)	−0.049** (−0.085 to −0.014)	−0.017 (−0.077 to 0.043)	−0.002 (−0.045 to 0.041)
	M2	0.001 (−0.003 to 0.004)	−0.054** (−0.088 to −0.019)	−0.012 (−0.069 to 0.045)	0.006 (−0.034 to 0.047)
	M3	0.001 (−0.003 to 0.004)	−0.026 (−0.060 to 0.008)	−0.006 (−0.051 to 0.040)	−0.001 (−0.040 to 0.038)
Chair rise speed	M1	−0.004 (−0.010 to 0.001)	−0.111** (−0.190 to −0.032)	−0.150*** (−0.233 to −0.067)	−0.061** (−0.102 to −0.020)
	M2	−0.004 (−0.010 to 0.002)	−0.115** (−0.188 to −0.043)	−0.152*** (−0.242 to −0.062)	−0.059* (−0.108 to −0.010)
	M3	−0.002 (−0.008 to 0.004)	−0.152**** (−0.224 to −0.079)	−0.188*** (−0.285 to −0.091)	−0.052* (−0.095 to −0.009)
Walk speed	M1	−0.006 (−0.016 to 0.005)	−0.129* (−0.242 to −0.016)	−0.254**** (−0.338 to −0.170)	−0.136**** (−0.171 to −0.102)
	M2	−0.006 (−0.015 to 0.002)	−0.142* (−0.250 to −0.033)	−0.247**** (−0.325 to −0.168)	−0.124**** (−0.159 to −0.089)
	M3	−0.001 (−0.013 to 0.010)	−0.185** (−0.304 to −0.066)	−0.266**** (−0.373 to −0.159)	−0.110**** (−0.144 to −0.075)
TUG speed	M1	0.000 (−0.010 to 0.011)	−0.075 (−0.161 to 0.011)	−0.233**** (−0.323 to −0.144)	−0.138** (−0.234 to −0.042)
	M2	0.001 (−0.009 to 0.012)	−0.102* (−0.186 to −0.017)	−0.236**** (−0.331 to −0.142)	−0.123* (−0.217 to −0.029)
	M3	0.007 (−0.007 to 0.021)	−0.166**** (−0.249 to −0.082)	−0.290**** (−0.383 to −0.197)	−0.132*** (−0.205 to −0.060)
OR (95% CI)					
Inability to balance on one leg for	M1	1.015 (0.994 to 1.036)	1.125 (0.937 to 1.351)	1.210 (0.948 to 1.543)	1.092 (0.957 to 1.246)
5 s	M2	1.013 (0.992 to 1.035)	1.155 (0.958 to 1.392)	1.232* (1.000 to 1.516)	1.064 (0.930 to 1.217)
	M3	1.007 (0.985 to 1.029)	1.361** (1.120 to 1.655)	1.415** (1.106 to 1.811)	1.074 (0.934 to 1.236)

*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.
†Models: (M1) age, sex and SEP adjusted, (M2) M1 + disease adjusted, (M3) M2 + height, BMI adjusted.
‡Association is for 1 CPD for comparison with genotypic analysis.
BMI, body mass index; CPD, cigarettes smoked per day; SEP, socioeconomic position; TUG, timed get up and go.

Table 3 Observational estimates for the associations between smoking and cognitive capabilities

		Regression coefficient (95% CI)			
Outcome	**Model†**	**Cigarette per day**	**Current vs ex-smoker**	**Current vs never smoker**	**Ever vs never smoker**
Crystallised measures					
Mill Hill	M1	−0.003 (−0.012 to 0.005)	−0.140*** (−0.216 to −0.063)	−0.192**** (−0.266 to −0.117)	−0.092*** (−0.142 to −0.043)
	M2	−0.004 (−0.013 to 0.005)	−0.146*** (−0.223 to −0.070)	−0.194**** (−0.269 to −0.119)	−0.090*** (−0.140 to −0.040)
NART	M1	−0.004 (−0.010 to 0.001)	−0.174**** (−0.261 to −0.088)	−0.159** (−0.262 to −0.056)	−0.041 (−0.097 to 0.016)
	M2	−0.004 (−0.010 to 0.002)	−0.194**** (−0.265 to −0.123)	−0.142* (−0.260 to −0.023)	−0.028 (−0.086 to 0.030)
Fluid measures					
Gf	M1	−0.002 (−0.011 to 0.007)	−0.173**** (−0.221 to −0.124)	−0.147**** (−0.205 to −0.088)	−0.036* (−0.067 to −0.006)
	M2	−0.002 (−0.011 to 0.007)	−0.178**** (−0.228 to −0.129)	−0.137**** (−0.199 to −0.074)	−0.029 (−0.059 to 0.002)
AH4	M1	0.000 (−0.008 to 0.009)	−0.139**** (−0.195 to −0.082)	−0.135**** (−0.196 to −0.073)	−0.047* (−0.090 to −0.003)
	M2	0.000 (−0.009 to 0.009)	−0.148**** (−0.206 to −0.090)	−0.129**** (−0.192 to −0.065)	−0.040 (−0.084 to 0.004)
Semantic fluency	M1	−0.002 (−0.008 to 0.003)	−0.139**** (−0.175 to −0.104)	−0.105*** (−0.162 to −0.047)	−0.019 (−0.059 to 0.021)
	M2	−0.002 (−0.007 to 0.003)	−0.136**** (−0.172 to −0.100)	−0.095** (−0.155 to −0.035)	−0.012 (−0.053 to 0.029)
Phonemic fluency	M1	−0.005 (−0.015 to 0.006)	−0.028 (−0.378 to 0.322)	0.041 (−0.289 to 0.372)	0.031 (−0.022 to 0.084)
	M2	−0.005 (−0.015 to 0.005)	−0.063 (−0.355 to 0.228)	0.056 (−0.295 to 0.407)	0.037 (−0.016 to 0.091)
Search speed‡	M1	−0.005* (−0.009 to −0.001)	−0.122*** (−0.188 to −0.057)	−0.148** (−0.239 to −0.057)	−0.059** (−0.099 to −0.019)
	M2	−0.006** (−0.010 to −0.002)	−0.122*** (−0.192 to −0.052)	−0.142** (−0.232 to −0.051)	−0.054** (−0.091 to −0.016)
Word recall	M1	−0.005 (−0.010 to 0.000)	−0.144*** (−0.222 to −0.067)	−0.138**** (−0.192 to −0.083)	−0.044*** (−0.071 to −0.018)
	M2	−0.005 (−0.010 to 0.000)	−0.143*** (−0.222 to −0.064)	−0.134**** (−0.191 to −0.078)	−0.042** (−0.069 to −0.016)

*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.
†Models: (M1) age, sex and SEP adjusted, (M2) M1 + disease adjusted.
‡Natural log transformed.
AH4, Alice-Heim 4-I test; GF, general fluid; NART, National Adult Reading Test; SEP, socioeconomic position.

day was associated with a slower search speed (p<0.01 in model 2) and there was a trend towards poorer word recall ability.

There were fewer associations apparent between smoking behaviour and cognitive decline (see online supplementary tables S7 and 8). Notably, current compared to ex-smokers were more likely to be in the quartile of greatest decliners in AH4 score, word recall ability, search speed and FCRT. Current smokers experienced worse decline than never smokers in word recall and FCRT, while ever smokers declined faster on the Raven's Progressive Matrices than never smokers. Adjustment for disease status had a small influence on the effect sizes for cognitive outcomes in general.

Instrumental variable analysis

The association between rs16969968 and smoking behaviour is summarised in figures 1 and 2 and online supplementary figures 1 and 2. Each minor allele predicts approximately 1 extra cigarette smoked per day (figure 1). The F statistic[21] for the strength of the association between each minor allele and CPD, obtained from the partial R^2 value from the first stage of the age and sex-adjusted 2SLS regression of CPD and log-transformed search speed, was 3.51 in NSHD, 7.18 in ELSA and 13.64 in NCDS. We also observed an association between each extra minor allele and an increased odds of being a current compared with an ex-smoker (figure 2, p=0.02). There was some evidence of an association between the SNP and being an ever versus a never smoker (see online supplementary figure 1, decreased odds of ever smoker p=0.05), but no evidence for an association with being a current versus a never smoker (see online supplementary figure 2).

There was no evidence for an association between rs16969968 and the covariates, although the age distribution was non-normal (see online supplementary table S9). We tested the association between rs16969968 and each outcome in current, never and in ever smokers (table 4, see online supplementary table S10). We observed an association between the smoking susceptibility allele and increased odds of being in the top 25% of decliners for FCRT in current smokers (see online supplementary table S10). In never smokers, we observed an association between the same allele and poorer search speed.

Given the strong association between rs16969968 and CPD, and the weaker association with this variant and current versus ex-smoker status, we progressed all observational associations for these exposures (arbitrary threshold of p<0.05) to an IV analysis. The results are described in table 5 (and see online supplementary table S11). The CIs from the observational estimates fall within the CIs from the IV estimates. The IV results, however, are not informative about the causality of the observational associations because the IV CIs are wide.

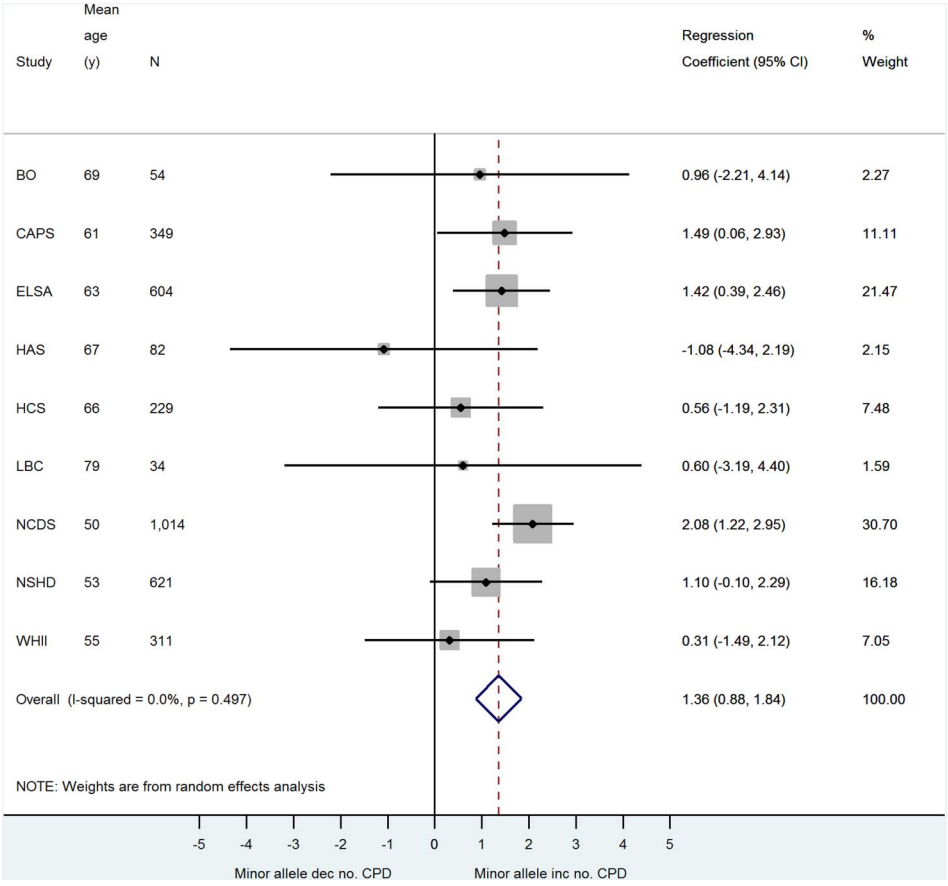


Figure 1 Meta-analysis of minor allele—cigarettes smoked per day association.

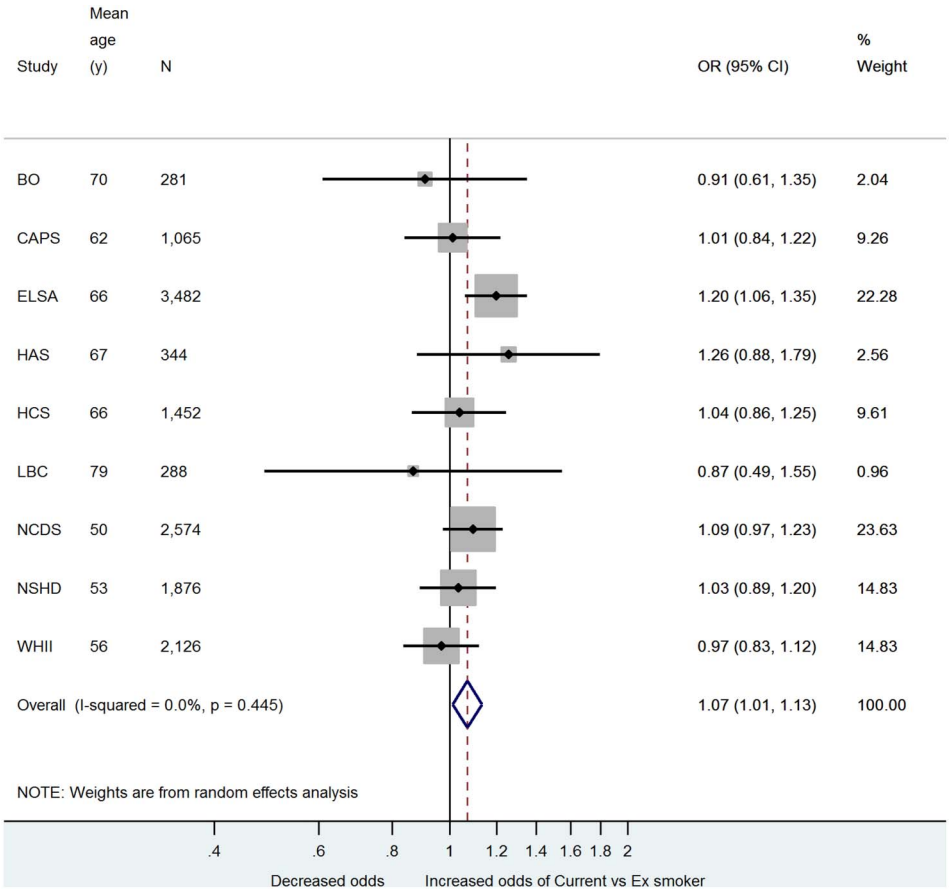


Figure 2 Meta-analysis of minor allele—current versus ex-smoker association.

DISCUSSION

This study has confirmed that smoking, whether it be current smoking or ever smoking, is associated with poorer grip strength, chair rise speed, TUG/walk speed and balance ability in a combined analysis of UK ageing cohorts. In addition, smoking was associated with poorer cognition across eight different cognitive tests and five measures of cognitive decline. The large sample size that HALCyon affords reveals some novel associations that, to our knowledge, have not been reported before. Notably, an association between cigarettes smoked per day and poorer search speed in current smokers. In addition, the high precision of the observational effect-estimates from this analysis are reflected in the narrow CIs. However, some of the observational association is due to residual confounding, as demonstrated by the association observed between smoking and measures of crystallised cognition which should be fairly robust to adverse environmental factors acting from young adulthood and later.

This study has analysed data across a wide range of cohorts from different geographical locations and with different age ranges. Many of the results should thus be generalisable to British individuals of European ancestry of middle to older ages. As with any cohort study, however, there may have been a healthy survivor effect and particularly the analysis of ever smokers may be applicable only to healthy smokers. All cohorts, however, demonstrated genotype frequencies within Hardy-Weinberg Equilibrium (HWE).[22] The HWE p value for all cohorts individually was >0.3 with the exception of LBC1921 (p=0.08), while the collective p value for all cohorts combined was 0.73 (see online supplementary table S3). If the analysis is indeed confined to healthy smokers, then the negative effect of smoking on physical and cognitive capability could be biased downwards here.

Adjustment for height and BMI in the physical capability observational models often resulted in an increased magnitude of effect of smoking status on outcome. BMI could both confound and mediate the association of smoking with physical capability, so the causal relationships are likely to be complex. BMI, in addition to disease status, could also be a collider, whereby covariate adjustment induces a false association between exposure and outcome. Such issues are the primary motivation for conducting an MR. Adjustment for history of disease did not substantially alter the effect estimates in general. While this could suggest that the diseases considered are not on the causal pathway between smoking and outcome, the lack of attenuation could be because these variables have not adequately captured disease history or because they were derived using history of ever having these conditions and thus do not capture smoking-induced incident disease.

The genetic analyses detected an association between rs16969968 and poorer search speed in never smokers. We caution that this could be a false positive due to

Table 4 Associations between rs16969968 and outcomes stratified by smoking status

Outcome category	Outcome	Current smokers	Never smokers	Ever smokers
Regression coefficient (95% CI)				
Physical capability	Grip strength	0.008 (−0.044 to 0.061)	0.016 (−0.030 to 0.062)	−0.015 (−0.037 to 0.008)
	Chair rise speed	0.003 (−0.123 to 0.130)	0.000 (−0.078 to 0.079)	−0.002 (−0.048 to 0.044)
	Walk speed	0.049 (−0.122 to 0.220)	0.026 (−0.015 to 0.068)	0.002 (−0.033 to 0.038)
	TUG speed	−0.043 (−0.154 to 0.067)	−0.015 (−0.083 to 0.054)	−0.024 (−0.077 to 0.029)
Cognitive capability	Mill Hill	0.062 (−0.058 to 0.182)	−0.031 (−0.096 to 0.034)	−0.042 (−0.214 to 0.129)
	NART	0.020 (−0.075 to 0.116)	−0.009 (−0.098 to 0.080)	−0.004 (−0.058 to 0.051)
	Gf	0.048 (−0.011 to 0.108)	−0.035 (−0.074 to 0.004)	0.016 (−0.015 to 0.047)
	AH4	0.024 (−0.069 to 0.118)	−0.040 (−0.102 to 0.021)	0.013 (−0.038 to 0.064)
	Semantic fluency	0.028 (−0.038 to 0.094)	−0.013 (−0.048 to 0.022)	0.014 (−0.014 to 0.042)
	Phonemic fluency	−0.033 (−0.160 to 0.094)	−0.018 (−0.085 to 0.048)	−0.126 (−0.398 to 0.145)
	Search speed†	0.024 (−0.045 to 0.093)	−0.060* (−0.116 to −0.003)	0.005 (−0.028 to 0.038)
	Word recall	0.045 (−0.009 to 0.099)	−0.018 (−0.058 to 0.023)	0.015 (−0.014 to 0.044)
OR (95% CI)				
Physical capability	Inability to balance on one leg for 5 s	1.010 (0.732 to 1.394)	1.052 (0.896 to 1.236)	0.921 (0.817 to 1.038)
Cognitive capability decline	Mill Hill	0.535 (0.063 to 4.559)	0.943 (0.798 to 1.114)	0.935 (0.802 to 1.089)
	NART	1.016 (0.678 to 1.522)‡	0.878 (0.559 to 1.379)	1.085 (0.848 to 1.388)
	AH4	1.007 (0.786 to 1.289)	1.127 (0.955 to 1.329)	0.925 (0.807 to 1.060)
	Semantic fluency	0.975 (0.748 to 1.271)	1.243 (0.926 to 1.670)	0.984 (0.891 to 1.086)
	Phonemic fluency	1.003 (0.735 to 1.369)§	1.097 (0.807 to 1.492)	1.163 (0.565 to 2.393)
	Word recall	0.999 (0.832 to 1.201)	1.024 (0.913 to 1.149)	0.978 (0.890 to 1.074)

Models adjusted for age and sex.
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.
†Natural log transformed.
‡Analysis in CAPS only.
§Analysis in WHII only.
AH4, Alice-Heim 4-I test; Gf, general fluid; NART, National Adult Reading Test; TUG, timed get up and go.

Table 5 Comparison of observational with instrumental variable estimates

Observational association of interest		Observational estimate, β_O†	IV estimate, β_{IV}
Smoking behaviour	Outcome	(95% CI)	(95% CI)
Cigarettes per day	Search speed‡	−0.005* (−0.009 to −0.001)	0.017 (−0.041 to 0.075)
Current vs ex-smoker	Grip strength	−0.049** (−0.085 to −0.014)	−0.417 (−1.344 to 0.510)
	Walk speed	−0.129* (−0.242 to −0.016)	−0.411 (−2.008 to 1.186)
	TUG speed[M3]	−0.166**** (−0.249 to −0.082)	−0.941 (−3.937 to 2.055)
	Chair rise speed	−0.111** (−0.190 to −0.032)	−0.039 (−1.398 to 1.319)
	NART	−0.174**** (−0.261 to −0.088)	−1.236 (−8.711 to 6.238)
	Mill Hill	−0.140*** (−0.216 to −0.063)	−2.785 (−7.107 to 1.537)
	Gf	−0.173**** (−0.221 to −0.124)	0.029 (−1.394 to 1.453)
	Semantic fluency	−0.139**** (−0.175 to −0.104)	−0.154 (−1.571 to 1.263)
	AH4	−0.139**** (−0.195 to −0.082)	−1.575 (−4.707 to 1.556)
	Word recall	−0.144*** (−0.222 to −0.067)	0.247 (−1.045 to 1.538)
	Search speed‡	−0.122*** (−0.188 to −0.057)	0.312 (−1.121 to 1.744)

As explained in the methods, sample used to calculate observed and IV estimates differs according to the availability of variables.
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.
†Observational estimates are model M1 unless stated otherwise.
‡Natural log transformed.
AH4, Alice-Heim 4-I test; Gf, general fluid; IV, instrumental variable; NART, National Adult Reading Test; TUG, timed get up and go.

multiple testing. Of the seven cross-sectional cognitive outcomes considered in table 4 and excluding the current smokers, there were 14 independent tests. The Bonferroni corrected p value is 0.004. However, the result is in support of a hypothesis described by Winterer *et al*[23] that suggested the mediation of the effect of the rs16969968 risk allele on greater nicotine dependence to be via poorer cognition. If the cognitive outcomes under consideration are on the causal pathway between the SNP and nicotine dependence, this violates the MR assumptions. In general, however, we conclude that the rs16969968 variant does not exert a direct effect on the outcomes that is large enough to be detectable in our sample, owing to the lack of an association in never smokers. The trend towards improved word recall ability per minor allele in current smokers complements a study on elderly Taiwanese individuals which found a protective effect of smoking on cognitive function.[24]

The association in a single cohort (CaPS) between current versus ex-smoker status and greater odds of decline in FCRT score complements the positive association of the minor allele of rs16969968 with greater odds of FCRT decline in current smokers but not never smokers. If this association is not spurious then it supports the causality of the observational association between continuing smoking and a decline in reaction ability over time.

The IV CIs did not refute the observational estimates. Our approach is a suitable framework for future studies with larger sample sizes. Taking the predictive utility of rs16969968 for CPD as an example, each additional allele predicts one extra cigarette smoked each day and accounts for approximately 1% of the variance in CPD among current smokers. Using mRnd,[25] an online sample size calculation tool for MR, and taking a mean CPD of 14 and a variance in CPD of 81 (as per ELSA current smokers), genetic association testing in current smokers would require 961 individuals to detect an effect of 1 extra CPD on outcome of 0.1 z-score units. This sample size, which assumes 80% power and a 5% type I error rate, was achieved for several of the outcomes in HALCyon (see online supplementary table S5). However, the observed point estimate for smoking 1 extra CPD on log-transformed search speed was −0.005 z-score units, which would require a sample size of 386 815 current smokers in a 2SLS IV analysis. This demonstrates that an MR approach in HALCyon with CPD as the exposure is underpowered to detect effect sizes comparable to the observational associations, but is powered to detect moderate effect sizes. The lack of associations observed in this MR analysis suggest that the true causal effect sizes are unlikely to be moderate and are more likely to be of small magnitude. A sample size of nearly 400 000 current smokers could only be achieved via meta-analysis of consortia and inclusion of large studies such as the UK BIOBANK study (http://www.ukbiobank.ac.uk/).

While the association between the minor allele of rs16969968 and smoking an extra cigarette per day reported here is in agreement with the literature,[12] our finding of a decreased odds of being an ever smoker with each extra minor allele is likely to be spurious. Lips *et al*[26] found no association of smoking initiation with this SNP while Sherva *et al*[27] found an association between the minor allele and increased odds of being a current versus a never smoker. In our analysis, the SNP appears to be associated with a decreased odds of initiating smoking but also with a decreased odds of quitting once smoking is initiated. Although there has been some evidence of an association between this variant and quitting ability in previous studies even after adjustment for smoking quantity,[28 29] this latter association has not been consistently replicated.[26] A recent MR study

which included several of the HALCyon cohorts reported a similar per allele OR[30] for current versus ex-smoker but no association for ever versus never. As discussed above, the lack of a genetic association with current versus never smoker status that we report here adds to a body of literature reporting conflicting associations of this SNP with smoking initiation.[12 27]

The F statistics extracted from the 2SLS IV analysis of CPD and natural log-transformed search speed suggest that rs16969968 could suffer from weak instrument bias. It has been suggested[21] that pooling the data from the individual studies and conducting an IPD MR can reduce this bias. A suitably powered IV analysis of CPD and search speed could explore these approaches further. However, it has also been noted[30 31] that IV estimates generated using CPD as the exposure variable and rs16969968 as the genetic instrument will be biased because this SNP predicts other measures of tobacco exposure independently of CPD, thus violating the statistical assumptions of MR. In light of this, future MR studies of CPD and physical and cognitive capability should focus on examination of the genetic associations in current and never smokers, rather than on generating precise IV estimates of the true observational association. Such an approach, however, may also be weakened if ever smokers incorrectly report themselves to be never smokers or current smokers do not report the true levels of cigarettes smoked per day. Objective measures of tobacco exposure like cotinine levels avoid some of the problems of inaccurate self-reporting and it has been shown that rs16969968 is a strong predictor of cotinine independent of CPD.[32] This biomarker, however, was not available across the HALCyon studies for meta-analysis. As recently highlighted,[30] a further limitation of MR is that collider bias can occur when we stratify the genetic associations by smoking status because rs16969968 is associated with smoking status. Given the few genetic associations observed in table 4 and online supplementary table S10, which are consistent with type I error, collider bias is unlikely to have affected this analysis.

Our study could be extended in several other respects. The observational analyses could consider change in smoking behaviour over time[3 33] and, as data becomes available, decline in physical capability. Further research using a longitudinal approach with repeat data is needed in the future. In addition, further covariates could be incorporated into the observational models. The association of smoking with physical and cognitive capability is likely to be confounded by other factors such as alcohol intake, IQ and stress.

Previous studies of smaller sample size than ours have been successful at implementing an IV approach using this SNP.[34] The success of using MR to infer causality depends on the predictive utility of the variant, in addition to the effect that smoking actually has on the outcome of interest which is less clearly understood. We have conducted an IPD meta-analysis of smoking and physical and cognitive capability in ageing UK cohorts. This is also the first study to date to use the *CHRNA5* rs16969968 variant to explore the causality of the relation between smoking and physical and cognitive capability. Although our results show that a larger sample size is required, this approach has demonstrated that MR analyses could be 'instrumental' in resolving the smoking-ageing question.

Author affiliations
[1]School of Social and Community Medicine, University of Bristol, Bristol, UK
[2]Division of Health Sciences, Warwick Medical School, University of Warwick, Warwick, UK
[3]Department of Mathematics and Statistics, Lancaster University, Lancaster, UK
[4]MRC Unit for Lifelong Health and Ageing at UCL, London, UK
[5]Population, Policy and Practice, UCL Institute of Child Health, University College London, London, UK
[6]Department of Psychology, University of Edinburgh, Edinburgh, UK
[7]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK
[8]University of Bristol/University Hospitals Bristol NHS Foundation Trust National Institute for Health Research Bristol Nutrition Biomedical Research Unit, University of Bristol, Bristol, UK
[9]MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK
[10]ISER, University of Essex, Essex, UK
[11]Department of Epidemiology and Public Health, UCL, London, UK
[12]National Institute for Health Research Nutrition Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK
[13]National Institute for Health Research Musculoskeletal Biomedical Research Unit, University of Oxford, Oxford, UK

Acknowledgements The authors thank Rebecca Hardy, Jon Heron and Mike Gardner for providing expertise during the manuscript preparation. The authors thank Kate Birnie, Karen Jameson, Holly Syddall, Vanessa Cox, Nikki Graham, Jorgen Engmann, Aida Sanchez, Jon Johnson, Sarah Harris and Andrew Wong for providing data.

Contributors INMD conceived the study. T-LN conducted the analyses. RC provided analysis scripts to harmonise the physical capability outcomes. SJL, TMP, T-LN, INMD and YB-S contributed to the design of the analysis and the interpretation of the data. All authors contributed to writing the manuscript and approved the final version. As cohort representatives, RC, CP, AP, JMS, IJD, RMM, AAS, MeK, CC, MiK, DK and YBS contributed to providing the data to analyse.

Funding The cohort collaboration was established by the HALCyon grant funded by the New Dynamics of Ageing (RES-353-25-0001) http://www.newdynamics.group.shef.ac.uk/. Details of individual funding and cohort funding are provided below. The funders had no role in this study design; in the analysis and interpretation of the data; in the writing of the paper; and in the decision to submit the paper for publication.

1958BC (NCDS) WTCCC: DNA collection was funded by MRC grant G0000934 and cellline creation by Wellcome Trust grant 068545/Z/02. This study makes use of data generated by the Wellcome Trust Case-Control Consortium. A full list of investigators who contributed to generation of the data is available from the Wellcome Trust Case-Control Consortium website. Funding for the project was provided by the Wellcome Trust under the award 076113. Great Ormond Street Hospital/University College London, Institute of Child Health receives a proportion of funding from the Department of Health's National Institute for Health Research (NIHR) ('Biomedical Research Centres' funding).

1958BC (NCDS) T1DGC: DNA collection was funded by MRC grant G0000934 and cellline creation by Wellcome Trust grant 068545/Z/02. This research used resources provided by the Type 1 Diabetes Genetics Consortium, a

collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases, National Human Genome Research Institute, National Institute of Child Health and Human Development, and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. Great Ormond Street Hospital/University College London, Institute of Child Health receives a proportion of funding from the Department of Health's National Institute for Health Research (NIHR) ('Biomedical Research Centres' funding).

The Whitehall II study has been supported by grants from the UK: the Medical Research Council (K013351); Economic and Social Research Council; British Heart Foundation; Health and Safety Executive; Department of Health; and from the US: the National Institute on Aging (NIA; grant no. AG13196; R01AG034454); National Heart Lung and Blood Institute (grant no. HL36310), NIH; and the John D. and Catherine T. MacArthur Foundation Research Networks on Successful Midlife Development and Socioeconomic Status and Health.

Samples from the English Longitudinal Study of Ageing (ELSA) DNA Repository (EDNAR), received support under a grant (AG1764406S1) awarded by the NIA. ELSA was developed by a team of researchers based at the National Centre for Social Research, University College London and the Institute of Fiscal Studies. The data were collected by the National Centre for Social Research.

BO: The Wellcome Trust funded R.M.M. to undertake the clinical third wave of follow-up of Boyd Orr (2002-03) as part of a research training fellowship in clinical epidemiology (grant GR063779FR).

The LBC1921 was funded by the Biotechnology and Biological Sciences Research Council (15/SAG09977), the Chief Scientist Office of the Scottish Executive Health Department (CZB/4/505, ETM/55, CZH/4/213, CZG/3/2/79), a Royal Society-Wolfson Research Merit Award to Ian Deary. The work was undertaken by The University of Edinburgh 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 (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged.

The MRC National Survey of Health and Development (NSHD) is supported by the UK Medical Research Council [MC_UU_12019/1, MC_UU12019/4]. Bona fide researchers can apply to access the NSHD data via a standard application procedure (further details available at: http://www.nshd.mrc.ac.uk/data.aspx).

The CaPS was conducted by the former MRC Epidemiology Unit (South Wales) and funded by the Medical Research Council of the United Kingdom. The Department of Social and Community Medicine, University of Bristol now maintains the archive.

HCS/HAS: The Hertfordshire studies were supported by the Medical Research Council, NIHR Southampton BRC, NIHR Musculoskeletal BRU (Oxford), Arthritis Research UK, International Osteoporosis Foundation, British Heart Foundation, and EU Framework 7 programme.

Competing interests All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/coi_disclosure.pdf and declare: TLN is an MRC-funded PhD student, IJD reports grants from the BBSRC and the Life Long Health and Wellbeing initiative, RC reports grants from the UK Medical Research Council (MC_UU12019/4) and the New Dynamics of Ageing (RES-353-25-0001), DK reports grants from the UK Medical Research Council (MC_UU12019/4) and the New Dynamics of Ageing (RES-353-25-0001), MiK reports grants from the Medical Research Council (K013351), the Economic and Social Research Council, the National Institute on Aging (R01AG034454) and the National Heart Lung and Blood Institute (R01HL36310) during the conduct of the study; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethics approval Ethical approval was obtained for all cohorts. Full details are provided in the supplement.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. Strand BH, Mishra G, Kuh D, *et al*. Smoking history and physical performance in midlife: results from the British 1946 birth cohort. *J Gerontol A Biol Sci Med Sci* 2011;66:142–9.
2. Kalmijn S, van Boxtel MP, Verschuren MW, *et al*. Cigarette smoking and alcohol consumption in relation to cognitive performance in middle age. *Am J Epidemiol* 2002;156:936–44.
3. Cadar D, Pikhart H, Mishra G, *et al*. The role of lifestyle behaviors on 20-year cognitive decline. *J Aging Res* 2012;2012:304014.
4. Corley J, Gow AJ, Starr JM, *et al*. Smoking, childhood IQ, and cognitive function in old age. *J Psychosom Res* 2012;73:132–8.
5. Galanis DJ, Petrovitch H, Launer LJ, *et al*. Smoking history in middle age and subsequent cognitive performance in elderly Japanese-American men: the Honolulu-Asia Aging Study. *Am J Epidemiol* 1997;145:507–15.
6. Elwood PC, Gallacher JE, Hopkinson CA, *et al*. Smoking, drinking, and other life style factors and cognitive function in men in the Caerphilly cohort. *J Epidemiol Community Health* 1999;53:9–14.
7. Nelson HD, Nevitt MC, Scott JC, *et al*. Smoking, alcohol, and neuromuscular and physical function of older women. *JAMA* 1994;272:1825–31.
8. Rapuri PB, Gallagher JC, Smith LM. Smoking is a risk factor for decreased physical performance in elderly women. *J Gerontol A Biol Sci Med Sci* 2007;62:93–9.
9. Cooper R, Kuh D, Hardy R, Mortality Review Group, FALCon and HALCyon Study Teams. Objectively measured physical capability levels and mortality: systematic review and meta-analysis. *BMJ* 2010;341:c4467.
10. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res* 2007;16:309–30.
11. Lawlor DA, Harbord RM, Sterne JA, *et al*. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008;27:1133–63.
12. Ware JJ, van den Bree M, Munafò MR. From men to mice: CHRNA5/CHRNA3, smoking behavior and disease. *Nicotine Tob Res* 2012;14:1291–9.
13. Kuh D, Cooper R, Hardy R, Richards M, Ben-Shlomo Y, eds. *A life course approach to healthy ageing*. Oxford: Oxford University Press, 2014.
14. Alfred T, Ben-Shlomo Y, Cooper R, *et al*. Absence of association of a single-nucleotide polymorphism in the TERT-CLPTM1L locus with age-related phenotypes in a large multicohort study: the HALCyon programme. *Aging Cell* 2011;10:520–32.
15. Gray JR, Thompson PM. Neurobiology of intelligence: science and ethics. *Nat Rev Neurosci* 2004;5:471–82.
16. Forrest KY, Zmuda JM, Cauley JA. Correlates of decline in lower extremity performance in older women: a 10-year follow-up study. *J Gerontol A Biol Sci Med Sci* 2006;61:1194–200.
17. Sabia S, Marmot M, Dufouil C, *et al*. Smoking history and cognitive function in middle age from the Whitehall II study. *Arch Intern Med* 2008;168:1165–73.
18. StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP, 2013.
19. Harris R, Bradburn M, Deeks J, *et al*. Metan: fixed- and random-effects meta-analysis. *Stata J* 2008;8:3–28.
20. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–88.
21. Burgess S, Thompson SG, CRP, CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* 2011;40:755–64.
22. Rodriguez S, Gaunt TR, Day IN. Hardy-weinberg equilibrium testing of biological ascertainment for Mendelian Randomization studies. *Am J Epidemiol* 2009;169:505–14.
23. Winterer G, Mittelstrass K, Giegling I, *et al*. Risk gene variants for nicotine dependence in the CHRNA5–CHRNA3–CHRNB4 cluster are associated with cognitive performance. *Am J Med Genet B Neuropsychiatr Genet* 2010;153B:1448–58.

24. Wang CC, Lu TH, Liao WC, *et al*. Cigarette smoking and cognitive impairment: a 10-year cohort study in Taiwan. *Arch Gerontol Geriatr* 2010;51:143–8.
25. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol* 2013;42:1497–501.
26. Lips EH, Gaborieau V, McKay JD, *et al*. Association between a 15q25 gene variant, smoking quantity and tobacco-related cancers among 17 000 individuals. *Int J Epidemiol* 2010;39:563–77.
27. Sherva R, Wilhelmsen K, Pomerleau CS, *et al*. Association of a single nucleotide polymorphism in neuronal acetylcholine receptor subunit alpha 5 (CHRNA5) with smoking status and with 'pleasurable buzz' during early experimentation with smoking. *Addiction* 2008;103:1544–52.
28. Munafò MR, Johnstone EC, Walther D, *et al*. CHRNA3 rs1051730 genotype and short-term smoking cessation. *Nicotine Tob Res* 2011;13:982–8.
29. Freathy RM, Ring SM, Shields B, *et al*. A common genetic variant in the 15q24 nicotinic acetylcholine receptor gene cluster (CHRNA5–CHRNA3–CHRNB4) is associated with a reduced ability of women to quit smoking in pregnancy. *Hum Mol Genet* 2009;18:2922–7.
30. Taylor AE, Fluharty ME, Bjørngaard JH, *et al*. Investigating the possible causal association of smoking with depression and anxiety using Mendelian randomisation meta-analysis: the CARTA consortium. *BMJ Open* 2014;4:e006141.
31. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, *et al*. Methodological challenges in Mendelian randomization. *Epidemiology* 2014;25:427–35.
32. Munafò MR, Timofeeva MN, Morris RW, *et al*. Association between genetic variants on chromosome 15q25 locus and objective measures of tobacco exposure. *J Natl Cancer Inst* 2012;104:740–8.
33. Sabia S, Elbaz A, Dugravot A, *et al*. Impact of smoking on cognitive decline in early old age: the Whitehall II cohort study. *Arch Gen Psychiatry* 2012;69:627–35.
34. Lewis SJ, Araya R, Davey Smith G, *et al*. Smoking is associated with, but does not cause, depressed mood in pregnancy—a mendelian randomization study. *PLoS ONE* 2011;6:e21689.

SUPPLEMENTARY MATERIAL

Effect of smoking on physical and cognitive capability in later life: A multi-cohort study using observational and genetic approaches

Teri-Louise North[*1], Tom M. Palmer[1,2,3], Sarah J. Lewis[1], Rachel Cooper[4], Chris Power[5], Alison Pattie[6], John M. Starr[7], Ian J. Deary[6,7], Richard M. Martin[1,8], Avan Aihie Sayer[9], Meena Kumari[10,11], Cyrus Cooper[9,12,13], Mika Kivimaki[11], Diana Kuh[4], Yoav Ben-Shlomo[1], Ian N. M. Day[1]

[1]School of Social and Community Medicine, University of Bristol, UK

[2]Division of Health Sciences, Warwick Medical School, University of Warwick, UK

[3]Department of Mathematics and Statistics, Lancaster University, Lancaster, UK

[4]MRC Unit for Lifelong Health and Ageing at UCL, London, UK

[5]Population, Policy and Practice, UCL Institute of Child Health, University College London, UK

[6]Department of Psychology, University of Edinburgh, UK

[7]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, UK

[8]University of Bristol / University Hospitals Bristol NHS Foundation Trust National Institute for Health Research Bristol Nutrition Biomedical Research Unit, University of Bristol, UK

[9]MRC Lifecourse Epidemiology Unit, University of Southampton, UK

[10]ISER, University of Essex, UK

[11]Department of Epidemiology and Public Health, UCL, London, UK

[12]National Institute for Health Research Nutrition Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, UK

[13]National Institute for Health Research Musculoskeletal Biomedical Research Unit, University of Oxford, UK

*Corresponding author:

Teri-Louise North

School of Social and Community Medicine

University of Bristol

Oakfield House

Oakfield Grove

Bristol

BS8 2BN

email: teri.north.2005@my.bristol.ac.uk

Wave of outcome assessment in this analysis

Boyd Orr (BO)(1): Physical capability measures were assessed at the third wave (2002-03).

Caerphilly Prospective Study (CAPS) : Cognitive function measures were assessed at phase III, with follow up measures to calculate decline taken from phase V. Physical capability measures were assessed at phase V.

English Longitudinal Study of Ageing (ELSA)(2): Physical and cognitive capability were assessed at wave 2 (2004/5). Follow up cognitive measures for decline calculations were taken from wave 5(2010/11).

Hertfordshire Ageing Study (HAS) (3): Cognitive function was assessed at wave 1 (1994/5) with follow up measures for decline taken from wave 2 (2003/05). Grip strength was assessed at wave 1 and all other physical capability measures at wave 2.

Hertfordshire Cohort Study (HCS) (4): Grip strength was assessed at wave 1 (1999-2004) while TUG speed, walk speed, balance ability and chair rise speed were assessed at both waves 1 (1999-2004) and 2 (2004/05) with partial overlap in some tests and no overlap in others. These latter measures were combined across waves, with priority given to wave 1, and the covariates tailored as such.

Lothian Birth Cohort 1921 (LBC1921) (5): Physical and cognitive capability were assessed at wave 1age 79 years, with follow up cognitive capability measures for decline calculations taken from wave 3 age 87 years.

National Child Development Study (NCDS) (6): Cognitive capability was assessed at the 2008 follow up when the study members were 50 years old.

MRC National Survey of Health and Development (NSHD) (7): All cognitive capability measures were taken from the 1999 wave when the study members were 53 years, with follow up measures for cognitive decline taken when the study members were 60-64 years. All physical capability measures were taken from the 1999 wave with the exception of TUG, which was analysed when the study members were 60-64 years.

Whitehall II Study (WHII) (8): Walking speed was analysed at phase 7 (2002-04), while all cognitive outcomes were analysed at phase 5 (1997-99) with follow up measures taken from phase 7.

Analyses of the genotype-covariate and genotype-smoking associations took the covariate and smoking outcomes from the earliest wave at which they were analysed in the observational associations.

Measures of physical capability

Details of the ascertainment and harmonisation of the five measures of physical capability used in analyses are described in detail elsewhere(9) and are summarised here. The approach to harmonise chair rise times (5 or 10 rises) was to calculate chair rise speed in the current study.

Grip strength was tested in ELSA, HAS, HCS, LBC AND NSHD using handheld dynamometers (the specific devices used in each study are described elsewhere (9)). The maximum measure was used in each study (extracted from 3 measures of each hand in ELSA, HAS and HCS, 3 measures of the dominant hand in LBC and 2 measures in each hand in NSHD). If repeat measures were missing the existing measures were used to derive the maximum.

Standing balance was assessed in BO, CaPS, ELSA, HAS, HCS and NSHD. Owing to the heterogeneity in the way the test was administered across cohorts, the outcome used in analyses was a derived binary variable for inability to balance on one leg with eyes open for five seconds. In ELSA the tests administered were more complex as described by Cooper *et al*(9) and we derived the outcome in the same way, namely, inability to balance in full tandem with eyes open for 5 seconds with individuals who were not progressed to the next phase of testing classed as unable. Individuals who did not complete the balance test for health reasons were classed as unable in all analyses. If tests were conducted more than once the best performance was used to derive the outcome variable.

The timed walk test was conducted in LBC (6 metres as fast as possible), HAS and HCS (3 metres at normal pace), ELSA (8 feet at normal pace with 2 trials) and WHII (8 feet at

normal pace with 3 trials). To normalise the distribution and to make a higher outcome a healthier outcome, times were converted to speeds in metres per second and then averaged where repeat trials were available.

The timed get up and go test was performed in BO, HAS, HCS, CaPS and NSHD. In all cohorts, study members had to rise from a chair, walk 3 metres at a normal pace and return to a seated position in the chair. The test was repeated in BO and CaPS. Again all times were converted to speeds in metres per second and then averaged where the trial was conducted more than once.

Timed chair rises were assessed in HAS, HCS, ELSA and NSHD. All times were converted to chair rise speed in stands per second. The cohorts measured time to complete 5 or 10 chair rises as fast as possible. In ELSA, individuals under 69 years performed 10 rises while those aged 70 and over performed 5 rises. Time to complete 5 rises was measured in both age groups and this was used to derive chair rise speed.

Some physical performance measures were conducted in part of the HCS cohort in one wave and in the remaining cohort in a later wave. To maximise sample size, measures were pooled across waves and covariates were tailored according to the wave at which the outcome had been performed.

Measures of cognitive capability

The measures of cognition across the HALCyon cohorts were categorised into measures of crystallised ability and measures of fluid cognition.

Measures of crystallised cognitive function

The National Adult Reading Test (NART)(10) was available in LBC, CaPS and NSHD. This requires study members to read aloud 50 words with irregular pronunciation and the number of words pronounced correctly is used in analyses here. NART should reflect pre-morbid IQ.

The Mill Hill vocabulary test(11) was administered in HAS and WHII. Study members had to choose the correct synonym for 33 words out of 6 multiple choice answers with increasing difficulty. The number of correct answers is used in analyses.

Measures of fluid cognitive function

Semantic fluency was tested in ELSA, NSHD, NCDS and WHII via a verbal or written test where study members were asked to name as many animals as possible in 1 minute. The number of unique animals named were used in analyses.

Verbal memory was tested in ELSA, NSHD, NCDS and WHII via a word recall test. The numbers of words correctly recalled was used in analyses. In NSHD, we summed the total score for remembering the same 15 words in writing over three consecutive trials. The sum of two trials with a delay for the second trial for remembering 10 words verbally was analysed in ELSA and NCDS. 20 words were recalled in writing in WHII.

Phonemic fluency was analysed in LBC and WHII. In LBC, study members were given three 1 minute trials to name as many words as possible beginning with F, L and C. The total

number of words is used in analyses. In WHII, study members wrote as many words as possible in 1 minute beginning with S.

Search speed was tested in ELSA (780 letters), NSHD (600 letters) and NCDS (780 letters) whereby participants must cross out particular letters in a large grid of letters. The number of letters searched per minute was used in analyses.

The Alice Heim 4-I test (AH4)(12) was available for analyses in CaPS, HAS and WHII. This involves 65 verbal and mathematical questions. The total score achieved in 10 minutes was used in analyses here.

Choice reaction time was assessed in CaPS via a computer test in which the study members had to press one of four key pads depending on which box a stimulus appeared in on screen.

Wechsler logical memory(13) was tested in LBC. The participants were asked to recall two stories immediately and following a delay for each. The total sum of the scores for each story were progressed to analysis.

Raven's Progressive Matrices(14) were used in LBC, in which study members were given 20 minutes to complete 60 multiple choice "complete the pattern" questions. The total score was used in the analysis.

Deriving a score for General Fluid Ability (Gf)

Where available, three fluid cognitive measures were included from each cohort to produce the factor. These were semantic fluency, AH4 and inverse transformed FCRT in CaPS; word recall, semantic fluency and natural log transformed search speed in NSHD and ELSA;

semantic fluency, ravens progressive matrices and logical memory in LBC; semantic fluency, word recall and AH4 in WHII.

Derivation of covariates

Disease status was defined as a binary variable. Individuals were assigned "case" status if they had a history of heart disease, stroke or diabetes. The definition of a case varied across cohorts depending on the availability of information. The numbers of cases by cohort are described in Supplementary Table 2.

Socio-economic position was classified according to the Registrar-General's Social Classes (RGSC) system and included in analyses as a categorical variable. This is based on a study member's current or most recent occupation. Individuals who did not have information were coded as missing while individuals with occupations beyond the classification system were coded in an "unclassifiable" category. As the classification differed across cohorts slightly, to assess the association between genotype and socio-economic position to test the Mendelian Randomization assumptions, individuals who had a valid RGSC coding (I-V) were binarised into Professional & Managerial or other.

Individuals aged 90 years or over are not assigned an exact age in ELSA data releases. As such, we estimated the age of these individuals using a representative estimate of the mean age of individuals aged 90 and over in England and Wales in 2005 (the year of wave 2 assessment). To calculate this estimate, we used the England and Wales Mid-Year Population Estimates of the Very Elderly, 2002-2010, demographic table "Mid-2010 Estimates of the very elderly (including centenarians) England and Wales; estimated resident population" which was part of the Population Estimates of the Very Elderly, 2010 Office for National Statistics release (release date 29 September 2011, date accessed 5 February 2014 from http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-223697). The estimated age used in analyses was 92.62 years.

Genotyping

The rs16969968 SNP was genotyped by KBioscience (http://www.lgcgenomics.com/) in CaPS, BO, HCS, HAS, ELSA and WHII. In LBC, the genotype data came from the Illumina 610-Quadv1 array (rs1051730) and in NCDS from a combination of the T1DGC(15) array (rs1051730) and the WTCCC2 array (rs16969968, Illumina 1.2M chip). The rs16969968 and the rs1051730 SNPs are used interchangeably and thus genotypes from these SNPs were pooled for the NCDS analysis. Genotype data for rs16969968 in NSHD came from a previous genotyping performed by KBioscience. For SNPs genotyped by HALCyon, call rate, clustering and duplicate concordance were examined where possible. Departures from HWE and MAF were examined in all cohorts. The whole analysis was restricted to unrelated individuals of European ancestry where this information was available. Further information on the genotyping quality is provided in Supplementary Table 1.

Supplementary Table 1: Genotype Quality

Cohort	SNP	Call Rate (%)	Duplicate Concordance Rate (%)
BO	rs16969968	97.50	97.87
CaPS	rs16969968	98.04	100
ELSA	rs16969968	99.38	99.76
HAS	rs16969968	94.79	Not available
HCS	rs16969968	97.06	Not available
LBC	rs1051730	NA	NA
NCDSa*	rs16969968	NA	NA
NCDSb*	rs1051730	NA	NA
NSHD	rs16969968	NA	NA
WHII	rs16969968	98.48	97.12

Notes:
*In NCDS we analysed a combination of the rs16969968 and the rs1051730 genotype depending on availability. As these SNPs are used interchangeably, each individual was classified according to their number of minor alleles of either SNP

Supplementary Table 2: Disease status by cohort

Cohort	% Individuals classed as a case
BO	29.14
CaPS	42.25
ELSA	19.94
HAS	23.75
HCS	15.33
LBC	35.51
NCDS	2.66
NSHD	8.14
WHII	10.31

Notes:
Based on individuals included in the observational analysis using the earliest wave analysed to calculate %

Supplementary Table 3: Genotype frequencies by cohort

Cohort	C/C	C/T	T/T	HWE p-value**	Total
BO	214	198	42	0.69	454
CaPS	593	569	125	0.50	1287
ELSA	2494	2404	589	0.79	5487
HAS	214	235	68	0.78	517
HCS	1191	1211	310	0.93	2712
LBC*	231	239	43	0.08	513
NCDS*	2099	2133	566	0.50	4798
NSHD	1222	1164	255	0.36	2641
WHII	1747	1750	423	0.62	3920
TOTAL	10005	9903	2421	0.73	22329

Notes:
Numbers based on all individuals with age, sex, smoking status and genotype. The earliest cohort phase or wave utilised in the analyses was used to extract data, and to perform the meta-analysis in Figures 1-4 (main paper)
*Cohort analysis uses rs1051730 as a proxy. NCDS analysis combines rs16969968 with rs1051730. Results represent pooled SNPs here.
**Based on chi-squared test with 1 degree of freedom

Supplementary Table 4: Observational sample size by outcome and cohort (Model M1)

	BO	CaPS	ELSA	HAS	HCS	LBC	NCDS	NSHD	WHII
Grip strength	0	0	5374(587)	635(87)	2803(239)	536(35)	0	2831(645)	0
Chair rise speed	0	0	4654(490)	244(16)	1516(105)	0	0	2721(608)	0
Walk speed	0	0	3442(312)	260(16)	2176(167)	534(35)	0	0	5319(278)
TUG speed	279(33)	999(123)	0	263(16)	2180(168)	0	0	1849(189)	0
Inability to balance on one leg for 5s	279(33)	1002(124)	5385(589)	265(17)	1569(106)	0	0	2860(648)	0
Mill Hill	0	0	0	633(86)	0	0	0	0	4563(354)
NART	0	1804(479)	0	0	0	540(36)	0	2812(646)	0
Gf	0	1698(450)	5306(582)	0	0	531(35)	0	2856(653)	4486(348)
AH4	0	1798(482)	0	621(85)	0	0	0	0	4552(357)
Semantic fluency	0	1820(485)	5418(594)	0	0	0	7652(1570)	2928(674)	4537(353)
Phonemic fluency	0	0	0	0	0	538(35)	0	0	4543(353)
Search speed	0	0	5306(582)	0	0	0	7521(1538)	2918(672)	0
Word recall	0	0	5411(591)	0	0	0	7600(1565)	2866(655)	4541(355)
Four choice reaction time	0	1722(459)	0	0	0	0	0	0	0
Logical memory	0	0	0	0	0	540(36)	0	0	0
Raven's Progressive Matrices	0	0	0	0	0	536(36)	0	0	0
Mill Hill decline	0	0	0	228(20)	0	0	0	0	4510(350)
NART decline	0	1021(218)	0	0	0	202(8)	0	0	0
AH4 decline	0	1003(211)	0	254(20)	0	0	0	0	4505(353)
Semantic fluency decline	0	1044(224)	3713(380)	0	0	0	0	0	4484(349)
Phonemic fluency decline	0	0	0	0	0	203(7)	0	0	4478(349)
Search speed decline	0	0	0	0	0	0	0	2055(383)	0
Word recall decline	0	0	3714(380)	0	0	0	0	2005(372)	4482(351)
Four choice reaction time decline	0	960(202)	0	0	0	0	0	0	0
Logical memory decline	0	0	0	0	0	204(8)	0	0	0
Raven's Progressive Matrices decline	0	0	0	0	0	198(8)	0	0	0

Notes: Numbers based on all individuals with age, sex, socio-economic position and the outcome measure. Numbers with smoking status are provided with numbers with CPD in brackets. Smoking status was analysed as binary variables in separate analyses comparing classes of smoker. Some regressions were not possible due to small sample size, and on some occasions individuals were removed from the analysis automatically by the software

Supplementary Table 5: Genetic sample size by outcome and cohort

	BO	CaPS	ELSA	HAS	HCS	LBC	NCDS	NSHD	WHII
Grip strength	0	0	764(587)/2647/1963	85(74)/233/143	324(224)/1090/1234	32(32)/252/223	0	593(588)/1200/727	0
Chair rise speed	0	0	634(490)/2279/1741	18(15)/101/54	126(96)/560/736	0	0	558(554)/1154/712	0
Walk speed	0	0	412(312)/1774/1256	19(15)/107/57	207(154)/828/1019	32(32)/251/222	0	0	495(214)/1877/1854
TUG speed	40(30)/108/106	129(91)/409/144	0	19(15)/109/58	208(155)/829/1021	0	0	171(167)/914/517	0
Inability to balance on one leg for 5s	40(30)/108/106	131(93)/411/146	768(589)/2643/1974	19(15)/109/59	127(97)/579/764	0	0	596(591)/1212/737	0
Mill Hill	0	0	0	84(73)/232/143	0	0	0	0	512(279)/1427/1656
NART	0	406(314)/571/203	0	0	0	33(33)/253/223	0	591(586)/1191/719	0
Gf	0	386(297)/542/191	759(582)/2617/1930	0	0	32(32)/248/222	0	599(594)/1206/738	506(273)/1406/1622
AH4	0	407(314)/569/201	0	84(73)/225/143	0	0	0	0	515(282)/1423/1647
Semantic fluency	0	410(317)/576/203	775(594)/2665/1978	0	0	0	960(957)/1516/2179	616(611)/1235/750	511(278)/1416/1648
Phonemic fluency	0	0	0	0	0	32(32)/251/224	0	0	511(278)/1419/1650
Search speed	0	0	759(582)/2617/1930	0	0	0	938(935)/1488/2140	614(609)/1232/748	0
Word recall	0	0	771(591)/2664/1976	0	0	0	959(956)/1505/2162	601(596)/1209/742	513(280)/1426/1641
Four choice reaction time	0	391(302)/550/195	0	0	0	0	0	0	0
Logical memory	0	0	0	0	0	33(33)/252/224	0	0	0
Raven's Progressive Matrices	0	0	0	0	0	33(33)/252/222	0	0	0
Mill Hill decline	0	0	0	19(15)/88/55	0	0	0	0	506(276)/1417/1633
NART decline	0	228(168)/359/151	0	0	0	6(6)/92/94	0	0	0
AH4 decline	0	217(161)/360/147	0	19(15)/99/62	0	0	0	0	509(279)/1413/1630
Semantic fluency decline	0	232(172)/369/153	492(380)/1785/1436	0	0	0	0	0	505(275)/1405/1625
Phonemic fluency decline	0	0	0	0	0	5(5)/92/96	0	0	505(275)/1405/1619

Search speed decline	0	0	0	0	0	0	0	360(356)/930/561	0
Word recall decline	0	0	492(380)/1786/1436	0	0	0	0	350(346)/907/550	506(276)/1412/1618
Four choice reaction time decline	0	213(158)/345/142	0	0	0	0	0	0	0
Logical memory decline	0	0	0	0	0	6(6)/92/96	0	0	0
Raven's Progressive Matrices decline	0	0	0	0	0	6(6)/90/92	0	0	0

Notes: Numbers based on all individuals with age, sex, genotype, smoking status and the outcome measure, and are restricted to those also in the observational analysis. Numbers provided are current smokers (number with CPD)/ex smokers/never smokers. Numbers reflect the data progressed to regression analyses, irrespective of whether the regression was possible.

Supplementary Table 6: Observational estimates for the associations between smoking and cognitive capabilities - single cohort analyses

Outcome	Cohort	Modelψ	Cigarette per day β (95% CI)	Current vs. ex smoker β (95% CI)	Current vs. never smoker β (95% CI)	Ever vs. never smoker β (95% CI)
Four choice reaction time#	CaPS	M1	0.008(-0.002,0.017)	-0.159**(-0.260,-0.058)	-0.050(-0.185,0.085)	0.044(-0.074,0.162)
		M2	0.007(-0.004,0.017)	-0.203***(-0.310,-0.096)	-0.063(-0.210,0.083)	0.061(-0.069,0.190)
Logical memory	LBC	M1	0.023(-0.049,0.095)	0.138(-0.220,0.496)	0.192(-0.161,0.545)	0.031(-0.140,0.202)
		M2	0.024(-0.049,0.098)	0.123(-0.240,0.485)	0.185(-0.170,0.540)	0.037(-0.137,0.210)
Raven's progressive matrices	LBC	M1	-0.015(-0.081,0.051)	-0.160(-0.501,0.180)	-0.094(-0.426,0.237)	0.009(-0.155,0.173)
		M2	-0.011(-0.077,0.056)	-0.163(-0.507,0.181)	-0.086(-0.420,0.248)	0.009(-0.157,0.175)

Notes:
ψModels: (M1) age, sex and SEP adjusted, (M2) M1 + disease adjusted
#Inverse transformed
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

Supplementary Table 7: Observational estimates for the associations between smoking and measures of cognitive decline.

	Outcome	Modelψ	Cigarette per day OR (95% CI)	Current vs. ex smoker OR (95% CI)	Current vs. never smoker OR (95% CI)	Ever vs. never smoker OR (95% CI)
Crystallized	Mill Hill	M1	1.003(0.980,1.027)	1.002(0.691,1.452)	1.121(0.915,1.375)	1.099(0.961,1.258)
		M2	1.004(0.980,1.028)	0.997(0.666,1.491)	1.117(0.911,1.371)	1.100(0.960,1.259)
	NART	M1	0.993(0.961,1.027)##	1.065(0.786,1.443)	1.134(0.777,1.655)	1.008(0.752,1.353)
		M2	0.998(0.962,1.036)##	1.030(0.742,1.430)	1.183(0.781,1.792)	1.053(0.765,1.451)
Fluid Measures	AH4	M1	0.998(0.979,1.017)	1.265*(1.047,1.527)	1.271(0.746,2.165)	1.058(0.778,1.439)
		M2	0.996(0.977,1.016)	1.283**(1.076,1.531)	1.370(0.806,2.328)	1.120(0.870,1.442)
	Semantic fluency	M1	1.019(0.992,1.046)	1.090(0.912,1.302)	1.167(0.883,1.544)	1.068(0.939,1.215)
		M2	1.023(0.992,1.056)	1.092(0.924,1.290)	1.188(0.898,1.573)	1.071(0.937,1.225)
	Phonemic fluency	M1	1.003(0.979,1.028)#	1.042(0.843,1.286)	0.990(0.805,1.218)	0.973(0.851,1.112)
		M2	1.002(0.978,1.027)#	1.038(0.840,1.283)	0.993(0.808,1.222)	0.984(0.860,1.125)
	Word recall	M1	0.997(0.969,1.026)	1.233**(1.077,1.412)	1.219**(1.061,1.401)	1.020(0.898,1.157)
		M2	0.997(0.969,1.025)	1.230**(1.074,1.409)	1.213**(1.055,1.395)	1.021(0.893,1.167)

Notes:
ψModels: (M1) age, sex and SEP adjusted, (M2) M1 + disease adjusted
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001
#analysis in WHII only
##analysis in CAPS only

Supplementary Table 8: Observational estimates for the associations between smoking and measures of cognitive decline - single cohort analyses

Outcome	Cohort	Modelψ	Cigarette per day OR (95% CI)	Current vs. ex smoker OR (95% CI)	Current vs. never smoker OR (95% CI)	Ever vs. never smoker OR (95% CI)
Search speed	NSHD	M1	1.001(0.978,1.025)	1.364*(1.043,1.784)	1.170(0.871,1.571)	0.934(0.751,1.161)
		M2	1.005(0.982,1.029)	1.359*(1.039,1.777)	1.174(0.874,1.577)	0.938(0.754,1.166)
Four choice reaction timeγ	CaPS	M1	1.021(0.990,1.054)	1.825***(1.295,2.572)	1.742*(1.128,2.690)	1.125(0.774,1.635)
		M2	1.017(0.981,1.054)	1.854**(1.281,2.684)	1.742*(1.082,2.803)	1.111(0.737,1.676)
Logical memory	LBC	M1	#	0.427(0.046,3.982)	0.409(0.046,3.650)	1.118(0.576,2.172)
		M2	#	0.419(0.044,4.026)	0.393(0.043,3.588)	1.236(0.624,2.448)
Raven's progressive matrices	LBC	M1	#	1.325(0.285,6.165)	2.668(0.545,13.052)	2.128*(1.074,4.219)
		M2	#	1.187(0.250,5.644)	2.990(0.588,15.191)	2.322*(1.153,4.677)

Notes:
ψModels: (M1) age, sex and SEP adjusted, (M2) M1 + disease adjusted
#Analysis not possible due to small sample size
γDecline in four choice reaction time was calculated using the bottom 25% decliners so that the outcome represents decline in cognitive ability over time
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

Supplementary Table 9: Association between rs16969968 and covariates

Covariate	Effect (95% CI)	I^2 (p-value$^{\tau}$)
Age (years)	-0.000(-0.010,0.010)	0.00(0.455)
BMI	-0.019(-0.139,0.101)	20.82(0.264)
Height (cm)	-0.003(-0.203,0.197)	0.00(0.957)
Sex	1.016(0.974,1.059)	0.00(0.600)
Disease status	1.029(0.971,1.092)	0.00(0.935)
SEP (binary RGSC codes I or II versus RGSC codes III-V)	1.031(0.987,1.077)	0.00(0.984)

Effect sizes are regression coefficients for continuous outcome measures and odds ratios for binary measures.
τ: I^2 is the percentage of the variation across studies that is due to heterogeneity rather than chance(16). The p-value is for Cochran's Q statistic.

Supplementary Table 10: Associations between rs16969968 and outcomes stratified by smoking status - single cohort analyses

Outcome Category	Outcome	Cohort	Current smokers β (95% CI)	Never smokers β (95% CI)	Ever smokers β (95% CI)
Cognitive capability	Four choice reaction time##	CaPS	0.053(-0.097,0.204)	-0.068(-0.283,0.148)	0.031(-0.065,0.127)
	Logical memory	LBC	0.240(-0.547,1.027)	0.095(-0.116,0.305)	0.125(-0.062,0.313)
	Raven's progressive matrices	LBC	0.375(-0.360,1.111)	-0.124(-0.328,0.080)	-0.010(-0.193,0.174)
			Current smokers OR (95% CI)	Never smokers OR (95% CI)	Ever smokers OR (95% CI)
Cognitive capability decline	Search speed	NSHD	1.029(0.712,1.486)	0.997(0.747,1.330)	1.003(0.824,1.222)
	Four choice reaction time γ	CaPS	1.736*(1.131,2.664)	1.008(0.528,1.924)	1.263(0.945,1.690)
	Logical memory	LBC	#	1.541(0.720,3.297)	1.630(0.759,3.500)
	Raven's progressive matrices	LBC	#	1.057(0.425,2.630)	1.401(0.696,2.821)

Notes:
#Analysis not possible due to small sample size
##Inverse transformed
γDecline in four choice reaction time was calculated using the bottom 25% decliners so that the outcome represents decline in cognitive ability over time
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001
Models adjusted for age and sex.

Supplementary Table 11: Comparison of observational with instrumental variable estimates - single cohorts analyses

Observational association of interest			Observational estimate, β_O (95% CI)	IV estimate, β_{IV} (95% CI)
Smoking behaviour	Outcome	Cohort		
Current vs ex smoker	Four choice reaction time#	CaPS	-0.159**(-0.260,-0.058)	-11.376(-212.414,189.661)

Notes: IV: instrumental variable;
As explained in the methods, sample used to calculate observed and IV estimates differs according to the availability of variables
#Inverse transformed
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

Heterogeneity statistics for meta-analysed estimates

Supplementary Table 12: Heterogeneity statistics for the observational estimates for the associations between smoking and physical capabilities (Table 2, main paper)

Outcome	Model$^{\psi}$	I^2 (p-value$^{\tau}$)			
		Cigarette per day$^{\delta}$	Current vs. ex smoker	Current vs. never smoker	Ever vs. never smoker
Grip strength	M1	0.00(0.576)	1.27(0.399)	57.83(0.050)	60.15(0.040)
	M2	0.00(0.710)	0.00(0.584)	54.30(0.068)	54.34(0.067)
	M3	0.00(0.684)	0.00(0.871)	34.55(0.191)	53.93(0.070)
Chair rise speed	M1	0.00(0.766)	30.01(0.232)	30.07(0.232)	0.00(0.403)
	M2	0.00(0.997)	22.13(0.278)	38.10(0.183)	18.51(0.298)
	M3	0.00(0.757)	20.50(0.287)	45.16(0.140)	5.02(0.368)
Walk speed	M1	37.14(0.174)	62.67(0.030)	36.02(0.181)	0.00(0.908)
	M2	17.97(0.300)	59.21(0.044)	29.22(0.227)	0.00(0.868)
	M3	56.98(0.054)	66.38(0.018)	57.23(0.053)	0.00(0.764)
TUG speed	M1	12.17(0.336)	6.29(0.371)	0.00(0.901)	57.69(0.051)
	M2	9.21(0.354)	0.00(0.740)	0.00(0.931)	52.83(0.076)
	M3	47.16(0.109)	0.00(0.791)	0.00(0.994)	29.08(0.228)
Inability to balance on one leg for 5s	M1	0.00(0.683)	0.00(0.672)	19.26(0.288)	0.00(0.709)
	M2	0.00(0.673)	0.00(0.833	0.00(0.423)	0.00(0.730)
	M3	0.00(0.778)	0.00(0.836)	13.18(0.330)	0.00(0.611)

Notes:
ψModels: (M1) age, sex and SEP adjusted, (M2) M1 + disease adjusted, (M3) M2 + height, BMI adjusted
δAssociation is for 1 CPD for comparison with genotypic analysis
τ I^2 is the percentage of the variation across studies that is due to heterogeneity rather than chance(16). The p-value is for Cochran's Q statistic.

Table 13: Heterogeneity statistics for the observational estimates for the associations between smoking and cognitive capabilities (Table 3, main paper)

	Outcome	Modelψ	I^2 (p-value$^{\tau}$)			
			Cigarette per day	Current vs. ex smoker	Current vs. never smoker	Ever vs. never smoker
Crystallized measures	Mill Hill	M1	0.00(0.871)	0.00(0.737)	0.00(0.686)	0.00(0.355)
		M2	0.00(0.840)	0.00(0.871)	0.00(0.555)	0.00(0.323)
	NART	M1	0.00(0.942)	38.31(0.198)	35.77(0.211)	0.00(0.897)
		M2	0.00(0.951)	13.95(0.313)	45.40(0.160)	0.00(0.856)
Fluid Measures	Gf	M1	72.99(0.005)	23.31(0.266)	34.84(0.189)	0.00(0.721)
		M2	70.30(0.009)	25.15(0.254)	40.81(0.149)	0.00(0.492)
	AH4	M1	40.59(0.186)	0.00(0.718)	0.00(0.562)	0.00(0.883)
		M2	45.68(0.159)	0.00(0.583)	0.22(0.367)	0.00(0.650)
	Semantic fluency	M1	55.00(0.064)	0.00(0.696)	55.18(0.063)	51.82(0.081)
		M2	52.97(0.075)	0.00(0.766)	57.00(0.054)	52.01(0.080)
	Phonemic fluency	M1	0.00(0.349)	71.96(0.059)	72.86(0.055)	0.00(0.954)
		M2	0.00(0.345)	61.60(0.107)	76.07(0.041)	0.00(0.688)
	Search speed#	M1	0.00(0.432)	52.72(0.121)	74.32(0.020)	32.00(0.230)
		M2	0.00(0.690)	58.19(0.091)	73.57(0.023)	22.52(0.275)
	Word recall	M1	43.43(0.151)	75.93(0.006)	49.89(0.112)	0.00(0.901)
		M2	40.73(0.167)	76.49(0.005)	52.38(0.098)	0.00(0.793)

Notes:
ψModels: (M1) age, sex and SEP adjusted, (M2) M1 + disease adjusted
#Natural log transformed
τ I^2 is the percentage of the variation across studies that is due to heterogeneity rather than chance(16). The p-value is for Cochran's Q statistic.

Table 14: Heterogeneity statistics for the observational estimates for the associations between smoking and measures of cognitive decline (Supplementary Table 7)

	Outcome	Modelψ	I^2 (p-value$^{\tau}$)			
			Cigarette per day	Current vs. ex smoker	Current vs. never smoker	Ever vs. never smoker
Crystallized	Mill Hill	M1	0.00(0.899)	14.36(0.280)	0.00(0.537)	0.00(0.790)
		M2	0.00(0.738)	17.00(0.272)	0.00(0.417)	0.00(0.758)
	NART	M1	##	0.00(0.523)	0.00(0.857)	0.00(0.484)
		M2	##	0.00(0.613)	0.00(0.692)	0.00(0.537)
Fluid Measures	AH4	M1	0.00(0.693)	6.08(0.345)	73.87(0.022)	59.08(0.087)
		M2	0.00(0.586)	0.00(0.379)	69.77(0.037)	38.88(0.195)
	Semantic fluency	M1	62.52(0.069)	33.15(0.224)	67.20(0.047)	31.28(0.233)
		M2	69.96(0.036)	22.29(0.276)	65.30(0.056)	33.50(0.222)
	Phonemic fluency	M1	#	0.00(0.741)	0.00(0.938)	0.00(0.775)
		M2	#	0.00(0.950)	0.00(0.871)	0.00(0.841)
	Word recall	M1	74.55(0.020)	0.00(0.824)	0.00(0.470)	41.37(0.182)
		M2	73.47(0.023)	0.00(0.857)	0.00(0.478)	46.72(0.153)

Notes:
ψModels: (M1) age, sex and SEP adjusted, (M2) M1 + disease adjusted
#analysis in WHII only
##analysis in CAPS only
τ I^2 is the percentage of the variation across studies that is due to heterogeneity rather than chance(16). The p-value is for Cochran's Q statistic.

Table 15: Heterogeneity statistics for the associations between rs16969968 and outcomes stratified by smoking status (Table 4, main paper)

Outcome Category	Outcome	I^2 (p-value$^{\tau}$)		
		Current smokers	Never smokers	Ever smokers
Physical capability	Grip strength	20.54(0.284)	55.79(0.060)	0.00(0.999)
	Chair rise speed	47.63(0.126)	49.81(0.113)	18.21(0.300)
	Walk speed	64.53(0.024)	0.00(0.623)	0.00(0.643)
	TUG speed	0.00(0.455)	0.00(0.681)	0.00(0.496)
Cognitive capability	Mill Hill	0.00(0.684)	0.00(0.619)	73.28(0.053)
	NART	0.00(0.885)	1.34(0.363)	0.00(0.484)
	Gf	0.00(0.803)	0.00(0.897)	0.00(0.702)
	AH4	0.00(0.784)	0.00(0.561)	0.00(0.511)
	Semantic fluency	40.45(0.152)	0.00(0.832)	0.00(0.551)
	Phonemic fluency	0.00(0.744)	0.00(0.739)	86.59(0.006)
	Search speed#	23.73(0.270)	44.80(0.163)	0.00(0.941)
	Word recall	0.00(0.509)	26.06(0.255)	0.00(0.813)
Physical capability	Inability to balance on one leg for 5s	26.86(0.233)	0.00(0.556)	0.00(0.987)
Cognitive capability decline	Mill Hill	33.46(0.220)	0.00(0.913)	0.00(0.603)
	NART	ψψ	0.00(0.674)	0.00(0.681)
	AH4	0.00(0.436)	0.00(0.444)	0.00(0.557)
	Semantic fluency	44.00(0.168)	75.69(0.016)	0.00(0.771)
	Phonemic fluency	ψ	22.18(0.257)	71.80(0.060)
	Word recall	0.00(0.819)	0.00(0.486)	0.00(0.708)

Notes: #Natural log transformed
ψanalysis in WHII only
ψψanalysis in CAPS only
Models adjusted for age and sex.
τ I^2 is the percentage of the variation across studies that is due to heterogeneity rather than chance(16). The p-value is for Cochran's Q statistic.

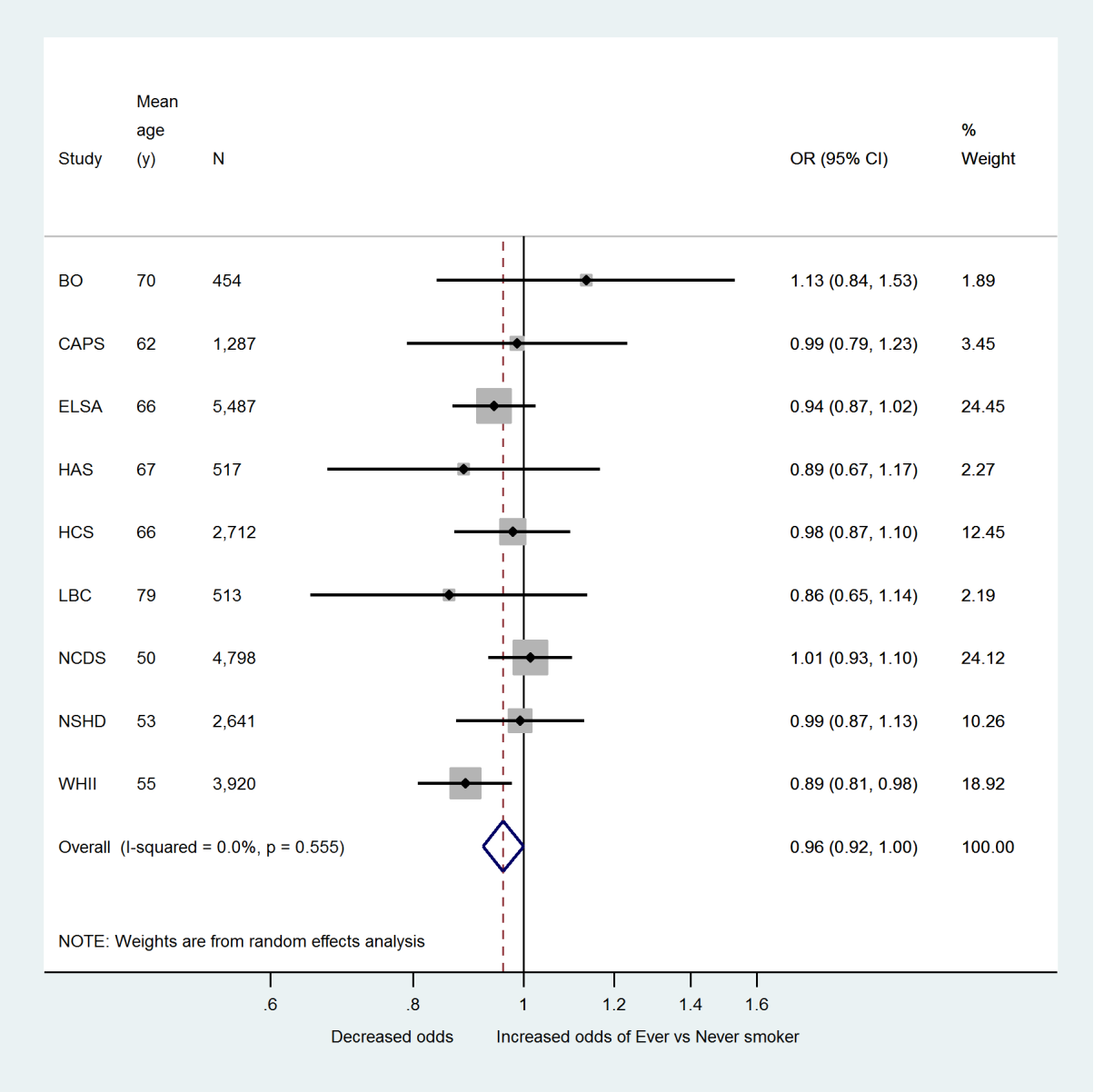
Table 16: Heterogeneity statistics for the IV estimates (Table 5, main paper)

Observational association of interest		I^2 (p-value$^{\tau}$)
Smoking behaviour	Outcome	
Cigarettes per day	Search speed#	27.95(0.250)
Current vs ex smoker	Grip strength	0.00(0.988)
	Walk speed	0.00(0.959)
	TUG speed	0.00(0.969)
	Chair rise speed	0.00(0.895)
	NART	0.00(0.818)
	Mill Hill	0.00(0.887)
	Gf	0.00(0.974)
	Semantic fluency	0.00(0.820)
	AH4	0.00(0.979)
	Word recall	0.00(0.973)
	Search speed#	0.00(0.984)

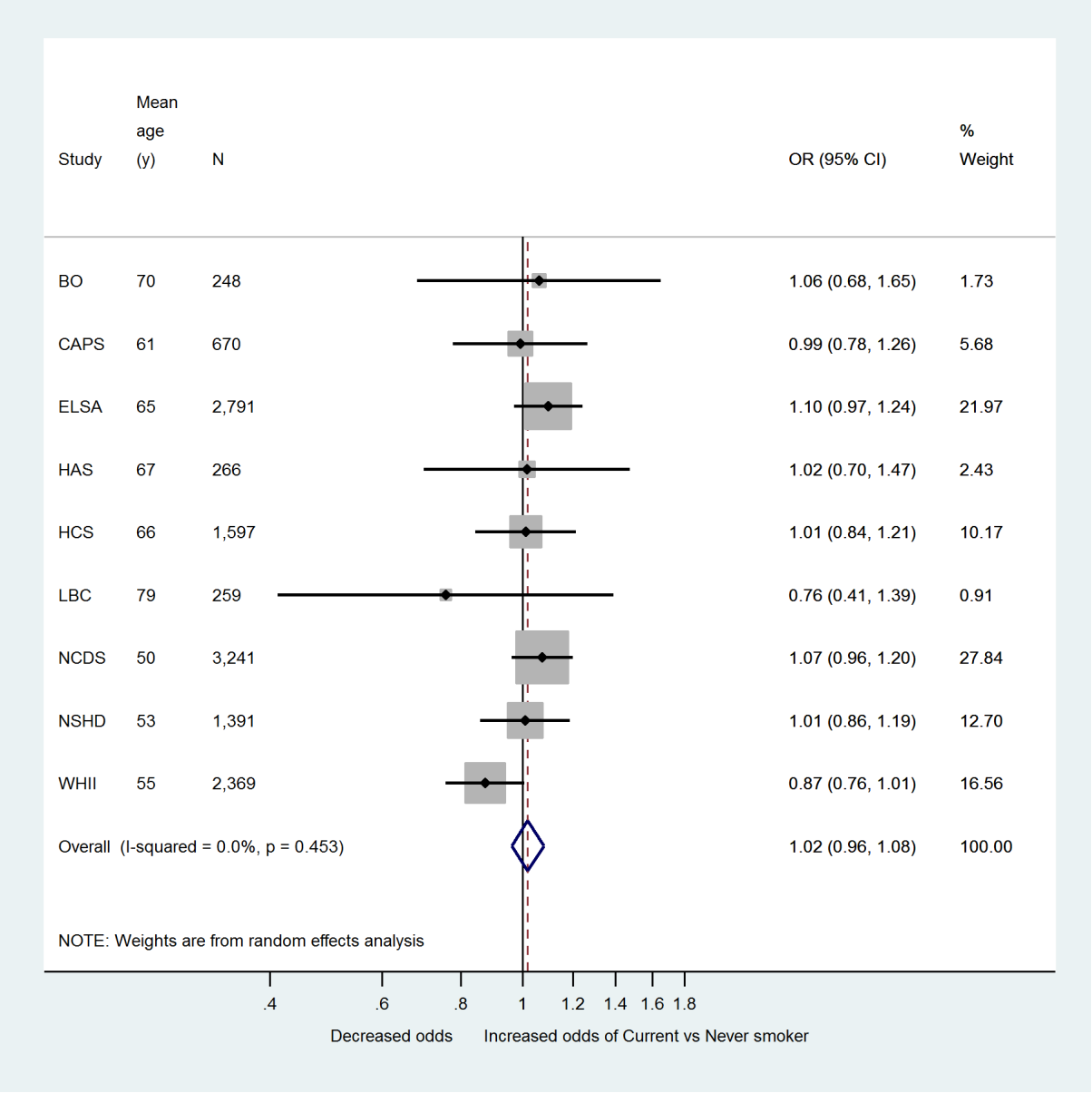
Notes:
#Natural log transformed
τ I^2 is the percentage of the variation across studies that is due to heterogeneity rather than chance(16). The p-value is for Cochran's Q statistic.

Supplementary Figures

Supplementary Figure 1: Meta-analysis of minor allele - ever versus never smoker association



Supplementary Figure 2: Meta-analysis of minor allele - current versus never smoker association



Cohort and other funding information

Cohorts

1958BC (NCDS) WTCCC: DNA collection was funded by MRC grant G0000934 and cell-line creation by Wellcome Trust grant 068545/Z/02. This study makes use of data generated by the Wellcome Trust Case-Control Consortium. A full list of investigators who contributed to generation of the data is available from the Wellcome Trust Case-Control Consortium website. Funding for the project was provided by the Wellcome Trust under the award 076113. Great Ormond Street Hospital/University College London, Institute of Child Health receives a proportion of funding from the Department of Health's National Institute for Health Research (NIHR) ('Biomedical Research Centres' funding).

1958BC (NCDS) T1DGC: DNA collection was funded by MRC grant G0000934 and cell-line creation by Wellcome Trust grant 068545/Z/02. This research used resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases, National Human Genome Research Institute, National Institute of Child Health and Human Development, and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. Great Ormond Street Hospital/University College London, Institute of Child Health receives a proportion of funding from the Department of Health's National Institute for Health Research (NIHR) ('Biomedical Research Centres' funding).

The Whitehall II study has been supported by grants from the
UK: the Medical Research Council (K013351); Economic and Social Research Council; British Heart Foundation; Health and Safety Executive; Department of Health; and from the US: the National Institute on Aging (NIA; grant no. AG13196; R01AG034454); National

Heart Lung and Blood Institute (grant no. HL36310), NIH; and the John D. and Catherine T. MacArthur Foundation Research Networks on Successful Midlife Development and Socio-economic Status and Health.

Samples from the English Longitudinal Study of Ageing (ELSA) DNA Repository (EDNAR), received support under a grant (AG1764406S1) awarded by the NIA. ELSA was developed by a team of researchers based at the National Centre for Social Research, University College London and the Institute of Fiscal Studies. The data were collected by the National Centre for Social Research.

BO: The Wellcome Trust funded R.M.M. to undertake the clinical third wave of follow-up of Boyd Orr (2002-03) as part of a research training fellowship in clinical epidemiology (grant GR063779FR).

The LBC1921 was funded by the Biotechnology and Biological Sciences Research Council (15/SAG09977), the Chief Scientist Office of the Scottish Executive Health Department (CZB/4/505, ETM/55, CZH/4/213, CZG/3/2/79), a Royal Society-Wolfson Research Merit Award to Ian Deary. The work was undertaken by The University of Edinburgh 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 (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged.

The MRC National Survey of Health and Development (NSHD) is supported by the UK Medical Research Council [MC_UU_12019/1, MC_UU12019/4]. Bona fide researchers can

apply to access the NSHD data via a standard application procedure (further details available at: http://www.nshd.mrc.ac.uk/data.aspx).

The CaPS was conducted by the former MRC Epidemiology Unit (South Wales) and funded by the Medical Research Council of the United Kingdom. The Department of Social and Community Medicine, University of Bristol now maintains the archive.

HCS/HAS: The Hertfordshire studies were supported by the Medical Research Council, NIHR Southampton BRC, NIHR Musculoskeletal BRU (Oxford), Arthritis Research UK, International Osteoporosis Foundation, British Heart Foundation, and EU Framework 7 programme.

Researchers

TLN is the recipient of an MRC PhD studentship in the Bristol Centre for Systems Biomedicine (BCSBmed) MRC doctoral training centre.

RC and DK are funded by the Medical Research Council [MC_UU_12019/1,MC_UU_12019/4].

Mika Kivimaki is supported by an Economic and Social Research Council professorial fellowship.

The NIHR Bristol Nutrition Biomedical Research Unit is funded by the National Institute for Health Research (NIHR) and is a partnership between the University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The funders had no role in the conduct or reporting of the study.

Ian Deary and John Starr are supported by the Centre for Cognitive Ageing and Cognitive Epidemiology, which is funded by the Biotechnology Sciences Research Council and the Medical Research Council and the University of Edinburgh as part of the cross council Lifelong Health and Wellbeing initiative (MR/K026992/1).

Cohort Ethical Approval

NCDS: Ethical approval for the 1958 birth cohort 45y survey (when DNA was collected) was obtained from South East Multi-centre Research Ethics Committee (ref. 01/1/44) and the Joint UCL/UCLH Committees on the Ethics of Human Research (Committee A) Ref: 08/H0714/40.

NSHD: Ethical approval for the NSHD data collection at 53 years was approved by the North Thames Multi-Centre Research Ethics Committee (ref. MREC 98/1/121). At 60–64 years ethical approval was obtained from the Central Manchester Local Research Ethics Committee (ref. 07/H1008/245) and the Scotland A Research Ethics Committee (ref. 08/MRE00/12). Written informed consent was obtained from study members at each stage of data collection.

English Longitudinal Study of Ageing (ELSA): ELSA has been approved by the National Research Ethics Service and all participants have given informed consent.

Whitehall II: All participants provided written consent and the University College London ethics committee approved the study.

BO: Ethical approval for the clinical third wave of follow-up of Boyd Orr (2002-03) was obtained from Multi-centre Research Ethics Committee Scotland. All participants gave informed consent.

LBC1921: Ethical approval for the Lothian Birth Cohort 1921 study was given by the Lothian Research Ethics Committee.

CAPS: Ethical approval for genotypic analyses was provided by South East Wales Local Research Ethics Committee Panel B (05/WSE02/131). The original CaPS project received ethical approval from the former South Glamorgan Area Health Authority.

HCS/HAS: Ethical approval for the Hertfordshire studies was obtained from the Hertfordshire Local Research Ethics Committee.

Supplementary References

1. Martin RM, Gunnell D, Pemberton J, Frankel S, Davey Smith G. Cohort Profile: The Boyd Orr cohort—an historical cohort study based on the 65 year follow-up of the Carnegie Survey of Diet and Health (1937–39). International Journal of Epidemiology. 2005;**34**:742-749.
2. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort Profile: The English Longitudinal Study of Ageing. International Journal of Epidemiology. 2013;**42**:1640-1648.
3. Syddall HE, Simmonds SJ, Martin HJ, Watson C, Dennison EM, Cooper C*, et al.* Cohort profile: The Hertfordshire Ageing Study (HAS). International Journal of Epidemiology. 2010;**39**:36-43.
4. Syddall H, Aihie Sayer A, Dennison E, Martin H, Barker D, Cooper C*, et al.* Cohort Profile: The Hertfordshire Cohort Study. International Journal of Epidemiology. 2005;**34**:1234-1242.
5. Deary IJ, Gow AJ, Pattie A, Starr JM. Cohort Profile: The Lothian Birth Cohorts of 1921 and 1936. International Journal of Epidemiology. 2012;**41**:1576-1584.
6. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). International Journal of Epidemiology. 2006;**35**:34-41.
7. Kuh D, Pierce M, Adams J, Deanfield J, Ekelund U, Friberg P*, et al.* Cohort Profile: Updating the cohort profile for the MRC National Survey of Health and Development: a new clinic-based data collection for ageing research. International Journal of Epidemiology. 2011;**40**:e1-e9.
8. Marmot M, Brunner E. Cohort Profile: The Whitehall II study. International Journal of Epidemiology. 2005;**34**:251-256.
9. Cooper R, Hardy R, Aihie Sayer A, Ben-Shlomo Y, Birnie K, Cooper C*, et al.* Age and Gender Differences in Physical Capability Levels from Mid-Life Onwards: The Harmonisation and Meta-Analysis of Data from Eight UK Cohort Studies. PLoS ONE. 2011;**6**:e27899.
10. Nelson HE, Willison JR. National Adult Reading Test (NART) test manual (Part II). Windsor, England: NFER-NELSON; 1991.
11. Raven JC. Guide to using the Mill Hill vocabulary test with progressive matrices. London, England: HK Lewis; 1965.
12. Heim AW. AH4 Group Test of General Intelligence. Windsor: The NFER-Nelson Publishing Company; 1970.
13. Wechsler D. Wechsler Memory Scale-Revised. New York, NY: Psychological Corp; 1987.
14. Raven JC, Court JH, Raven J. Manual for Raven's Progressive Matrices and Vocabulary Scales. London, England: H. K. Lewis; 1977.
15. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA*, et al.* Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet. 2009;**41**:703-707.
16. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics in Medicine. 2002;**21**:1539-1558.