Title: Estimating the genetically predicted effects of lifestyle risk factors, educational attainment and Alzheimer's disease liability on weight change during midlife

Grace M. Power[1*], Jessica Tyrrell[2], Apostolos Gkatzionis[1], Si Fang[1], Jon Heron[1], George Davey Smith[1*†], Tom G. Richardson[1,3*†]

[1] MRC Integrative Epidemiology Unit at the University of Bristol, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, BS8 2BN, UK

[2] Genetics of Complex Traits, College of Medicine and Health, University of Exeter, Exeter, UK.

[3] Department of Genetics, Novo Nordisk Research Centre Oxford, Oxford, UK

*Corresponding authors: Grace M Power (grace.power@bristol.ac.uk), George Davey Smith (kz.davey-smith@bristol.ac.uk) and Tom G Richardson (tom.g.richardson@bristol.ac.uk)

†Contributed equally as last authors

Word Count: 3,528

Abstract (290):

Background: Weight change is a major indicator of adverse health outcomes. This study aims to examine factors contributing to weight change over a one-year interval in midlife.

Methods: Observational and one-sample Mendelian randomisation (MR) analyses were conducted to estimate effects on weight change compared to one year previously (mean age: 56.5 years) using data from the UK Biobank study (n=453,169). Risk factors included alcohol consumption, smoking intensity, body mass index (BMI), educational attainment and Alzheimer's disease liability.

Results: Observational analyses indicated strong evidence of an association between greater educational attainment and family history of Alzheimer's disease with weight loss. In contrast, smoking intensity and higher BMI were associated with weight gain. MR analyses were consistent with observational estimates for educational attainment and Alzheimer's disease liability on weight loss and provided strong evidence of a genetically predicted effect between higher overall BMI and weight gain. Whilst there was little evidence of a genetically predicted effect between smoking intensity and weight change in those who had ever smoked, when stratified, smoking intensity was associated with weight loss in current smokers and weight gain in previous smokers. There was little evidence of an association between alcohol consumption and weight change in the one-year period. Inverse probability weighting was used to account for non-random selection on smoking and alcohol status and further stratification by smoking status, indicating that our results were largely robust to collider bias.

Conclusions: Individuals who have been in education for longer, may have more opportunity to reduce their weight in midlife. The effect of Alzheimer's disease liability on weight loss could be

43 indicative of early signs of dementia. The relationship between smoking intensity and weight
44 change is complex, reinforcing the importance of combining interventions aimed at controlling
45 weight and smoking cessation among cigarette smokers.

Key Messages:

- Observational and Mendelian randomisation analyses used in this study provide consistent evidence that educational attainment and Alzheimer's disease liability have an effect on weight loss over one year during midlife.
- This suggests that individuals experiencing unintentional weight loss in midlife may be at high risk of developing Alzheimer's disease in later life.
- These results additionally demonstrate the value of Mendelian randomisation in minimising spurious associations obtained through observational research when assessing lifestyle risk factors prone to residual confounding, such as smoking intensity.
- Our findings show that the relationship between smoking and weight change is complex, endorsing the notion that strategies to tackle weight control and smoking cessation should be combined.

Key words: weight change, midlife, Alzheimer's disease, educational attainment, smoking intensity, alcohol consumption, lifestyle risk factors, observational epidemiology, genetic epidemiology

63 **Abbreviations**

64

BMI	Body mass index
IPW	Inverse Probability Weighting
MR	Mendelian randomisation
OLR	Ordinal logistic regression
UKB	UK Biobank

65

66 **Background:**

67

68 Weight change across the lifecourse can have major health implications (1-3). Weight gain in
69 adulthood has been associated with increased risk of cardiometabolic disease, including type 2
70 diabetes and coronary heart disease (4), whereas weight loss in midlife may be an early indicator
71 of dementia and mortality (5-9). Furthermore, lifestyle risk factors, including increased alcohol
72 consumption and smoking intensity, have been shown to associate with weight change. Alcohol
73 has been reported to increase weight as a consequence of excess calorie intake (10). Evidence
74 concerning the relationship between smoking and weight change, however, is both conflicting and
75 complex (11, 12). Such inconsistencies are likely due to other lifestyle factors, since physical
76 inactivity, unhealthy diet and alcohol consumption are positively correlated with both adiposity
77 and smoking behaviour (13). There is evidence indicating that smoking intensity reduces weight
78 within current smokers, whilst having been a heavier smoker previously has been reported to lead
79 to greater weight gain amongst previous smokers (14). Furthermore, higher body mass index
80 (BMI) has been shown to influence lifetime smoking, smoking initiation and smoking heaviness
81 (15). In addition, a 2013 systematic review comprising 289 observational studies assessing
82 educational attainment and obesity, identified an inverse relationship in studies from higher-
83 income countries (16). This review revealed an association between lower education and obesity,
84 whilst other studies have revealed an association between higher educational attainment and
85 reduced weight gain (17). Evident within this review is a sparsity of observational investigations
86 that adjusted for potential confounders, further highlighting the inherent issues of residual
87 confounding within observational research (18).

88

Mendelian randomisation (MR) is a technique that can be used to generate more reliable evidence about causal effects of exposures on disease outcomes compared with observational studies (19). This is achieved by harnessing genetic variants associated with modifiable exposures, an approach that can be more robust to confounding factors and reverse causation than naïve observational analyses (20, 21). MR has been successfully applied to estimate the genetically predicted effect of biological, behavioural and socioeconomic exposures on disease risk (22-25). MR investigations have also been extended to evaluate the genetically predicted effects of disease liability on outcomes (26), which includes diseases that typically have a late-onset in the lifecourse, such as dementia. Whilst latent liability to dementia can be thought of as probabilistically underlying the binary diagnosis measure, results are therefore interpreted on the liability scale (27, 28). Within the data employed for our investigation, participants are likely to be at different stages of disease development and some who would never develop disease may still show influences of liability on outcomes, emphasising the need for interpretation to be the effect of liability, not disease.

This study aims to estimate the observational and genetically predicted effects of alcohol intake, smoking, BMI, educational attainment and Alzheimer's disease liability on weight change compared to one year previously during midlife, using data on participants from the UK Biobank (UKB) study (1). Initially, observational analyses will be conducted using ordinal logistic regression (OLR). MR analyses will then be applied to compare the effects of genetically predicted exposures with observational estimates. In doing so, we sought to develop a better understanding of the factors that contribute to weight change in midlife to improve effective strategies in the prevention of related adverse outcomes (29, 30).

112 **Methods:**

113

114 Data resources

115

116 Data were collected from the UKB study; a population-based cohort aged between 40-69 years at
117 recruitment, enrolled in the United Kingdom between 2006 and 2010 (1). In depth details on data
118 collection within UKB during clinic visits and through electronic health records, as well as
119 genotyping, quality control and imputation have been described in detail previously (31).

120

121 All observational exposure data were collected cross-sectionally (Table S1). Smoking intensity by
122 current smokers was assessed by number of cigarettes per day. Alcohol intake by ever drinkers
123 was assessed by units per day, which was calculated from the NHS estimates of pure alcohol
124 consumption per day using data collected via the UKB (32). Weekly and monthly average intake
125 was converted to daily intake by dividing by 7 and 30, respectively. Educational attainment was
126 derived as a continuous score using the criteria described previously by the SSGAC consortium
127 (33). Family history of Alzheimer's disease was analysed as a binary variable indicating either or
128 both parents reportedly having Alzheimer's disease, as well as subsequent analyses of mothers and
129 fathers separately. The outcome variable used in this investigation was obtained from participants
130 being asked the question "Compared with one year ago, has your weight changed?" with the
131 response options "Yes – lost weight", "No – weight about the same" and "Yes – gained weight"
132 which were coded 0, 1 and 2, respectively. We used to the variable containing data on previous
133 and current smokers to stratify on smoking status in MR analyses.

134

135 Genetic risk scores

136

137 Instruments for each exposure were constructed as genetic risk scores coded additively in UKB
138 using the software PLINK (34). Genetic variants robustly associated with alcohol intake (35), BMI
139 (36), educational attainment (37) and Alzheimer's disease (38) were identified from large-scale
140 genome-wide association studies which did not include the UKB study. Variants were selected
141 based on $P<5x10^{-8}$ and $r^2<0.001$ using a reference panel of Europeans from the 1000 genomes
142 project phase 3 to ensure they were independent (39). Influence on smoking is poorly captured by
143 the observational variable indicating number of cigarettes smoked per day (40). Thus, for smoking
144 intensity, we used a single variant as our instrument (rs16969968) located at the *CHRNA5* locus
145 which has been shown to be robustly associated with smoking heaviness (41).

146

147 Statistical analysis

148

149 OLR analyses were undertaken to calculate observational estimates of each of the five exposures
150 in turn on weight change compared to one year previously. Analyses were adjusted for age, sex
151 and the top ten principal components to account for residual population stratification in the UKB
152 study.

153

154 Univariable MR analyses were conducted using individual-level data by applying OLR with the
155 genetic instruments described above in turn on weight change compared to one year previously.
156 MR analyses were adjusted for the same covariates as in the observational analyses. Using one-
157 sample MR, as opposed to two-sample MR, allowed us to remove participants who had reportedly

never smoked or who were never drinkers. This ensured analyses were conducted accurately on appropriate populations: ever smoked (n=150,812) and ever drunk alcohol (n=312,948), making the modelling more intuitive.

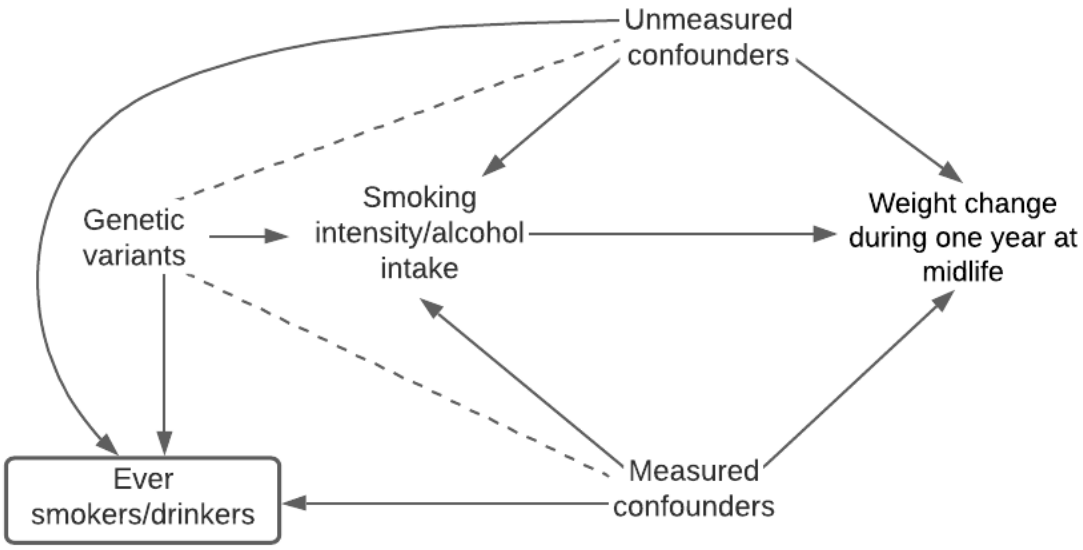
Previous evidence has found that the effect of ever smoking on BMI weakens as time since smoking cessation increases (41, 42). Therefore, to get an accurate estimate of smoking on weight related outcomes, analyses should be undertaken in current smokers as a separate group. Consequently, MR analyses of the relationship between the genetic instrument used for smoking intensity and weight change were completed after stratifying on current (n=33,098) and previous (n=117,714) smoking. Furthermore, conducting MR analyses in current smokers ensures comparisons can made between MR and observational estimates, since the observational smoking intensity variable is measured by number of cigarettes smoked per day and is therefore collected for current smokers only.

Collider bias

We used Inverse Probability Weighting (IPW) to account for both (i) selection on smoking and alcohol use and (ii) stratification by smoking status, both of which may have induced collider bias into analyses. This was to ensure our estimates had accounted for unmeasured confounding and non-random selection (43). As illustrated in Figure 1, the exclusion of never-smokers (or never-drinkers) opens a pathway between the genetic instrument used and the outcome, weight change compared to one year previously, via both measured and unmeasured confounders (Figure 1A).

180 Likewise, stratification within the ever smokers' group to investigate the relationship between

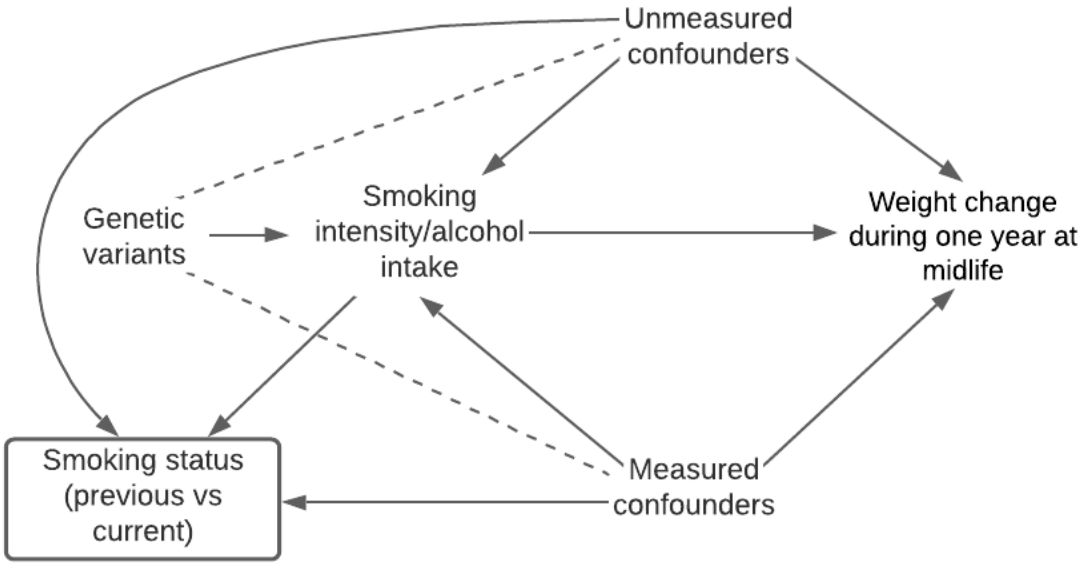
181 previous or current smoking on weight change may also induce bias (Figure 1B).



182 Figure 1A: A directed acyclic graph indicating potential collider bias encountered in this study

183 when selecting participants based on smoking and alcohol drinking status in the UK Biobank.

184



185 Figure 1B: A directed acyclic graph indicating potential collider bias encountered in this study
186 when stratifying on participants' smoking status in the UK Biobank.
187
188 IPW, an approach that uses weights to attempt to make selected participants a representative
189 sample of the study population, has been shown to ameliorate bias (43, 44). Here, we explain how
190 we have applied this method to investigate the relationship between smoking intensity and weight
191 change. The first stage was additionally computed to assess the association between alcohol
192 consumption and weight change in MR analysis.
193
194 IPW weights were derived from a logit model estimated on the whole population. The logit model
195 included the collider (never vs ever smokers) as the binary outcome, and risk factors for smoking
196 initiation as independent variables. These weights were then included in the analysis assessing the
197 effect of the genetic instrument used for smoking intensity on weight change in ever smokers. A
198 second logit model was estimated within the ever (previous vs current) smoking population. This
199 included risk factors for current smoking. Weights from this second model were multiplied by the
200 previous weights described above. These new weights were then included in the stratified analysis
201 assessing the effect of the genetic instrument for smoking intensity on weight change within
202 previous (n=106,666) and current (n=7,669) smokers. Each selected participant is now not only
203 accounting for themselves but for those with similar characteristics who were not selected (45).
204 Reduced sample sizes were a result of including additional variables in our models as data on all
205 variables was necessary for analyses. The first stage was additionally computed to assess the
206 association between alcohol consumption and weight change in MR analysis.
207

208 *Sensitivity analyses*

209

210 Additional sensitivity analyses were conducted using MR to apply the pleiotropy robust weighted
211 median and MR-Egger methods using the 'TwoSampleMR' R package (46-48). This was
212 computed for the exposures measuring alcohol consumption, educational attainment, BMI and
213 genetically predicted Alzheimer's disease to examine the robustness of results to horizontal
214 pleiotropy; the phenomenon where genetic variants influence multiple traits or disease outcomes
215 via independent biological pathways. Pleiotropy robust methods were not undertaken for smoking
216 intensity due to only using a single SNP instrument at the *CHRNA5* locus.

217

218 **Results:**

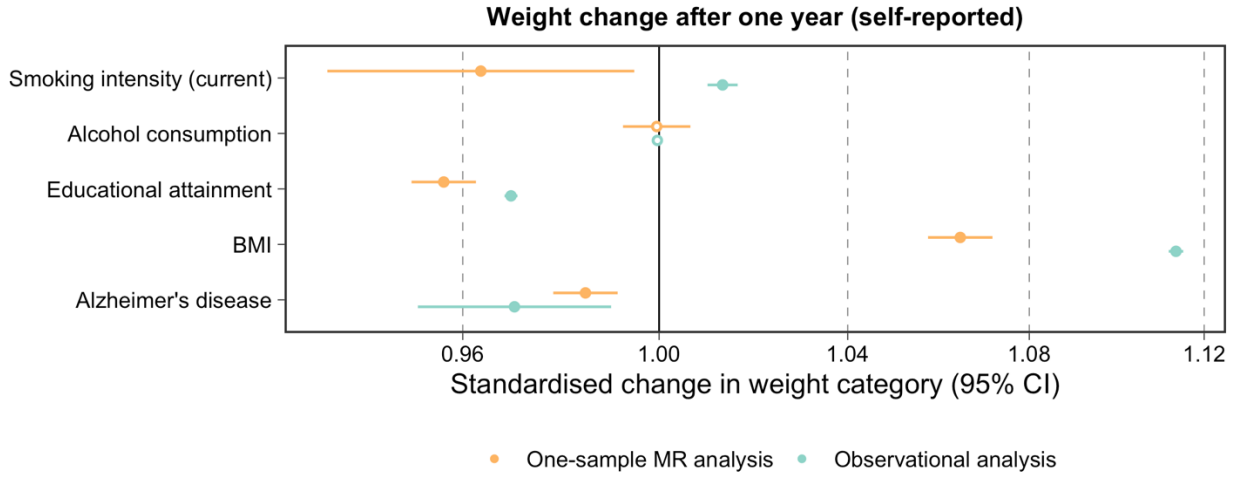
219

220 Observational analyses adjusted for age, sex and the top ten principal components to account for
221 residual population stratification indicated strong evidence of association between smoking
222 intensity in current smokers and weight gain (OR, 95% CI: 1.013, 1.010 to 1.016, P= $1x10^{-16}$),
223 higher overall BMI and weight gain (OR, 95% CI: 1.113, 1.112 to 1.115, P<$1x10^{-300}$), educational
224 attainment and weight loss (OR, 95% CI: 0.969, 0.968 to 0.971, P<$1x10^{-300}$) and Alzheimer's
225 disease in either or both parents and weight loss (OR, 95% CI: 0.970, 0.951 to 0.990, P=0.003)
226 (Table 1).

227

228 Causal effect estimates from one-sample MR were similar to the observational estimates for most
229 exposures (Figure 2). For instance, there was strong evidence of a genetically predicted effect
230 between higher overall BMI and weight gain (OR, 95% CI: 1.065, 1.058 to 1.072, P=$5x10^{-75}$),

231 educational attainment and weight loss (OR, 95% CI: 0.956, 0.950 to 0.963, P=3x10^{-39}) and
232 Alzheimer's disease liability and weight loss (OR, 95% CI: 0.985, 0.978 to 0.991, P=7x10^{-6}). In
233 contrast, MR analyses provided little evidence of an effect between the genetic instrument used
234 for smoking intensity and weight change in those that had ever smoked (i.e., previous and current
235 smokers) (OR, 95% CI: 1.006, 0.991 to 1.021, P=0.5). However, upon stratifying ever smokers
236 into previous and current, there was evidence that smoking intensity has a genetically predicted
237 effect on weight loss in current smokers (OR, 95% CI: 0.964, 0.933 to 0.995, P=0.02) and weight
238 gain in previous smokers (OR, 95% CI: 1.017, 1.000 to 1.035, P=0.04), with strong evidence of
239 effect modification (P=0.004). There was little evidence of a genetically predicted effect of alcohol
240 intake on weight change (OR, 95% CI: 1.00, 0.993 to 1.007, P=0.9) (Table 2).
241



242

243 Figure 2: A forest plot comparing results from one-sample mendelian randomization (yellow) and
244 observational (blue) analyses with the outcome variable comprising three tiers coded 0 ("Yes –
245 lost weight", 1 ("No – weight about the same") and 2 ("Yes – gained weight"). Dots filled-in
246 indicate some to very strong statistical evidence of an association and the exposure variable
247 Alzheimer's disease presented on the continuous liability scale in one-sample MR and coded 0

248 (“Alzheimer’s disease in neither parent”) and 1 (“Alzheimer’s disease in either or both parents”)
249 in observational analyses.
250
251 *Collider bias*
252
253 IPW analyses were conducted for estimates that may be biased by non-random selection. Initially
254 these were computed for estimates calculated within the ever smoker and alcohol consumer
255 groups. There was evidence of an effect between the genetic instrument used for smoking intensity
256 and weight change in those that had ever smoked (i.e., previous and current smokers) (OR, 95%
257 CI: 1.009, 1.002 to 1.016, P=0.009) and alcohol consumption and weight change in those that had
258 ever drunk alcohol (OR, 95% CI: 0.997, 0.989 to 1.004, P=0.371) (Table 3). Upon stratifying,
259 there was strong evidence of an effect between smoking intensity and weight change in those that
260 had previously smoked (OR, 95% CI: 1.019, 1.008 to 1.030, P=0.001) and little evidence between
261 smoking intensity and weight change in those that were current smokers (OR, 95% CI: 0.988,
262 0.948 to 1.029, P=0.6) (Table 3). Including additional variables into IPW models reduced sample
263 sizes and therefore impacted the precision of these estimates.
264
265 *Sensitivity analyses*
266
267 Findings from the two-sample MR weighted median and MR-Egger method revealed similar
268 trends to one-sample MR, suggesting that effects detected are unlikely to be prone to horizontal
269 pleiotropy (Table 4). MR leave one out sensitivity analysis showed that when ‘rs12972156’,
270 located at the *APOE* locus on chromosome 19, is left out, the effect estimates for genetically

271 predicted Alzheimer's disease liability on weight change compared to one year previously
272 attenuated but still supported an effect on weight loss (OR, 95% CI: 0.993, 0.988 to 0.998,
273 P=0.008).
274
275 **Discussion:**
276
277 In this study, results from observational analyses indicated strong evidence of an association
278 between four exposures (smoking intensity, BMI, educational attainment and Alzheimer's disease
279 family history) and weight change compared to one year previously in midlife (mean age: 56.5
280 years). Conversely, there was little evidence of a relationship between alcohol intake and weight
281 change within this lifecourse interval. Using one-sample MR, there was strong evidence of an
282 effect between genetically predicted BMI and weight gain, whereas genetically predicted
283 educational attainment and Alzheimer's disease liability provided evidence of an effect on weight
284 loss. MR analyses found weak evidence of a genetically predicted effect of smoking intensity and
285 alcohol consumption on weight change compared to one year previously in midlife, however, when
286 those who had ever smoked were stratified by previous and current smoking, smoking intensity
287 had a genetically predicted effect on weight gain and weight loss, respectively. Results from IPW
288 analyses, conducted to account for potential collider bias, reveal consistent directions of effect
289 with stronger evidence and weaker statistical evidence, respectively. This is likely a reflection of
290 sample size reduction in the latter.
291
292 Educational attainment in our study was associated with weight loss in midlife based on evidence
293 from both observational and MR analyses. An inverse association between educational attainment

294 and obesity has been consistently observed in studies conducted in high-income countries (16).
295 For instance, a prospective cohort study comprising children (n=12,150) born in Aberdeen
296 between 1950 and 1956 concluded educational attainment was inversely associated with adult BMI
297 (mean age 48 years) after controlling for childhood intelligence and other measured covariates
298 (49). Findings from this study were explained by fixed family factors, such as parental
299 socioeconomic background, intelligence and education, and school and neighbourhood
300 characteristics. Results for educational attainment in our study may also be partially interpreted as
301 a proxy for socioeconomic position. In addition to overall BMI, higher educational attainment, has
302 been associated with a lower increase in weight gain during adulthood in a multicentre cohort
303 study conducted between 1992 to 2000 across Europe (n=361,467) (17). Education, as well as
304 other socioeconomic factors, may influence weight-related behaviours such as diet and physical
305 activity, thus affecting energy balance. If health behaviours of this kind persist, this imbalance has
306 the potential to accumulate over time, leading to accelerated increase in weight, and therefore,
307 continued and greater weight gain in midlife.
308
309 Alzheimer's disease liability was a strong predictor of weight loss in both observational and MR
310 analyses. Importantly, using the *APOE* variant as a genetic instrument for Alzheimer's disease
311 liability may be less prone to misclassification than reported family history, which may explain
312 the narrower confidence intervals when comparing observational and MR estimates for this
313 exposure. This result was supported using a two-sample MR leave one out analysis undertaken.
314 Here we found removal of the 'rs12972156' genetic variant, located at the *APOE* locus, attenuated
315 the association between Alzheimer's disease liability and weight loss. Nevertheless, findings
316 consistently show a genetically predicted effect between Alzheimer's disease liability and weight

317 loss in midlife. These results are in line with previous observational research revealing that weight
318 loss precedes the diagnosis of dementia (5) and MR research indicating an association between a
319 higher polygenic risk score for Alzheimer's disease and lower BMI and body fat (50).
320
321 Findings related to alcohol consumption were not consistent with *a priori* expectations derived
322 from previous studies assessing long-term weight change (10). Specifically, evidence from the
323 literature suggesting that increase in alcohol use is associated with weight gain. This may be in
324 part due to our outcome measuring weight change compared to just one year previously at midlife,
325 assuming that on average, those that drink in this age group have been doing so for a considerable
326 length of time over the lifecourse. Thus, weight change as a result of drinking is likely to have had
327 a substantial effect already, resulting in evidence of an impact within one year challenging to
328 ascertain. In addition, observational analyses indicated that increased smoking intensity resulted
329 in weight gain. Whilst there was no evidence of an association between smoking intensity and
330 weight change compared to one year previously in midlife within those who had ever smoked in
331 one-sample MR, when this group was stratified by smokers who previously smoked and who
332 currently smoke, results were in line with evidence pointing to an inverse association between
333 current cigarette smoking and body weight, followed by weight gain after smoking cessation (11,
334 14, 51, 52). Again, it is important to reiterate that careful consideration is required if stratifying a
335 population into subgroups when conducting MR analyses, as well as observational analyses, to
336 evaluate when collider bias may have been introduced. We express the importance of adjusting for
337 this using IPW (43). These findings also highlight the potential for MR to better mitigate against
338 distortions resulting from residual confounding, often present in observational epidemiology (19).
339 It has also been established that there is a bi-directional association between heavy smoking and

340 BMI, outlining the added complexity of this relationship (13). Higher levels of BMI have been
341 reported to increase risk of smoking uptake as well as intensity of smoking (13). This highlights
342 the need to combine weight control and smoking cessation strategies within interventions aimed
343 at tackling these important public health concerns.
344
345 The association between BMI and weight gain during one year in midlife in observational and one-
346 sample MR analyses suggests that individuals with a higher BMI will typically continue to gain
347 weight at this stage in the lifecourse. Reduced metabolism levels in those with higher BMI are
348 likely to lead to low energy, which may discourage physical activity, further enhanced by midlife
349 body changes as a result of aging (53).
350
351 *Strengths and limitations*
352
353 Observational analyses have previously been conducted alongside MR to compare findings;
354 however, this study is distinctive in that it assesses a time varying outcome; weight change
355 compared to one year previously at midlife. This study also utilises additional sensitivity analyses
356 through weighted median and MR-Egger methods in two-sample MR to ensure the robustness of
357 univariable results to horizontal pleiotropy.
358
359 Important limitations exist in this investigation, however. Firstly, the outcome is a self-reported
360 measure of perceived weight change over a one-year period. Individuals with certain
361 characteristics, including higher overall BMI and educational attainment, may be more likely to
362 misreport their weight, indicating the potential for recall bias and thus differential

363 misclassification. For instance, those with higher educational history have been shown to be more
364 likely to underreport their weight, which would result in overestimated effect sizes (54). Secondly,
365 the outcome used is restricted since it does not demonstrate weight change over a substantial period
366 of time (i.e., more than one year). As discussed previously, this was a potential issue when
367 assessing alcohol consumption, however, additionally has an added benefit in being able to
368 examine recent and rapid weight change during midlife. Additionally, we do not have data on
369 whether an individual changed their smoking or drinking throughout the year in which we examine
370 their weight change. Another key limitation of this study is selection bias and thus limited
371 generalisability. Participants in the UKB are older, more likely to be female and less likely to live
372 in socially deprived areas, compared to participants in nationally representative data sources (55).
373 In addition, previous MR studies have revealed that longer educational duration and older
374 menarche have increased participation, whilst a genetic liability of neuroticism, Alzheimer's
375 disease and schizophrenia have reduced participation, providing further evidence for participation
376 bias (56). Furthermore, (i) selection on smoking and alcohol use, and (ii) stratification by smoking
377 status, may have induced collider bias (43, 57). Whilst IPW was conducted to account for this,
378 including additional variables to generate weights resulted in smaller sample sizes and thus
379 reduced power and precision in our IPW estimates.
380
381 **Conclusion:**
382
383 This investigation provides novel evidence of the effects of lifestyle risk factors, educational
384 attainment and Alzheimer's disease liability on weight change compared to one year previously in
385 midlife. Results suggest that individuals who attended higher education may have more

386 opportunity to help reduce weight gain during this stage in the lifecourse. At the same time,
387 findings may have implications for early testing for Alzheimer's disease when weight loss is
388 observed at midlife. In addition, the relationship between smoking and weight change is complex
389 with findings endorsing the idea that strategies to tackle weight control and smoking cessation
390 should be combined.
391
392 **Acknowledgments:**
393
394 We would like to thank the UK Biobank study and all participants who contributed to it, as well
395 as the authors of all the genome-wide association studies who made their summary statistics
396 available for the benefit of this work.

397 **References:**

398

399 1. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An
400 Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of
401 Middle and Old Age. PLOS Medicine. 2015;12(3):e1001779.

402 2. Yarnell JWG, Patterson CC, Thomas HF, Sweetnam PM. Comparison of weight in
403 middle age, weight at 18 years, and weight change between, in predicting subsequent 14 year
404 mortality and coronary events: Caerphilly Prospective Study. Journal of Epidemiology and
405 Community Health. 2000;54(5):344.

406 3. Corrada MM, Kawas CH, Mozaffar F, Paganini-Hill A. Association of Body Mass Index
407 and Weight Change with All-Cause Mortality in the Elderly. American Journal of Epidemiology.
408 2006;163(10):938-49.

409 4. Wilson PW, Kannel WB. Obesity, diabetes, and risk of cardiovascular disease in the
410 elderly. Am J Geriatr Cardiol. 2002;11(2):119-23,25.

411 5. Knopman DS, Edland SD, Cha RH, Petersen RC, Rocca WA. Incident dementia in
412 women is preceded by weight loss by at least a decade. Neurology. 2007;69(8):739-46.

413 6. Johnson DK, Wilkins CH, Morris JC. Accelerated Weight Loss May Precede Diagnosis
414 in Alzheimer Disease. Archives of Neurology. 2006;63(9):1312-7.

415 7. Chen C, Ye Y, Zhang Y, Pan XF, Pan A. Weight change across adulthood in relation to
416 all cause and cause specific mortality: prospective cohort study. BMJ. 2019;367:l5584.

417 8. Hamer M, O'Donovan G. Sarcopenic obesity, weight loss, and mortality: the English
418 Longitudinal Study of Ageing. Am J Clin Nutr. 2017;106(1):125-9.

419 9. Wong CJ. Involuntary weight loss. Med Clin North Am. 2014;98(3):625-43.

420 10. Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. Changes in diet and lifestyle and
421 long-term weight gain in women and men. N Engl J Med. 2011;364(25):2392-404.
422 11. Sneve M, Jorde R. Cross-sectional study on the relationship between body mass index
423 and smoking, and longitudinal changes in body mass index in relation to change in smoking
424 status: the Tromso Study. Scand J Public Health. 2008;36(4):397-407.
425 12. Dare S, Mackay DF, Pell JP. Relationship between smoking and obesity: a cross-
426 sectional study of 499,504 middle-aged adults in the UK general population. PLoS One.
427 2015;10(4):e0123579.
428 13. Carreras-Torres R, Johansson M, Haycock PC, Relton CL, Davey Smith G, Brennan P, et
429 al. Role of obesity in smoking behaviour: Mendelian randomisation study in UK Biobank. BMJ.
430 2018;361:k1767-k.
431 14. Piirtola M, Jelenkovic A, Latvala A, Sund R, Honda C, Inui F, et al. Association of
432 current and former smoking with body mass index: A study of smoking discordant twin pairs
433 from 21 twin cohorts. PLOS ONE. 2018;13(7):e0200140.
434 15. Taylor AE, Richmond RC, Palviainen T, Loukola A, Wootton RE, Kaprio J, et al. The
435 effect of body mass index on smoking behaviour and nicotine metabolism: a Mendelian
436 randomization study. Human molecular genetics. 2019;28(8):1322-30.
437 16. Cohen AK, Rai M, Rehkopf DH, Abrams B. Educational attainment and obesity: a
438 systematic review. Obes Rev. 2013;14(12):989-1005.
439 17. Rohrmann S, Steinbrecher A, Linseisen J, Hermann S, May A, Luan J, et al. The
440 association of education with long-term weight change in the EPIC-PANACEA cohort. Eur J
441 Clin Nutr. 2012;66(8):957-63.

442 18. Smith GD, Ebrahim S. Data dredging, bias, or confounding. BMJ. 2002;325(7378):1437-
443 8.
444 19. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology
445 contribute to understanding environmental determinants of disease?*. International Journal of
446 Epidemiology. 2003;32(1):1-22.
447 20. Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable
448 behavioural and environmental exposures? BMJ. 2005;330(7499):1076-9.
449 21. Davey Smith G, Holmes MV, Davies NM, Ebrahim S. Mendel's laws, Mendelian
450 randomization and causal inference in observational data: substantive and nomenclatural issues.
451 European Journal of Epidemiology. 2020;35(2):99-111.
452 22. Swerdlow DI, Holmes MV, Kuchenbaecker KB, Engmann JE, Shah T, Sofat R, et al. The
453 interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian
454 randomisation analysis. Lancet. 2012;379(9822):1214-24.
455 23. Holmes MV, Dale CE, Zuccolo L, Silverwood RJ, Guo Y, Ye Z, et al. Association
456 between alcohol and cardiovascular disease: Mendelian randomisation analysis based on
457 individual participant data. BMJ. 2014;349:g4164.
458 24. Nguyen TT, Tchetgen Tchetgen EJ, Kawachi I, Gilman SE, Walter S, Liu SY, et al.
459 Instrumental variable approaches to identifying the causal effect of educational attainment on
460 dementia risk. Ann Epidemiol. 2016;26(1):71-6.e1-3.
461 25. Tillmann T, Vaucher J, Okbay A, Pikhart H, Peasey A, Kubinova R, et al. Education and
462 coronary heart disease: mendelian randomisation study. BMJ. 2017;358:j3542.
463 26. Tudball MJ, Bowden J, Hughes RA, Ly A, Munafò MR, Tilling K, et al. Mendelian
464 randomisation with coarsened exposures. Genet Epidemiol. 2021;45(3):338-50.

465 27. Davey Smith G. Post–Modern Epidemiology: When Methods Meet Matter. American
466 Journal of Epidemiology. 2019;188(8):1410-9.
467 28. Edwards JH. Familial predisposition in man. Br Med Bull. 1969;25(1):58-64.
468 29. Strand BH, Wills AK, Langballe EM, Rosness TA, Engedal K, Bjertness E. Weight
469 Change in Midlife and Risk of Mortality From Dementia up to 35 Years Later. The Journals of
470 Gerontology: Series A. 2017;72(6):855-60.
471 30. Janssen I, Bacon E. Effect of Current and Midlife Obesity Status on Mortality Risk in the
472 Elderly. Obesity. 2008;16(11):2504-9.
473 31. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank
474 resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203-9.
475 32. NHS. Calculating Alcohol Units [Available from: https://www.nhs.uk/live-well/alcohol-
476 support/calculating-alcohol-units/.
477 33. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, et al. Gene discovery and
478 polygenic prediction from a genome-wide association study of educational attainment in 1.1
479 million individuals. Nature Genetics. 2018;50(8):1112-21.
480 34. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation
481 PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7-.
482 35. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to
483 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use.
484 Nature Genetics. 2019;51(2):237-44.
485 36. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of
486 body mass index yield new insights for obesity biology. Nature. 2015;518(7538):197-206.

487 37. Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, et al. Genome-
488 wide association study identifies 74 loci associated with educational attainment. Nature.
489 2016;533(7604):539-42.
490 38. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-
491 analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat
492 Genet. 2013;45(12):1452-8.
493 39. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An
494 integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491(7422):56-65.
495 40. Munafò MR, Timofeeva MN, Morris RW, Prieto-Merino D, Sattar N, Brennan P, et al.
496 Association Between Genetic Variants on Chromosome 15q25 Locus and Objective Measures of
497 Tobacco Exposure. JNCI: Journal of the National Cancer Institute. 2012;104(10):740-8.
498 41. Millard LAC, Munafò MR, Tilling K, Wootton RE, Davey Smith G. MR-pheWAS with
499 stratification and interaction: Searching for the causal effects of smoking heaviness identified an
500 effect on facial aging. PLoS Genet. 2019;15(10):e1008353.
501 42. Freathy RM, Kazeem GR, Morris RW, Johnson PC, Paternoster L, Ebrahim S, et al.
502 Genetic variation at CHRNA5-CHRNA3-CHRNB4 interacts with smoking status to influence
503 body mass index. Int J Epidemiol. 2011;40(6):1617-28.
504 43. Hughes RA, Davies NM, Davey Smith G, Tilling K. Selection Bias When Estimating
505 Average Treatment Effects Using One-sample Instrumental Variable Analysis. Epidemiology.
506 2019;30(3).
507 44. Gkatzionis A, Burgess S. Contextualizing selection bias in Mendelian randomization:
508 how bad is it likely to be? International Journal of Epidemiology. 2019;48(3):691-701.

509 45. Nohr EA, Liew Z. How to investigate and adjust for selection bias in cohort studies. Acta
510 Obstetricia et Gynecologica Scandinavica. 2018;97(4):407-16.
511 46. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base
512 platform supports systematic causal inference across the human phenome. Elife. 2018;7.
513 47. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid
514 instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol.
515 2015;44(2):512-25.
516 48. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian
517 Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet
518 Epidemiol. 2016;40(4):304-14.
519 49. Lawlor DA, Clark H, Davey Smith G, Leon DA. Childhood intelligence, educational
520 attainment and adult body mass index: findings from a prospective cohort and within sibling-
521 pairs analysis. International Journal of Obesity. 2006;30(12):1758-65.
522 50. Korologou-Linden R, Howe LD, Millard LAC, Ben-Shlomo Y, Williams DM, Smith
523 GD, et al. The causes and consequences of Alzheimer's disease: phenome-wide evidence from
524 Mendelian randomization. medRxiv. 2020:2019.12.18.19013847.
525 51. Munafò MR, Tilling K, Ben-Shlomo Y. Smoking status and body mass index: a
526 longitudinal study. Nicotine Tob Res. 2009;11(6):765-71.
527 52. Kaufman A, Augustson EM, Patrick H. Unraveling the Relationship between Smoking
528 and Weight: The Role of Sedentary Behavior. J Obes. 2012;2012:735465.
529 53. Khandelwal S. Obesity in midlife: lifestyle and dietary strategies. Climacteric.
530 2020;23(2):140-7.

531 54. Stommel M, Schoenborn CA. Accuracy and usefulness of BMI measures based on self-
532 reported weight and height: findings from the NHANES & NHIS 2001-2006. BMC Public
533 Health. 2009;9:421.
534 55. Sekula P, Del Greco M F, Pattaro C, Köttgen A. Mendelian Randomization as an
535 Approach to Assess Causality Using Observational Data. Journal of the American Society of
536 Nephrology. 2016;27(11):3253-65.
537 56. Tyrrell J, Zheng J, Beaumont R, Hinton K, Richardson TG, Wood AR, et al. Genetic
538 predictors of participation in optional components of UK Biobank. Nat Commun.
539 2021;12(1):886.
540 57. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when
541 selection bias can substantially influence observed associations. International Journal of
542 Epidemiology. 2018;47(1):226-35.
543