Exploring causal effects of smoking and alcohol related lifestyle factors on self-report tiredness: a Mendelian randomization study

Heshan Li[1¶], Junru Zhao[2¶], Jing Liang[3¶], Xiaoyu Song[1*]

[1]School of Life Sciences and Technology, Harbin Institute of Technology, Harbin, China

[2]School of Mathematical Sciences, Harbin Engineering University, Harbin, China

[3]Harbin Huaqiang Power Automation Engineering Company Limited, Harbin, China

***** Corresponding Author
Email: songxyhit@hit.hit.edu.cn (XS)

¶These authors contributed equally to this work and share first authorship

Abstract

Self-reported tiredness or low energy, often referred to as fatigue, has been linked to lifestyle factors, although data from randomized–controlled trials are lacking. We investigate whether modifiable lifestyle factors including smoking and alcohol intake related exposures (SAIEs) are causal factors for fatigue using Mendelian randomization (MR). A two-sample MR study was performed by using genome-wide association summary results from UK Biobank (UKBB), and each of the sample size is more than 100,000. We used the inverse variance weighted method, and sensitivity analyses (MR Egger, weighted median and penalized median estimators) to account for pleiotropy. The two-sample MR analyses showed inverse causal effect of never-smoking status and positive effect of current smoking status on the risk of fatigue. Similarly, genetically predicted alcoholic intake was positively associated with fatigue. The results were consistent across the different MR methods. Our Mendelian randomization analyses do support that the cessation of smoking and alcohol can decrease the risk of fatigue, and limit alcohol intake frequency can also reduce the risk.

Author summary

Many lifestyle factors have been associated with the risk of fatigue, but we cannot ascertain the causality between lifestyle factors and the risk of fatigue; whether the modification of lifestyles will reduce the risk. Another challenge is that fatigue is usually caused by various physiological and pathological factors, so most epidemiological data which examined risk factor modification have not studied the relationship between modifiable risk factors and self-reported tiredness in extensive conditions. SAIEs are the ones of the most influential lifestyle factors for human health and wellbeing. We performed MR analyses to estimate the causal effect of SAIEs on fatigue. In our study, we initially identified genetic variants which are significantly associated with SAIEs. We

found SAIEs are causally involved in fatigue. The results could be extremely useful in the context of lifestyle - health relationships.

Introduction

Self-reported tiredness and low energy are often called fatigue which is widespread in the population [1–5]. It is a common presentation in primary care [6-9]. While many lifestyle factors have been associated with fatigue, the causality of the association of lifestyle factors with fatigue remains unestablished. The first challenge is that demonstrating the causality with observational epidemiological studies is infeasible due to biases such as confounding and reverse causation. One way to assess causality is with Mendelian randomization (MR), which uses genetic variants as instrumental variables in an approach analogous to a randomized controlled trial, to assess whether risk factors have causal associations with an outcome of interest [10-12].

Another challenge is that fatigue is both heterogeneous (occurring in different conditions) and multifactorial, and it is commonly unexplained by underlying disease [7, 13]. It is the complexity of fatigue that makes it infeasible to discern its causal link with lifestyle factors via observational studies. Most epidemiological data examining risk factor modification have studied the relationships between modifiable risk factors and fatigue in a specific sample and under a certain condition.

We performed a two-sample Mendelian randomization analysis – including inverse variance weighted (IVW), MR-Egger regression, weighted median estimator, and penalized weighted median estimator – to investigate the etiological role of the two most important modifiable lifestyle factors (smoking and alcohol intake related exposures) on fatigue with unlimited samples.

Methods

We undertook Mendelian randomization analyses to estimate effects of SAIEs on fatigue in a two-sample Mendelian randomization framework. The Mendelian randomization approach was based on the following assumptions, which have been widely described in recent studies [14-16]:

1. The genetic variants used as instrumental variables (IVs) are predictive of the exposure.
2. IVs are independent of any confounders of the exposure-outcome relationship.
3. IVs are associated with outcomes only through the clinical risk factors, that is, a lack of pleiotropy (Fig 1).

Fig 1. A directed acyclic graph illustrating core instruments variable assumptions of the Mendelian randomization approach. IV assumption 1: IVs are strongly correlated with exposures. IV assumption 2: IVs are independent of outcomes (i.e., IVs can only affect outcomes through exposures). IV assumption 3: IVs are not related to confounding factors. IV: instrument variable.

These assumptions imply that there is only one causal pathway from the genetic variant to the outcome via the risk factor, and no other causal pathway either directly or indirectly to the outcome [16]. A diagram corresponding to these assumptions is presented in Fig 1.

Publicly available summary results of SAIE – effect estimates and their standard errors for each single-nucleotide polymorphism (SNP) 's effect on exposures and fatigue – were used for the main Mendelian randomization analysis and therefore no additional ethics approval was required.

Two-sample MR analyses

Data sources and processing

Genome-wide significant ($P < 5 \times 10^{-8}$) genetic variants of predicted SAIE were extracted to satisfy IV assumption 1, and the insignificant genetic variants of self-reported fatigue were extracted to satisfy IV assumption 3. Consequently, their intersection was analyzed by MR analyses.

Genetic variants should be further processed into smaller datasets that meet IV assumptions. First of all, we removed genetic variants with linkage disequilibrium (LD). LD induces correlation between two genetic variants , and destroys the randomness of genetic variants [17, 18]. After SNPs with potential LD were removed, the parts with longer physical distance (more than 10,000kb) and less possibility of LD ($R^2 < 0.001$) were retained. Besides, genetic variants of palindromic and incompatible alleles should be removed when we harmonized the exposure and outcome datasets.

Smoking related exposures. 70 SNPs that were significantly associated with never-smoking status were extracted from summary results data, which belongs to a GWAS of UKBB (N =359,706 individuals of European ancestry). We also extracted 13 SNPs significantly associated with current-smoking status from summary results data, which belongs to a GWAS of UKBB (N =336,024 individuals of European ancestry). The SNPs significantly associated with smoking status are listed in Table A in S1 Table.

Alcohol intake related exposures. Only one SNP that was significantly associated with never-drinking alcohol status was extracted from summary results, which belongs to a GWAS of UKBB (N =336,965 individuals of European ancestry). We extracted three SNPs significantly associated with current-drinking alcohol status, which belong to summary results from a GWAS of UKBB (N =360,726 individuals of European ancestry). Besides this, we extracted 90 SNPs significantly associated with alcohol intake frequency (AIF) from summary results, which belong to a GWAS of

UKBB (N =462,346 individuals of European ancestry). The SNPs significantly associated with alcohol intake are listed in Table C in S1 Table.

self-reported tiredness. We obtained summary results statistics from the GWAS of UKBB. The GWAS of self-reported tiredness included 449,019 individuals, who reported their own tiredness or lethargy in last weeks.

Statistical analysis

To investigate the causal relationship between SAIE and fatigue, we initially performed two-sample inverse variance weighted MR analyses by the SNPs extracted from GWAS summaries [19]. The magnitude of the causal effect ($\hat{\beta}_{IVW}$) was estimated as the average of the SNP-outcome effect ($\hat{\beta}_{ZY}$) divided by the SNP-exposure effect ($\hat{\beta}_{ZX}$). The regression slope (ratio) in IVW analysis is forced through a zero intercept. Analyses were performed with the Two-Sample MR package [20].

To avoid the violation of the IV assumptions 2 and 3, we assessed instrument strength for the standard IVW MR analysis via calculating the approximate F statistics for each of exposures that will be used in our two-sample MR studies [i. e. $F \approx (\beta/SE)^2$ shown in Table A and C in S1 Table].

To test whether all the IVs satisfy the IV assumptions, we performed a test of heterogeneity of causal effect estimates across each of the SNPs via Cochran's Q value quantifying heterogeneity and detecting outliers. If an individual SNP's Q contribution is extremely large (i.e. , above the 5% threshold of 3.84 or instead of a Bonferroni-corrected threshold) , it might imply heterogeneity, including horizontal pleiotropy [21,22].Then we used MR Egger regression to further investigate the possibility of directional pleiotropy in our data (i.e. where some SNPs influence the outcome via additional paths other than via the exposure), and further verified such pleiotropy with funnel plots, which plot instrument strength against the causal estimates for all the IVs. Asymmetry in funnel plots suggests that pleiotropic effects are not balanced and may indicate directional pleiotropy.

Sensitivity analyses

As the resulting IVW analysis suffers from bias and invalid IVs [23], we performed sensitivity analyses via three additional MR models, i.e., MR Egger regression, the weighted median estimator, penalized weighted median estimator, to obtain more robust estimates [24]. All the following sensitivity analyses were subsequently performed upon exclusion and inclusion of outliers respectively, which were identified through Cochran's Q statistics.

In the case of MR Egger regression analysis, we also assessed instrument strength using an I^2_{GX} statistic because F statistics from the individual marks are not sufficient indicators of instrument strength [25]. Bowden showed that an I^2_{GX} statistic can be used to quantify the expected dilution of MR□Egger regression estimates [25]. He showed that a high value of I^2_{GX} (i.e. close to one) would indicate no dilution. On the contrary, if I^2_{GX} is less than 0.9, inference from MR Egger should be interpreted with caution, and some alternative sensitivity analyses should be considered [26,27].

In contrast to the weighted mean used in the IVW analysis, the weighted median estimator uses the weighted median of the ratio (i.e., the ratio of $\hat{\beta}_{ZY}$ and $\hat{\beta}_{ZX}$). In the weighted regression model , the penalized weighted median approach downweighs (or penalizes) the contribution of some candidate variants with heterogeneous ratio estimates, and may have better finite sample properties than weighted median, particularly if there is directional pleiotropy [24].

If the four MR models (IVW, MR Egger regression, weighted median estimator, penalized weighted median), which make different assumptions regarding instrument validity, produce similar estimates of causal effects, then we will be more confident in the robustness of our findings.

We used the Two-Sample MR package (version 0.5.6) in R (version 4.1.3) to perform IVW, MR Egger regression, weighted median, and penalized median estimator. We used the codes provided on https://mrcieu.github.io/TwoSampleMR/ to run analyses.

Results

Smoking status

In this section, we analysed two types of smoking status (never-smoking status vs current-smoking status) to investigate the causal effects of smoking exposures on fatigue outcome.

Never-smoking status. The two-sample MR IVW analysis showed evidence for a negative causal effect of never-smoking status on fatigue. This was supported by median-based model and penalized MR model both before and after outliers extraction (Fig 2 and 3).

Fig 2. The Mendelian Randomization analysis of smoking status (never vs current) on fatigue including outliers. OR = odds ratio per unit decreases in smoking status (never vs current). Forest plot comparing results from inverse variance weighted, weighted median, penalized median and MR Egger methods. CI: confidence interval.

There was evidence of heterogeneity in the causal effect estimates of never-smoking status on fatigue across the individual SNPs (Q=172.4, $P = 8.21 \times 10^{-11}$). On the one hand, the estimate of the intercept of MR Egger regression showed no evidence of directional horizontal pleiotropy for never-smoking status exposure (Egger intercept was -0.0012 and its p-value was 0.53). But MR Egger result from never-smoking status analysis would be unreliable due to the low I^2_{GX} statistic of 0.67 which was below the suggested cut-off of 0.9, suggesting that the MR Egger results could be influenced by measurement error or weak instruments bias. On the other hand, there was an abundance of asymmetry of the funnel plots suggesting there was directional horizontal pleiotropy among our data (Fig A in S1 Note).

current smoking status. The two-sample MR model showed evidence for a positive casual effect of current smoking status on fatigue both before and after the outliers extraction (Fig 2 and 3). This was also supported by median-based and penalized MR models.

Fig 3. The Mendelian Randomization analysis of SST and AIF on fatigue after outliers extraction. OR = odds ratio per unit decreases in SST and AIF. Forest plot comparing results from inverse variance weighted, weighted median, penalized median and MR Egger methods. CI: confidence interval; SST: smoking status: never vs current; AIF: alcohol intake frequency.

Our analysis suggested no significant evidence of horizontal pleiotropy (as indicated by MR-Egger regression intercept was -0.0035, with a *P* value larger than 0.05 shown in Table J in S1 Table). But the low I^2_{GX} statistic of 0.86, below the suggested cut-off of 0.9, indicates that the MR Egger results would be unreliable. There was evidence of heterogeneity in the causal effect estimates of current-smoking status on fatigue across the individual SNPs (Q=29.9, $P = 0.0029$). There was strong evidence of directionally horizontal pleiotropy because of an abundance of asymmetry of the funnel plots (Fig B in S1 Note).

Alcohol intake-related exposures

Alcohol intake status. The two-sample MR IVW analysis showed strong evidence for a negative causal effect of never alcohol intake status on fatigue while there was no evidence for causal effect of current alcohol intake status on fatigue (Fig 4).

Fig 4. The Mendelian Randomization analysis of alcohol intake-related exposures on fatigue including outliers. OR = odds ratio per unit decreases in alcohol intake related phenotypes. Forest plot comparing results from inverse variance weighted, weighted median, penalized median and MR Egger methods. CI: confidence interval.

Alcohol intake frequency. The two-sample MR IVW analysis showed evidence for a positive causal effect of AIF on fatigue. This was supported by median-based and penalized MR models (Figs 3 and 4).

There was evidence of heterogeneity in the causal effect estimates of AIF on fatigue across the individual SNPs (Q=232.6, $P = 1.15 \times 10^{-15}$). The estimate of the intercept of MR Egger regression showed strong evidence of directionally horizontal pleiotropy for AIF exposure (Egger intercept was 0.003 and its p-value was 8.43×10^{-5}). There was an abundant asymmetry of the funnel plots (Fig C in S1 Note), suggesting there was directionally horizontal pleiotropy among our data.

Discussions

Smoking status

We found strong evidence of negative causal effect of never-smoking status on fatigue and positive causal effect of current-smoking status on fatigue. The causal effect size estimates were consistent across the different sensitivity methods applied, and did not depend on whether or not there was direct pleiotropy in the analyses. Sensitivity analysis was performed with MR Egger regression, the weighted median estimator, and penalized median estimator. These analyses were consistent with the results from the IVW analyses. Thus, we generated that smoking is a risk factor to fatigue.

In former observational studies, smoking status has been associated with self-reported tiredness. Subjects who were heavy cigarette smokers were prone to fatigue. But this effect is only specific to the limited age range of participants rather than a broad population [28-30]. The effect estimates in observational study and our MR study are similar, providing additional confidence in our results. Our MR analysis goes one step further than observational studies. The former observational

studies focused on healthy young males; our MR studies extended the sample to a wider range without age and gender limits.

Potential complication in interpreting the results of these sorts of MR analyses is the extent to which one can be sure that the causal estimates reflect the effect of smoking on energy levels. Low energy level refers to fatigue. Nicotine exposure from cigarette smoking can cause a negative energy state which is characterized by reducing energy intake and increasing energy expenditure. Energy intake and expenditure depend on brain feeding regulation [31, 32]. Nicotine can influence caloric consumption and energy expenditure by promoting the release of norepinephrine, serotonin, and other factors. These factors influence the brain to alter appetite and metabolic rate [32-35].

It is also possible that the causal estimates could reflect the effect of nicotine exposure on the increase of myocardial contractility and vasoconstriction, which results in increasing myocardial work and oxygen requirement and reducing coronary and cerebral blood flow [36]. Poor blood flow causes low energy level and then fatigue. Also, the heart must pump harder when circulation is poor, which can lead to further fatigue.

Our study has proved that smoking increases the risk of fatigue, and never-smoking reduces the risk. Different smoking status will produce different gene expression patterns [37]. Smokers are prone to fatigue.

Alcohol intake-related exposures

Alcohol intake status. Evidence is lacking for causal effect of current-drinking status on fatigue. In contrast, we found some evidence of an inverse effect of never-drinking status on fatigue. Such inconsistence of causal effect detection for alcohol intake may result from small number of SNPs associated with alcohol drinker statuses, which limits capacity to test causal effect [38].

Alcohol intake frequency. Evidence is enough for a causal relationship between AIF and fatigue. We found there is a statistically significant causal effect of AIF on fatigue. The causal effect size estimates of MR Egger regression were inconsistent across other different sensitivity methods applied. The low I^2_{GX} for AIF partly indicated that sensitivity analysis of the resulting MR Egger regression was unreliable. The analyses from alternative pleiotropy-robust estimation strategies (including the weighted median estimator and penalized weighted model), complementary to MR Egger regression [39], were consistent with the results from the IVW analyses. The causal effect size estimate did not depend on whether or not there is direct pleiotropy in the analyses.

Results from this study are consistent with previous conventional epidemiological studies which have associated alcohol intake with an increased risk of fatigue. But little evidence shows these apparently association may be causal [40]. Also, the subjects were drawn from a special fatigue clinic, so results cannot be generalised to other settings [40, 41]. Our MR study has extended this associate relationship to a causal one for a broad population without age and gender limits.

A significant causal relationship between alcohol intake and fatigue would reflect the detrimental effect level of alcohol on human physiology. Alcohol use actually inhibits the absorption and usage of vital nutrients such as vitamin B1, vitamin B12, folic acid, and zinc. Zinc is also essential to energy metabolic processes [42, 43]. Drinking alcohol constricts aerobic metabolism and decreases in hepatic ATP synthesis [44]. Thus the physical responses to alcohol in the body can lead to a feeling of fatigue and weakness.

Strengths and limitations

One of the strengths of our MR study was the ability to select a large-scale SAIE GWAS data set as shown in Table J in S1 Table and fatigue (N =449,019 individuals reporting their tiredness) in two-sample MR approach to estimate the effect of smoking status on fatigue, which helps overcome

power limitations of MR. Second, participants in all the GWAS datasets are of European decent, which may reduce the influence on potential association caused by population stratification. The third strength is that we performed a series of sensitivity analyses to explore potential bias due to horizontal pleiotropy.

Due to the lack of relevant data, an important limitation with this study is that confounders associated with exposures in our MR study were not explored. Besides, the small number of genetic instruments for some traits may have introduced weak instrument bias.

Our findings support several potential recommendations. In spite of the complexity of fatigue, individuals can still prevent it by changing lifestyles to reduce fatigue risks, such as smoking and alcohol intake cession.

Conclusion

We found evidence for a causal effect of heavier alcohol intake on increasing the risk of fatigue. In addition, evidence is sufficient for a positive causal relationship of cigarette smoking on fatigue. The resulting MR analyses are consistent with previous observational studies with their relatively small sample size and the limited age range of participants.

Acknowledgments

This research has been conducted using the UK Biobank Resource and IEU open GWAS summary data. We would like to thank Ph.D. candidate Tao Wang from Peking University for help with statistical analyses.

Reference

1. Boter H, Mänty M, Hansen AM, Hortobágyi T, Avlund K. Self-reported fatigue and physical function in late mid-life. J Rehabil Med. 2014 Jul;46(7):684-90. doi: 10.2340/16501977-1814.
2. Deary V, Hagenaars SP, Harris SE, Hill WD, Davies G, Liewald DCM; International Consortium for Blood Pressure GWAS; CHARGE Consortium Aging and Longevity Group; CHARGE Consortium Inflammation Group, McIntosh AM, Gale CR, Deary IJ. Genetic contributions to self-reported tiredness. Mol Psychiatry. 2018 Mar;23(3):609-620. doi: 10.1038/mp.2017.5.
3. Schwarz R, Krauss O, Hinz A. Fatigue in the general population. Onkologie. 2003 Apr;26(2):140-4. doi: 10.1159/000069834.
4. Hagelin CL, Wengström Y, Runesdotter S, Fürst CJ. The psychometric properties of the Swedish Multidimensional Fatigue Inventory MFI-20 in four different populations. Acta Oncol. 2007;46(1):97-104. doi: 10.1080/02841860601009430.
5. Watt T, Groenvold M, Bjorner JB, Noerholm V, Rasmussen NA, Bech P. Fatigue in the Danish general population. Influence of sociodemographic factors and disease. J Epidemiol Community Health. 2000 Nov;54(11):827-33. doi: 10.1136/jech.54.11.827.
6. Cullen W, Kearney Y, Bury G. Prevalence of fatigue in general practice. Ir J Med Sci. 2002 Jan-Mar;171(1):10-2. doi: 10.1007/BF03168931.
7. Ridsdale L, Evans A, Jerrett W, Mandalia S, Osler K, Vora H. Patients with fatigue in general practice: a prospective study. BMJ. 1993 Jul 10;307(6896):103-6. doi: 10.1136/bmj.307.6896.103.
8. Mota DD, Pimenta CA. Self-report instruments for fatigue assessment: a systematic review.

Res Theory Nurs Pract. 2006 Spring;20(1):49-78. doi: 10.1891/rtnp.20.1.49.

9. Engberg I, Segerstedt J, Waller G, Wennberg P, Eliasson M. Fatigue in the general population- associations to age, sex, socioeconomic status, physical activity, sitting time and self-rated health: the northern Sweden MONICA study 2014. BMC Public Health. 2017 Aug 14;17(1):654. doi: 10.1186/s12889-017-4623-y.

10. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003 Feb;32(1):1-22. doi: 10.1093/ije/dyg070.

11. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014 Sep 15;23(R1):R89-98. doi: 10.1093/hmg/ddu328.

12. Evans DM, Davey Smith G. Mendelian Randomization: New Applications in the Coming Age of Hypothesis-Free Causality. Annu Rev Genomics Hum Genet. 2015;16:327-50. doi: 10.1146/annurev-genom-090314-050016.

13. Nijrolder I, van der Windt D, de Vries H, van der Horst H. Diagnoses during follow-up of patients presenting with fatigue in primary care. CMAJ. 2009 Nov 10;181(10):683-7. doi: 10.1503/cmaj.090647.

14. Emdin CA, Khera AV, Natarajan P, Klarin D, Zekavat SM, Hsiao AJ, et al. Genetic Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. JAMA. 2017 Feb 14;317(6):626-634. doi: 10.1001/jama.2016.21042.

15. Larsson SC, Burgess S, Michaëlsson K. Association of Genetic Variants Related to Serum Calcium Levels With Coronary Artery Disease and Myocardial Infarction. JAMA. 2017 Jul

25;318(4):371-380. doi: 10.1001/jama.2017.8981.

16. Clarke PS, Windmeijer F. Instrumental variable estimators for binary outcomes. J Am Stat Assoc. 2012;107:1638–1652.DOI: 10.1080/01621459.2012.734171.
17. Latvala A, Ollikainen M. Mendelian randomization in (epi)genetic epidemiology: an effective tool to be handled with care. Genome Biol. 2016 Jul 14;17(1):156. doi: 10.1186/s13059-016-1018-9.
18. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013 Nov;37(7):658-65. doi: 10.1002/gepi.21758.
19. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. Int J Epidemiol. 2016 Jun;45(3):908-15. doi: 10.1093/ije/dyw127.
20. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife. 2018 May 30;7:e34408. doi: 10.7554/eLife.34408.
21. Bowden J, Del Greco M F, Minelli C, Zhao Q, Lawlor DA, Sheehan NA, Thompson J, Davey Smith G. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. Int J Epidemiol. 2019 Jun 1;48(3):728-742. doi: 10.1093/ije/dyy258.
22. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. Epidemiology. 2017 Jan;28(1):30-42. doi: 10.1097/EDE.0000000000000559.
23. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. Am J Epidemiol. 2013 Oct 1;178(7):1177-84. doi: 10.1093/aje/kwt084.

24. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016 May;40(4):304-14. doi: 10.1002/gepi.21965.
25. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015 Apr;44(2):512-25. doi: 10.1093/ije/dyv080.
26. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol. 2016 Dec 1;45(6):1961-1974. doi: 10.1093/ije/dyw220.
27. Moen GH, Beaumont RN, Grarup N, Sommer C, Shields BM, Lawlor DA, et al. Investigating the causal effect of maternal vitamin B12 and folate levels on offspring birthweight. Int J Epidemiol. 2021 Mar 3;50(1):179-189. doi: 10.1093/ije/dyaa256.
28. Corwin EJ, Klein LC, Rickelman K. Predictors of fatigue in healthy young adults: moderating effects of cigarette smoking and gender. Biol Res Nurs. 2002 Apr;3(4):222-33. doi: 10.1177/109980040200300407.
29. Juster RP, McEwen BS, Lupien SJ. Allostatic load biomarkers of chronic stress and impact on health and cognition. Neurosci Biobehav Rev. 2010 Sep;35(1):2-16. doi: 10.1016/j.neubiorev.2009.10.002.
30. Bremander A, Jacobsson LT, Bergman S, Haglund E, Löfvendahl S, Petersson IF. Smoking is associated with a worse self-reported health status in patients with psoriatic arthritis: data from a Swedish population-based cohort. Clin Rheumatol. 2015 Mar;34(3):579-83. doi: 10.1007/s10067-014-2742-8.

31. Hellerstein MK, Benowitz NL, Neese RA, Schwartz JM, Hoh R, Jacob P 3rd, et al. Effects of cigarette smoking and its cessation on lipid metabolism and energy expenditure in heavy smokers. J Clin Invest. 1994 Jan;93(1):265-72. doi: 10.1172/JCI116955.
32. Chen H, Saad S, Sandow SL, Bertrand PP. Cigarette smoking and brain regulation of energy homeostasis. Front Pharmacol. 2012 Jul 25;3:147. doi: 10.3389/fphar.2012.00147.
33. Churg A, Dai J, Tai H, Xie C, Wright JL. Tumor necrosis factor-alpha is central to acute cigarette smoke-induced inflammation and connective tissue breakdown. Am J Respir Crit Care Med. 2002 Sep 15;166(6):849-54. doi: 10.1164/rccm.200202-097OC.
34. Audrain-McGovern J, Benowitz NL. Cigarette smoking, nicotine, and body weight. Clin Pharmacol Ther. 2011 Jul;90(1):164-8. doi: 10.1038/clpt.2011.105.
35. Stojakovic A, Espinosa EP, Farhad OT, Lutfy K. Effects of nicotine on homeostatic and hedonic components of food intake. J Endocrinol. 2017 Oct;235(1):R13-R31. doi: 10.1530/JOE-17-0166.
36. Rigotti NA, Clair C. Managing tobacco use: the neglected cardiovascular disease risk factor. Eur Heart J. 2013 Nov;34(42):3259-67. doi: 10.1093/eurheartj/eht352.
37. Vink JM, Jansen R, Brooks A, Willemsen G, van Grootheest G, de Geus E, et al. Differential gene expression patterns between smokers and non-smokers: cause or consequence? Addict Biol. 2017 Mar;22(2):550-560. doi: 10.1111/adb.12322.
38. Brion MJ, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. Int J Epidemiol. 2013 Oct;42(5):1497-501. doi: 10.1093/ije/dyt179.
39. Bowden J. Misconceptions on the use of MR-Egger regression and the evaluation of the InSIDE assumption. Int J Epidemiol. 2017 Dec 1;46(6):2097-2099. doi: 10.1093/ije/dyx192.
40. Woolley J, Allen R, Wessely S. Alcohol use in chronic fatigue syndrome. J Psychosom Res. 2004 Feb;56(2):203-6. doi: 10.1016/S0022-3999(03)00077-1.

41. Vella LD, Cameron-Smith D. Alcohol, athletic performance and recovery. Nutrients. 2010 Aug;2(8):781-9. doi: 10.3390/nu2080781.

42. Skalny AV, Skalnaya MG, Grabeklis AR, Skalnaya AA, Tinkov AA. Zinc deficiency as a mediator of toxic effects of alcohol abuse. Eur J Nutr. 2018 Oct;57(7):2313-2322. doi: 10.1007/s00394-017-1584-y.

43. Liuzzi JP, Narayanan V, Doan H, Yoo C. Effect of zinc intake on hepatic autophagy during acute alcohol intoxication. Biometals. 2018 Apr;31(2):217-232. doi: 10.1007/s10534-018-0077-7.

44. Mitchell MC, Herlong HF. Alcohol and nutrition: caloric value, bioenergetics, and relationship to liver damage. Annu Rev Nutr. 1986;6:457-74. doi: 10.1146/annurev.nu.06.070186.002325.

Supporting information

S1 Table. Supplementary tables.

Table A: Effect of SNPs on alcohol intake frequency and alcohol drinker status (never vs current). The statistics are derived from linear regression. The GWAS summary data in the table is from UK Biobank. The GWAS identifiers of alcohol intake frequency and alcohol drinker status (never vs current) are corresponding to ukb-b-5779, ukb-a-226 and ukb-d-20117_2 respectively .1The effect is the mean effect of the increaser allele estimated on a quantile normalized scale. The outlying genetic instruments which had large contribution to Cochran 's Q statistics are shown in bold. EAF: effect allele frequency; SNP: single nucleotide polymorphism; SE: Standard Error; GWAS: Genome-wide association study.

Table B: The effect of the alcohol intake frequency and alcohol drinker status (never vs current) SNPs on self-reported tiredness. The statistics are derived from linear regression. The

GWAS summary data in the table is from UK Biobank. The GWAS identifiers of alcohol intake frequency and alcohol drinker status (never vs current) are corresponding to ukb-b-5779, ukb-a-226 and ukb-d-20117_2 respectively .1The effect is the mean effect of the increaser allele estimated on a quantile normalized scale. EAF: effect allele frequency; SNP: single nucleotide polymorphism; SE: Standard Error; GWAS: Genome-wide association study.

Table C: Effect of SNPs on smoking status (never vs current). The statistic are derived from linear regression. The GWAS summary data in the table is from UK Biobank. The GWAS identifiers of smoking status (never vs current) are corresponding to ukb-d-20116_0 and ukb-a-225 respectively .1The effect is the mean effect of the increaser allele estimated on a quantile normalized scale. The outlying genetic instruments which had large contribution to Cochran 's Q statistics are shown in bold. EAF: effect allele frequency; SNP: single nucleotide polymorphism; SE: Standard Error; GWAS: Genome-wide association study.

Table D: The effect of the smoking status (never vs current) SNPs on self-reported tiredness. 1 The effect is the mean effect of the increaser allele estimated on a quantile normalized scale. EAF: effect allele frequency; SE: Standard Error; SNP: single nucleotide polymorphism; GWAS: Genome-wide association study.

Table E: Mendelian Randomization estimates of the causal effect of alcohol intake frequency on self-reported tiredness. Causal effects are estimated using four MR models: inverse variance weighted (IVW), weighted median, penalized median and MR Egger regression. Causal effect estimates are the difference in mean self-reported tiredness (in standard deviation; SD) per 1SD higher alcohol intake frequency. Significant results from Mendelian Randomization analysis are shown in bold. SE: Standard Error; IVs: Instrumental variables.

Table F: Mendelian Randomization estimates of the causal effect of alcohol status (current) on self-reported tiredness. Causal effects are estimated using four MR models: inverse variance weighted (IVW), weighted median, penalized median and MR Egger regression. Causal effect estimates are the difference in mean self-reported tiredness (in standard deviation; SD) per 1SD higher alcohol status (current).

Table G: Mendelian Randomization estimates of the causal effect of smoking status (never) on self-reported tiredness. Causal effects are estimated using four MR models: inverse variance weighted (IVW), weighted median, penalized median and MR Egger regression. Causal effect estimates are the difference in mean self-reported tiredness (in standard deviation; SD) per 1SD higher alcohol status (never). Significant results from Mendelian Randomization analysis are shown in bold. SE: Standard Error; IVs: Instrumental variables.

Table H: Mendelian Randomization estimates of the causal effect of smoking status (current) on self-reported tiredness. Causal effects are estimated using four MR models: inverse variance weighted (IVW), weighted median, penalized median and MR Egger regression. Causal effect estimates are the difference in mean self-reported tiredness (in standard deviation; SD) per 1SD higher alcohol status (current).

Table I: Results of heterogeneity tests. Significant results from the heterogeneity tests are shown in bold.

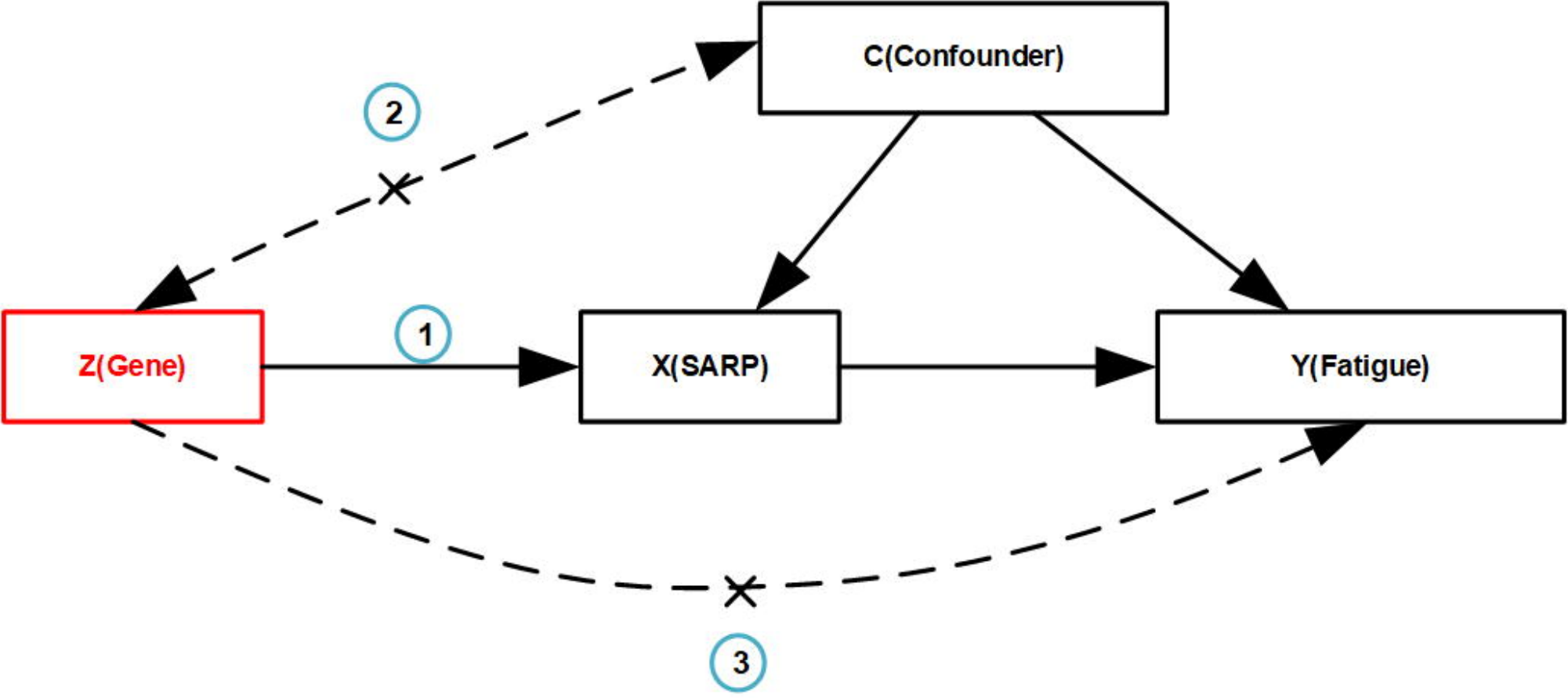
Table J: Results of test for directional horizontal pleiotropy. Significant results from the directional horizontal pleiotropy tests are shown in bold. SE: Standard Error.

S1 Note. Supplementary results.

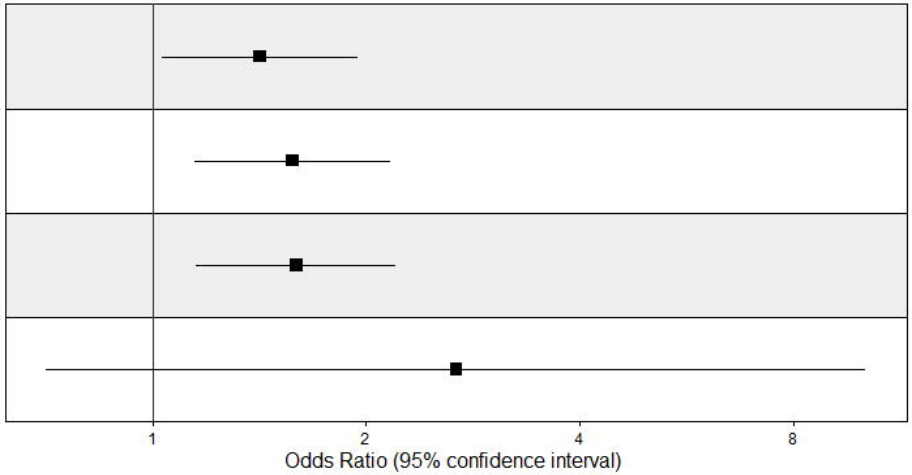
Fig A. Funnel plots for the effect of never smoking status on risk of self-reported fatigue for each single-nucleotide polymorphism (SNP), the resulting Mendelian randomization (MR) estimate is plotted against the inverse of the standard error of the MR estimate. Symmetry noted in this plot provides evidence against the presence of directional horizontal pleiotropy. The inverse-variance weighted and MR Egger causal estimates are represented by a red and blue line respectively.

Fig B. Funnel plots for the effect of current smoking status on risk of self-reported fatigue for each single-nucleotide polymorphism (SNP), the resulting Mendelian randomization (MR) estimate is plotted against the inverse of the standard error of the MR estimate. Symmetry noted in this plot provides evidence against the presence of directional horizontal pleiotropy. The inverse-variance weighted and MR Egger causal estimates are represented by a red and blue line respectively.

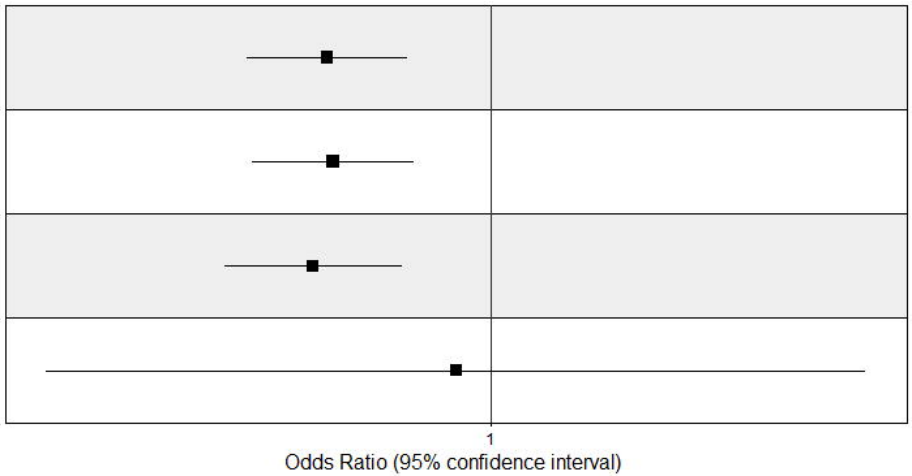
Fig C. Funnel plots for the effect of alcohol intake frequency on risk of self-reported fatigue for each single-nucleotide polymorphism (SNP), the resulting Mendelian randomization (MR) estimate is plotted against the inverse of the standard error of the MR estimate. Symmetry noted in this plot provides evidence against the presence of directional horizontal pleiotropy. The inverse-variance weighted and MR Egger causal estimates are represented by a red and blue line respectively.


C(Confounder)
2
Z(Gene)
1
X(SARP)
Y(Fatigue)
3

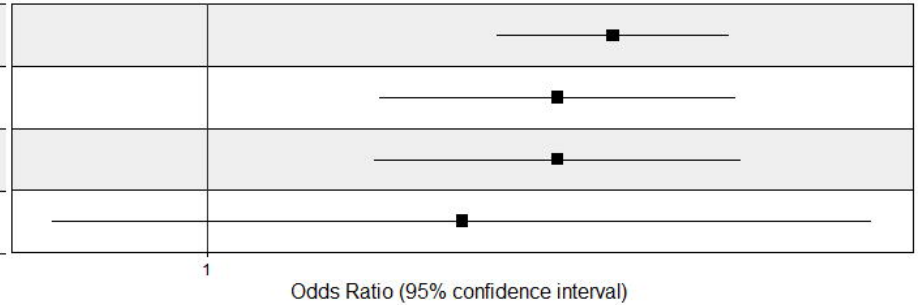
Smoking status: Current	OR	CI
Inverse variance weighted	1.417	[1.033,1.944]
Weighted median	1.576	[1.146,2.168]
Penalised weighted median	1.594	[1.153,2.203]
MR Egger	2.678	[0.708,10.122]



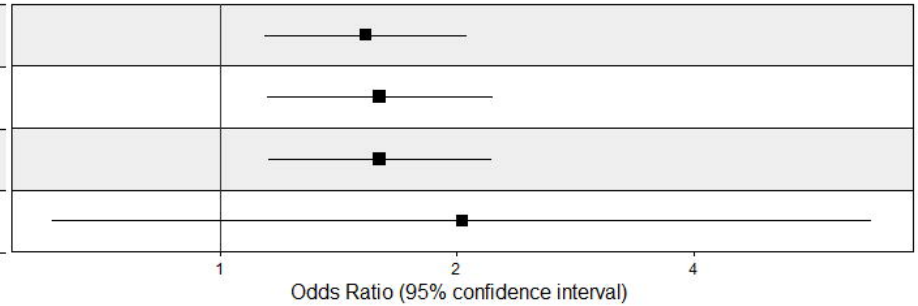
Smoking status: Never	OR	CI
Inverse variance weighted	0.834	[0.763,0.912]
Weighted median	0.84	[0.768,0.919]
Penalised weighted median	0.822	[0.745,0.907]
MR Egger	0.963	[0.612,1.515]



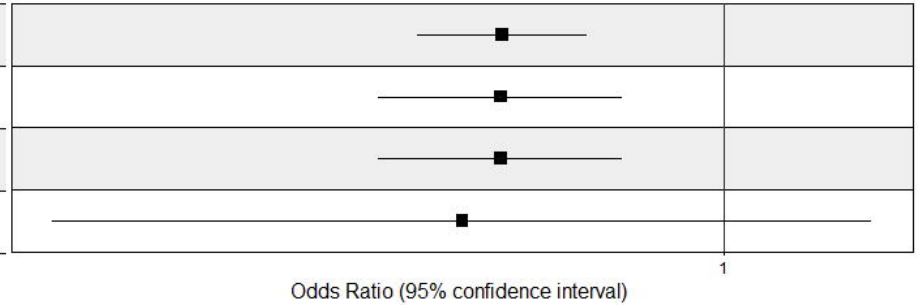
Alcohol intake frequency.	OR	CI
Inverse variance weighted	1.081	[1.057,1.105]
Weighted median	1.069	[1.034,1.106]
Penalised weighted median	1.069	[1.032,1.108]
MR Egger	1.05	[0.971,1.135]



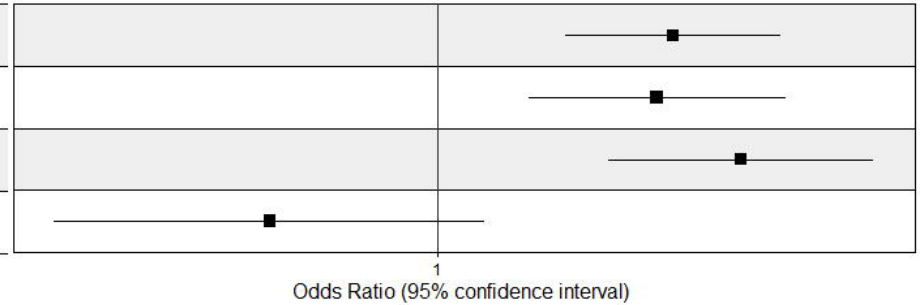
Smoking status: Current	OR	CI
Inverse variance weighted	1.53	[1.138,2.057]
Weighted median	1.595	[1.146,2.219]
Penalised weighted median	1.595	[1.15,2.21]
MR Egger	2.03	[0.611,6.744]



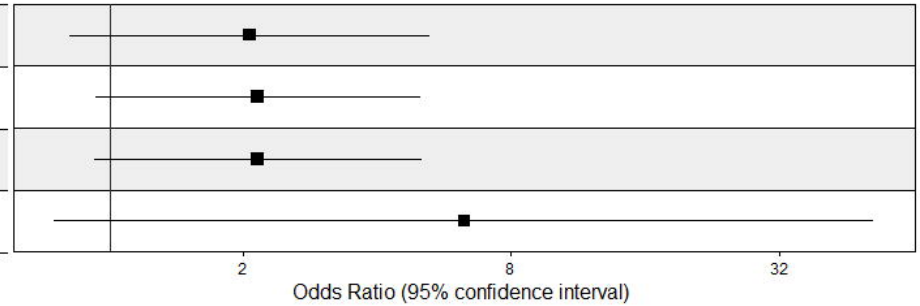
Smoking status: Never	OR	CI
Inverse variance weighted	0.84	[0.786,0.897]
Weighted median	0.839	[0.762,0.924]
Penalised weighted median	0.839	[0.762,0.924]
MR Egger	0.814	[0.589,1.124]



Alcohol intake frequency.	OR	CI
Inverse variance weighted	1.062	[1.033,1.091]
Weighted median	1.057	[1.023,1.092]
Penalised weighted median	1.08	[1.044,1.117]
MR Egger	0.958	[0.907,1.012]



Alcohol drinker status: Current	OR	CI
Inverse variance weighted	2.068	[0.812,5.267]
Weighted median	2.154	[0.928,4.999]
Penalised weighted median	2.154	[0.922,5.03]
MR Egger	6.268	[0.752,52.266]



Alcohol drinker status: Never	OR	CI
Wald ratio	0.283	[0.127,0.631]