

ARTICLE

Open Access

Polygenic risk score, healthy lifestyles, and risk of incident depression

Zhi Cao^{1,2}, Hongxi Yang^{1,3}, Yixuan Ye³, Yuan Zhang¹, Shu Li¹, Hongyu Zhao³ and Yaogang Wang¹ 

Abstract

Genetic factors increase the risk of depression, but the extent to which this can be offset by modifiable lifestyle factors is unknown. We investigated whether a combination of healthy lifestyles is associated with lower risk of depression regardless of genetic risk. Data were obtained from the UK Biobank and consisted of 339,767 participants (37–73 years old) without depression between 2006 and 2010. Genetic risk was categorized as low, intermediate, or high according to polygenic risk score for depression. A combination of healthy lifestyles factors—including no current smoking, regular physical activity, a healthy diet, moderate alcohol intake and a body mass index <30 kg/m²—was categorized into favorable, intermediate, and unfavorable lifestyles. The risk of depression was 22% higher among those at high genetic risk compared with those at low genetic risk (HR = 1.22, 95% CI: 1.14–1.30). Participants with high genetic risk and unfavorable lifestyle had a more than two-fold risk of incident depression compared with low genetic risk and favorable lifestyle (HR = 2.18, 95% CI: 1.84–2.58). There was no significant interaction between genetic risk and lifestyle factors (*P* for interaction = 0.69). Among participants at high genetic risk, a favorable lifestyle was associated with nearly 50% lower relative risk of depression than an unfavorable lifestyle (HR = 0.51, 95% CI: 0.43–0.60). We concluded that genetic and lifestyle factors were independently associated with risk of incident depression. Adherence to healthy lifestyles may lower the risk of depression regardless of genetic risk.

Introduction

Depression is a major cause of death in the UK population, and incidence rate has been on the rise, as 19.7% UK people aged 16 years and older showed symptoms of depression or anxiety in 2014¹. According to data from the 2014 National Health Service (NHS), 39% of individuals with a diagnosis of depression or anxiety were taking related medications². Depression is a complex disease that is affected by both genetic and environmental factors. The effect of genetic predisposition on the risk of depression has been examined in the previous studies^{3,4}. A meta-analysis showed that the heritability for depression was 37% (95% CI: 31–42%), and data from family studies have

reported a two- to three-fold increase in the risk of depression in the first-degree offspring of patients with depression⁵. In the last ten years, several genome-wide association studies (GWASs) have identified polymorphisms associated with the development of depression^{6–8}. These single nucleotide polymorphisms (SNPs) provided quantitative measures of genetic susceptibilities and may be predictive of the incidence of depression⁹.

Lifestyle is a common and crucial modifiable risk factor for depression. A large number of evidence has established that many individual lifestyle factors—such as non-smoking¹⁰, a healthy diet¹¹, physical activity¹², moderate alcohol intake¹³, and a healthy weight¹⁴—are associated with decreased risk of depression. However, because many of the lifestyle factors often coexist, exploring the joint impact of these lifestyle behaviors on the risk of depression is extremely relevant. Yet, there is limited study on this topic. As for other diseases, recent studies have shown that adherence to a combined healthy

Correspondence: Yaogang Wang (wangyg@tmu.edu.cn)

¹School of Public Health, Tianjin Medical University, Tianjin, China

²Department of Big Data in Health Science, School of Public Health, Zhejiang University School of Medicine, Hangzhou, China

Full list of author information is available at the end of the article

These authors contributed equally: Zhi Cao, Hongxi Yang

© The Author(s) 2021



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

lifestyles plays a key role in attenuating the impacts of genetic risks of several chronic diseases, including coronary heart disease¹⁵, stroke¹⁶, hypertension¹⁷, diabetes¹⁸, dementia^{19,20}, breast cancer²¹, and colorectal cancer^{22,23}. However, to date, no studies have examined whether an advantageous combined lifestyle factors attenuate on the risk of developing depression.

Therefore, in the present study, to further evaluate depression risk and lifestyle factors, we created a combination of healthy lifestyle scores based on recommendations of five potentially modifiable lifestyle factors: smoking, diet, physical activity, alcohol intake, and body fat. Our present study included the following aims: (1) to explore the associations of genetic risk and combined lifestyle factors with the risk of incident depression; and (2) to investigate whether adherence to a healthy lifestyle might counteract genetic risk for depression, using data from a large-scale cohort of the UK population.

Methods

Study design and population

Our study derived from UK Biobank, a prospective population-based cohort study. The UK Biobank recruited 502,528 adults (37–73 years old) from the general population between 2006 and 2010. Participants came from 22 assessment centers across England, Scotland, and Wales, where they completed touchscreen and nurse-led questionnaires, had physical measurements taken, and provided biological samples²⁴. The exposures of interest in our study included imputed genetic data and lifestyle (alcohol intake, smoking, diet, body mass index [BMI], and physical activity). In the current study, we included all participants who were classified as white British and without a history of depression that were self-reported and/or medical records, and with complete data on their genetics and lifestyle factors.

Lifestyle factors

We defined the following five healthy lifestyle factors on basis of the American Heart Association (AHA) guidelines: moderate alcohol intake, no current smoking; healthy diet; BMI < 30 kg/m² and regular physical activity^{16,25}.

In our study, a healthy diet was determined according to the increased consumption of fruit, vegetables, and fish, as well as the decreased consumption of processed meats and red meats. A healthy diet was defined if a individual meets at least two of the healthy food items listed above. Supplemental text S1 provides additional details on our construction of a healthy diet score. Moderate alcohol intake was defined as 0–28 g/d for men and 0–4 g/d for women. Moderate physical activity was defined as at least 150 min of moderate-intensity activity weekly or 75 min of vigorous activity weekly.

Participants scored 1 point if they meet an healthy lifestyle factor. The total lifestyle scores ranged from 0 to 5, with higher scores meaning a healthier lifestyle. Lifestyle behaviors were further categorized as favorable (4 or 5 healthy lifestyle factors), intermediate (2 or 3), and unfavorable (0 or 1) lifestyles. Additionally, a weighted-standardized healthy lifestyle score was established based on the β coefficient of each lifestyle factor in the Cox model with all five lifestyle factors, as well as adjustments for other covariates. The initial dichotomy lifestyle factors were multiplied by the β coefficients and were then summed, divided by the sum of the β coefficients, and multiplied by 100. According to the distribution of unweighted lifestyle scores, the weighted-standardized lifestyle scores were grouped by favorable, intermediate, and unfavorable²⁶.

Polygenic risk score

Polygenic risk score (PRS) was constructed to assess the cumulative genetic risk for depression. The score was constructed based on the GWAS summary-statistic data of depression in the Genetic Epidemiology Research on Adult Health and Aging (GERA)²⁷. The summary-statistic data were based on people of European ancestry; therefore, the present study was restricted to European samples in the UK biobank. The posterior effect sizes of SNPs on depression were inferred using PRS-CS, a polygenic prediction method²⁸. An individual-level polygenic score was defined as the sum of the number of risk alleles present at each SNP weighted by the corresponding posterior effect sizes across all available SNPs ($N=1,065,182$) in the UK biobank, and it was produced using the plink “-score” command. The PRS for all individuals in UK Biobank were z-standardized and then categorized into low, intermediate, and high risks.

Incidence of depression

Depression was ascertained using hospital inpatient records containing data on admissions and diagnoses. Participants were followed up until the date of incident depression, death or end of study (March 31, 2017). Diagnoses were recorded using the International Classification of Diseases (version 10; code ICD-10) coding system. Participants with depression were identified as having a primary/secondary diagnosis (hospital records) using ICD-10 codes.

Covariates

All analyses were adjusted for age, sex, qualifications (college degree; A levels/AS levels, O levels/General Certificate of Secondary Education (GCESs); Certificate of Secondary Education (CSEs); National Vocational Qualification (NVQ); Higher National Diploma (HND) or Higher National Certificate (HNC); other professional

qualifications), socioeconomic status (categories derived from Townsend deprivation index, combining information on social class, employment, car availability, and housing), and the first ten principal components of ancestry.

Statistical analyses

Baseline characteristics of the participants were summarized across depression status as the percentage for categorical variables and the mean and standard deviation (SD) for continuous variables. Person-year were calculated from the date of recruitment to date of developing depression or censoring time (whichever event occurred first). Incident rates and 95% confidence intervals (95% CIs) per 1000 person-years were estimated by the exact Poisson test, and analyses were stratified by genetic risk.

Cox proportional hazard regression models were utilized to explore the association of PRS, lifestyle, and combined genetic and lifestyle categories with risk of depression. The proportional hazard assumption was checked by tests based on Schoenfeld residuals, and the results suggested that the assumptions had not been violated (Supplemental Fig. S1). The primary model was adjusted for age, sex, socioeconomic status, qualifications, and the first 10 principal components of ancestry. Then, we fitted an additional model such that PRS and lifestyle category were adjusted for one another. To further investigate whether and to what extent healthy lifestyle practices may counteract the harmful effect of high genetic risk on incident depression, the associations between lifestyle factors and depression stratified by genetic risk were explored. An interaction term in the Cox model was included to test for a statistical interaction between lifestyle and genetic risk in relation to depression.

Sensitivity analysis

We undertook a series of analyses to evaluate the robustness of our findings. First, the risk of depression was examined by using a weighted lifestyle score. Second, potential differences in risk of depression were explored by sociodemographic factors in subgroup analyses, including age (middle-aged [37–59 years old] or older adults [≥ 60 years old]), sex (male, female) and socioeconomic status (low, intermediate, high). Third, to minimize the possibility of spurious associations due to reverse causation, the associations of genetic risk and lifestyle with depression were re-analyzed after excluding the people who died during the first two years of follow-up. Fourth, waist-to-hip ratio (WHR) is often used as an indicator of body fat, BMI was replaced with WHR to include the lifestyle criteria. Fifth, a competing risk regression was also fitted using the sub-distribution method proposed by Fine and Grey, with the competing event being the death from all causes other than

developing depression. Finally, E-value was calculated to assess the effect of unmeasured confounding. The E-value evaluates the minimum strength that an unmeasured confounding variable that must offset the observed association of exposure with outcomes when considering all other measured covariates²⁹. The E-value contributed to evaluate the plausibility of unmeasured confounding and towards the evidence for causality³⁰.

To maximize the likelihood of reporting true findings, we set the α at 0.05 and used Bonferroni correction to adjust for multiple testing. We considered two-sided P values less than 0.05 (P value of less than 0.05 divided by the number of tests, i.e., 0.05/3) statistically significant. All analyses were performed using STATA 15 statistical software (StataCorp).

Results

At baseline, 502,528 participants were assessed. After excluding non-White participants ($n = 29,811$) with prevalent depression ($n = 29,239$), as well as those without genetic information ($n = 12,709$) or lifestyle information ($n = 90,648$), a total of 339,767 participants was finally included in our study. Baseline characteristics of the participants are provided in Table 1. During a total of 2,722,320 person-year (a median follow-up of 8.1 years), there were 5739 cases of developing depression. PRS was normally distributed (Fig. 1a), and almost half of participants engaged in 4 or 5 (43.6%) of the 5 healthy lifestyle factors (Fig. 1b).

Table 2 presents the incidence rates and HRs of depression across genetic risk categories. The incidence rates per 1000 person-years were 1.88 (1.79–1.97), 2.13 (2.04–2.23), and 2.32 (2.22–2.42) in participants with low, intermediate, and high genetic risk, respectively. The cumulative incidence rate of depression raised with increasing genetic risk category (log-rank test < 0.001) (Supplementary Fig. S2). The risk of incident depression also increased monotonically across genetic risk categories (P for trend < 0.001). The HRs of depression were 1.13 (95% CI: 1.06–1.20) and 1.22 (95% CI: 1.14–1.30) for intermediate and high genetic risk categories compared to those with low genetic risk after adjustment for potential confounders. Additional adjustments for lifestyle factors did not essentially change these results, indicating that genetic risk for depression was statistically independent of lifestyle factors. Restricted spline curve showed a linear and dose–response association between PRS and depression (P for non-linear = 0.395), with HR being 1.09 (95% CI: 1.06–1.12) per 1-SD increment (Supplementary Fig. S3).

The incidence rate and risk of depression also increased monotonically across lifestyle categories (Table 2). Compared with a favorable lifestyle, an unfavorable lifestyle was associated with a high risk of depression (HR = 1.81,

Table 1 Characteristics of participants at baseline.

Characteristics	All (N = 339,767)	Incident depression (N = 5,739)	No depression (N = 334,028)	P value
Age, mean (SD)	56.6 (8.1)	56.8 (8.2)	56.6 (8.1)	0.176
Sex				<0.001
Male	161,463 (47.5)	2229 (38.8)	159,234 (47.7)	
Female	178,304 (52.5)	3510 (61.2)	174,794 (52.3)	
Townsend deprivation index, mean (SD)	-1.62 (2.89)	-0.85 (3.25)	-1.63 (2.88)	<0.001
Qualifications				<0.001
College degree	118,288 (34.8)	1417 (24.7)	116,871 (35.0)	
A levels/AS levels	39,459 (11.6)	587 (10.2)	38,872 (11.6)	
O levels/GCESs	72,547 (21.4)	1,293 (22.5)	71,254 (21.3)	
CSEs	17,629 (5.19)	413 (7.2)	17,216 (5.2)	
NVQ or HND or HNC	22,505 (6.6)	426 (7.4)	22,079 (6.6)	
Other	17,475 (5.1)	322 (5.6)	17,153 (5.1)	
None of the above	49,616 (14.6)	1228 (21.4)	48,388 (14.5)	
Healthy lifestyle factors				
No current smoking	308,615 (90.8)	4776 (83.2)	303,839 (90.9)	<0.001
Healthy diet	149,811 (44.1)	2531 (44.1)	147,280 (44.1)	0.988
BMI < 30 kg/m ²	266,971 (78.6)	3971 (69.2)	263,000 (78.7)	<0.001
Regular physical activity	211,798 (62.3)	3453 (60.2)	208,345 (62.4)	0.001
Moderate alcohol intake	174,359 (51.3)	2535 (44.2)	17,1824 (51.4)	<0.001
Healthy lifestyle scores				<0.001
0 or 1 healthy lifestyle	17,655 (5.2)	498 (8.7)	17,157 (5.1)	
2 or 3 healthy lifestyles	174,175 (51.3)	3238 (56.4)	170,937 (51.2)	
4 or 5 healthy lifestyles	147,937 (43.5)	2003 (34.9)	145,934 (43.7)	
Genetic risk category				<0.001
Low	113,333 (33.4)	1707 (29.7)	111,626 (33.4)	
Intermediate	113,197 (33.3)	1931 (33.7)	111,266 (33.3)	
High	113,237 (33.3)	2101 (36.6)	111,136 (33.3)	

Values are numbers (participants) unless stated otherwise.

95% CI: 1.64–2.00). Likewise, additional adjustment for genetic risk resulted in an HR of 1.83 (95% CI: 1.65–2.02), in line with the independence of genetic and lifestyle factors. A risk gradient was noted across the number of healthy lifestyle factors, such that the participants with no healthy lifestyle were at significantly higher risk of depression than those with five healthy lifestyle factors (Supplemental Table S1). Moreover, there were a dose–response relationship between the increasing lifestyle factors and depression, regardless of genetic risk (all *P* for trend <0.001) (Fig. 1c). In addition, we calculated correlations between individual lifestyle factors and found only weak correlations (*r* < 0.1) (Supplemental Table S2). WHR, rather than BMI, was included in lifestyle factors,

the association between lifestyle and depression was not essentially changed (Supplemental Table S3).

In both sex and age subgroup, the effect of high genetic risk on depression was stronger in female than male. Results were consistent in that genetic risk was monotonically associated with the risk of depression in subgroups of age. A monotonic association between lifestyle and depression was also observed for both sex and age subgroups (Supplemental Table S4). The modification effect of socioeconomic status on the association of lifestyle and genetic risk with depression were also explored, we found that high genetic risk was not significantly associated with depression in high socioeconomic status category, the beneficial effect of lifestyle on depression

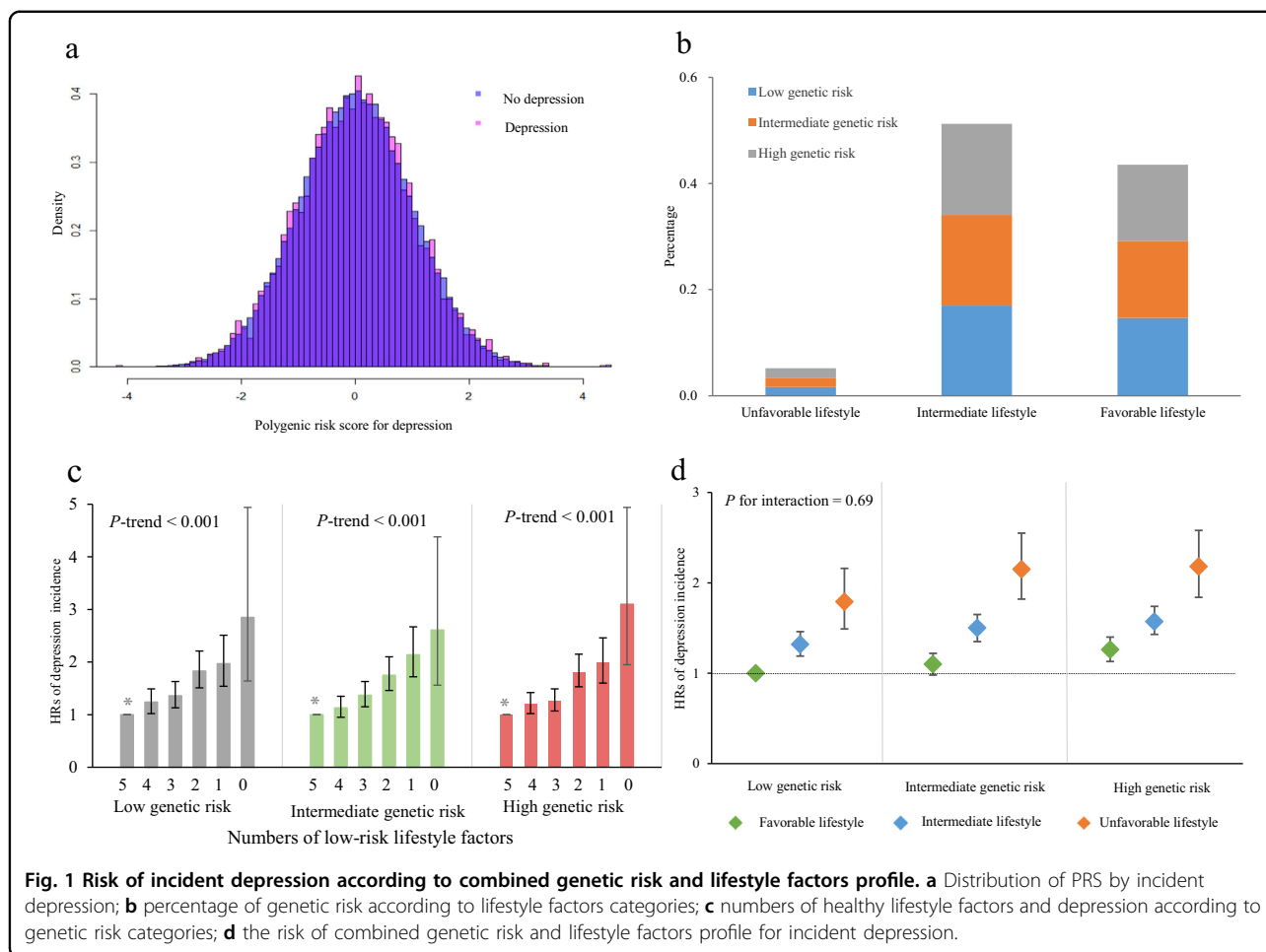


Fig. 1 Risk of incident depression according to combined genetic risk and lifestyle factors profile. **a** Distribution of PRS by incident depression; **b** percentage of genetic risk according to lifestyle factors categories; **c** numbers of healthy lifestyle factors and depression according to genetic risk categories; **d** the risk of combined genetic risk and lifestyle factors profile for incident depression.

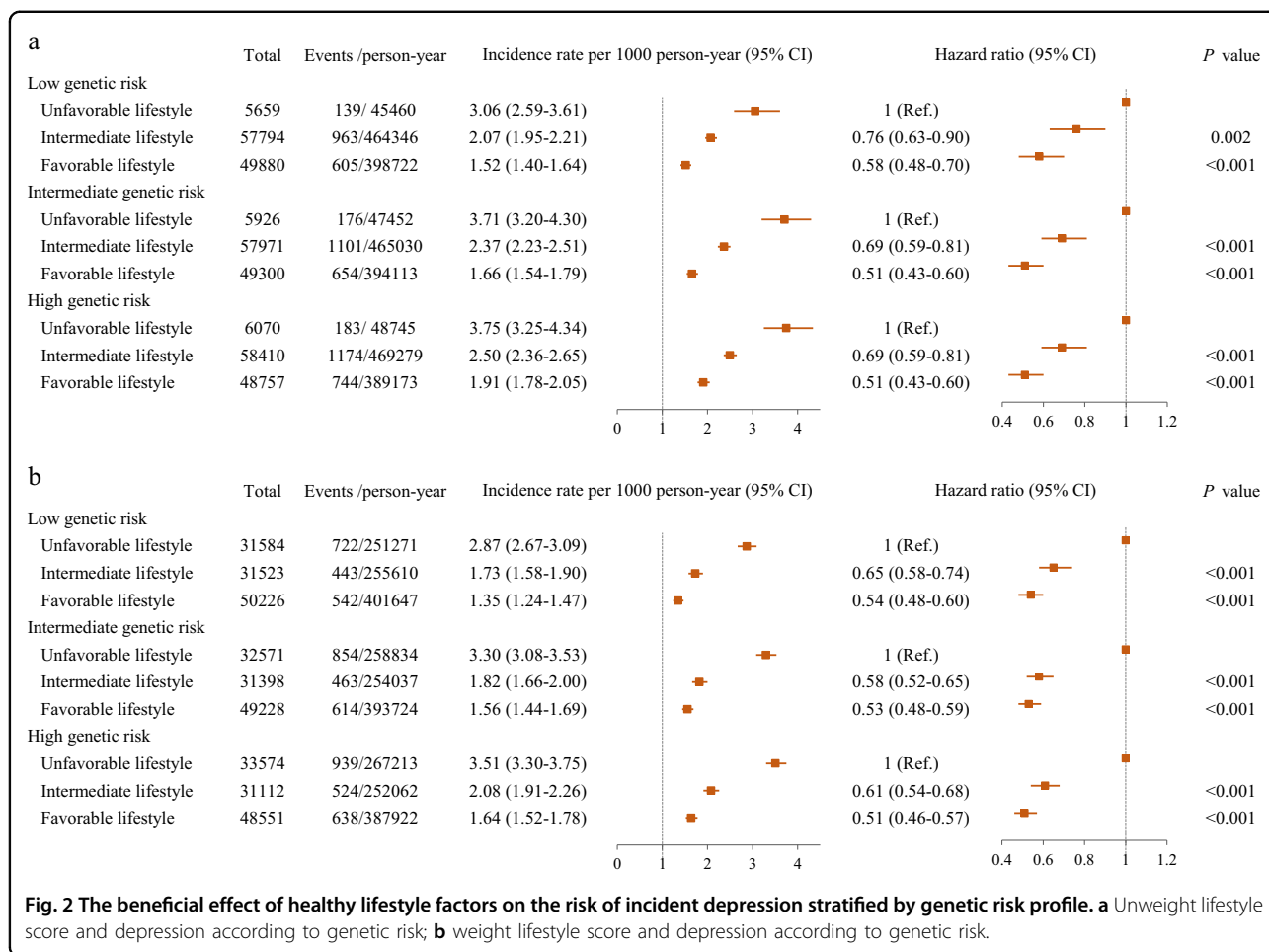
Table 2 Risk of incident depression according to genetic risk and lifestyle categories.

	Events/person-year	Incidence rate per 1000 person-year	Model 1 ^a	Model 2 ^b
Genetic risk				
Low	1707/908,528	1.88 (1.79–1.97)	1 (Ref.)	1 (Ref.)
Intermediate	1931/906,595	2.13 (2.04–2.23)	1.13 (1.06–1.20)	1.12 (1.05–1.20)
High	2101/907,197	2.32 (2.22–2.42)	1.22 (1.14–1.30)	1.22 (1.14–1.30)
<i>P</i> for trend ^c			<0.001	<0.001
Lifestyle categories				
Favorable	2003/1182,008	1.69 (1.62–1.77)	1 (Ref.)	1 (Ref.)
Intermediate	3238/1398,655	2.32 (2.24–2.40)	1.30 (1.23–1.38)	1.31 (1.24–1.39)
Unfavorable	498/141,658	3.52 (3.22–3.84)	1.81 (1.64–2.00)	1.83 (1.65–2.02)
<i>P</i> for trend ^c			<0.001	<0.001

^aModel 1 was adjusted for age, sex, qualifications, socioeconomic status and first 10 principal components of ancestry.

^bModel 2 was adjusted for Model 1 and lifestyle categories (for genetic risk model) and genetic risk (for lifestyle factors model).

^c*P* value for trend was calculated treating the genetic risk score or lifestyle factors as a continuous variable.



was stronger in high socioeconomic status than lower socioeconomic status (P for interaction <0.05) (Supplemental Table S5).

When we conducted analyses of combined genetic risk and lifestyle categories, there was an overall monotonic association with an increasing genetic risk and an unfavorable lifestyle (Fig. 1d). People with a high genetic risk and an unfavorable lifestyle had a highest risk of depression compared with participants with low genetic risk and a favorable lifestyle (HR = 2.18, 95% CI: 1.84–2.58). No significant interaction between lifestyle and genetic risk in relation to depression was observed (P for interaction = 0.69) (Supplemental Table S6), meaning that the association with lifestyle factors did not vary substantially across genetic risk categories. Sensitivity analyses showed similar results among participants with at least two years of follow-ups (Supplemental Table S7). The competing risk analysis was also in accordance with the main analysis on the association of combined genetic risk and lifestyle factors with the incident depression (Supplementary Table S8).

Further analyses stratified by genetic risk category, confirmed that a favorable lifestyle was related to a lower

risk of depression across genetic risk categories (Fig. 2a). Among participants who had a low genetic risk of depression, intermediate and favorable lifestyles were associated with 24% (HR = 0.76, 95% CI: 0.63–0.90) and 42% (HR = 0.58, 95% CI: 0.48–0.70) lower risks of depression, respectively. Among those at high genetic risk, intermediate and favorable lifestyles separately reduced the risk of depression by 31% (HR = 0.69, 95% CI: 0.59–0.81) and 49% (HR = 0.51, 95% CI: 0.43–0.60), respectively. These results were not essentially altered much when weight lifestyle scores were used (Fig. 2b). In addition, among individual components of lifestyle factors, no current smoking and non-obesity contributed most to reduce the risk of developing depression, regardless of genetic risk (Table 3). The E-values analysis showed that the point estimate and lower confidence bound were 1.74 and 1.54 for high genetic risk, 3.06 and 2.69 for unfavorable lifestyle, respectively. To offset the observed association, an unmeasured confounding variable must be correlated with both exposure and outcomes, at least the risk ratio of the E-value at each time point, and conditional on other confounding variables.

Table 3 Multivariable Cox regression analysis of individual lifestyle factors in relation to risk of depression, stratified by genetic risk.

Lifestyle factors	Low genetic risk	Intermediate genetic risk	High genetic risk
No current smoking	0.53 (0.47–0.61)	0.54 (0.48–0.62)	0.56 (0.49–0.63)
BMI < 30 kg/m ²	0.69 (0.62–0.76)	0.66 (0.60–0.73)	0.63 (0.57–0.69)
Healthy diet	1.02 (0.93–1.13)	0.97 (0.89–1.07)	1.07 (0.98–1.17)
Regular physical activity	0.88 (0.80–0.97)	0.90 (0.82–0.98)	0.95 (0.87–1.04)
Moderate alcohol intake	0.93 (0.85–1.03)	0.92 (0.84–1.00)	0.85 (0.78–0.93)

The analyses were adjusted for age, sex, qualifications, socioeconomic status and first 10 principal components of ancestry, individual lifestyle factor was adjusted for each other.

Discussion

In this large-scale prospective study, we found that genetic risk and lifestyle factors were independently associated with the risk of depression. Participants with high genetic risk and unfavorable lifestyle profile had a more than two-fold increased risk of incident depression compared with that of participants with a low genetic risk of depression and a favorable lifestyle. No significant interaction between genetic risk and healthy lifestyle was observed. The beneficial impact of healthy lifestyle on depression was found across all genetic risk categories, suggesting that the genetically predetermined rise in depression can be counteracted, at least to some extent, by a healthy lifestyle. These findings provide evidence for the necessity of a combination of healthy lifestyle factors in preventing depression and reinforce the considerable potential of primary prevention.

Our present study adds to previously reported studies on individual lifestyle factors and risk of depression. Although previous studies found that individual lifestyle factor, such as non-smoking¹⁰, a healthy diet¹¹, physical activity¹², moderate alcohol intake¹³, and a healthy weight¹⁴, was related to depression, the samples of these studies were relatively smaller. Moreover, our present study is the first to incorporate information of PRS. Recently, GWASs have identified several common genetic loci that are significantly related to risk of depression^{31,32}. The contribution of an inherited genetic predisposition to the occurrence of sporadic depression has also previously been reported. A clinical and epidemiological study of a youth cohort suggested that PRS may serve as an early indicator of clinically significant levels of depression⁹. A recent study also showed that PRS was associated with depression³³. Similar to this previous study, our present study demonstrated that a higher genetic risk increased the risk of depression by 22%.

A reduced risk of depression by adherence to a healthy lifestyle have been well documented^{34,35}. To the best of our knowledge, however, only a limited number of studies have examined the associations of combined healthy

lifestyle factors with risk of developing depression in a general UK population. The risk reduction linked to healthy lifestyle in the current study was consistent across all strata of genetic risks, which emphasizes the benefits for entire populations in adherence to a healthy lifestyle, regardless of genetic risk. Among these lifestyle factors, the most statistically significant associations were discovered for no current smoking and for a BMI < 30 kg/m². These findings further supplemented the current recommendation that people with high genetic predisposition of depression should clinically practice a more healthier lifestyle factors to maintain a lower risk of developing depression.

The precise mechanisms by which the combination of lifestyle and genetic factors affect the risk for depression remain unclear. The plausible biological effects of each factor are pleiotropic in nature and likely involve overlapping influences via pathways that were related to the development of depression. Common genetic variants related to depression might have an effect on both personality and psychological regulation³⁶. Both neuroendocrine disruption and inflammation have also been associated with the etiology of depression. Obesity, poor diet, exposure to chemicals or pollutants, and high stress levels may potentially break the hypothalamic-pituitary-adrenal axis and increase cortisol, low-grade systemic inflammation, and oxidative stress^{36,37}. Although the judicious use of medications and psychological techniques was recommended, the emerging evidence encouraged a more integrative approach for managing depression and acknowledged that modifiable risk factors, such as lifestyle intervention, should be a routine part for treatments and preventative efforts.

Strengths and limitations

The strengths of our present study included its large sample size of UK Biobank participants, its prospective design, information on a wide-range of depression-associated SNPs and the use of standardized protocols for data collection. Another distinctive feature of our

study compared with prior studies is that all SNPs associated with depression were included to calculate PRS. Despite these strengths, several limitations of the current study need to be considered. First, behavioral changes before or after baseline examinations may affect our results. Second, although all models were adjusted for known potential biases and participants were followed up for a median of 8.1 years, it is possible that unmeasured confounds and reverse causation remained. Third, depression was diagnosed by hospital inpatient records in our study, although diagnosed by a doctor is a more common and precise way. Milder depression or severely depressed people who very hardly go to the hospital unless they are in danger of life may easily be underestimated. Finally, this study was primarily restricted to participants of European ancestry aged 37–73 years at baseline and, therefore, further research is warranted to examine to what degree these results generalize to other populations.

Implications for further research

We found that healthy lifestyle factors were associated with a substantially lower risk of depression regardless of genetic risk, suggesting that genetic risk might be attenuated by adherence to a healthy lifestyle. Our results provide encouraging support of lifestyle recommendations for depression prevention that highlight a combination of healthy lifestyle factors for everyone. From a public health viewpoint, our findings should be used to strengthen the importance of primary prevention to policymakers, as well as to support population-wide efforts to lower the risk of incident depression through combined lifestyle modifications. Future research is needed to investigate the potential mechanism of depression in relation to lifestyle factors and genetic risk.

Acknowledgements

We thank the participants of the UK Biobank. This research has been conducted using the UK Biobank Resource under the project number of 45676. This study was supported by the National Natural Science Foundation of China (Grant No. 91746205, grants No. 71910107004).

Author details

¹School of Public Health, Tianjin Medical University, Tianjin, China. ²Department of Big Data in Health Science, School of Public Health, Zhejiang University School of Medicine, Hangzhou, China. ³Department of Biostatistics, School of Public Health, Yale University, New Haven, CT, USA

Author contributions

Z.C. and H.X.Y. contributed equally to this study. Y.G.W. and Z.C., and H.X.Y. conceived the study. Z.C. and H.X.Y. wrote the first and successive drafts of the manuscript. Z.C., H.X.Y., Y.X.Y., and S.L. analyzed the data. Z.C., Y.X.Y., S.L., Y.Z., and X.Y.Z. had full access to the data and take responsibility for the integrity of the data and the accuracy of the data analysis. Y.G.W. and H.Y.Z. reviewed the manuscript. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Data availability

The data that support the findings of this study are available from UK Biobank project site, subject to successful registration and application process. Further details can be found at <https://www.ukbiobank.ac.uk/>.

Conflict of interest

The authors declare no competing interests.

Ethics approval

All participants gave informed consent for data provision and linkage. UK Biobank has full ethical approval from the NHS National Research Ethics Service (16/NW/0274).

Publisher's note

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

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41398-021-01306-w>.

Received: 27 October 2020 Revised: 16 February 2021 Accepted: 24 February 2021

Published online: 29 March 2021

References

1. Foundation, M.H. *Mental Health Statistics: Depression*. (Mental Health Foundation, London, 2018).
2. Mcmanus, S. B. P., Jenkins, R. & Brugha, T. *Mental Health and Wellbeing in England Adult Psychiatric Morbidity Survey 2014: Executive Summary*. (Leeds NHS Digital, Leeds, England, 2016).
3. Howard, D. M. et al. Genetic stratification of depression in UK Biobank. *Transl. Psychiatry* **10**, 163 (2020).
4. Shen, X. Y. et al. A phenome-wide association and Mendelian Randomisation study of polygenic risk for depression in UK Biobank. *Nat. Commun.* **11**, 2301 (2020).
5. Sullivan, P. F., Neale, M. C. & Kendler, K. S. Genetic epidemiology of major depression: review and meta-analysis. *Am. J. Psychiatry* **157**, 1552–1562 (2000).
6. Shadrina, M., Bondarenko, E. A. & Slominsky, P. A. Genetics factors in major depression disease. *Front. Psychiatry* **9**, 334 (2018).
7. Hall, L. S. & Adams, M. J. Genome-wide meta-analyses of stratified depression in Generation Scotland and UK Biobank. *Transl. Psychiatry* **8**, 9 (2018).
8. Amare, A. T. et al. Bivariate genome-wide association analyses of the broad depression phenotype combined with major depressive disorder, bipolar disorder or schizophrenia reveal eight novel genetic loci for depression. *Mol. Psychiatry* **25**, 1420–1429 (2019).
9. Halldorsdottir, T. et al. Polygenic risk: predicting depression outcomes in clinical and epidemiological cohorts of youths. *Am. J. Psychiatry* **176**, 615–625 (2019).
10. Wootton, R. E. et al. Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. *Psychol. Med* **50**, 1–9 (2019).
11. Masana, M. F. et al. Mediterranean diet and depression among older individuals: the multinational MEDIS study. *Exp. Gerontol.* **110**, 67–72 (2018).
12. Choi, K. W. et al. Assessment of bidirectional relationships between physical activity and depression among adults: a 2-sample Mendelian randomization study. *JAMA Psychiatry* **76**, 399–408 (2019).
13. Gémes, K., Forsell, Y., Janszky, I. & László, K. D. Moderate alcohol consumption and depression—a longitudinal population-based study in Sweden. *Acta Psychiatr. Scand.* **139**, 526–535 (2019).
14. Muhlig, Y., Antel, J., Fockler, M. & Hebebrand, J. Are bidirectional associations of obesity and depression already apparent in childhood and adolescence as based on high-quality studies? A systematic review. *Obes. Rev.* **17**, 235–249 (2016).
15. Khara, A. V. et al. Genetic Risk, Adherence to a Healthy Lifestyle, and Coronary Disease. *N. Engl. J. Med* **375**, 2349–2358 (2016).
16. Rutten-Jacobs, L. C. et al. Genetic risk, incident stroke, and the benefits of adhering to a healthy lifestyle: cohort study of 306 473 UK Biobank participants. *BMJ* **363**, k4168 (2018).

17. Pazoki, R. et al. Genetic predisposition to high blood pressure and lifestyle factors: associations with midlife blood pressure levels and cardiovascular events. *Circulation* **137**, 653–661 (2018).
18. Han, X., et al. Genetic risk, a healthy lifestyle, and type 2 diabetes: the Dongfeng-Tongji cohort study. *J. Clin. Endocrinol. Metab.* **105**, dgz325 (2020).
19. Lourida, I. et al. Association of lifestyle and genetic risk with incidence of dementia. *JAMA* **322**, 430–437 (2019).
20. Licher, S., Ahmad, S., Karamujic-Comic, H., Voortman, T. & Leening, M. J. G. Genetic predisposition, modifiable-risk-factor profile and long-term dementia risk in the general population. *Nat. Med.* **25**, 1364–1369 (2019).
21. Arthur, R. S., Wang, T., Xue, X., Kamensky, V. & Rohan, T. E. Genetic factors, adherence to healthy lifestyle behavior, and risk of invasive breast cancer among women in the UK Biobank. *J. Natl Cancer Inst.* **112**, 893–901 (2020).
22. Carr, P. R. et al. Healthy Lifestyle Factors Associated With Lower Risk of Colorectal Cancer Irrespective of Genetic Risk. *Gastroenterology* **155**, 1805–1815. e1805 (2018).
23. Cho, Y. A. et al. Genetic Risk Score, Combined Lifestyle Factors and Risk of Colorectal Cancer. *Cancer Res Treat.* **51**, 1033–1040 (2019).
24. Sudlow, C. et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
25. Benjamin, E. J. et al. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. *Circulation* **135**, e146–e603 (2017).
26. Jiao, L. et al. A combined healthy lifestyle score and risk of pancreatic cancer in a large cohort study. *Arch. Intern Med* **169**, 764–770 (2009).
27. Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
28. Ge, T. & Chen, C. Y. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* **10**, 1776 (2019).
29. VanderWeele, T. J. & Ding, P. Sensitivity analysis in observational research: introducing the E-value. *Ann. Intern Med.* **167**, 268–274 (2017).
30. Haneuse, S., VanderWeele, T. J. & Arterburn, D. Using the E-Value to assess the potential effect of unmeasured confounding in observational Studies. *Jama* **321**, 602–603 (2019).
31. Wray, N. R. et al. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Mol. Psychiatry* **17**, 36–48 (2012).
32. Howard, D. M. & Adams, M. J. Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat. Commun.* **9**, 1470 (2018).
33. Shen, H., Gelaye, B. & Huang, H. Polygenic prediction and GWAS of depression, PTSD, and suicidal ideation/self-harm in a Peruvian cohort. *Neuropsychopharmacology* **45**, 1595–1602 (2020).
34. van Dammen, L. et al. A systematic review and meta-analysis of lifestyle interventions in women of reproductive age with overweight or obesity: the effects on symptoms of depression and anxiety. *Obes. Rev.* **19**, 1679–1687 (2018).
35. Garcia-Toro, M. et al. Adherence to lifestyle recommendations by patients with depression. *Rev. Psiquiatr Salud Ment.* **5**, 236–240 (2012).
36. Sarris, J., O'Neil, A., Coulson, C. E., Schweitzer, I. & Berk, M. Lifestyle medicine for depression. *BMC Psychiatry* **14**, 107 (2014).
37. Maes, M. et al. New drug targets in depression: inflammatory, cell-mediated immune, oxidative and nitrosative stress, mitochondrial, antioxidant, and neuroprogressive pathways. And new drug candidates-Nrf2 activators and GSK-3 inhibitors. *Inflammopharmacology* **20**, 127–150 (2012).