Associations of Combined Genetic and Lifestyle Risks with Incident Type 2 Diabetes in the UK Biobank

Chi Zhao[1], Konstantinos Hatzikotoulas[2], Raji Balasubramanian, ScD[1], Elizabeth Bertone-Johnson, ScD[1], Na Cai, PhD[3,4,5], Lianyun Huang[3,4,5], Alicia Huerta-Chagoya[6], Margaret Janiczek[1], Chaoran Ma, MD, PhD[7], Ravi Mandla[8,9], Amanda Paluch, PhD[10], Nigel W Rayner[2], Lorraine Southam[2], Susan R. Sturgeon, DrPH[1], Ken Suzuki, PhD, MD[12,13], Henry J Taylor[14,15,16], Nicole VanKim, PhD[1], Xianyong Yin[17,18], Chi Hyun Lee, PhD[19], Francis Collins, PhD, MD[20], Cassandra N. Spracklen, PhD[1]

[1]Department of Biostatistics and Epidemiology, University of Massachusetts Amherst, Amherst, MA, USA
[2]Institute of Translational Genomics, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany
[3]Helmholtz Pioneer Campus, Helmholtz Munich, Neuherberg, Germany
[4]Computational Health Centre, Helmholtz Munich, Neuherberg, Germany
[5]School of Medicine and Health, Technical University of Munich, Munich, Germany
[6]Programs in Metabolism and Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA
[7]Department of Nutrition, University of Massachusetts Amherst, Amherst, MA, USA
[8]Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA
[9]Graduate Program in Genomics and Computational Biology, University of Pennsylvania, Philadelphia, PA, USA
[10]Department of Kinesiology, Institute for Applied Life Sciences, University of Massachusetts Amherst, Amherst, MA, USA
[12]Department of Diabetes and Metabolic Diseases, Graduate School of Medicine, University of Tokyo, Tokyo, Japan
[13]Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita, Japan
[14]Center for Precision Health Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA
[15]British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
[16]Heart and Lung Research Institute, University of Cambridge, Cambridge, UK
[17]Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing, China
[18]Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA
[19]Department of Applied Statistics, Yonsei University, Seoul, South Korea
[20]National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

[*]Corresponding author:
Cassandra N. Spracklen

Department of Biostatistics and Epidemiology, 715 North Pleasant Street, University of Massachusetts Amherst, Amherst, MA, USA
Phone: (413) 545-2379
Email: cspracklen@umass.edu

Abstract

Background

Type 2 diabetes (T2D) results from a complex interplay between genetic predisposition and lifestyle factors. Both genetic susceptibility and unhealthy lifestyle are known to be associated with elevated T2D risk. However, their combined effects on T2D risk are not well studied. We aimed to determine whether unhealthy modifiable health behaviors were associated with similar increases in the risk of incident T2D among individuals with different levels of genetic risk.

Methods

We performed a genetic risk score (GRS) by lifestyle interaction analysis within 332,251 non-diabetic individuals at baseline from the UK Biobank. Multi-ancestry GRS were calculated by summing the effects of 783 T2D-associated variants and ranked into tertiles. We used baseline self-reported data on smoking, BMI, physical activity level, and diet quality to categorize participants as having a healthy, intermediate, or unhealthy lifestyle. Cox proportional hazards regression models were used to generate adjusted hazards ratios (HR) of T2D risk and associated 95% confidence intervals (CI).

Results

During follow-up (median 13.6 years), 13,128 (4.0%) participants developed T2D. GRS ($P <$ 0.001) and lifestyle classification ($P < 0.001$) were independently associated with increased risk for T2D. Compared with healthy lifestyle, unhealthy lifestyle was associated with increased T2D risk in all genetic risk strata, with adjusted HR ranging from 7.11 (low genetic risk) to 16.33 (high genetic risk).

Conclusions

High genetic risk and unhealthy lifestyle were the most significant contributors to the development of T2D. Individuals at all levels of genetic risk can greatly mitigate their risk for T2D through lifestyle modifications.

Introduction

Type 2 diabetes (T2D) is among the leading causes of morbidity and mortality worldwide and has become one of the most challenging and concerning chronic diseases in public health. The prevalence of T2D has been steadily increasing over the past few decades: currently, > 28 million and > 500 million individuals have been diagnosed with T2D in the US and globally.[1,2] Additionally, the prevalence of T2D is higher in certain populations in the US, ranging from 7.5% in non-Hispanic Whites to 14.7% in Native Americans/Alaska Natives, and both Hispanic Americans and non-Hispanic Blacks are almost twice as likely to have T2D as non-Hispanic Whites.[3] The high prevalence and incidence of T2D results in a substantial disease burden, including treatment, reduced quality of life, disease complications, and premature death.

It is well-established that T2D is mainly caused by a complex interplay between genetic predisposition and lifestyle factors. Genome-wide association studies (GWAS) have identified more than 1,200 independent genetic variants associated with T2D, of which many appear to be related to insulin secretion and/or pancreatic β-cell development,[4,5] and explain approximately 20-40% of the overall heritability for T2D.[6] Lifestyle factors also play an important role in modulating T2D risk, and epidemiologic studies have identified increased T2D risk among individuals with a higher body mass index (BMI), low physical activity level, and unhealthy diet quality, as well as those who smoke.[7–10] It is also known that lifestyle interventions can reduce the risk of development of T2D and improve cardiovascular health, especially among individuals at high risk of T2D.[11]

Several studies have attempted to examine the potential joint effects between genetic risk and overall behavioral and/or lifestyle factors on the risk for T2D.[12–19] While most of the studies demonstrate the strongest T2D risk among those with the highest genetic risk and unhealthiest lifestyle factors, differences by biological sex were not often considered. Additionally, prior analyses were predominantly performed in East Asian-[14,15,17,18] or European-ancestry[11–13,16,19] individuals, limiting the generalizability to multi-ancestry and other non-European populations. Comparability across studies is further hindered by differences in behavioral and/or lifestyle factors considered as well as their classification and measurement. Finally, in terms of the GRS, the majority of the studies included fewer than 100 genetic variants in calculating the GRS,[11,14–19] used effect size weights from a mismatched ancestry population,[15,18] or used variants and effect size weights that were not independent of the study population,[12,17] all of which limit the accuracy in measuring the genetic risk of T2D.

Therefore, the goal of this study was to determine whether unhealthy modifiable health behaviors, as determined by the American Heart Association, were associated with similar increases in the risk of incident T2D among individuals with different levels of genetic risk, utilizing the most up-to-date list of independent T2D-associated genetic variants, across all individuals in the UK Biobank.

Methods

Data Source

Details of the UK Biobank (UKB) study design and population have been described elsewhere.[20–22] Briefly, the UKB is a population-based prospective cohort of > 500,000 participants designed to examine environmental, lifestyle, and genetic determinants of adult-onset diseases. Individuals aged 40-69 years old were recruited from 22 assessment centers throughout the United Kingdom from 2006-2010.[21] At enrollment, participants provided extensive information on their demographics, health, and lifestyle through baseline questionnaires, interviews, and physical assessments. Blood samples were collected for genotyping; genotypes were then imputed to a merged UK10K and 1000 Genomes phase 3 panel.[22] Follow-up of participants is ongoing through linked health records. The UKB study has approval from the North West Multi-center Research Ethics Committee, and all UKB participants provided written informed consent.

Study Population

The study population was comprised of individuals from the UKB for which complete data were available for their genotypes, lifestyle factors at enrollment (BMI, smoking status, diet, and physical activity), and covariates. Participants were excluded if they: had a mismatch between their genetic and self-reported sex (n = 372), were missing genotypes for ≥ 7.5% of the included variants (n = 498), had a BMI at enrollment < 18.5 kg/m^2 (n = 2,505), or had prevalent diabetes of any kind at enrollment (n = 15,250). To limit our analyses to those who developed incident T2D, we further excluded individuals who developed incident type 1 diabetes (T1D; ICD-10 E10), malnutrition-related diabetes (ICD-10 E12), or other specified diabetes (ICD-10 E13) (n = 699) during the follow-up period. Our final analytical sample size was n = 332,251 (Figure 1).

Genetic and Lifestyle Exposure groups

To estimate the genetic predisposition to T2D, genetic risk scores (GRS) were created following an additive model using 783 genome-wide significant variants identified from the most recent multi-ancestry GWAS meta-analysis (Table S1) after excluding results from the UKB, to avoid potential effect overestimation.[4] The GRS were calculated using a weighted method in which, for each variant, the number of T2D risk-increasing alleles a person has is multiplied by the effect size estimate from the multi-ancestry fixed-effect meta-analysis; using PLINK 1.9,[23] the products for each variant are then summed together for each individual into a continuous multi-ancestry GRS. GRS were divided into tertiles and categorized as low, moderate, and high genetic risk groups. We also generated ancestry-specific GRS for each individual using the effect size estimates from the global ancestry meta-analyses (African/African American, East Asian, European, South Asian) that most closely matched an individual's self-reported ethnic background; individuals for which their self-reported ethnic background was "mixed", "other", or missing were omitted from the ancestry-specific analyses. To avoid potential effectmoverestimation, index variants and variant effect sizes were generated without inclusion of the UKB data in the meta-analyses. Genetic variants identified in the meta-analyses but missing in the UKB were excluded from the GRS.

We summarized an individual's overall lifestyle into one of three categories (healthy, intermediate, and unhealthy) following the American Heart Association 2020 Strategic Impact Goal guidelines (termed ideal, intermediate, and poor, respectively) for smoking, BMI, and physical activity.[24] Dietary priorities for cardiometabolic health were used to define a healthy or unhealthy category for diet quality.[25] Definitions for healthy, intermediate, and unhealthy classifications for each of the lifestyle components can be found in the Supplementary Appendix (Table S2 and Table S3). Based on the categorization of the four lifestyle factors, we assigned participants to an overall lifestyle category: healthy (having at least 3 healthy lifestyle factors), unhealthy (having at least 3 unhealthy lifestyle factors), or intermediate (all other combinations).

Ascertainment of Incident T2D Outcomes

Incident cases of T2D within the UKB were identified using the first occurrences data. The first occurrences dataset indicates the first occurrence of any disease (mapped to ICD-10 diagnosis codes) from primary care, hospital inpatient, death register, and self-reported data.[21] For this study, we used the first record of ICD-10 diagnosis code E11 (type 2 diabetes mellitus) and the corresponding date to define incident cases.

Statistical Analysis

Descriptive statistics for participants were generated using baseline data and compared between censored observations and incident T2D cases using t-tests for continuous variables and chi-squared test for categorical variables. For this analysis, participants were followed from enrollment until diagnosis of T2D, death, lost to follow-up, or censoring date (by the time of analysis, the censoring date is 2022-10-31 for individuals in England, 2021-07-31 for individuals in Scotland, and 2018-02-28 for individuals in Wales), whichever came first. Several multivariable Cox regression models were used to test both the independent (both genetic risk and lifestyle as predictor variables) and joint associations of genetic risk and lifestyle groups with incident T2D; hazards ratios (HR) and associated 95% confidence intervals (95% CI) were calculated using individuals with low GRS and healthy lifestyle as the reference group. Additionally, we tested the independent association of genetic risk and all the individual lifestyle factors. The proportional hazards assumption was tested based on visualization of the survival probabilities over time and the scaled Schoenfeld residuals; the assumptions were met (Supplementary Figures S1-S10). Adjusted models included the following covariates: age at baseline, biological sex, years in education,[26] Townsend Deprivation Index (TDI),[27] income, and the first 16 genetic principal components (to adjust for population subtructure).[23] Sex-stratified analyses were also conducted including the same covariates except for biological sex. We also tested for statistical interaction between the GRS and lifestyle factors. Finally, we calculated the population attributable fraction (PAF), to evaluate the proportion of incident T2D that would have been prevented if participants with intermediate or unhealthy lifestyle had been in the healthy category. All analyses were additionally conducted using the ancestry-specific GRS that most closely matched the self-reported ethnic background. All statistical tests were two-sided, and P-values < 0.05 were considered statistically significant. All analyses were conducted using R 4.3.0.

Results

Baseline characteristics of study participants can be found in Table 1. Overall, the mean (SD) age was 55.19 (8.06) years, and 177,869 (54%) were female. During a median follow-up of 13.56 (IQR: 12.74-14.25) years, 13,128 (4%) participants developed incident T2D during a median time to onset of 7.98 years (IQR: 4.92, 10.82), with higher incidence rates among the high GRS tertile and unhealthy lifestyle classifications (0.75 per 1,000 person-years for those with the lowest GRS and healthy lifestyle to 12.53 per 1,000 person-year for those with the highest GRS and unhealthy lifestyle; Table 2; Supplementary Figure S11). At baseline, participants who later developed T2D were older, more likely to be male, have fewer years of education, lower income, and more severe social deprivation than those who did not develop T2D during the follow-up period. Individuals who developed incident T2D also had a significantly higher mean GRS and were more likely classified in the moderate or high GRS tertile. Finally, compared to the censored observations, incident T2D cases were more likely to have a higher BMI, lower physical activity level, lower diet quality, and to be a current smoker, which resulted in a significantly higher proportion of incident T2D cases also falling into the "unhealthy" lifestyle category.

GRS tertiles (moderate, HR: 1.59, 95% CI 1.52-1.67; high, HR: 2.58, 95%CI: 2.47-2.70) and overall lifestyle categories (intermediate, HR: 2.38, 95% CI 2.24-2.52; unhealthy, HR: 6.83, 95% CI 6.32-7.38) were independently associated with T2D risk (Supplementary Table S4). For the standardized GRS, a 1-SD increase was associated with a 53% increased risk of T2D (HR: 1.53, 95% CI: 1.50-1.55; Supplementary Table S5). Results were similar for individual lifestyle factors and in models stratified by sex (Supplementary Tables S6-S7); BMI had the strongest independent association with incident T2D (intermediate, HR: 2.81, 95% CI 2.63-3.00; unhealthy, HR: 8.84, 95% CI 8.29-9.42).

Across all GRS tertiles, individuals classified as having a unhealthy lifestyle were at substantially increased risk for T2D compared to those classified as having an healthy lifestyle, with HR ranging from 7.11 (low GRS tertile) to 16.33 (high GRS tertile; Figure 2). Compared to those in the healthy lifestyle group, individuals in the intermediate lifestyle group were also at increased risk for T2D, with a 2-, 4-, and 6- fold increased risk for those in the low, moderate, and high GRS tertile. When focusing within a single genetic risk tertile (e.g., low GRS), individuals in the unhealthy lifestyle category were at a 6- to 8-fold increased risk of T2D compared to the healthy lifestyle. Results were similar in sex-stratified analyses (Supplemental Table S8). While the effect estimates were slightly stronger in females than in males, the difference was not statistically significant ($P = 0.39$). We did not detect a significant interaction between the GRS tertiles and lifestyle classification, however we did detect a significant interaction when considering the standardized (continuous) GRS and lifestyle classifications (GRS*intermediate lifestyle, $p<0.20$; GRS*unhealthy lifestyle, $p=0.004$; Supplementary Tables

S9-S10). Lastly, results showed similar trends across the self-reported ancestry groups but were generally underpowered in many GRS tertile/lifestyle category combinations for analyses of non-European-ancestry individuals (Supplemental Table S11).

When calculating the ancestry-specific GRS using the same set of genetic variants, but with the ancestry-specific weights, we found similar trends in the associations between combined GRS and lifestyle risk with incident T2D (Supplemental Tables S12-16; Supplementary Figures S12-14).

To evaluate the proportion of incident T2D that would have been prevented if subjects with intermediate or unhealthy lifestyle (also considered "non-healthy") had instead been in the healthy category, we calculated the population attributable fraction (PAF; Supplementary Table S17). Regardless of the GRS, more than 55% of incident T2D cases in the UKB would have been prevented if all individuals in the "non-healthy" lifestyle categories would have been in the healthy lifestyle category (Year 1: 95% CI 0.53-0.58). The PAF proportions were consistent across each time point during the 15 years of follow-up.

Discussion

In this large population-based prospective cohort study with over 332,000 multi-ancestry participants from the UK Biobank, both high GRS and unhealthy lifestyle were independently associated with increased risk of T2D. Across and within different GRS tertiles, adherence to an intermediate or unhealthy lifestyle was associated with substantially increased risk of T2D compared to an healthy lifestyle. Overall, while our analyses support the notion that while genetics play a large role in the risk for developing T2D and T2D etiology, lifestyle factors play a substantially larger role, particularly BMI. Further, we demonstrated that individuals with any level of genetic risk could greatly reduce their risk for T2D through modifiable healthy lifestyles.

To our knowledge, this study is the first to test the effect of combined lifestyle factors in different genetic risk level for T2D based on nearly 800 genetic variants and the first to consider both a multi-ancestry and ancestry-matched GRS. Consistent with findings from prior studies[11–19], we found high GRS and unhealthy lifestyle factors were independently and jointly associated with increased risk of developing T2D. However, there is wide variability in effect sizes across the prior studies, most likely due to fundamental differences in study design and methodology, including sample size, T2D GRS composition and calculation, and consideration of behavioral and lifestyle factors. Most similar to this study, Said, et al. previously used the UKB study population to examine the combined effects of genetic and lifestyle risk of T2D.[11] Among 322,014 individuals, the study also found strong effects of unhealthy lifestyle across different GR tertiles, with adjusted HR ranging from 10.82 to 15.46 in sex-combined analyses. While both studies used a similar approach to categorize lifestyle factors based on the American Heart Association guidelines, the prior study included only the European-ancestry individuals within the UKB, had a less-restrictive definition for incident T2D that likely resulted in outcome misclassification, and calculated the GRS based on only 38 variants. Our study included a much more comprehensive measure of GRS (783 variants), excluded individuals from the analyses if they did not have confirmed T2D, and did not exclude individuals based on genetic ancestry.

Based on the results presented, it is clear that individuals who have either moderate or higher genetic risk for T2D with intermediate or unhealthy lifestyle are at substantially increased risk for T2D. These findings indicate the strong potential benefits of adherence to multiple healthy lifestyle factors to mitigate disease risk, regardless of genetic risk. In fact, our analysis suggests that 55% of incident T2D risk in the population that could be theoretically eliminated if individuals with non-healthy lifestyle were to be shifted to be having healthy lifestyle, highlighting the potential impact of shifting individuals from the non-healthy to the healthy lifestyle category. Although challenges remain in communicating individual genetic risk information to patients that is understandable and interpretable by the general population, knowledge of the strong impact a healthy lifestyle can have to mitigate genetic or familial risk for T2D may motivate patients to change behaviors

Strengths and Limitations

To our knowledge, this is the first study to investigate the associations of combined genetic and lifestyle risk of T2D using the most up-to-date set of T2D-associated variants.[4] Major strengths of the study were the prospective cohort design, large sample size, and comprehensive measure of genetic risk. The list of T2D-associated variants and their effect sizes used for our GRS calculation were determined independently of the UKB study population. Our study also utilized all individuals in the UKB, regardless of self-reported ancestry, and used both combined and ancestry-specific genetic effect sizes when calculating the GRS, which improves the external validity of our findings. Further, we classified all of the lifestyle factors into healthy, intermediate, and unhealthy based on guidelines from the American Heart Association,[24] which allows for a more direct clinical interpretation of our results.

There are also several limitations. Measurements for all lifestyle factors were obtained at study entry, of which three were based on self-reported data. Because they are all potentially time-varying covariates, misclassification of exposure is possible. However, due to the prospective design of the UKB, any misclassification would be nondifferential and would bias the result towards the null, resulting in an underestimation of the true association. Second, incident T2D cases were identified using the first occurrence data in the UK Biobank, which includes self-reported outcomes. Further, the suspected rate of undiagnosed T2D in the UK is estimated to be around 2%.[28] Thus, misclassification of some T2D cases as non-cases is possible; however, we would expect this bias the results toward the null. The genetic variants used in the GRS calculation may also have pleiotropic effects on lifestyle factors, including BMI. Although our study included individuals with diverse ethnic backgrounds, the generalizability of the findings remains somewhat limited due to the predominance of European participants in the UKB.

In conclusion, both genetic risk and lifestyle were independently associated with elevated T2D risk, but individuals with the unhealthiest lifestyle were at the highest risk for incident disease. Comprehensive and multifactorial lifestyle modifications should be encouraged in individuals at all levels of genetic risk to greatly mitigate their risk of developing T2D, though individuals at the highest levels of genetic risk will gain the most benefit. Further studies investigating the joint effects of lifestyle changes over time and their interplay with genetics for the T2D risk is warranted.

Acknowledgments

The present study was conducted under application #81009 of the UK Biobank resource.

Funding

CNS and CZ were supported by the NIH (R01DK118011) and the American Diabetes Association (11-22-JDFPM-06). CNS was further supported by the NIH (R01DK136671). EBJ was supported by the NIH (R34MH118396 and K01DK123193) and the Commonwealth of Massachsutts (INTF4107H78241733154). NVK was support by the NIH (K01DK123193). AH-C is supported by the American Diabetes Association grant 11-23-PDF-35.

References (≤40 references)

1. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Kaabi JA. Epidemiology of Type 2 Diabetes – Global Burden of Disease and Forecasted Trends. *J Epidemiol Glob Health*. 2020;10(1):107. doi:10.2991/jegh.k.191028.001
2. Bullard KM, Cowie CC, Lessem SE, et al. Prevalence of Diagnosed Diabetes in Adults by Diabetes Type - United States, 2016. *MMWR Morb Mortal Wkly Rep*. 2018;67(12):359-361. doi:10.15585/mmwr.mm6712a2
3. Ma Y, Hébert JR, Manson JE, et al. Determinants of Racial/Ethnic Disparities in Incidence of Diabetes in Postmenopausal Women in the U.S. *Diabetes Care*. 2012;35(11):2226-2234. doi:10.2337/dc12-0412
4. Suzuki K, Hatzikotoulas K, Southam L, et al. Genetic drivers of heterogeneity in type 2 diabetes pathophysiology. *Nature*. 2024;627(8003):347-357. doi:10.1038/s41586-024-07019-6
5. Chen J, Spracklen CN, Marenne G, et al. The trans-ancestral genomic architecture of glycemic traits. *Nat Genet*. 2021;53(6):840-860. doi:10.1038/s41588-021-00852-9
6. Prasad RB, Groop L. Genetics of Type 2 Diabetes—Pitfalls and Possibilities. *Genes*. 2015;6(1):87-123. doi:10.3390/genes6010087
7. Ye C, Niu J, Zhao Z, et al. Genetic susceptibility, family history of diabetes and healthy lifestyle factors in relation to diabetes: A gene-environment interaction analysis in Chinese adults. *J Diabetes Investig*. 2021;12(11):2089-2098. doi:10.1111/jdi.13577
8. Langenberg C, Sharp SJ, Franks PW, et al. Gene-lifestyle interaction and type 2 diabetes: the EPIC interact case-cohort study. *PLoS Med*. 2014;11(5):e1001647. doi:10.1371/journal.pmed.1001647
9. Klimentidis YC, Chen Z, Arora A, Hsu CH. Association of physical activity with lower type 2 diabetes incidence is weaker among individuals at high genetic risk. *Diabetologia*. 2014;57(12):2530-2534. doi:10.1007/s00125-014-3380-z
10. Lim JE, Kang JO, Ha TW, et al. Gene-environment interaction in type 2 diabetes in Korean cohorts: Interaction of a type 2 diabetes polygenic risk score with triglyceride and cholesterol on fasting glucose levels. *Genet Epidemiol*. 2022;46(5-6):285-302. doi:10.1002/gepi.22454
11. Said MA, Verweij N, van der Harst P. Associations of Combined Genetic and Lifestyle Risks With Incident Cardiovascular Disease and Diabetes in the UK Biobank Study. *JAMA Cardiol*. 2018;3(8):693-702. doi:10.1001/jamacardio.2018.1717
12. Zhao Y, Li Y, Zhuang Z, et al. Associations of polysocial risk score, lifestyle and genetic factors with incident type 2 diabetes: a prospective cohort study. *Diabetologia*. 2022;65(12):2056-2065. doi:10.1007/s00125-022-05761-y
13. Schnurr TM, Jakupović H, Carrasquilla GD, et al. Obesity, unfavourable lifestyle and genetic risk of type 2 diabetes: a case-cohort study. *Diabetologia*. 2020;63(7):1324-1332. doi:10.1007/s00125-020-05140-5
14. Yu H, Wang T, Zhang R, et al. Alcohol consumption and its interaction with genetic variants are strongly associated with the risk of type 2 diabetes: a prospective cohort study. *Nutr Metab*. 2019;16:64. doi:10.1186/s12986-019-0396-x
15. Villegas R, Delahanty R, Gao YT, et al. Joint effect of genetic and lifestyle risk factors on type 2 diabetes risk among Chinese men and women. *PloS One*. 2012;7(11):e49464. doi:10.1371/journal.pone.0049464

16. Szczerbiński Ł, Gościk J, Bauer W, et al. Efficacy of family history, genetic risk score, and physical activity in assessing the prevalence of type 2 diabetes. *Pol Arch Intern Med*. 2019;129(7-8):442-450. doi:10.20452/pamw.14866
17. Li H, Khor CC, Fan J, et al. Genetic risk, adherence to a healthy lifestyle, and type 2 diabetes risk among 550,000 Chinese adults: results from 2 independent Asian cohorts. *Am J Clin Nutr*. 2020;111(3):698-707. doi:10.1093/ajcn/nqz310
18. Han X, Wei Y, Hu H, et al. Genetic Risk, a Healthy Lifestyle, and Type 2 Diabetes: the Dongfeng-Tongji Cohort Study. *J Clin Endocrinol Metab*. 2020;105(4):dgz325. doi:10.1210/clinem/dgz325
19. Ericson U, Hindy G, Drake I, et al. Dietary and genetic risk scores and incidence of type 2 diabetes. *Genes Nutr*. 2018;13:13. doi:10.1186/s12263-018-0599-1
20. Allen NE, Lacey B, Lawlor DA, et al. Prospective study design and data analysis in UK Biobank. *Sci Transl Med*. 2024;16(729):eadf4428. doi:10.1126/scitranslmed.adf4428
21. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;12(3):e1001779. doi:10.1371/journal.pmed.1001779
22. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203-209. doi:10.1038/s41586-018-0579-z
23. Privé F, Luu K, Blum MGB, McGrath JJ, Vilhjálmsson BJ. Efficient toolkit implementing best practices for principal component analysis of population genetic data. *Bioinforma Oxf Engl*. 2020;36(16):4449-4457. doi:10.1093/bioinformatics/btaa520
24. Lloyd-Jones DM, Hong Y, Labarthe D, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic Impact Goal through 2020 and beyond. *Circulation*. 2010;121(4):586-613. doi:10.1161/CIRCULATIONAHA.109.192703
25. Mozaffarian D. Dietary and Policy Priorities for Cardiovascular Disease, Diabetes, and Obesity: A Comprehensive Review. *Circulation*. 2016;133(2):187-225. doi:10.1161/CIRCULATIONAHA.115.018585
26. International Standard Classification of Education (ISCED). March 16, 2017. Accessed July 15, 2024. https://uis.unesco.org/en/topic/international-standard-classification-education-isced
27. Townsend P. Deprivation*. *J Soc Policy*. 1987;16(2):125-146. doi:10.1017/S0047279400020341
28. Whicher CA, O'Neill S, Holt RIG. Diabetes in the UK: 2019. *Diabet Med J Br Diabet Assoc*. 2020;37(2):242-247. doi:10.1111/dme.14225

Table 1. Baseline characteristics of 332,251 participants from the UK Biobank

Variable	Overall N = 332,251[1]	Censored observations N = 319,123[1]	Incident T2D N = 13,128[1]
Age at baseline (years)	55.19 (8.06)	55.09 (8.06)	57.56 (7.63)
Sex			
Female	177,869 (54)	172,507 (54)	5,362 (41)
Follow-up time (years)			
Median (IQR)	13.56 (12.74, 14.25)	13.61 (12.88, 14.27)	7.98 (4.92, 10.82)
Multi-ancestry GRS	21.54 (0.56)	21.53 (0.56)	21.76 (0.55)
Multi-ancestry GRS tertile			
Low	110,753 (33)	108,130 (34)	2,623 (20)
Moderate	110,755 (33)	106,670 (33)	4,085 (31)
High	110,743 (33)	104,323 (33)	6,420 (49)
BMI (kg/m^2)			
Healthy (18.5-24.9)	118,246 (36)	117,038 (37)	1,208 (9.2)
Intermediate (25.0-29.9)	143,047 (43)	138,249 (43)	4,798 (37)
Unhealthy (≥30)	70,958 (21)	63,836 (20)	7,122 (54)
Smoking status			
Healthy (non-smoker)	188,546 (57)	182,341 (57)	6,205 (47)
Intermediate (past smoker)	111,190 (34)	106,143 (33)	5,047 (39)
Unhealthy (current smoker)	31,959 (9.6)	30,104 (9.4)	1,855 (14)
Physical activity level[2]			
Healthy (regular physical activity)	93,866 (29)	90,532 (29)	3,334 (27)
Intermediate (some physical activity)	192,820 (60)	185,844 (60)	6,976 (57)
Unhealthy (no regular physical activity)	33,588 (10)	31,599 (10)	1,989 (16)
Diet quality[3]			
Healthy (adequate dietary intake)	159,553 (48)	153,928 (48)	5,625 (43)
Unhealthy (inadequate dietary intake)	172,686 (52)	165,183 (52)	7,503 (57)
Overall lifestyle			
Healthy (≥3 healthy factors)	70,854 (21)	69,669 (22)	1,185 (9.0)
Intermediate (all other combinations)	248,927 (75)	238,434 (75)	10,493 (80)
Unhealthy (≥3 unhealthy factors)	12,470 (3.8)	11,020 (3.5)	1,450 (11)
Years in education	15.54 (4.44)	15.56 (4.44)	14.97 (4.38)
Income (pound sterling)			
<18,000	54,537 (16)	51,007 (16)	3,530 (27)
18,000-30,999	82,294 (25)	78,494 (25)	3,800 (29)
31,000-51999	95,351 (29)	91,947 (29)	3,404 (26)
52,000-100,000	78,705 (24)	76,717 (24)	1,988 (15)
>100,000	21,364 (6.4)	20,958 (6.6)	406 (3.1)
Self-reported Ancestry			
AFR	4,385 (1.3)	3,963 (1.2)	422 (3.2)

Variable	Overall N = 332,251[1]	Censored observations N = 319,123[1]	Incident T2D N = 13,128[1]
EAS	970 (0.3)	914 (0.3)	56 (0.4)
EUR	316,747 (95)	304,982 (96)	11,765 (90)
Missing	840 (0.3)	796 (0.2)	44 (0.3)
Mixed	1,984 (0.6)	1,888 (0.6)	96 (0.7)
Other	2,494 (0.8)	2,299 (0.7)	195 (1.5)
SAS	4,831 (1.5)	4,281 (1.3)	550 (4.2)

[1]Mean (SD); n (%); all p-value <0.001

[2]Healthy: ≥150 min/week moderate or ≥75 min/week vigorous or ≥150 min/week mixed; intermediate: 1-149 min/week moderate or 1-74 min/week vigorous or 1-149 min/week mixed; unhealthy; not performing any moderate or vigorous

[3]Healthy: adequate intake of at least half of certain dietary components; unhealthy: less than half

Abbreviations: AFR, African and African American; BMI, body mass index; EAS, East Asian; EUR, European; GRS, genetic risk score; IQR, interquartile range; SAS, South Asian; T2D, type 2 diabetes

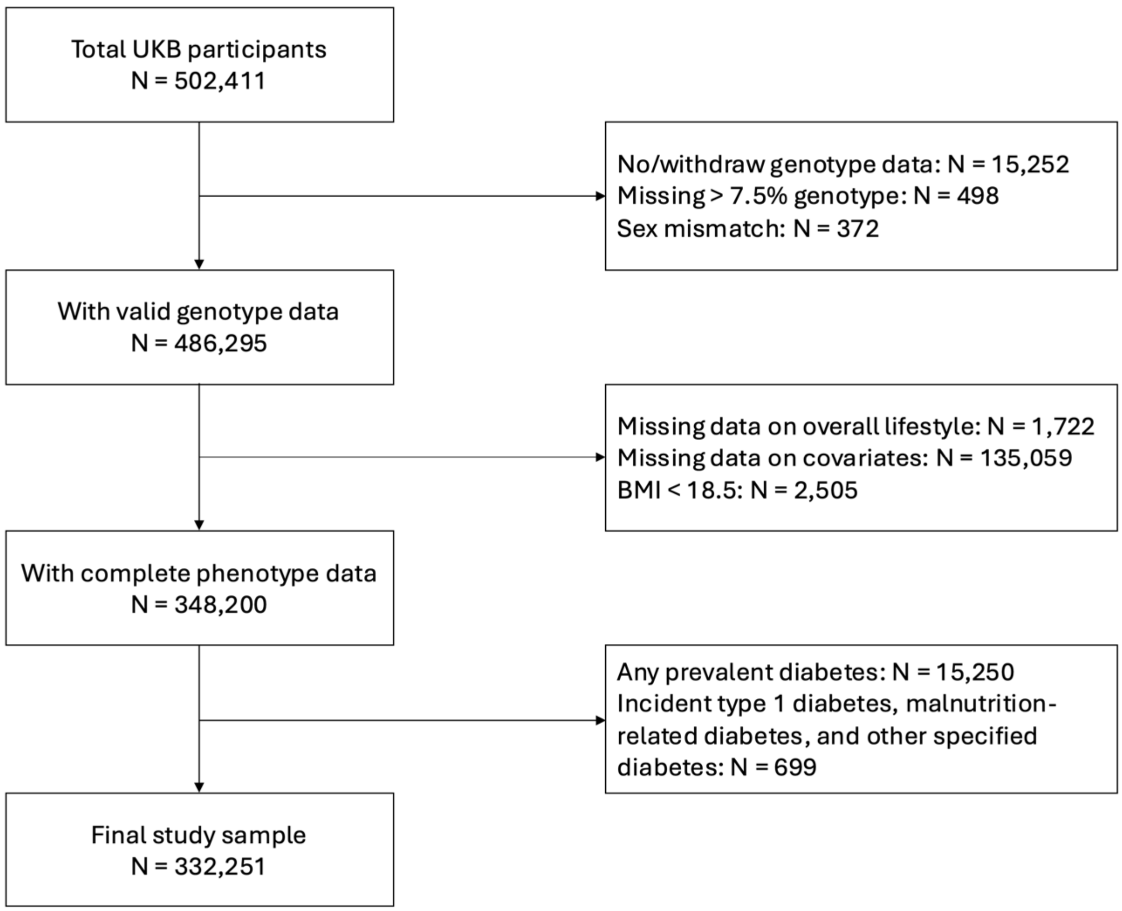
Table 2. Incidence rates of T2D by the category of combined genetic and lifestyle risk

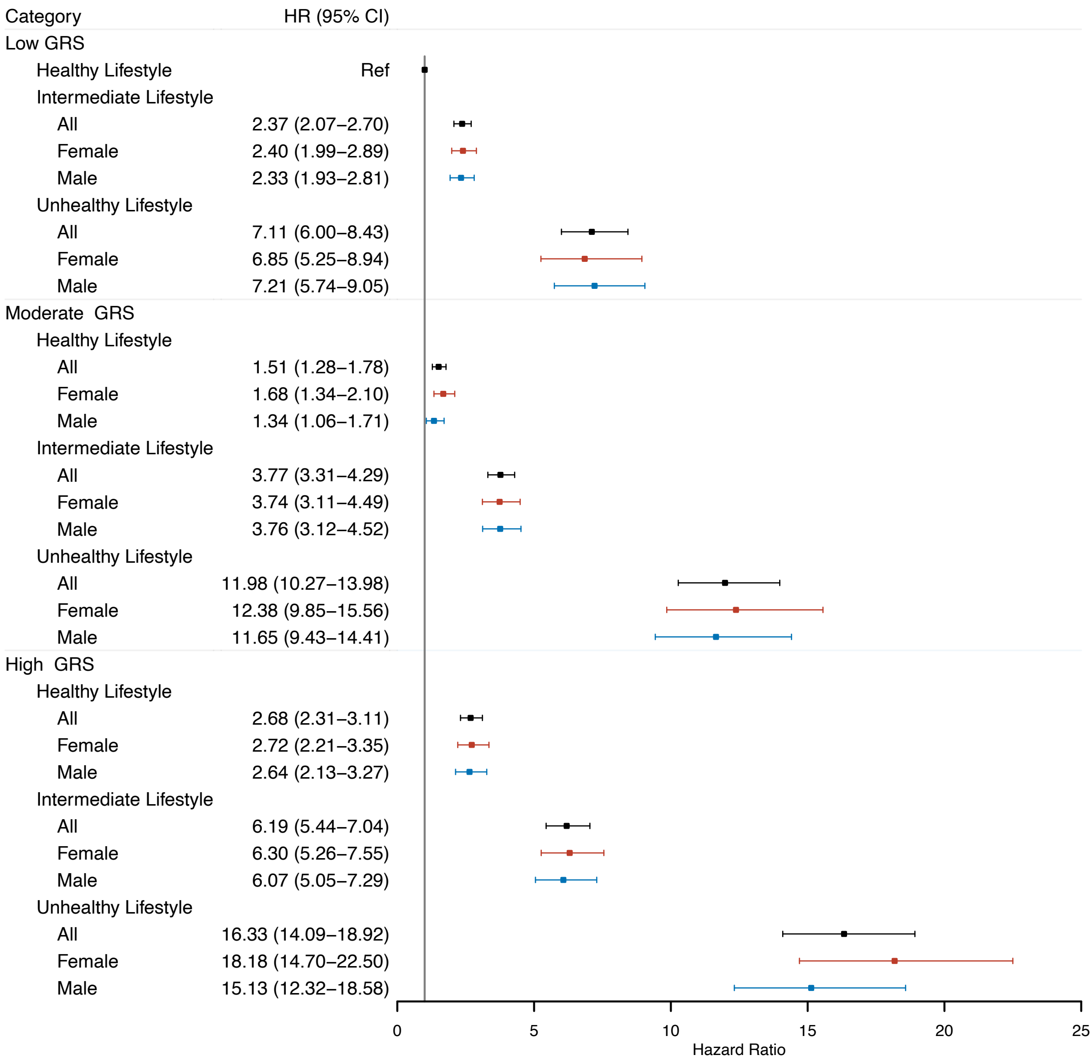
Multi-ancestry GRS tertile	Lifestyle	N total (%)	N case (%[1])	Person-year	Incidence rate[2]
Low	Healthy	24,408 (7.4)	245 (1.0)	325,304.3	0.75
	Intermediate	82,402 (24.8)	2,082 (2.5)	1,083,869.8	1.92
	Unhealthy	3,943 (1.2)	296 (7.5)	50,042.0	5.92
Moderate	Healthy	23,373 (7.0)	347 (1.5)	311,018.5	1.12
	Intermediate	83,224 (25.1)	3,250 (3.9)	1,088,448.2	2.99
	Unhealthy	4,158 (1.3)	488 (11.7)	51,563.5	9.46
High	Healthy	23,073 (6.9)	593 (2.6)	305,394.7	1.94
	Intermediate	83,301 (25.1)	5,161 (6.2)	1,078,391.9	4.79
	Unhealthy	4,369 (1.3)	666 (15.2)	53,161.2	12.53

[1]Cumulative incidence

[2]The incidence rate is provided per 1,000 person-years.

Figure 1. Flowchart of study population inclusion and exclusion





Category
HR (95% CI)
Low GRS
Healthy Lifestyle
Ref
Intermediate Lifestyle
All 2.37 (2.07–2.70)
Female 2.40 (1.99–2.89)
Male 2.33 (1.93–2.81)
Unhealthy Lifestyle
All 7.11 (6.00–8.43)
Female 6.85 (5.25–8.94)
Male 7.21 (5.74–9.05)
Moderate GRS
Healthy Lifestyle
All 1.51 (1.28–1.78)
Female 1.68 (1.34–2.10)
Male 1.34 (1.06–1.71)
Intermediate Lifestyle
All 3.77 (3.31–4.29)
Female 3.74 (3.11–4.49)
Male 3.76 (3.12–4.52)
Unhealthy Lifestyle
All 11.98 (10.27–13.98)
Female 12.38 (9.85–15.56)
Male 11.65 (9.43–14.41)
High GRS
Healthy Lifestyle
All 2.68 (2.31–3.11)
Female 2.72 (2.21–3.35)
Male 2.64 (2.13–3.27)
Intermediate Lifestyle
All 6.19 (5.44–7.04)
Female 6.30 (5.26–7.55)
Male 6.07 (5.05–7.29)
Unhealthy Lifestyle
All 16.33 (14.09–18.92)
Female 18.18 (14.70–22.50)
Male 15.13 (12.32–18.58)
0
5
10
15
20
25
Hazard Ratio