Phenotypic vs. Genetic Mismatch of BMI and Type 2 Diabetes: Evidence from Two Perspective Cohort Studies

Short running title: Phenotypic vs. Genetic Mismatch of BMI and Type 2 Diabetes

Aolin Li, MD[1], Shuo Gong, MD[1], Canqing Yu, PhD[1,2,3], Pei Pei, BSc[2], Ling Yang, PhD[4], Iona Y. Millwood, DPhil[4], Robin G. Walters, PhD[4], Yiping Chen, DPhil[4], Huaidong Du, PhD[4], Xiaoming Yang, PhD[4], Wei Hou, MD[5], Junshi Chen, MD[6], Zhengming Chen, DPhil[4], Jun Lv, PhD[1,2,3,7], Liming Li, MPH[1,2,3], Dianjianyi Sun, PhD[1,2,3*], on behalf of the China Kadoorie Biobank Collaborative Group[†]

1. Department of Epidemiology & Biostatistics, School of Public Health, Peking University, Beijing, 100191, China;

2. Peking University Center for Public Health and Epidemic Preparedness & Response, Beijing, 100191, China;

3. Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing 100191, China;

4. Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, United Kingdom;

5. NCDs Prevention and Control Department, Licang CDC, Qingdao, 266041, China;

6. China National Center for Food Safety Risk Assessment, Beijing 100022, China;

7. State Key Laboratory of Vascular Homeostasis and Remodeling, Peking University, Beijing 100191, China

[*] Corresponding author

† The members of the steering committee and collaborative group are listed in the Supplementary material

Corresponding author

Dianjianyi Sun, MD, PhD

Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, 38 Xueyuan Road, Beijing 100191, China

Tel: +86 10-82801528 ext. 322

Email: dsun1@bjmu.edu.cn

Key words

body mass index; obesity; mismatch of phenotype and genotype; type 2 diabetes

Word count:

4,087 (2 tables and 3 figures)

Abstract

Little is known about the population-based mismatch between phenotypic and genetic BMI (BMI-PGM) and its association with type 2 diabetes. We therefore used data from the China Kadoorie Biobank and UK Biobank and calculated BMI-PGM for each participant as the difference between the percentile for adjusted BMI at baseline and the percentile for adjusted polygenic risk score for BMI. Participants were categorized into discordantly low (BMI-PGM<the 1st quartile), concordant (the 1st quartile ≤BMI-PGM<the 3rd quartile), and discordantly high (BMI-PGM≥the 3rd quartile) groups. We calculated adjusted hazard ratios (aHRs) for the association of BMI-PGM and type 2 diabetes using Cox proportional hazard models in each cohort, and combined HRs using random-effects meta-analyses. During a median follow-up of 12 years for both cohorts, BMI-PGM was associated with the risk of type 2 diabetes, with the discordantly low group showing reduced risk and the discordantly high group showing elevated risk compared to the concordant group, independent of BMI and other conventional risk factors. In addition, normal-weight individuals with discordantly high BMI-PGM faced a higher risk of type 2 diabetes than overweight individuals. These findings suggest that BMI-PGM may play a potential role in reassessing the risk of type 2 diabetes, particularly among normal-weight populations.

Article Highlights

- Social developments have fostered an "obesogenic environment" that exacerbated phenotypic vs. genetic mismatch of BMI (BMI-PGM) and the risk of type 2 diabetes.
- The study quantified BMI-PGM and examined its association with type 2 diabetes.

- The risk of type 2 diabetes was lower in discordantly low BMI-PGM group and higher in discordantly high BMI-PGM group, with concordant BMI-PGM group as reference.
- These findings indicate the potential to reassess type 2 diabetes risk by quantifying the mismatch between phenotypic and genetic BMI on individual levels.

Introduction

Diabetes is now a growing concern worldwide. The International Diabetes Federation estimated that the global prevalence of diabetes reached 10.5% in 2021, and deaths attributed to diabetes and its complications accounted for 12.2% of all-cause deaths (1). Obesity has been widely recognized as a strong risk factor for developing type 2 diabetes (2). A global burden of disease study on type 2 diabetes in 2019 reported that the portion of disability-adjusted life years (DALYs) attributable to high body mass index (BMI) in type 2 diabetes was 51.9%, much higher than other risk factors (3).

BMI is a complex trait influenced by genetics and environment. Polygenic risk score for BMI (BMI-PRS) has been widely used to measure the genetic susceptibility to BMI (4, 5). However, PRS among unrelated individuals only predicts up to 8.5% of the variance in BMI (6). There is also a large portion of BMI's variance that cannot be explained by genetics, which causes a mismatch between phenotypic BMI and genetic BMI. Such mismatch is likely due to environmental factors and the potential interactions between environmental factors and genetics. In the past century, social and economic developments have fostered an "obesogenic environment" likely to have exacerbated this mismatch, such as the low cost of fast food and less physical activity (7). Given the increasing disease burden of obesity and type 2 diabetes, it is necessary to quantify this mismatch and explore the relationship between the mismatch and type 2 diabetes.

In the present study, we use data from the China Kadoorie Biobank (CKB) and the UK Biobank (UKB) to generate a new variable called “BMI-PGM” based on BMI and BMI-PRS, which quantitatively reflects one’s mismatch of phenotypic vs. genetic BMI. We aim to explore the sociodemographic characteristics and lifestyle that may be associated with BMI-PGM and examine its association with the risk of incident type 2 diabetes in the two populations.

Research Design and Methods

Study participants in CKB and UKB

In CKB, 512 723 participants aged 30-79 years, with written informed consent, were recruited from five urban and five rural areas in China during 2004-2008. The participants completed the questionnaires including sociodemographic characteristics, lifestyle, and personal and family medical history, underwent comprehensive physical examination, and provided biological samples for genotyping. Ethics approvals were obtained from the ethical committee and research council of the Chinese Center for Disease Control and Prevention (Beijing, China) and the Oxford Tropical Research Ethics Committee at the University of Oxford (Cambridge, UK). More details of the CKB could be found elsewhere (8, 9).

UKB is a large prospective cohort of 502 628 participants aged 40-69 years recruited from 22 centres in the UK between 2006 and 2010 (10). Participants provided written informed consent and underwent a series of sociodemographic, physical, and medical assessments. UKB followed the Declaration of Helsinki guidelines and obtained ethical approvals from the

North West Multi-Centre Research Ethics Committee (Manchester, UK) and the NHS National Research Ethics Service. The present study was conducted under Application Number 86473 of the UKB data resource.

In CKB and UKB, we conducted cross-sectional analyses of the BMI-PGM distribution and its risk factors among the participants who were genotyped and with valid BMI measurements in the testing set (Supplementary Fig. 1). In the prospective analysis of the association between BMI-PGM and type 2 diabetes, we further excluded patients with diabetes at baseline (self-reported diabetes, fasting blood glucose ≥7.0mmol/l, or random blood glucose ≥11.0mmol/l in CKB; self-reported diabetes, glycated hemoglobin ≥6.5% (48.0mmol/mol) (11), or those with the date of developing diabetes [ICD-10: E10-E14, O24] before the date of attending the centres in UKB).

BMI in CKB and UKB

Anthropometric measurements in CKB and UKB were conducted by trained investigators following their respective standard procedures. Standing height was measured by stadiometers (CKB: Unknown; UKB: Seca 202, Seca Inc., Hamburg, Germany) and weight was measured by body composition analyzers (CKB: TANITATBF-300 GS, Tanita Inc., Tokyo, Japan; UKB: Tanita BC-418 MA, Tanita Inc., Tokyo, Japan). BMI was calculated as measured weight in kilograms divided by the standing height in meters squared (kg/m^2).

BMI-PRS in CKB and UKB

Before calculating BMI-PRS for CKB, quality control (QC) was conducted for SNPs and samples, and the remaining participants were divided into a training set (N = 22 177) and a testing set (N = 64 029). GWAS summary statistics of BMI from BioBank Japan (BBJ) and Genetic Investigation of ANthropometric Traits (GIANT) were chosen as the base data 1 and base data 2, respectively (12, 13). Three methods were used to compute BMI-PRS: (1) Using PRSice-2 and base data 1; (2) Using PRS-CS and base data 1; (3) Using PRS-CSx, base data 1, and base data 2 (PRS-CSx can utilize multi-ancestry GWAS data to enhance the performance of PRS (14). BMI-PRS generated by PRS-CSx performed best (R^2 = 0.058, Supplementary Table 1) in the training set. Subsequent analyses were conducted in the testing set using BMI-PRS calculated by this method. UKB has generated PRS for 28 diseases and 25 quantitative traits including BMI. Following QC of the genetic variants and samples, UKB conducted a meta-analysis of the phenotypes' pre-existing GWAS and generated centred and standardized PRS for the corresponding phenotype using Bayesian regression and Khera et al.'s approach (15). The process for calculating PRS can be found elsewhere (16). In this study, we retrieved the standard PRS for BMI and conducted further analyses in UKB's testing set of 104 592 participants. Details on the genotyping data for CKB and UKB, and the construction of BMI-PRS in CKB, are provided in the Supplementary Methods section of the Supplementary Materials.

BMI-PGM in CKB and UKB

Previous studies reported different methods to assess discrepancies between phenotype and genotype (17, 18). We attempted a novel approach in the present study. Participants with missing BMI from the testing sets of CKB and UKB were excluded, leaving 64 028 and 102 514 participants in each cohort. A linear regression was performed in both the two cohorts with BMI as the dependent variable and age, age squared, sex, study area, and their interaction terms as independent variables, i.e.,

$$BMI=\beta_0+\beta_1\ Age+\beta_2Age^2+\beta_3Sex+\beta_4Study\ Area+\beta_5(Age\times Sex)+\beta_6(Age\times Study\ Area)+\beta_7(Age^2\times Sex)+\beta_8(Age^2\times Study\ Area)+\beta_9(Sex\times Study\ Area)+\beta_{10}(Age\times Sex\times Study\ Area)+\beta_{11}(Age^2\times Sex\times Study\ Area)+\epsilon.$$

The residuals from this model represented the variance of individuals' BMI with age, age squared, sex, study area, and their interactions adjusted. Again, we ran a linear regression in both the two cohorts with BMI-PRS as the dependent variable and the first ten genetic principal components (PCs), batch, study area, and their interaction terms as independent variables, i.e.,

$$BMI\text{-}PRS=\beta_0+\sum_{i=1}^{10}\beta_iPC_i+\beta_{11}Batch+\beta_{12}Study\ Area+\sum_{i=1}^{10}\beta_{13+i}(PC_i\times Batch)+\sum_{i=1}^{10}\beta_{23+i}(PC_i\times Study\ Area)+\beta_{34}(Batch\times Study\ Area)+\sum_{i=1}^{10}\beta_{35+i}(PC_i\times Batch\times Study\ Area)+\epsilon.$$

The residuals from this model represented the variance of individual's BMI-PRS with 10 PCs, batch, study area, and their interactions adjusted. Ultimately, the value of individual's BMI-PGM was generated by subtracting the percentile of residual for BMI-PRS from the percentile of residual for BMI.

The range of BMI-PGM in both CKB and UKB was (-100, 100) (Supplementary Fig. 2). A positive value indicated that the individual's actual BMI ranked higher than the BMI-PRS ranked, while a negative value was the opposite. We categorized participants into three groups according to BMI-PGM quartiles: discordantly low (BMI-PGM<the 1st quartile), concordant (the 1st quartile≤BMI-PGM<the 3rd quartile), and discordantly high (BMI-PGM≥the 3rd quartile). The concordant group was defined as the reference group.

Ascertainment of incident type 2 diabetes events in CKB and UKB

Incident cases of type 2 diabetes in CKB were identified by linking with local disease and death registries, with health insurance databases, and by active follow-up until 21 December 2018. In UKB, they were identified using the first occurrences of disease via linkage with the Hospital Episode Statistics of England (up to 30 September 2021), Scottish Morbidity Record of Scotland (up to 31 July 2021), Patients Episode Database of Wales (up to 28 February 2018), and the death cause registry records from the National Health Service (NHS) Information Centre (England and Wales, up to 30 November 2022) and the NHS Central Register, National Records of Scotland (Scotland, up to 30 November 2022). ICD-10 was used to define type 2 diabetes (E11) in both CKB and UKB.

Other covariates in CKB and UKB

When analyzing the risk factors of BMI-PGM, we adjusted for covariates including personal history of hypertension, cardiovascular diseases (CVD), respiratory diseases (including

chronic obstructive pulmonary disease, pulmonary emphysema, chronic bronchitis, tuberculosis, and asthma), cancer, and diabetes (see Supplementary Table 2 for definitions of the history of different diseases). In examining the association between BMI-PGM and type 2 diabetes, we considered sociodemographic characteristics, lifestyle behaviors, family history of diabetes (yes, no), and BMI (continuous) as covariates. Sociodemographic characteristics included age (continuous), sex (men, women), urban-rural status (urban, rural), education (low, medium, high), and household income (low, medium, high). Lifestyle behaviors included smoking status (never, previous, current), alcohol consumption (never, previous, current), physical activity (divided into four groups based on quartiles of sex-specific metabolic equivalent tasks), sedentary duration(divided into three groups based on tertiles), dietary habits (healthy, unhealthy, which were adapted from the American Heart Association Guidelines (19), and sleep duration (<7h, 7-9h, ≥9h). Education, household income, and the variables related to dietary habits were evaluated differently between the questionnaires of CKB and UKB but harmonized in a reasonably uniform way, as shown in Supplementary Table 2. The proportions of missing data in CKB and UKB for covariates were 0% and <7%. Continuous and categorical covariates in UKB were imputed using predictive mean matching or logistic regression, respectively.

Statistical analysis

Baseline characteristics of included participants were represented by BMI-PGM groups, with mean (SD) for normally distributed continuous variables, median, tertiles, or IQR for non-normally distributed continuous variables, and n (%) for categorical variables. Pearson

correlation coefficients assessed the relationships between BMI, BMI-PRS, and BMI-PGM. After categorizing the CKB and UKB participants according to the corresponding BMI cut-offs defining underweight, normal weight, overweight and obese (<18.5, 18.5~23.9, 24.0~27.9, and ≥28.0kg/m^2 for CKB; <18.5, 18.5~24.9, 25.0~29.9, and ≥30.0kg/m^2 for UKB) (20, 21), the proportions of three BMI-PGM groups within each BMI category were calculated. A multiple linear regression model with BMI-PGM as the dependent variable was used to identify sociodemographic characteristics and lifestyle behaviors associated with BMI-PGM. The model adjusted for personal history of hypertension, CVD, respiratory diseases, cancer, and diabetes. Variance inflation factors (VIF) confirmed the absence of collinearity among independent variables, with all VIF values below 2.0 in both cohorts.

After excluding participants with diabetes at baseline from CKB and UKB, Kaplan–Meier curves were plotted, and the log-rank test was used to compare the cumulative risks of developing type 2 diabetes between different BMI-PGM groups. Cox proportional hazards regression models were used to evaluate the hazard risk of incident type 2 diabetes related to BMI-PGM (per SD increase, and the discordantly low/high group vs. the concordant group). The Cox models, with age as the time scale, were stratified by age at risk (five-year bands in CKB and UKB) and study areas (ten groups in CKB and 22 groups in UKB). Four models were stepwise adjusted. Model 1 adjusted for sex, education, and household income; model 2 additionally adjusted for smoking status, alcohol consumption, physical activity, sedentary duration, dietary habits, and sleep duration; model 3 included family history of diabetes; and model 4 further adjusted for BMI (as a continuous variable). Restricted cubic spline curves of

adjusted hazard ratios (aHRs) with 95% confidence intervals (CIs) from model 4 were plotted to visualize the association between BMI-PGM and incident risk of type 2 diabetes. Analyses were also stratified by BMI categories and tertiles of BMI-PRS and followed the adjustments in Model 4. The interaction between sociodemographic characteristics, lifestyle behaviors, and BMI-PGM groups on the risk of type 2 diabetes was further examined through likelihood-ratio tests, contrasting models with and without cross-product interaction terms. Subgroup analyses were also conducted for these characteristics, with the same covariates as model 4 adjusted. Schoenfeld residuals were used to evaluate the proportional hazard assumptions, and no apparent violation was found.

Meta-analyses of model 4's hazard ratios (HRs) for BMI-PGM (per SD increase) and BMI-PGM groups on type 2 diabetes risk in CKB and UKB were performed using random effect model. Heterogeneity was assessed using the Q statistic and inconsistency using I^2. Sensitivity analyses were performed by excluding participants with fewer than two 2 years of follow-up, as well as those with cancer or respiratory disease at baseline, respectively. To assess the association's robustness between BMI-PGM and type 2 diabetes under different classifications of BMI-PGM, it was additionally grouped using tertiles and the 1^{st} and 4^{th} quintiles as cut-off values, respectively. All analyses were conducted using R 4.0.3. The level of significance was set at a two-sided p-value <0.05.

Data and Resource Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request. Details of how to access China Kadoorie Biobank data and details of the data release schedule are available from www.ckbiobank.org/site/Data+Access. UK Biobank data are available to all researchers for health-related research. Data can be accessed through the Access Management System (details at https://www.ukbiobank.ac.uk/enable-your-research/register.

Results

A total of 64 028 participants from CKB and 102 514 participants from UKB were included in this study. Both in CKB and UKB, participants in the discordantly high BMI-PGM group were more physically inactive, more overweight on average (CKB: 26.3kg/m^2; UKB: 30.9kg/m^2), and had a higher prevalence of hypertension, CVD, and diabetes (Table 1). Notably, in CKB, individuals with high education and high household income were more prevalent in the discordantly high BMI-PGM group, in contrast to UKB, where such individuals were more commonly found in the discordantly low BMI-PGM group (Table 1). In UKB, more of the discordantly low BMI-PGM group slept 7-9h, and more of the discordantly high BMI-PGM group slept <7h or ≥9h, which was not obvious in CKB (Table 1). BMI-PGM values in both cohorts increased progressively from negative to positive as BMI categories were underweight, normal weight, overweight, and obese (Fig. 1*C*, 1*D*, 1*E*, and 1*F*). Concordant BMI-PGM was observed in approximately half of the participants across different BMI subgroups, with no individuals showing discordantly high BMI-PGM in the underweight

category and only a few showing discordantly low BMI-PGM in the obese category (Fig. 1*G* and 1*H*).

Sociodemographic characteristics and lifestyle behaviors associated with BMI-PGM

In Table 2, increasing age, current smoking, higher levels of physical activity, and less frequent meat intake were negatively associated with BMI-PGM in both CKB and UKB. Conversely, long sedentary duration was positively associated with BMI-PGM. In CKB, men, rural residency, high household income, and frequent fresh fruit intake were associated with higher BMI-PGM. In UKB, high education, current alcohol consumption, and 7-9 hours of sleep duration were negatively associated with BMI-PGM, which differed from CKB.

BMI-PGM and the risk of incident type 2 diabetes

During a median follow-up of 12 years among 59 873 CKB and 94 398 UKB participants, 1 765 and 4 870 participants developed type 2 diabetes, respectively. Kaplan-Meier curves showed that, compared to the concordant BMI-PGM group, the cumulative risk of type 2 diabetes onset was lower in the discordantly low BMI-PGM group and higher in the discordantly high BMI-PGM group during the follow-up period in both cohorts (Fig. 2*C* & 2*D*). After adjusting for BMI and other conventional risk factors, a linear relationship between BMI-PGM and the risk of developing type 2 diabetes was observed in CKB. In contrast, UKB had a deceleration in the increase of morbidity risk as BMI-PGM reached 25 (Fig. 2*A* & 2*B*). With the same covariate adjustments, each SD increase in BMI-PGM corresponded to an 18% higher risk of incident type 2 diabetes (Fig. 3). Relative to the

concordant BMI-PGM group, the discordantly high BMI-PGM group exhibited a higher risk of type 2 diabetes (HR [95% CI]: 1.22 [1.06-1.40]), while the discordantly low BMI-PGM group showed a lower risk (HR [95% CI]: 0.77 [0.71-0.83])(Fig. 3). Risk estimates were robust to different sensitivity analyses (Supplementary Table 5). Even among participants with normal weight, the discordantly high BMI-PGM group was associated with a 1.36-fold higher risk of type 2 diabetes compared with the concordant BMI-PGM group, while the association was stronger in Chinese adults vs. UK adults (HRs [95%CI]: 1.40[1.11-1.79] vs. 1.01[0.50-2.06] in Fig. 3). Moreover, the increased risk of type 2 diabetes associated with discordantly high BMI-PGM was more prominent among participants in the middle and upper BMI-PRS tertiles (HRs [95%CI]: 1.27[1.14-1.42] and 1.34[1.18-1.53]) vs. the bottom BMI-PRS tertile (HR [95% CI]: 1.12 [0.99-1.27]) (Fig. 3). In subgroup analyses in both cohorts, most characteristics showed no interaction with BMI-PGM on the risk of type 2 diabetes (Supplementary Fig. 3 & 4).

Discussion

In the present study, we created a quantile-based score (BMI-PGM) quantitatively measuring the deviation between one's percentile of adjusted BMI and one's percentile of adjusted BMI-PRS among two ancestry populations. In either CKB or UKB participants, approximately half of the individuals across different BMI categories were concordant for BMI-PGM, while discordantly low and high BMI-PGM groups were more prevalent among the underweight and the obese, respectively. Notably, BMI-PGM was associated with incident risk of type 2

diabetes, independent of BMI and other conventional risk factors. Moreover, adults with normal weight but discordantly high BMI-PGM had a higher incident risk of type 2 diabetes.

The thrifty gene hypothesis, first proposed by Neel et al.(22) in 1962, posits that "thrifty genes" once conferred an evolutionary advantage by promoting fat storage and reducing fat metabolism during times of famine, thereby enhancing survival. However, in the present day, these genes have become a mismatch with our “toxic environment” characterized by abundant food, reduced physical activity, and increasingly sedentary lifestyles, etc. (7, 23), contributing to the rising prevalence of obesity and its complications. BMI-PGM offers a novel approach to exploring the differences between phenotypic and genetic obesity. The deviation of BMI from an individual’s genetic set point may serve as a new, precise marker for obesity prevention and monitoring type 2 diabetes risk.

Our finding is consistent with the previous study (18), showing that even after adjusting for BMI, the difference between phenotypic and genetic BMI remains associated with an increased risk of type 2 diabetes in populations from different countries. This suggests that the mismatch between phenotypic and genetic BMI may be linked to a distinct pathophysiological mechanism underlying type 2 diabetes risk, independent of BMI itself. Rhee et al.(18) found that among participants whose observed BMI exceeded their genetically predicted BMI, there was a significant decline in insulin sensitivity that could not be compensated by β-cell function. This suggests that the interaction between insulin resistance and β-cell function might be one of the pathways mediating the relationship between BMI-

PGM and type 2 diabetes. We also found that this association was more pronounced in individuals with normal weight. Given that a significant proportion of newly diagnosed type 2 diabetes patients are not obese (24), BMI-PGM may have potential value in identifying high-risk individuals within the normal-weight population. The association between BMI-PGM and type 2 diabetes in CKB and UKB was similar to the relationship between BMI and type 2 diabetes in native Chinese and other ethnicities (25). It was found that native Chinese could achieve a similar incident rate of type 2 diabetes to other ethnicities at lower BMIs (25). In both cohorts, most characteristics showed no interaction with BMI-PGM regarding type 2 diabetes risk. However, in UKB, significant interactions were noted for alcohol consumption and urban-rural status. A systematic review reported that moderate drinking might reduce the risk of type 2 diabetes among non-Asian populations, which could partially explain this result (26). Those living in rural areas had a higher risk of type 2 diabetes when their BMI-PGM was discordantly high. This indicates a need for targeted policies in the UK to encourage weight management among rural populations, preventing them from becoming more obese than genetically predicted.

Our findings of the two cohorts support that low-frequency intake of red and processed meats may move BMI-PGM in a more healthful direction. Previous studies have discovered that red and processed meat intake was associated with central obesity (27). This kind of meat is higher in saturated fatty acids, has a less thermogenic effect (28), and alters the microbiota of the gut (29). The consistent direction of associations between physical activity, sedentary duration, and BMI-PGM across both populations reinforces the important impact of these

lifestyles on BMI. The World Health Organization advocates that adults should do at least 150–300 min of moderate-intensity aerobic physical activity or at least 75–150 min of vigorous-intensity aerobic physical activity throughout the week (30). However, the levels of physical activity in both the Chinese and British populations are not encouraging (31, 32). The observed inverse trend between age and BMI-PGM may be partially related to the decline in muscle mass with aging (33). Certain factors showed different directions of association with BMI-PGM between the Chinese and British populations. Chinese men were more likely to have higher BMI-PGM than women, whereas it was not observed in the UK. This may stem from gender roles in Chinese society, where women often focus on domestic responsibilities and men on career and social status, potentially giving men greater access to food resources (34). Variations in urban-rural status, education, and household income's influence on BMI-PGM align with existing evidence (35-38), highlighting different socio-economic dynamics between developed and developing countries. In developed countries, individuals with higher education and income typically consume healthier diets, whereas, in developing countries, the same groups often have greater access to excess food. Moreover, low-educated and low-income groups in developing countries typically engage in physically demanding manual labor (39). The differing associations between fresh fruit intake frequency and BMI-PGM may be attributed to China's high glycemic index staple foods like rice and wheat (40), where increased fruit consumption might elevate glycemic load, thus increasing BMI-PGM.

This study has several strengths. The large sample size and comprehensive data collection of CKB enabled us to construct BMI-PRS for Chinese. The extensive assessment of various factors related to BMI and type 2 diabetes in CKB and UKB enabled reasonable sensitivity and stratified analyses and adequate consideration of confounding factors. The similar results from two different populations, partly validate the generalizability of the results.

However, several limitations should be considered. First, the possibility of reverse causality and residual confounding cannot be completely ruled out, even though the sensitivity analyses demonstrated the robustness of our results and the key confounding factors were adjusted. Second, the data on sociodemographic characteristics and lifestyle was self-reported in both CKB and UKB, which might cause measurement errors. Third, the follow-up data in CKB did not include outpatient records. Given that type 2 diabetes patients with lower severity are typically treated in outpatient and emergency settings, this may limit the applicability of our findings to the estimation of associations. Fourth, in analyzing the association between BMI-PGM and type 2 diabetes, we only adjusted for BMI at a single time point, which may not fully account for the confounding effects of lifetime exposure to high BMI or changes in BMI over time. Finally, the observed associations in this study should be interpreted with caution. BMI-PRS aggregates various genetic variants linked to BMI-related traits. So, the pleiotropy may lead to an overestimation of the association between BMI-PGM and type 2 diabetes. Additionally, individuals with a significant mismatch between phenotypic and genetic BMI may be influenced by monogenic variants or extreme environmental factors. Although previous studies have explored the proteomic

signature of BMI and other obesity-related biomarkers, such as fat mass ratio, in relation to type 2 diabetes, further transcriptomic and proteomic analyses are needed to clarify the biological mechanisms underlying individual obesity and the associated risk of type 2 diabetes. Although previous studies have explored the proteomic features of BMI and its association with type 2 diabetes (41), as well as other obesity-related biomarkers such as fat mass ratio (42), further transcriptome and proteome analyses are needed to elucidate the biological mechanisms underlying the obesity and type 2 diabetes risk in these individuals.

Conclusions

In this study, BMI-PGM was generated for the Chinese and the British using BMI and BMI-PRS to quantify the non-genetically explained variance of BMI and it showed a strong association with type 2 diabetes. Monitoring one’s BMI-PGM might have the potential value in precision prevention for type 2 diabetes, particularly in individuals with normal weight. Various sociodemographic characteristics and lifestyle factors were found to be associated with BMI-PGM. China and the UK need to implement targeted policies and interventions to foster an "anti-obesogenic" environment and promote healthy lifestyles, reflecting the unique characteristics of their respective populations.

Acknowledgments

The most important acknowledgment is to the participants in the study and the members of the survey teams in each of the 10 regional centers, as well as to the project development and

management teams based at Beijing, Oxford, and the 10 regional centers This research has also been conducted using the UK Biobank Resource under Application Number 86473.

Author Contributions

DiS conceived of and designed the paper. LL, ZC, and JC, as the members of the CKB steering committee, designed and supervised the CKB study obtained funding, and, together with JL, CY, PP, LY, IM, RW, YC, HD, XY, and WH acquired the data. AL and SG analyzed the data, AL wrote the first draft of the manuscript. DiS helped to interpret the results, contributed to the critical revision of the manuscript for important intellectual content, and approved the final version. All authors reviewed and approved the final manuscript.

Guarantor Statement

DiS is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of interest statement

The authors declare that they have no conflict of interests.

Funding

This work was supported by the National Key R&D Program of China (2023YFC2509400) and the National Natural Science Foundation of China (82103920, 82388102, 82192901, 82192904, 82192900). The CKB baseline survey and the first re-survey were supported by the Kadoorie Charitable Foundation in Hong Kong. The long-term follow-up has been

supported by Wellcome grants to Oxford University (212946/Z/18/Z, 202922/Z/16/Z, 104085/Z/14/Z, 088158/Z/09/Z) and grants (2016YFC0900500) from the National Key R&D Program of China, National Natural Science Foundation of China (82388102, 81390540, 91846303, 81941018), and Chinese Ministry of Science and Technology (2011BAI09B01). The UK Medical Research Council (MC_UU_00017/1, MC_UU_12026/2, MC_U137686851), Cancer Research UK (C16077/A29186, C500/A16896) and the British Heart Foundation (CH/1996001/9454), provide core funding to the Clinical Trial Service Unit and Epidemiological Studies Unit at Oxford University for the project.

Reference

1. Magliano DJ, Boyko EJ, IDF Diabetes Atlas 10th edition scientific committee: *IDF Diabetes Atlas*. Brussels, International Diabetes Federation, 2021
2. Piché M-E, Tchernof A, Després J-P: Obesity Phenotypes, Diabetes, and Cardiovascular Diseases. Circ Res 2020;126:1477-1500
3. Safiri S, Karamzad N, Kaufman JS, Bell AW, Nejadghaderi SA, Sullman MJM, Moradi-Lakeh M, Collins G, Kolahi A-A: Prevalence, Deaths and Disability-Adjusted-Life-Years (DALYs) Due to Type 2 Diabetes and Its Attributable Risk Factors in 204 Countries and Territories, 1990-2019: Results From the Global Burden of Disease Study 2019. Front Endocrinol (Lausanne) 2022;13:838027
4. Lyall DM, Celis-Morales C, Ward J, Iliodromiti S, Anderson JJ, Gill JMR, Smith DJ, Ntuk UE, Mackay DF, Holmes MV, Sattar N, Pell JP: Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank: A Mendelian Randomization Study. JAMA Cardiol 2017;2:882-889
5. Zhu J, Zong G, Lu L, Gan W, Ji L, Hu R, Ye X, Sun L, Loos RJF, Li H, Lin X: Association of genetic predisposition to obesity with type 2 diabetes risk in Han Chinese individuals. Diabetologia 2014;57:1830-1833
6. Bann D, Wright L, Hardy R, Williams DM, Davies NM: Polygenic and socioeconomic risk for high body mass index: 69 years of follow-up across life. PLoS Genet 2022;18:e1010233
7. Kahan LG, Mehrzad R: Chapter 10 - Environmental factors related to the obesity epidemic. In *Obesity* Mehrzad R, Ed., Elsevier, 2020, p. 117-139
8. Chen Z, Chen J, Collins R, Guo Y, Peto R, Wu F, Li L: China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. Int J Epidemiol 2011;40:1652-1666

9. Chen Z, Lee L, Chen J, Collins R, Wu F, Guo Y, Linksted P, Peto R: Cohort profile: the Kadoorie Study of Chronic Disease in China (KSCDC). Int J Epidemiol 2005;34:1243-1249
10. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R: 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:e1001779
11. de Jong M, Woodward M, Peters SAE: Diabetes, Glycated Hemoglobin, and the Risk of Myocardial Infarction in Women and Men: A Prospective Cohort Study of the UK Biobank. Diabetes Care 2020;43:2050-2059
12. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, Narita A, Konuma T, Yamamoto K, Akiyama M, Ishigaki K, Suzuki A, Suzuki K, Obara W, Yamaji K, Takahashi K, Asai S, Takahashi Y, Suzuki T, Shinozaki N, Yamaguchi H, Minami S, Murayama S, Yoshimori K, Nagayama S, Obata D, Higashiyama M, Masumoto A, Koretsune Y, Ito K, Terao C, Yamauchi T, Komuro I, Kadowaki T, Tamiya G, Yamamoto M, Nakamura Y, Kubo M, Murakami Y, Yamamoto K, Kamatani Y, Palotie A, Rivas MA, Daly MJ, Matsuda K, Okada Y: A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet 2021;53:1415-1424
13. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, Frayling TM, Hirschhorn J, Yang J, Visscher PM: Meta-analysis of genome-wide association studies for height and body mass index in ～700000 individuals of European ancestry. Hum Mol Genet 2018;27:3641-3649
14. Ruan Y, Lin Y-F, Feng Y-CA, Chen C-Y, Lam M, Guo Z, He L, Sawa A, Martin AR, Qin S, Huang H, Ge T: Improving polygenic prediction in ancestrally diverse populations. Nat Genet 2022;54:573-580
15. Khera AV, Chaffin M, Zekavat SM, Collins RL, Roselli C, Natarajan P, Lichtman JH, D'Onofrio G, Mattera J, Dreyer R, Spertus JA, Taylor KD, Psaty BM, Rich SS, Post W, Gupta N, Gabriel S, Lander E, Ida Chen Y-D, Talkowski ME, Rotter JI, Krumholz HM, Kathiresan S: Whole-Genome Sequencing to Characterize Monogenic and Polygenic Contributions in Patients Hospitalized With Early-Onset Myocardial Infarction. Circulation 2019;139:1593-1602
16. Deborah JT, Daniel W, Saskia S, Iliana P, Rachel M, Kevin S, William AT, Edward JB, Fernando R-M, Carla G-D, Duncan P, Priyanka S, James H, Marta F, Genomics England Research C, Gil M, Vincent P, Peter D, Michael EW: UK Biobank release and systematic evaluation of optimised polygenic risk scores for 53 diseases and quantitative traits. medRxiv 2022:2022.2006.2016.22276246
17. Hawkes G, Yengo L, Vedantam S, Marouli E, Beaumont RN, Tyrrell J, Weedon MN, Hirschhorn J, Frayling TM, Wood AR: Identification and analysis of individuals who deviate from their genetically-predicted phenotype. PLoS Genet 2023;19:e1010934
18. Rhee T-M, Choi J, Lee H, Merino J, Park J-B, Kwak SH: Discrepancy Between Genetically Predicted and Observed BMI Predicts Incident Type 2 Diabetes. Diabetes Care 2024;47:1826-1833
19. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, de Ferranti SD, Floyd J, Fornage M, Gillespie C, Isasi CR, Jiménez MC, Jordan LC, Judd SE, Lackland D, Lichtman JH, Lisabeth L, Liu S, Longenecker CT, Mackey RH, Matsushita K, Mozaffarian

D, Mussolino ME, Nasir K, Neumar RW, Palaniappan L, Pandey DK, Thiagarajan RR, Reeves MJ, Ritchey M, Rodriguez CJ, Roth GA, Rosamond WD, Sasson C, Towfighi A, Tsao CW, Turner MB, Virani SS, Voeks JH, Willey JZ, Wilkins JT, Wu JH, Alger HM, Wong SS, Muntner P: Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. Circulation 2017;135:e146-e603
20. Zhou B-F: Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults--study on optimal cut-off points of body mass index and waist circumference in Chinese adults. Biomed Environ Sci 2002;15:83-96
21. World Health Organization: *Obesity: preventing and managing the global epidemic: report of a WHO consultation*. Geneva, Word Health Organization, 2000
22. Neel JV: Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? Am J Hum Genet 1962;14:353-362
23. Kyrou I, Tsigos C, Mavrogianni C, Cardon G, Van Stappen V, Latomme J, Kivelä J, Wikström K, Tsochev K, Nanasi A, Semanova C, Mateo-Gallego R, Lamiquiz-Moneo I, Dafoulas G, Timpel P, Schwarz PEH, Iotova V, Tankova T, Makrilakis K, Manios Y: Sociodemographic and lifestyle-related risk factors for identifying vulnerable groups for type 2 diabetes: a narrative review with emphasis on data from Europe. BMC Endocr Disord 2020;20:134
24. UK Prospective Diabetes Study (UKPDS). VIII. Study design, progress and performance. Diabetologia 1991;34:877-890
25. Wang S, Shen J, Koh W-P, Yuan J-M, Gao X, Peng Y, Xu Y, Shi S, Huang Y, Dong Y, Zhong VW: Comparison of race- and ethnicity-specific BMI cutoffs for categorizing obesity severity: a multicountry prospective cohort study. Obesity (Silver Spring) 2024;32:1958-1966
26. Knott C, Bell S, Britton A: Alcohol Consumption and the Risk of Type 2 Diabetes: A Systematic Review and Dose-Response Meta-analysis of More Than 1.9 Million Individuals From 38 Observational Studies. Diabetes Care 2015;38:1804-1812
27. Rouhani MH, Salehi-Abargouei A, Surkan PJ, Azadbakht L: Is there a relationship between red or processed meat intake and obesity? A systematic review and meta-analysis of observational studies. Obes Rev 2014;15:740-748
28. Casas-Agustench P, López-Uriarte P, Bulló M, Ros E, Gómez-Flores A, Salas-Salvadó J: Acute effects of three high-fat meals with different fat saturations on energy expenditure, substrate oxidation and satiety. Clin Nutr 2009;28:39-45
29. Maruvada P, Leone V, Kaplan LM, Chang EB: The Human Microbiome and Obesity: Moving beyond Associations. Cell Host Microbe 2017;22:589-599
30. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, Carty C, Chaput J-P, Chastin S, Chou R, Dempsey PC, DiPietro L, Ekelund U, Firth J, Friedenreich CM, Garcia L, Gichu M, Jago R, Katzmarzyk PT, Lambert E, Leitzmann M, Milton K, Ortega FB, Ranasinghe C, Stamatakis E, Tiedemann A, Troiano RP, van der Ploeg HP, Wari V, Willumsen JF: World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020;54:1451-1462
31. Doherty A, Jackson D, Hammerla N, Plötz T, Olivier P, Granat MH, White T, van Hees VT, Trenell MI, Owen CG, Preece SJ, Gillions R, Sheard S, Peakman T, Brage S, Wareham NJ: Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. PLoS One 2017;12:e0169649

32. Li X, Zhang W, Zhang W, Tao K, Ni W, Wang K, Li Z, Liu Q, Lin J: Level of physical activity among middle-aged and older Chinese people: evidence from the China health and retirement longitudinal study. BMC Public Health 2020;20:1682
33. Pan X-F, Wang L, Pan A: Epidemiology and determinants of obesity in China. Lancet Diabetes Endocrinol 2021;9:373-392
34. Fang C, Liang Y: Social disparities in body mass index (BMI) trajectories among Chinese adults in 1991-2011. Int J Equity Health 2017;16:146
35. Popkin BM, Gordon-Larsen P: The nutrition transition: worldwide obesity dynamics and their determinants. Int J Obes Relat Metab Disord 2004;28 Suppl 3:S2-S9
36. NCD Risk Factor Collaboration: Rising rural body-mass index is the main driver of the global obesity epidemic in adults. Nature 2019;569:260-264
37. Kim TJ, Roesler NM, von dem Knesebeck O: Causation or selection - examining the relation between education and overweight/obesity in prospective observational studies: a meta-analysis. Obes Rev 2017;18:660-672
38. Dinsa GD, Goryakin Y, Fumagalli E, Suhrcke M: Obesity and socioeconomic status in developing countries: a systematic review. Obes Rev 2012;13:1067-1079
39. Sobal J, Stunkard AJ: Socioeconomic status and obesity: a review of the literature. Psychol Bull 1989;105:260-275
40. Salari-Moghaddam A, Keshteli AH, Haghighatdoost F, Esmaillzadeh A, Adibi P: Dietary glycemic index and glycemic load in relation to general obesity and central adiposity among adults. Clin Nutr 2019;38:2936-2942
41. Wang X, Ma H, Kou M, Heianza Y, Fonseca V, Qi L: Proteomic Signature of Body Mass Index and Risk of Type 2 Diabetes. Diabetes 2024;
42. Agrawal S, Luan Ja, Cummings BB, Weiss EJ, Wareham NJ, Khera AV: Relationship of Fat Mass Ratio, a Biomarker for Lipodystrophy, With Cardiometabolic Traits. Diabetes 2024;73:1099-1111

Tables

Table 1 Basic characteristics of the study participants in CKB and UKB by three BMI-PGM groups

	CKB				UKB			
	Discordantly low	Concordant	Discordantly high	p value	Discordantly low	Concordant	Discordantly high	p value
NO. (%)	16 006	32 014	16 008		25 629	51 256	25 629	
Age, years	52.3 (10.6)	52.3 (10.6)	52.7 (10.6)	<0.001	55.1 (8.3)	55.3 (8.3)	55.4 (8.3)	0.001
Men (%)	62.3	61.3	60.7	0.010	41.4	45.7	45.4	<0.001
Rural (%)	49.3	49.3	49.3	0.999	9.7	9.9	9.1	0.004
Education level (%)*				<0.001				<0.001
Low	48.5	47.3	48.1		18.4	21.2	24.2	
Medium	44.7	45.3	44.0		38.8	40.5	42.0	
High	6.8	7.3	7.9		42.8	38.3	33.8	
Household income (%)*				<0.001				<0.001
Low	29.7	28.1	25.5		47.9	50.0	53.8	
Medium	30.9	30.5	31.3		24.9	24.3	23.7	
High	39.4	41.4	43.2		27.3	25.7	22.5	
Smoking (%)				<0.001				<0.001
Never	63.5	64.4	64.7		56.2	55.4	55.7	
Previous	4.7	6.2	7.3		30.6	32.9	34.2	
Current	31.8	29.4	28.0		13.2	11.7	10.0	
Alcohol consumption (%)				0.001				<0.001
Never	46.3	46.1	46.6		8.1	8.6	10.8	
Previous	3.7	3.6	4.3		3.9	3.9	4.7	
Current	50.0	50.3	49.1		88.0	87.5	84.5	
Physical activity (%)				<0.001				<0.001

Q1	24.0	24.4	26.1		21.4	24.8	28.7	
Q2	23.9	25.3	26.4		24.5	24.7	24.0	
Q3	25.4	25.2	24.4		27.0	25.6	24.2	
Q4	26.7	25.1	23.1		27.1	24.9	23.1	
Sedentary duration (%)				<0.001				<0.001
Q1	18.6	16.9	15.8		54.4	47.2	41.3	
Q2	49.8	49.3	49.0		18.8	20.2	20.9	
Q3	31.6	33.8	35.2		26.9	32.6	37.8	
Healthy dietary habits (%)[†]	91.5	92.1	92.5	0.002	57.8	56.1	55.8	<0.001
Sleep hours (%)				0.002				<0.001
<7	24.0	22.9	22.3		25.5	27.7	30.2	
7-9	60.0	60.4	61.3		67.8	64.9	61.4	
≥9	16.0	16.6	16.4		6.7	7.5	8.4	
Prevalent diseases at baseline (%)[‡]								
Diabetes	5.0	6.5	8.0	<0.001	4.5	7.9	11.4	<0.001
Hypertension	27.2	35.6	44.9	<0.001	41.4	51.9	62.0	<0.001
CVD	4.6	5.5	6.2	<0.001	3.9	5.5	6.9	<0.001
Respiratory diseases	10.7	8.6	7.8	<0.001	40.7	43.1	47.4	<0.001
Cancer	0.6	0.5	0.5	0.021	6.8	6.6	6.9	0.337
BMI, kg/m^2	21.5 (2.3)	23.9 (3.4)	26.3 (2.9)	<0.001	23.9 (2.5)	27.5 (4.7)	30.9 (4.4)	<0.001
BMI-PRS	0.7 (0.7)	-0.0 (1.0)	-0.8 (0.7)	<0.001	0.5 (0.8)	-0.1 (1.0)	-0.7 (0.8)	<0.001

[*] Education level and household income were evaluated differently between the questionnaires of CKB and UKB. We reclassified these characteristics. The details of the classification were shown in the Supplementary Table 2.

[†] The details of the definitions for healthy dietary habits in CKB and UKB were shown in the Supplementary Table 2.

‡The details of the definitions for different diseases were shown in the Supplementary Table 2.

CVD=cardiovascular diseases. BMI=body mass index. BMI-PRS=polygenic risk score for body mass index.

Table 2 Regression coefficients (β) and 95% CI of the risk factors on BMI-PGM in CKB and UKB

	CKB		UKB	
	β (95% CI)*	p value	β (95% CI)*	p value
Age (years)	-0.16 (-0.19, -0.12)	<0.001	-0.28 (-0.31, -0.26)	<0.001
Sex				
Women	Ref.		Ref.	
Men	3.16 (2.19, 4.12)	<0.001	-0.02 (-0.47, 0.43)	0.930
Urban-rural status				
Rural	Ref.		Ref.	
Urban	-2.66 (-3.38, -1.94)	<0.001	-0.13 (-0.86, 0.61)	0.736
Education level†				
Low	Ref.		Ref.	
Medium	-0.70 (-1.36, -0.03)	0.041	-1.05 (-1.65, -0.44)	<0.001
High	0.81 (-0.42, 2.04)	0.195	-4.05 (-4.69, -3.40)	<0.001
p value for trend		0.910		<0.001
Household income†				
Low	Ref.		Ref.	
Medium	1.26 (0.52, 1.99)	<0.001	0.31 (-0.25, 0.87)	0.275
High	2.01 (1.26, 2.77)	<0.001	0.06 (-0.52, 0.65)	0.840
p value for trend		<0.001		0.820
Smoking status				
Never	Ref.		Ref.	
Previous	2.14 (0.75, 3.54)	0.003	1.39 (0.90, 1.87)	<0.001
Current	-4.56 (-5.54, -3.58)	<0.001	-5.34 (-6.06, -4.63)	<0.001
Alcohol consumption				
Never	Ref.		Ref.	
Previous	0.12 (-1.40, 1.64)	0.878	-1.52 (-2.82, -0.22)	0.022
Current	-0.25 (-0.87, 0.38)	0.442	-2.85 (-3.64, -2.06)	<0.001
Physical activity				
Q1	Ref.		Ref.	
Q2	-0.02 (-0.80, 0.76)	0.956	-2.51 (-3.13, -1.89)	<0.001
Q3	-1.16 (-1.98, -0.35)	0.005	-4.14 (-4.75, -3.52)	<0.001
Q4	-2.72 (-3.58, -1.86)	<0.001	-4.81 (-5.43, -4.19)	<0.001
p value for trend		<0.001		<0.001
Sedentary duration				
Q1	Ref.		Ref.	
Q2	1.99 (1.22, 2.76)	<0.001	4.53 (3.95, 5.11)	<0.001
Q3	3.55 (2.71, 4.39)	<0.001	6.78 (6.28, 7.28)	<0.001
p value for trend		<0.001		<0.001
Fresh vegetables intake†				
Less frequent	Ref.		Ref.	

Frequent	0.98 (-0.30, 2.27)	0.135	0.04 (-1.14, 1.22)	0.946
Fresh fruit intake[†]				
Less frequent	Ref.		Ref.	
Frequent	2.71 (2.09, 3.33)	<0.001	0.45 (-0.14, 1.05)	0.135
Fish intake[†]				
Less frequent	Ref.		Ref.	
Frequent	0.48 (-0.18, 1.15)	0.154	-0.03 (-0.47, 0.41)	0.888
Meat intake[†]				
Less frequent	-1.53 (-2.13, -0.92)	<0.001	-3.01 (-4.37, -1.64)	<0.001
Frequent	Ref.		Ref.	
Sleep duration (h)				
<7	Ref.		Ref.	
7-9	0.98 (0.31, 1.65)	0.004	-2.13 (-2.80, -1.82)	<0.001
≥9	0.80 (-0.09, 1.69)	0.079	-0.01 (-0.90, 0.88)	0.975
p value for trend		0.051		<0.001

[*] Models adjusted for the history of diabetes, hypertension, cardiovascular diseases, respiratory diseases and cancer at baseline.

[†] Education level, household income, fresh vegetables intake frequency, fresh fruit intake frequency, fish intake frequency, and meat intake frequency were evaluated differently between the questionnaires of CKB and UKB. We reclassified these characteristics. The details of the classification were shown in the Supplementary Table 2.

Figures Legend

Figure 1 Correlation between BMI and BMI-PRS, BMI and BMI-PGM, and distribution of BMI-PGM by BMI in CKB and UKB

Panel (A) and Panel (B): Correlation between BMI and BMI-PRS in CKB (A) and UKB (B). Panel (C) and Panel (D): Correlation between BMI and BMI-PGM in CKB (C) and UKB (D). Panel (E) and Panel (F): Quartiles of BMI-PGM by BMI in CKB (E) and UKB (F). Panel (G) and Panel (H): Proportions of BMI-PGM groups within each BMI category in CKB (G) and UKB (H). In Panel (E) and Panel (G), the criteria for BMI classification in CKB was based on that published by the Working Group on Obesity in China: <18.5 kg/m^2 for the underweight category, 18.5-24.0 kg/m^2 for the normal weight category, 24.0-28.0 kg/m^2 for the overweight category, and ≥28.0 kg/m^2 for the obese category. In Panel (F) and Panel (H), the criteria for BMI classification in UKB was based on that published by the World Health Organization: <20.0 kg/m^2 for the underweight category, 20.0-25.0 kg/m^2 for the normal weight category, 25.0-30.0 kg/m^2 for the overweight category, and ≥30.0 kg/m^2 for the obese category.

Figure 2 Results of the restricted cubic spline and Kaplan-Meier estimates for the association between BMI-PGM and type 2 diabetes in CKB and UKB

Panel (A) and Panel (B): Restricted cubic spline plots for the association between BMI-PGM and the risk of type 2 diabetes among CKB's (A) and UKB's (B) participants without diabetes at baseline (CKB: n=59 873; UKB: n=94 398). Models, with age as the time scale, were stratified by age at risk (five-year bands) and study areas (ten groups in CKB and 22

groups in UKB), and further adjusted for sex, education level, household income, smoking, alcohol consumption, dietary habits, physical activity, sedentary duration, sleep duration, the family history of diabetes, and BMI. The purple dashed line represents the line where the point corresponding to BMI-PGM in the curve was located when HR=0. *P* values of nonlinearity and significance of the association were described in the graph. Panel (C) and Panel (D): Kaplan-Meier estimates of the risk of type 2 diabetes were shown in CKB's (C) and UKB's (D), respectively (CKB: n=59 873; UKB: n=94 398). In each analysis, participants with concordant BMI-PGM were taken as the reference group.

Figure 3 Adjusted hazard ratios for incident type 2 diabetes associated with BMI-PGM in CKB and UKB

Adjusted hazard ratios were calculated among participants without diabetes at baseline (CKB: n=59 873; UKB: n=94 398). Models, with concordant BMI-PGM group as the reference and age as the time scale, were stratified by age at risk (five-year bands) and study areas (ten groups in CKB and 22 groups in UKB), and further adjusted for sex, education level, household income, smoking, alcohol consumption, dietary habits, physical activity, sedentary duration, sleep duration, the family history of diabetes, and BMI. Incident rates were per 1000 person-year. Weights were from random-effects meta-analyses. The criteria for BMI classification in CKB was based on that published by the Working Group on Obesity in China: <18.5 kg/m^2 for the underweight category, 18.5-24.0 kg/m^2 for the normal weight category, 24.0-28.0 kg/m^2 for the overweight category, and ≥28.0 kg/m^2 for the obese category. Due to the low number of type 2 diabetes events in the discordantly high group of

people with underweight and the discordantly low group of people with obesity in the CKB and UKB, HR values could not be calculated, so only the HRs of the normal-weight and overweight individuals were shown.

Figure 1

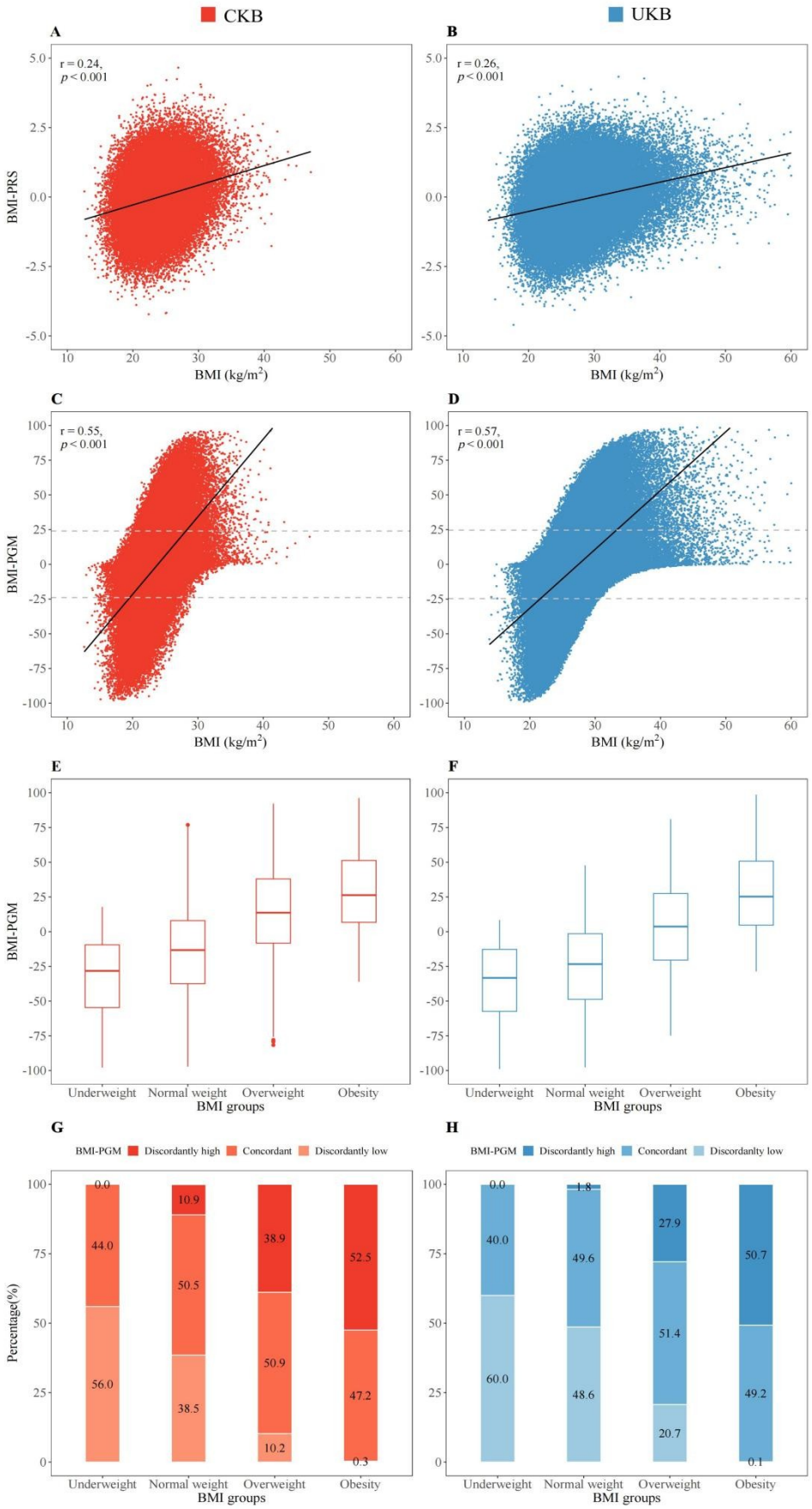

CKB
UKB
A
B
C
D
E
F
G
H
r = 0.24, p < 0.001
r = 0.26, p < 0.001
r = 0.55, p < 0.001
r = 0.57, p < 0.001
BMI-PRS
BMI-PGM
BMI (kg/m2)
BMI groups
Underweight
Normal weight
Overweight
Obesity
BMI-PGM Discordantly high Concordant Discordantly low
Discordanlty low
Percentage(%)
0.0
44.0
56.0
10.9
50.5
38.5
38.9
50.9
10.2
52.5
47.2
0.3
1.8
40.0
60.0
49.6
48.6
27.9
51.4
20.7
50.7
49.2
0.1

Figure 2

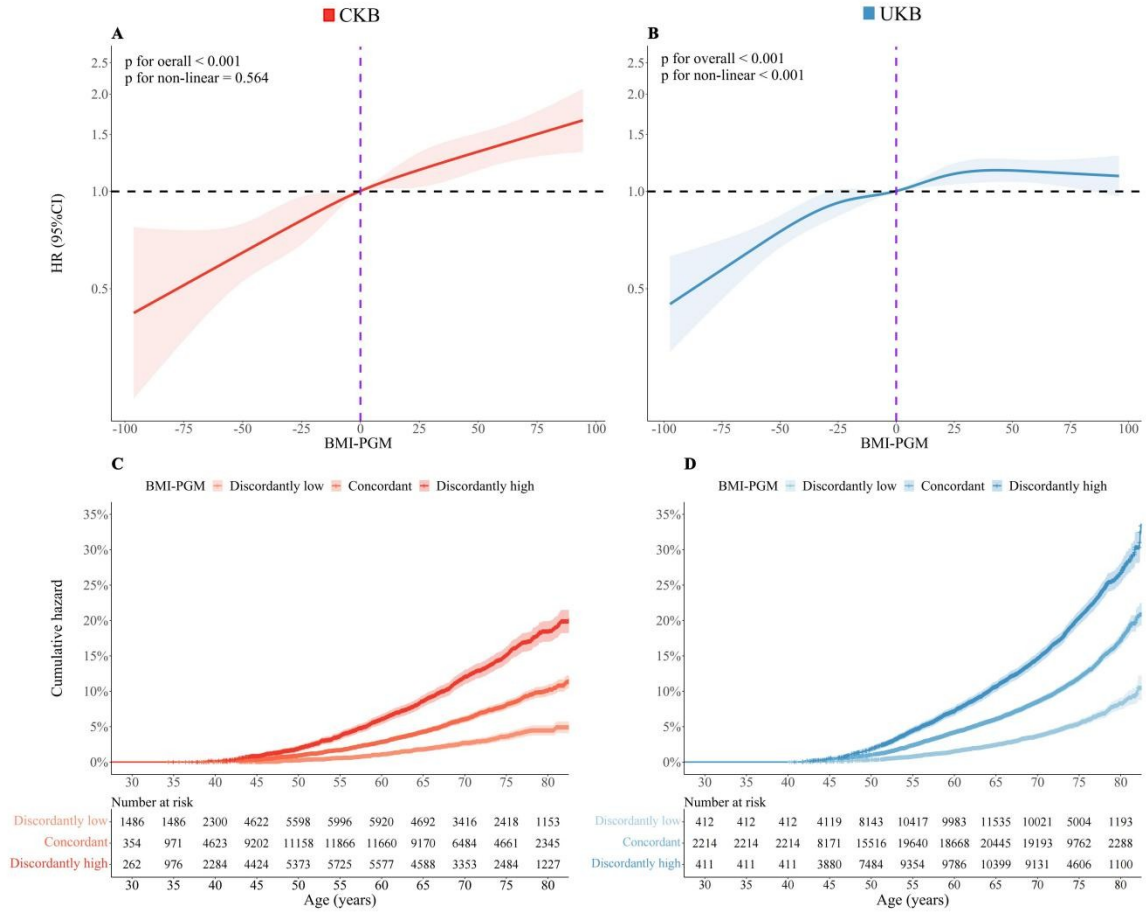
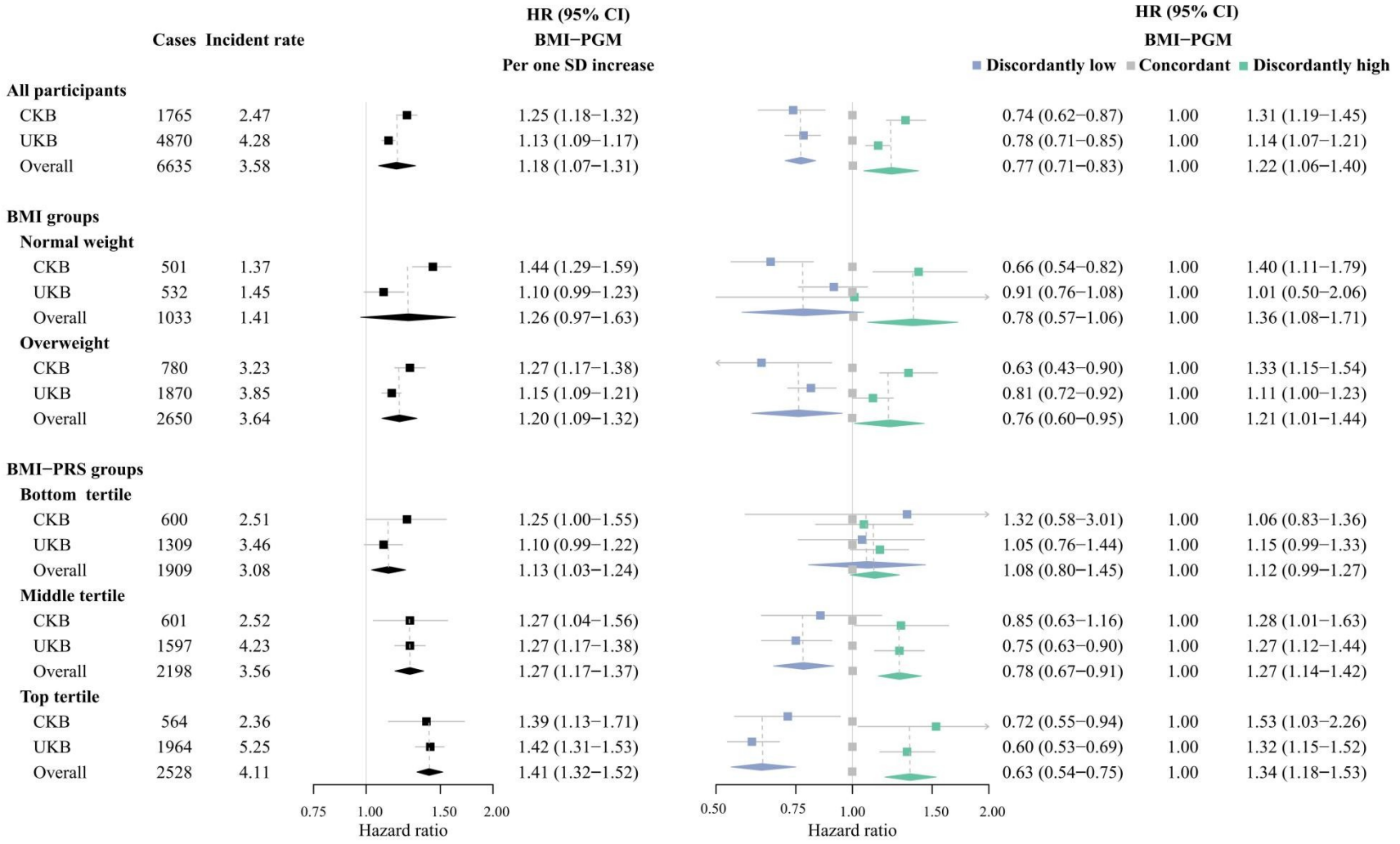

CKB
UKB
A
B
p for oerall < 0.001
p for non-linear = 0.564
p for overall < 0.001
p for non-linear < 0.001
HR (95%CI)
BMI-PGM
C
D
BMI-PGM Discordantly low Concordant Discordantly high
Cumulative hazard
Number at risk
Discordantly low 1486 1486 2300 4622 5598 5996 5920 4692 3416 2418 1153
Concordant 354 971 4623 9202 11158 11866 11660 9170 6484 4661 2345
Discordantly high 262 976 2284 4424 5373 5725 5577 4588 3353 2484 1227
Discordantly low 412 412 412 4119 8143 10417 9983 11535 10021 5004 1193
Concordant 2214 2214 2214 8171 15516 19640 18668 20445 19193 9762 2288
Discordantly high 411 411 411 3880 7484 9354 9786 10399 9131 4606 1100
Age (years)

Figure 3


Cases
Incident rate
HR (95% CI)
BMI−PGM
Per one SD increase
HR (95% CI)
BMI−PGM
Discordantly low
Concordant
Discordantly high
All participants
CKB 1765 2.47 1.25 (1.18−1.32) 0.74 (0.62−0.87) 1.00 1.31 (1.19−1.45)
UKB 4870 4.28 1.13 (1.09−1.17) 0.78 (0.71−0.85) 1.00 1.14 (1.07−1.21)
Overall 6635 3.58 1.18 (1.07−1.31) 0.77 (0.71−0.83) 1.00 1.22 (1.06−1.40)
BMI groups
Normal weight
CKB 501 1.37 1.44 (1.29−1.59) 0.66 (0.54−0.82) 1.00 1.40 (1.11−1.79)
UKB 532 1.45 1.10 (0.99−1.23) 0.91 (0.76−1.08) 1.00 1.01 (0.50−2.06)
Overall 1033 1.41 1.26 (0.97−1.63) 0.78 (0.57−1.06) 1.00 1.36 (1.08−1.71)
Overweight
CKB 780 3.23 1.27 (1.17−1.38) 0.63 (0.43−0.90) 1.00 1.33 (1.15−1.54)
UKB 1870 3.85 1.15 (1.09−1.21) 0.81 (0.72−0.92) 1.00 1.11 (1.00−1.23)
Overall 2650 3.64 1.20 (1.09−1.32) 0.76 (0.60−0.95) 1.00 1.21 (1.01−1.44)
BMI−PRS groups
Bottom tertile
CKB 600 2.51 1.25 (1.00−1.55) 1.32 (0.58−3.01) 1.00 1.06 (0.83−1.36)
UKB 1309 3.46 1.10 (0.99−1.22) 1.05 (0.76−1.44) 1.00 1.15 (0.99−1.33)
Overall 1909 3.08 1.13 (1.03−1.24) 1.08 (0.80−1.45) 1.00 1.12 (0.99−1.27)
Middle tertile
CKB 601 2.52 1.27 (1.04−1.56) 0.85 (0.63−1.16) 1.00 1.28 (1.01−1.63)
UKB 1597 4.23 1.27 (1.17−1.38) 0.75 (0.63−0.90) 1.00 1.27 (1.12−1.44)
Overall 2198 3.56 1.27 (1.17−1.37) 0.78 (0.67−0.91) 1.00 1.27 (1.14−1.42)
Top tertile
CKB 564 2.36 1.39 (1.13−1.71) 0.72 (0.55−0.94) 1.00 1.53 (1.03−2.26)
UKB 1964 5.25 1.42 (1.31−1.53) 0.60 (0.53−0.69) 1.00 1.32 (1.15−1.52)
Overall 2528 4.11 1.41 (1.32−1.52) 0.63 (0.54−0.75) 1.00 1.34 (1.18−1.53)
0.75
1.00
1.50
2.00
Hazard ratio
0.50
0.75
1.00
1.50
2.00
Hazard ratio