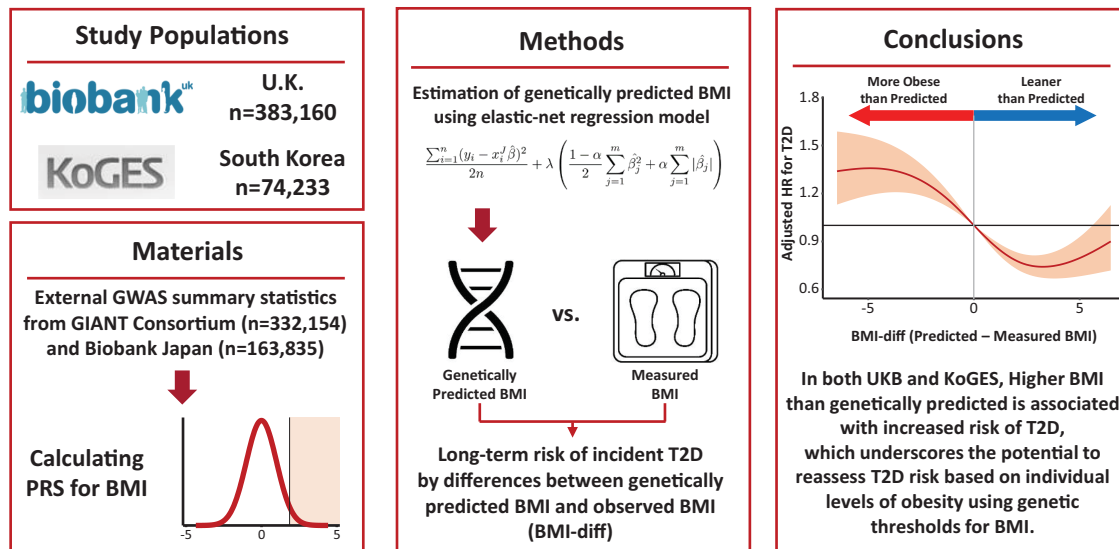


Discrepancy Between Genetically Predicted and Observed BMI Predicts Incident Type 2 Diabetes

Tae-Min Rhee, Jaewon Choi, Hyunsuk Lee, Jordi Merino, Jun-Bean Park, and Soo Heon Kwak

Diabetes Care 2024;47(10):1826–1833 | <https://doi.org/10.2337/dc24-0879>

Discrepancy Between Genetically Predicted and Observed BMI Predicts Incident Type 2 Diabetes



GWAS, genome-wide association study; GIANT, Genetic Investigation of ANthropometric Traits; HR, hazard ratio; KoGES, Korean Genome and Epidemiology Study; PRS, polygenic risk scores; T2D, type 2 diabetes; UKB, UK Biobank.

ARTICLE HIGHLIGHTS

- **Why did we undertake this study?**

We aimed to explore whether deviations from genetically predicted BMI affect type 2 diabetes (T2D) risk, independent of actual BMI.

- **What is the specific question we wanted to answer?**

Can BMI exceeding genetic set points predict T2D risk and be associated with insulin sensitivity?

- **What did we find?**

Individuals with higher observed BMI than genetically predicted BMI showed an increased T2D risk and impaired insulin sensitivity.

- **What are the implications of our findings?**

Our results support personalized BMI targets and suggest that tailored lifestyle modifications could be important in T2D prevention. Further research is needed on the underlying pathophysiology of T2D and its remission after tailored weight reduction.



Discrepancy Between Genetically Predicted and Observed BMI Predicts Incident Type 2 Diabetes

Diabetes Care 2024;47:1826–1833 | <https://doi.org/10.2337/dc24-0879>

Tae-Min Rhee,¹ Jaewon Choi,²
Hyunsuk Lee,^{3,4,5} Jordi Merino,⁶
Jun-Bean Park,³ and Soo Heon Kwak^{2,3}

OBJECTIVE

Obesity is a key predictor of type 2 diabetes (T2D). However, metabolic complications are not solely due to increased BMI. We hypothesized that differences between genetically predicted BMI and observed BMI (BMI-diff) could reflect deviation from individual set point and may predict incident T2D.

RESEARCH DESIGN AND METHODS

From the UK Biobank cohort, we selected participants of European ancestry without T2D ($n = 332,154$). The polygenic risk score for BMI was calculated via Bayesian regression and continuous shrinkage priors (PRS-CS). According to the BMI-diff, the 10-year risk of T2D was assessed using multivariable Cox proportional hazards model. Independent data from the Korean Genome and Epidemiology Study (KoGES) cohort from South Korea ($n = 7,430$) were used for replication.

RESULTS

Participants from the UK Biobank were divided into train ($n = 268,041$) and test set ($n = 115,119$) to establish genetically predicted BMI. In the test set, the genetically predicted BMI explained 7.1% of the variance of BMI, and there were 3,599 T2D cases (3.1%) during a 10-year follow-up. Participants in the higher quintiles of BMI-diff (more obese than genetically predicted) had significantly higher risk of T2D than those in the lowest quintile after adjusting for observed BMI: the adjusted hazard ratio of the 1st quintile (vs. 5th quintile) was 1.61 (95% CI 1.26–2.05, $P < 0.001$). Results were consistent among individuals in the KoGES study. Moreover, higher BMI than predicted was associated with impaired insulin sensitivity.

CONCLUSIONS

Having a higher BMI than genetically predicted is associated with an increased risk of T2D. These findings underscore the potential to reassess T2D risk based on individual levels of obesity using genetic thresholds for BMI.

The global rise in obesity has resulted in a significant increase in the prevalence of type 2 diabetes (T2D), posing a major health concern worldwide (1). BMI is strongly associated with adverse metabolic consequences, such as T2D (2). However, relying solely on BMI does not provide an accurate representation of individual's risk of developing metabolic diseases (3). Some individuals with a normal BMI may still have an elevated risk for obesity-related health issues, whereas others with a high BMI may remain resilient to common metabolic complications (4). In this regard, the concept of the personal fat threshold, suggesting that each individual has a unique fat

¹Department of Internal Medicine, Seoul National University Hospital Healthcare System Gangnam Center, Seoul, Republic of Korea

²Innovative Biomedical Technology Research Institute, Seoul National University Hospital, Seoul, Republic of Korea

³Department of Internal Medicine, Seoul National University College of Medicine and Seoul National University Hospital, Seoul, Republic of Korea

⁴Department of Translational Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

⁵Genomic Medicine Institute, Medical Research Center, Seoul National University College of Medicine, Seoul, Republic of Korea

⁶Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Copenhagen, Denmark

Corresponding author: Soo Heon Kwak, shkwak@snu.ac.kr

Received 30 April 2024 and accepted 19 July 2024

This article contains supplementary material online at <https://doi.org/10.2337/figshare.26351821>.

© 2024 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/journals/pages/license>.

set point, has emerged as a tool for understanding the development of T2D in individuals without obesity and the remission of diabetes after substantial weight loss in persistently individuals with obesity (5,6). However, little evidence is available in the context of the development of metabolic diseases.

A genetic epidemiology approach could be leveraged to test the personal fat threshold hypothesis in the development of T2D and provide insights into T2D pathophysiology. It is now possible to generate polygenic risk scores (PRS) based on common genetic variants derived from genome-wide association studies (GWAS) that can predict BMI at the individual level (7). By examining the discrepancy between genetically predicted BMI (BMI-pred) and observed BMI, we could obtain a quantitative measure representing the deviation from the biological set point and provide an opportunity to unveil the relationship between BMI and T2D at the individual level (8).

Thus, in this study, we hypothesized that BMI-pred could represent a threshold for an individual's susceptibility to obesity and sought to investigate the effect of the difference between an individual's observed BMI and BMI-pred (i.e., BMI-diff), on the occurrence of incident T2D.

RESEARCH DESIGN AND METHODS

Study Participants: UK Biobank Cohort

The UK Biobank is a nationwide longitudinal cohort study that included >500,000 individuals aged 40–69 years between 2006 and 2010. The current study received approval from the UK Biobank Review Committee under application number 44196. Ethical approval for the UK Biobank study was granted by the North West Multi-Center Research Ethics Committee (Manchester, U.K.). The full protocols are described elsewhere (9). All participants provided written informed consent.

A total of 502,413 participants were initially eligible, and exclusions were made for individuals 1) of non-European ancestry ($n = 92,859$; field ID 22006), 2) with high missingness or excess heterozygosity on genetic quality control ($n = 731$; field ID 22027), 3) with significant genetic kinship to other participants ($n = 164$; field ID 22021), 4) with missing values on BMI

data and covariates ($n = 1,298$), and 5) with previous diagnosis of T2D ($n = 20,501$). Additionally, participants with BMI values >3 SDs away from the mean of the initial population ($n = 3,700$) were excluded to minimize the influence of rare monogenic causes of obesity, as commonly defined in other genetic epidemiology studies (10). The final analysis included 332,154 participants (Supplementary Fig. 1A).

Genotyping: UK Biobank

Genotyping of the UK Biobank participants was conducted as described elsewhere (11), using the custom UK Biobank Lung Exome Variant Evaluation Axiom (Affymetrix; $n = 49,950$; including 807,411 single nucleotide polymorphisms [SNPs]) or the UK Biobank Axiom array (Affymetrix; $n = 438,427$; including 825,927 SNPs). Genotype data were imputed based on reference panels, including the Haplotype Reference Consortium, UK10K, and 1000 Genomes Phase 3 panels (12–14). SNPs not in Hardy-Weinberg equilibrium ($P < 1.0 \times 10^{-6}$), with minor allele frequency <1%, or with genotyping failure rate >5% were excluded.

Study Participants: Korean Genome and Epidemiology Study Population-Based Cohort

The Korean Genome and Epidemiology Study (KoGES) cohort, part of the National Biobank of Korea, is a prospective study that collected comprehensive phenotypic measures and biological samples including genetic data. Participants were recruited from the national health examinee registry, with an age of ≥ 40 at baseline. KoGES data consist of three cohorts: community-based (Ansan/Ansung cohort), urban, and rural cohorts. Anthropometric and clinical data were collected through physical examinations, laboratory tests, and interviews about disease status (15). From a total of 92,707 participants ($n = 82,677$ in the urban/rural cohort; $n = 10,030$ in the Ansan/Ansung cohort), those who did not pass the genotype quality control (QC) were excluded, as described below. For the Ansan/Ansung cohort, which was used as the test set, 1) those who had missing BMI data, 2) had previous T2D diagnosis, and 3) had extreme BMI values over ± 3 SD were also excluded (Supplementary Fig. 1B).

Genotyping: KoGES

All participants in the KoGES analysis were genotyped using the KoreanChip,

a customized array optimized for the Korean population. The details of the KoreanChip contents (833,000 variants) and the genotyping in the KoGES cohorts can be found elsewhere (16). QC measures were applied to exclude samples with low call rate (<97%), excessive heterozygosity, excessive singletons, sex discrepancy, and cryptic first-degree relatives. SNPs not in Hardy-Weinberg equilibrium ($P < 1.0 \times 10^{-6}$) or low call rate (<95%) were also excluded. After QC, data were phased and imputed using the Eagle v2.3 and IMPUTE 4 algorithms, respectively, with the 1000 Genomes Project Phase 3 data and the Korean reference genome as the reference panel. Variants with an imputation quality score <0.8 and a minor allele frequency <1% were excluded after imputation. We analyzed 6,430,431 variants in total after these processes.

Calculating PRS Using the Continuous Shrinkage Method

PRS were calculated using the continuous shrinkage (PRS-CS) method (17). The PRS-CS method is a polygenic prediction method that infers posterior effect sizes of each SNP from genome-wide association summary statistics and external linkage disequilibrium reference panel data (for UK Biobank, 1000 Genomes Project Phase 3 European reference; for KoGES, 1000 Genomes Project Phase 3 East Asian reference). After acquiring the posterior effect sizes calculated by Bayesian regression and CS priors, each SNP was coded as 0, 1, or 2 according to the number of risk alleles, and the PRS-CS was calculated by standardizing the cumulative effect of each SNP for each participant. The effect sizes and P values of each SNP for BMI were obtained from previously published GWAS summary statistics data from the Genetic Investigation of Anthropometric Traits (GIANT) consortium for the UK Biobank data, which analyzed 332,154 participants with 2,554,637 variants (18). The summary statistics of the GWAS data from the Biobank Japan, including 163,835 participants with 13,236,464 variants, was used for the KoGES analysis (19).

Estimating Genetically BMI-Pred From PRS-CS: UK Biobank

The final study population for analysis in the UK Biobank cohort was randomly divided into a train set ($n = 268,041$) and a test set ($n = 115,119$) in a seven-to-three

ratio. The PRS-CS-auto method, which does not require independent validation data, was used to calculate the PRS-CS for BMI in the entire cohort. The train set was used to develop a genetic prediction model for BMI using the PRS-CS, and the test set was reserved to evaluate the effect of BMI-diff on incident T2D. An optimal elastic-net regression model was built using the calculated PRS-CS and observed BMI values in the train set. The genetically BMI-pred for each participant in the test set was obtained using the model derived from the train set. We then calculated the numerical difference (BMI-diff = BMI-pred – observed BMI) between the predicted and observed BMI in the test set. Higher or positive values represent leaner than genetically predicted, and lower or negative values more obese than predicted.

Estimating Genetically BMI-Pred From PRS-CS: KoGES

For the KoGES cohort, the urban/rural subcohort was randomly split into train and test sets at a seven-to-three ratio. We used the train set ($n = 45,809$) to create an optimal PRS-CS model. This model was then applied to both the test set ($n = 20,994$) of the Urban/Rural cohort and the Ansan/Ansung cohort ($n = 7,430$; termed the analysis set). An elastic-net regression model was developed in the test set to estimate genetically BMI-pred from the PRS-CS. Using this model, the BMI-pred was obtained from the PRS-CS in the analysis set, and the absolute difference (BMI-diff = BMI-pred – observed BMI) between the predicted and measured BMI was calculated.

Definitions of Covariates and Outcomes

The primary outcome of this study was the development of T2D during the follow-up. T2D was defined using inpatient hospital registry data in the UK Biobank, using the corresponding ICD-10 code E11. Detailed definitions of covariates and outcomes are provided in Supplementary Table 1. The covariates included quintile groups of BMI-diff, age, sex, BMI as a continuous value, enrollment center, Townsend deprivation index, current smoking, moderate-to-vigorous physical activity, hypertension, dyslipidemia, genotype array, and genetic principal component (PC) 1 to 10.

In the KoGES cohort, patients with T2D were defined as participants with self-reported T2D, on T2D medications,

or meeting the American Diabetes Association diagnostic criteria (fasting plasma glucose ≥ 7.0 mmol/L, 2-h postload glucose ≥ 11.1 mmol/mol or glycated hemoglobin [HbA_{1c}] ≥ 48 mmol/mol) (20). The covariates for the KoGES cohort included quintile groups of BMI-diff, age, sex, BMI as a continuous value, current smoking, moderate-to-vigorous physical activity, hypertension, dyslipidemia, and genetic PC 1 to 10.

Plasma insulin and glucose levels and 75-g oral glucose tolerance test (OGTT) data from the KoGES were used to estimate parameters of pancreatic β -cell function and insulin sensitivity and compared among the quintile groups of BMI-diff in KoGES data (21). Pancreatic β -cell function was estimated using the 60-min insulinogenic index (IGI₆₀) derived from plasma insulin and glucose levels at 0 min and 60 min of the OGTT. IGI₆₀ was calculated as $(\text{insulin}_{60\text{ min}} - \text{insulin}_{0\text{ min}} [\mu\text{U/mL}]) / (\text{glucose}_{60\text{ min}} - \text{glucose}_{0\text{ min}} [\text{mmol/L}])$. Insulin sensitivity was assessed using the composite (Matsuda) insulin sensitivity index, which was calculated as below, using 0-min, 60-min, and 120-min values of OGTT. The OGTT-derived disposition index was calculated by multiplying the IGI₆₀ and Matsuda index to account for β -cell function adjusted for the insulin sensitivity.

Statistical Analyses

Continuous variables are presented as mean $\pm$ SD and categorical variables as numbers (percentages). The χ^2 test was used for categorical variables and the ANOVA test for continuous variables to examine differences between baseline features. For the main results, Cox regression models were used to calculate unadjusted and adjusted hazard ratios (HRs) and CIs for T2D by BMI-diff, using the 5th quintile of BMI-diff as the reference. The same models were also used to investigate the associations between BMI-diff and the risk of T2D in the continuous scale. The restricted cubic spline curves of adjusted HRs with 95% CIs from these models were plotted to visualize the association between BMI-diff and the risk of T2D. The incidence rate per 1,000 person-years of T2D was calculated and compared among the quintile groups of BMI-diff. Participants were censored on the earliest date among the date of the end of follow-up, that of the outcome

occurrence, or that of death. The median follow-up duration was 12.6 years in the test set of the UK Biobank and 13.7 years in the analysis set of KoGES cohort. We censored participants based on the earliest occurrence of death, the outcome of interest, or 10 years of follow-up. We evaluated the predictive ability of the prediction models for T2D risk using the Harrell C-index and assessed the additive predictive effect of BMI-diff by the Somers D statistic. The category-free net reclassification index (NRI) was calculated to evaluate the degree of T2D reclassification for the addition of BMI-diff in the previous models. For both cohorts, we conducted sensitivity analyses to assess the robustness of our findings. These included 1) participants with extreme BMI over ± 3 SDs, 2) BMI < 30 kg/m², 3) BMI 25–30 kg/m², and 4) male or female sex. A P value < 0.05 was considered statistically significant. All statistical analyses were performed using R 4.2.3 software and Stata 17.0 software.

Data and Resource Availability

For UK Biobank data, all data used in this study are publicly available (<https://biobank.ndph.ox.ac.uk/>). The KoGES data are available from Korea National Institute of Health (<https://www.nih.go.kr/>), but restrictions apply to the availability of these data. Data may be available upon reasonable request and with permission of Korea National Institute of Health.

RESULTS

Clinical Characteristics of the Study Cohort

The characteristics of the 115,119 UK Biobank participants (mean age 56.8 ± 8.0 years, 45.2% men) without diabetes included in the test set are provided in Supplementary Table 2. More than half of the participants were classified as overweight, with a mean BMI of 27.0 ± 4.2 kg/m². The distribution of predicted and observed BMI is provided in Supplementary Fig. 2. Common genetic variation explained 7.1% of variance of predicted BMI. The BMI-pred was normally distributed. The clinical characteristics based on categories of BMI-diff are summarized in Table 1. Individuals in the lower quintile group, representing those with higher BMI than genetically predicted, were more likely to have a higher prevalence of hypertension and dyslipidemia,

Table 1—Baseline characteristics of test set from UK Biobank according to the discrepancy between measured and genetically BMI-pred

	1st quintile of BMI-diff (n = 23,024)	2nd quintile of BMI-diff (n = 23,024)	3rd quintile of BMI-diff (n = 23,024)	4th quintile of BMI-diff (n = 23,024)	5th quintile of BMI-diff (n = 23,023)	P
Demographics						
Age, years	57.1 ± 7.8	57.4 ± 7.9	57.2 ± 7.9	56.6 ± 8.1	55.6 ± 8.2	<0.001
Male sex	10,788 (46.9)	12,731 (55.3)	11,799 (51.2)	10,024 (43.5)	6,685 (29.0)	<0.001
Townsend deprivation index	−1.2 ± 3.1	−1.6 ± 2.9	−1.8 ± 2.8	−1.8 ± 2.8	−1.7 ± 2.9	<0.001
Measurements						
BMI, kg/m ²	33.2 ± 3.0	28.7 ± 1.6	26.5 ± 1.5	24.6 ± 1.5	22.1 ± 1.8	<0.001
Genetically BMI-pred, kg/m ²	27.4 ± 1.4	27.2 ± 1.4	27.3 ± 1.4	27.3 ± 1.3	27.7 ± 1.3	<0.001
BMI-diff, kg/m ²	−5.8 ± 2.5	−1.4 ± 0.7	0.8 ± 0.6	2.7 ± 0.6	5.6 ± 1.5	<0.001
Waist circumference, cm	103.4 ± 10.0	94.4 ± 8.4	88.9 ± 8.2	83.6 ± 8.2	76.2 ± 8.0	<0.001
Hip circumference, cm	113.0 ± 7.6	105.3 ± 5.1	101.9 ± 4.8	99.0 ± 4.7	94.9 ± 5.0	<0.001
Systolic blood pressure, mmHg	144.5 ± 18.9	143.0 ± 19.1	141.0 ± 19.3	138.1 ± 19.6	133.5 ± 19.7	<0.001
Diastolic blood pressure, mmHg	86.3 ± 10.2	84.3 ± 10.2	82.3 ± 10.2	80.6 ± 10.2	77.7 ± 10.2	<0.001
Lifestyle factors and comorbidities						
Current smoker	2,063 (9.0)	2,203 (9.6)	2,174 (9.4)	2,319 (10.1)	2,735 (11.9)	<0.001
Daily drinking	3,973 (17.3)	5,003 (21.7)	5,352 (23.2)	5,345 (23.2)	5,174 (22.5)	<0.001
Moderate-to-vigorous physical activity	8,480 (47.3)	9,998 (53.3)	10,572 (56.1)	11,117 (58.6)	11,243 (59.7)	<0.001
IPAQ physical activity classification						<0.001
Low	4,447 (19.3)	3,653 (15.9)	3,213 (14.0)	2,799 (12.2)	2,577 (11.2)	
Moderate	7,248 (31.5)	7,689 (33.4)	7,766 (33.7)	7,762 (33.7)	7,627 (33.1)	
High	6,240 (27.1)	7,408 (32.2)	7,857 (34.1)	8,404 (36.5)	8,614 (37.4)	
Hypertension	9,763 (42.4)	7,497 (32.6)	6,045 (26.3)	4,689 (20.4)	3,497 (15.2)	<0.001
Dyslipidemia	5,526 (24.0)	4,808 (20.9)	3,889 (16.9)	2,986 (13.0)	1,952 (8.5)	<0.001
Laboratory results						
Total cholesterol, mmol/L	5.8 ± 1.2	5.8 ± 1.2	5.8 ± 1.1	5.8 ± 1.1	5.7 ± 1.1	<0.001
LDL-cholesterol, mmol/L	3.7 ± 0.9	3.7 ± 0.9	3.7 ± 0.9	3.6 ± 0.8	3.4 ± 0.8	<0.001
HDL-cholesterol, mmol/L	1.3 ± 0.3	1.4 ± 0.3	1.4 ± 0.4	1.5 ± 0.4	1.7 ± 0.4	<0.001
Triglyceride, mmol/L	2.2 ± 1.2	2.0 ± 1.1	1.8 ± 1.0	1.5 ± 0.8	1.2 ± 0.6	<0.001
Fasting glucose, mmol/L	5.1 ± 0.9	5.0 ± 0.8	5.0 ± 0.7	4.9 ± 0.7	4.9 ± 0.7	<0.001
HbA _{1c} , mmol/L	36.3 ± 5.1	35.4 ± 4.5	34.9 ± 3.9	34.5 ± 3.7	34.3 ± 3.6	<0.001

Data are presented as n (%) or mean ± SD. IPAQ, International Physical Activity Questionnaire.

higher socioeconomic deprivation, and be less physically active than those in higher quintile group. Conversely, the higher quintile groups, representing individuals who were leaner than genetically predicted, were younger, had a higher proportion of women, and had a higher proportion of current smokers and daily drinkers.

Risk of Incident T2D According to Discrepancy Between BMI-pred and Observed BMI

During a 10-year follow-up period, there were 3,599 T2D incident cases. The cumulative incidences at 10 years was 3.20%, and the incidence rate per 1,000 person-years was 3.23 (Supplementary Table 3). Investigating the associations between BMI-diff and the risk of T2D, we showed that a higher BMI-diff, indicating lower BMI than genetically predicted, was significantly associated with a lower

risk of incident T2D (Fig. 1). These associations were consistent when discrepancy between BMI-pred and observed BMI was categorized into quintiles of BMI-diff distribution. Compared with individuals in the top BMI-diff distribution, indicating individuals with lower BMI than genetically predicted, those with higher BMI than genetically predicted had an adjusted HR for T2D of 1.61 (95% CI 1.26–2.05, $P < 0.001$) (Table 2). These results were consistent in a range of sensitivity analyses including the population without excluding those with extreme BMI over ± 3 SDs (Supplementary Table 4), participants with BMI < 30 kg/m² (Supplementary Table 5), those with BMI between 25 and 30 kg/m² (Supplementary Table 6), and in both men and women (Supplementary Table 7).

Adding BMI-diff significantly improved the T2D predictive ability and reclassification index in models that included only BMI (the Harrell C-index improved from

0.701 [95% CI 0.693–0.709] to 0.718 [95% CI 0.710–0.726], $P < 0.001$; category-free NRI 4.28% [95% CI 40.97–7.59], $P = 0.012$) and selected clinical variables (the Harrell C-index improved from 0.776 [95% CI 0.768–0.784] to 0.781 [95% CI 0.773–0.789], $P < 0.001$; category-free NRI 6.56% [95% CI 2.78–10.34], $P < 0.001$) (Supplementary Table 8).

Independent Replication in the KoGES Cohort

The KoGES Ansan/Ansung cohort was analyzed for replication in the East Asian ethnicity, comprising 7,430 individuals (mean age 52.0 ± 8.9 years, 47.3% men). The KoGES population was younger and less obese compared with the UK Biobank cohort, with a mean BMI of 24.5 ± 3.0 kg/m² (Supplementary Table 9). The distribution of BMI-pred and observed BMI is provided in Supplementary Fig. 2. Common genetic variation explained 5.7%

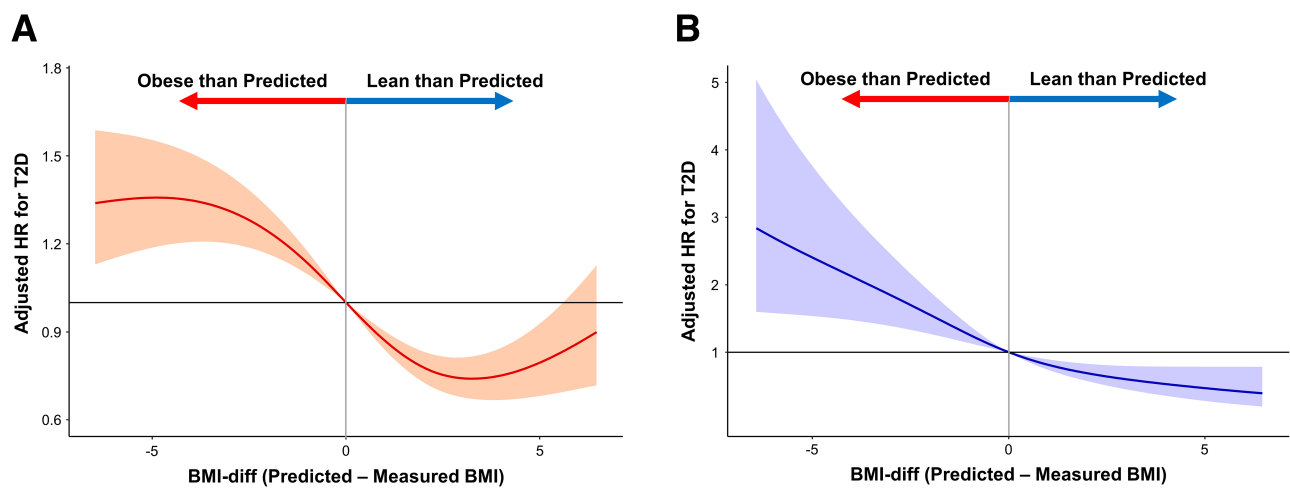


Figure 1—Adjusted risk plots of cubic spline analysis of 10-year risk of T2D by discrepancy between measured and genetically BMI-pred in UK Biobank (A) and KoGES Ansan-Ansung cohort (B). The shaded areas indicate the 95% CI.

of the predicted BMI variance. The clinical characteristic based on categories of BMI-diff in the KoGES cohort showed similar patterns as in the UK Biobank (Supplementary Table 10), with the higher quintile groups having higher proportion of women and a higher proportion of individuals with high-income status.

There were 658 T2D incident cases during the follow-up. The cumulative incidences at 10 years was 9.76%, and the incidence rate per 1,000 person-years was 10.44 (Supplementary Fig. 3). In the KoGES cohort, the difference in T2D incidence was even greater, with the 1st quintile of

BMI-diff having an incidence of 20.25 per 1,000 person-years compared with only 4.61 per 1,000 person-years in the 5th quintile. The adjusted HR for the 1st quintile was 3.05 (95% CI 1.67–5.58, $P < 0.001$). We observed a notable sex-based divergence, with a significant trend observed only among women. The adjusted HR for the 1st quintile group was 4.02 (95% CI 1.59–10.13, $P = 0.003$), using the 5th quintile group as the reference.

Similar to the UK Biobank cohort, incorporating BMI-diff significantly improved the predictive ability and reclassification index in models with BMI (the Harrell C-index from 0.618 [95% CI 0.599–0.637] to

0.641 [95% CI 0.620–0.662], $P < 0.001$; category-free NRI 14.21% [95% CI 6.21–22.21], $P < 0.001$) as well as selected clinical variables (the Harrell C-index from 0.669 [95% CI 0.649–0.689] to 0.682 [95% CI 0.662–0.702], $P = 0.002$; category-free NRI 18.06% [95% CI 10.06–26.06], $P < 0.001$) (Supplementary Table 11).

Characteristics Associated With Discrepancy in BMI

Given the detailed glycemic phenotyping in the KoGES cohort, we further investigated the association between BMI-diff and the indicators of insulin secretion and resistance. We showed significant differences

Table 2—Risk of type 2 diabetes according to difference between genetically BMI-pred and measured BMI						
UK Biobank cohort	Event <i>n</i> /total <i>N</i>	IR (95% CI) per 1,000 patient-years	Unadjusted HR (95% CI)	<i>P</i>	Adjusted HR (95% CI)*	<i>P</i>
1st quintile of BMI-diff	1,758/23,024	8.06 (7.69–8.44)	8.96 (7.75–10.36)	<0.001	1.61 (1.26–2.05)	<0.001
2nd quintile of BMI-diff	881/23,024	3.96 (3.71–4.23)	4.38 (3.76–5.10)	<0.001	1.48 (1.21–1.81)	<0.001
3rd quintile of BMI-diff	477/23,024	2.13 (1.94–2.33)	2.35 (2.00–2.77)	<0.001	1.09 (0.90–1.34)	0.383
4th quintile of BMI-diff	279/23,024	1.24 (1.10–1.39)	1.37 (1.14–1.64)	0.001	0.89 (0.73–1.10)	0.295
5th quintile of BMI-diff	204/23,023	0.91 (0.79–1.04)	1 (reference)	—	1 (reference)	—
KoGES Ansan/Ansung cohort	Event <i>n</i> /Total <i>N</i>	IR (95% CI) per 1,000 patient-years	Unadjusted HR (95% CI)	<i>P</i>	Adjusted HR (95% CI)†	<i>P</i>
1st quintile of BMI-diff	242/1,486	20.25 (17.86–22.97)	4.37 (3.29–5.81)	<0.001	3.05 (1.67–5.58)	<0.001
2nd quintile of BMI-diff	162/1,486	12.82 (10.99–14.95)	2.78 (2.06–3.75)	<0.001	2.15 (1.36–3.42)	0.001
3rd quintile of BMI-diff	110/1,486	8.62 (7.15–10.39)	1.87 (1.37–2.57)	<0.001	1.62 (1.08–2.44)	0.019
4th quintile of BMI-diff	85/1,486	6.60 (5.33–8.16)	1.43 (1.03–2.00)	0.034	1.36 (0.94–1.96)	0.103
5th quintile of BMI-diff	59/1,486	4.61 (3.57–5.95)	1 (reference)	—	1 (reference)	—

IR, incident rate. *Used multivariable Cox regression model including age, sex, and BMI as a continuous value, and enrollment center, Townsend deprivation index, current smoking, moderate-to-vigorous physical activity, hypertension, dyslipidemia, genotype array, and genetic principal components 1–10. †Used multivariable Cox regression model including age, sex, and BMI as a continuous value, and current smoking, moderate-to-vigorous physical activity, hypertension, dyslipidemia, and genetic principal components 1–10.

in the Matsuda index and the disposition index when individuals in the lower versus higher BMI-diff quintile groups were compared (Matsuda index, 8.5 ± 5.9 vs. 13.7 ± 7.8 ; disposition index, 85.8 ± 180.5 vs. 137.4 ± 318.7), indicating a decline in insulin sensitivity among individuals with higher BMI than genetically predicted (Table 3). There were no differences in IGI_{60} across the BMI-diff quintile groups.

Subgroup Analysis

The interactions between BMI-diff quintiles and age by median, sex, smoking status, hypertension, and dyslipidemia were assessed (Supplementary Table 12). The increase in T2D risk associated with BMI-diff was more predominant in relatively low-risk groups, such as younger age-groups, women, non/former smokers, and those without dyslipidemia. Statistically significant interactions were observed in the UK Biobank cohort for age-group, hypertension, and dyslipidemia. Although similar trends were observed in the KoGES cohort, they did not achieve statistical significance due to insufficient statistical power.

CONCLUSIONS

In the current study, we used data from two nationwide biobanks, from U.K. and South Korea, to investigate the discrepancy between observed and genetically BMI-pred on T2D risk. We found that individuals with a higher BMI than genetically predicted had an increased risk of T2D regardless of their own BMI. Deviation from genetically BMI-pred improved the ability to predict and reclassify T2D risk in both cohorts. We also showed that a higher-than-predicted BMI was associated with a significant decline in an estimate of insulin sensitivity and that the adverse impact of discrepant BMI on

glycemic outcomes was more likely to be relevant in women from East Asia.

The intricate relationship between T2D and weight gain, coupled with subsequent excessive fat accumulation in nonadipose tissue, such as the liver and pancreas, is widely acknowledged (6). However, a notable proportion of individuals newly diagnosed with T2D do not have obesity (22), highlighting the substantial variability in individual susceptibility to T2D, regardless of the extent of fat deposition within nonadipose tissues (23). In general, in cases of T2D with a normal BMI, rapid β -cell failure is suggested as the major pathophysiology rather than insulin resistance (24). However, there are genetic conditions, such as lipodystrophy, where severe insulin resistance can lead to diabetes, even at low BMI levels (25). In addition, the higher prevalence of T2D in men without obesity and those of Asian ethnicity implies distinct genetic predisposition to T2D (24). Currently, there is limited understanding of how these genotype-phenotype discordance in BMI might affect the risk of developing T2D. Our data suggest that deviation of BMI from a personal set point could be a novel precision marker for incident T2D.

Notably, the observed association in our study results remained significant even after adjusting for individual BMI levels. This suggests that the deviation from genetically determined BMI might hold closer relevance to the underlying pathophysiology driving T2D beyond the mere BMI measurements themselves. This observation is in line with the personal fat threshold hypothesis, postulating that each individual possesses a distinct set point of personal fat; even among individuals with normal body weights, achieving weight reduction below a certain threshold could culminate in T2D remission (5).

However, interpreting the associations observed in this study requires caution due to pleiotropy. BMI-diff as well as PRS for BMI encompasses genetic variations related to various BMI-related traits, which may lead to associations with T2D. Multiomic profiling of individuals with extreme differences between genetically predicted and observed BMI may further elucidate the mechanisms behind the association between obesity and T2D risk and help tailor therapeutic interventions.

We observed a substantial decrease in the insulin sensitivity among participants whose actual BMI exceeded their genetically BMI-pred. Those who had higher BMI than genetically predicted might not have had a sufficient compensatory increase in β -cell function to compensate for the marked decrease in insulin sensitivity. The interaction between the extent of insulin resistance and the level of β -cell function could contribute to the diverse spectrum of T2D presentations, especially within individuals without obesity (26). Our results imply the possibility that the discrepancy between an individual's genetically BMI-pred and observed BMI may be more closely associated with environmental factors, diet-related actions of the central nervous system, or peripheral energy expenditure than with the BMI itself. Therefore, our findings could provide a basis for further research on the underlying pathophysiology of T2D and its remission after tailored weight reduction.

The increase in T2D risk due to BMI-diff is not independent of the influence of various other confounding factors involved in T2D risk. Notably, our subgroup analysis showed that the impact of BMI-diff was more pronounced in relatively low-risk groups. In contrast, the increase in T2D risk due to BMI-diff was attenuated in current smokers and in individuals with

Table 3—Parameters of insulin secretion and resistance in KoGES Ansan/Ansung cohort

	1st quintile of BMI-diff (n = 1,486)	2nd quintile of BMI-diff (n = 1,486)	3rd quintile of BMI-diff (n = 1,486)	4th quintile of BMI-diff (n = 1,486)	5th quintile of BMI-diff (n = 1,486)	P
Matsuda index (composite insulin sensitivity index)	8.5 ± 5.9	9.8 ± 6.3	10.9 ± 6.7	11.9 ± 7.2	13.7 ± 7.8	<0.001
IGI_{60}	12.5 ± 17.4	12.9 ± 24.9	13.4 ± 23.7	13.4 ± 23.6	11.9 ± 23.8	0.45
Disposition index ($IGI_{60} \times$ Matsuda index)	85.8 ± 180.5	105.2 ± 247.3	123.9 ± 265.3	139.8 ± 285.0	137.4 ± 318.7	<0.001

Data are presented as the mean $\pm$ SD.

hypertension or dyslipidemia, particularly in the UK Biobank cohort. Therefore, deviation from the genetically BMI-pred may be more useful in assessing T2D risk primarily in younger, low-risk populations.

Although the results were generally consistent across Europeans and East Asians in this study, there were differences in the characteristics of two populations. Notably, in the East Asian population, a significant association between the discrepancy and T2D risk was solely observed among women. Given the predominant normal range of BMI among East Asian women (27), this result may offer an important clue in understanding the mechanisms driving T2D development within this subset. Meanwhile, the distribution of measured BMI by sex in our study differed between the European and East Asian populations (in UK Biobank: men 27.5 ± 3.8 kg/m² and women 26.6 ± 4.5 kg/m²; in KoGES: men 24.2 ± 2.9 kg/m² and women 24.7 ± 3.1 kg/m²). Therefore, the sex difference might be partly influenced by the fact that the model used to estimate genetically BMI-pred was not optimized for each sex separately. Because East Asians tend to develop T2D with lesser weight gain (28), even a modest weight gain could significantly affect T2D development. Given the relatively narrow range of BMI distribution in East Asians compared with Europeans, an approach with a personalized threshold of BMI rather than a universal cutoff value should be emphasized. Furthermore, our results should be validated in populations with diverse ethnic backgrounds, such as Hispanic, African, and South Asian groups, which have distinct genetic characteristics compared with the populations used in this study.

Clinical Implications

Based on our results, we could include deviation from genetically BMI-pred as a novel predictor to reclassify the future risk of T2D. Interventions targeting this deviation may potentially prevent the onset of T2D. Excessive weight loss attempts in individuals with obesity might not only have low success rates but could also reduce the likelihood of maintaining the weight loss (29). Moreover, the clinical significance of weight loss or lifestyle modifications in populations without obesity is often overlooked (5). Thus, maintaining lifestyle in line with one's personalized BMI threshold, whether in overweight or

normal-weight populations, could lead to favorable long-term metabolic outcomes. Although this concept requires clinical validation, it holds potential to inform clinical practice for weight loss and the prevention and treatment of T2D. On the other hand, the significance of BMI-diff and its pathophysiologic mechanisms in the development of T2D in individuals with normal or low genetically BMI-pred should also be investigated.

Study Limitations

Several limitations should be discussed. First, as an observational study, the results should not be interpreted as establishing a causal relationship. Although we provide consistent evidence in two independent data sets that higher than genetically BMI-pred is associated with increased T2D risk, it is possible that reverse causation or selection bias influence our results. Furthermore, the results should be interpreted with caution considering the possible effect of pleiotropy. Second, it is possible that additional anthropometric measurements, such as central obesity or muscle mass, could influence our results. Third, although we tried to adjust for numerous potential confounding factors, there may be unmeasured or residual confounding factors.

Conclusion

Our data provide evidence that the discrepancy between genetically BMI-pred and observed BMI is associated with increased T2D risk and early glycemic alterations. Our results offer valuable insights into genetic susceptibility to T2D and underscore the significance of genetically determined fat threshold in individuals' risk for metabolic disorders.

Acknowledgments. The authors express their sincere thanks to all participants in the UK Biobank and KoGES as well as staff of these studies. During the course of preparing this work, the authors used ChatGPT (OpenAI) for the purpose of grammar check. Following the use of this tool/service, the authors formally reviewed the content for its accuracy and edited it as necessary. The authors take full responsibility for all the content of this publication.

Funding. S.H.K. is supported by the National Research Foundation of Korea grant funded by the Korean Ministry of Science and ICT (RS-2023-00262002) and by a grant (23212MFDS202) from Korean Ministry of Food and Drug Safety in 2023. S.H.K. and J.C. are supported by National Human Genome Research Institute, grant FAIR# U01HG011723. J.M. is supported by a grant from the Novo Nordisk Foundation, which partially

supports the Novo Nordisk Foundation Center for Basic Metabolic Research. H.W. received a grant from the MD-PhD/Medical Scientist Training Program through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea. This study was conducted with bioresources from National Biobank of Korea, the Centers for Disease Control and Prevention, and Republic of Korea (NBK-2021-015).

Duality of Interest. S.H.K. serves as a chief medical officer at SNUH venture. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. T.-M.R., J.C., H.L., J.M., and S.H.K. contributed to analysis and interpretation of data. J.C., H.L., J.M., and J.-B.P. contributed to administrative, technical, or material supports. T.-M.R. and S.H.K. contributed to study conception and design, data acquisition, and writing the manuscript. S.H.K. contributed to study supervision. S.H.K. 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.

Prior Presentation. Parts of this study were presented as an abstract at the 83rd Scientific Sessions of the American Diabetes Association, San Diego, CA, 23–26 June 2023.

Handling Editors. The journal editors responsible for overseeing the review of the manuscript were Elizabeth Selvin and Jennifer E. Posey.

References

1. Khan MAB, Hashim MJ, King JK, Govender RD, Mustafa H, Al Kaabi J. Epidemiology of type 2 diabetes - global burden of disease and forecasted trends. *J Epidemiol Glob Health* 2020;10:107–111
2. Lee CMY, Woodward M, Pandeya N, et al.; Obesity, Diabetes and Cardiovascular Disease Collaboration. Comparison of relationships between four common anthropometric measures and incident diabetes. *Diabetes Res Clin Pract* 2017; 132:36–44
3. Lee DH, Keum N, Hu FB, et al. Comparison of the association of predicted fat mass, body mass index, and other obesity indicators with type 2 diabetes risk: two large prospective studies in men and women. *Eur J Epidemiol* 2018;33:1113–1123
4. Tsatsoulis A, Paschou SA. Metabolically healthy obesity: criteria, epidemiology, controversies, and consequences. *Curr Obes Rep* 2020;9:109–120
5. Taylor R, Holman RR. Normal weight individuals who develop type 2 diabetes: the personal fat threshold. *Clin Sci (Lond)* 2015;128: 405–410
6. Lean ME, Leslie WS, Barnes AC, et al. Primary care-led weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial. *Lancet* 2018;391:541–551
7. Khera AV, Chaffin M, Wade KH, et al. Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell* 2019;177:587–596.e9
8. Ottosson F, Smith E, Ericson U, et al. Metabolome-defined obesity and the risk of future type 2 diabetes and mortality. *Diabetes Care* 2022;45:1260–1267
9. 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:e1001779

10. Fairbrother U, Kidd E, Malagamuwa T, Walley A. Genetics of severe obesity. *Curr Diab Rep* 2018; 18:85
11. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–209
12. Huang J, Howie B, McCarthy S, et al.; UK10K Consortium. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nat Commun* 2015;6:8111
13. Walter K, Min JL, Huang J, et al.; UK10K Consortium. The UK10K project identifies rare variants in health and disease. *Nature* 2015;526: 82–90
14. McCarthy S, Das S, Kretschmar W, et al.; Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016;48:1279–1283
15. Kim Y, Han B-G, KoGES group. Cohort profile: the Korean Genome and Epidemiology Study (KoGES) Consortium. *Int J Epidemiol* 2017; 46:e20
16. Nam K, Kim J, Lee S. Genome-wide study on 72,298 individuals in Korean biobank data for 76 traits. *Cell Genom* 2022;2:100189
17. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun* 2019; 10:1776
18. Locke AE, Kahali B, Berndt SI, et al.; International Endogene Consortium. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;518:197–206
19. Sakaue S, Kanai M, Tanigawa Y, et al.; FinnGen. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet* 2021;53:1415–1424
20. Hahn S-J, Kim S, Choi YS, Lee J, Kang J. Prediction of type 2 diabetes using genome-wide polygenic risk score and metabolic profiles: a machine learning analysis of population-based 10-year prospective cohort study. *EBioMedicine* 2022;86:104383
21. Ohn JH, Kwak SH, Cho YM, et al. 10-year trajectory of β -cell function and insulin sensitivity in the development of type 2 diabetes: a community-based prospective cohort study. *Lancet Diabetes Endocrinol* 2016;4:27–34
22. UK Prospective Diabetes Study (UKPDS). VIII. Study design, progress and performance. *Diabetologia* 1991;34:877–890
23. Perry RJ, Samuel VT, Petersen KF, Shulman GI. The role of hepatic lipids in hepatic insulin resistance and type 2 diabetes. *Nature* 2014;510: 84–91
24. George AM, Jacob AG, Fogelfeld L. Lean diabetes mellitus: an emerging entity in the era of obesity. *World J Diabetes* 2015;6:613–620
25. Angelidi AM, Filippaios A, Mantzoros CS. Severe insulin resistance syndromes. *J Clin Invest* 2021;131:e142245
26. Eckel RH, Kahn SE, Ferrannini E, et al.; Endocrine Society, American Diabetes Association, European Association for the Study of Diabetes. Obesity and type 2 diabetes: what can be unified and what needs to be individualized? *Diabetes Care* 2011;34:1424–1430
27. Park B, Kim S, Kim H, Cha C, Chung MS. Associations between obesity, metabolic health, and the risk of breast cancer in East Asian women. *Br J Cancer* 2021;125:1718–1725
28. Ma RCW, Chan JCN. Type 2 diabetes in East Asians: similarities and differences with populations in Europe and the United States. *Ann N Y Acad Sci* 2013;1281:64–91
29. Leibel RL, Seeley RJ, Darsow T, Berg EG, Smith SR, Ratner R. Biologic responses to weight loss and weight regain: report from an American Diabetes Association Research Symposium. *Diabetes* 2015;64:2299–2309