



Published in final edited form as:

Biodemography Soc Biol. 2020 ; 66(1): 40–49. doi:10.1080/19485565.2020.1806032.

What Your Genes Can (and Can't) Tell You About BMI and Diabetes

Carmen D. Ng^{1, #a, *}, Jordan Weiss, PhD^{2, #b, *}

¹Hubert Department of Global Health and the Emory Global Diabetes Research Center, Emory University, Atlanta, GA, USA

²Population Studies Center and the Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, USA

Abstract

Body mass index (BMI) is commonly used as a proxy for adiposity in epidemiological and public health studies. However, BMI may suffer from issues of misreporting and, because it fluctuates over the life course, its association with morbidities such as diabetes is difficult to measure. We examined the relative associations between actual BMI, genetic propensity for high BMI, and diabetes to better understand whether a BMI polygenic score (PGS) explained more variation in diabetes than self-reported BMI. We used a sample of non-Hispanic white adults from the longitudinal Health and Retirement Study (1992–2016). Structural equation models were used to determine how much variation in BMI could be explained by a BMI PGS. Then, we used logistic regression models (n=12,086) to study prevalent diabetes at baseline and Cox regression models (n=11,129) to examine incident diabetes with up to 24 years of follow-up. We observed that while both actual BMI and the BMI PGS were significantly associated with diabetes, BMI had a stronger association than its genetic counterpart and resulted in better model performance. Moreover, actual BMI explained more variation in baseline and incident diabetes than its genetic counterpart which may suggest that actual BMI captures more than just adiposity as intended.

Keywords

body mass index; diabetes; genes; polygenic scores

Introduction

Obesity is a growing public health concern in the United States, both because of its high prevalence and its role as a risk factor in many leading causes of death. An estimated 40% of American adults were classified as obese (Body Mass Index [BMI] ≥ 30) in 2015–2016.¹ Elevated levels of BMI are associated with many health morbidities, including heart disease, stroke, and diabetes,² and as a result, studies investigating these health outcomes adjust for BMI in their analyses. However, BMI, which is typically derived from self-reported

*Correspondence to: ngcarmen68@gmail.com, jordanmnw@gmail.com.

#aCurrent affiliation: Genentech, South San Francisco, CA, USA

#bCurrent affiliation: Department of Demography, UC Berkeley, Berkeley, CA, USA

height and weight information in epidemiological studies, is often misreported.³⁻⁵ Such discrepancies could be intentional due to a desire to be taller/thinner, or could be the result of recall bias. The issue of recall could be especially problematic for people at older ages who suffer from memory loss and height shrinkage. Furthermore, BMI could fluctuate over the life course in response to changes in health status, environment, and life stage transitions; thus, its association with health outcomes can be challenging to measure.⁶ For example, a patient who is pre-diabetic may seek to lose weight to protect against diabetes, making it challenging to untangle how these dynamics unfold over the lifespan. A more stable measure of BMI could provide useful information in determining an individual's morbidity risk profile and obtain more precise estimates of the associations between BMI and a range of health outcomes.

Prior research has demonstrated that high BMI is a risk factor for multiple comorbidities,⁷⁻⁹ but less is known about the association between genetic BMI risk and chronic conditions, such as diabetes. Discerning whether elevated BMI is a precursor to diabetes, or whether diabetes and insulin-resistance affect BMI presents a challenge. The study of genetic risk has been made possible with increased collection of genetic material over the past decade and the advent of polygenic scores (PGSs) which combine the estimated effect sizes for millions of genetic variants, referred to as single nucleotide polymorphisms (SNPs), into a single measure of genetic risk for an outcome of interest. PGSs can capture genetic variance in complex traits such as BMI, and the aggregation of many different SNPs provide a better understanding of genetic risk for high BMI compared to any one specific genetic marker.¹⁰ Furthermore, this is a stable measure of BMI that may not fluctuate in the same way that actual BMI is known to.

The objective of this study was to contrast the association between diabetes and actual BMI against the association between diabetes and the polygenic risk for high BMI, measured in the form of a BMI PGS. We used data from the nationally representative and longitudinal Health and Retirement Study (HRS). We first examined the extent to which BMI can be attributed to genetic propensity for high BMI and how much is due to other sociodemographic, economic, and behavioral components. We then examined prevalent and incident diabetes among older adults and compared covariate-adjusted models using actual BMI with models using a BMI PGS with an aim to understand associational (i.e., magnitude of association) and explanatory differences (i.e., variation explained) between the two measures. As a measure of genetic propensity for high BMI, the BMI PGS is not subject to misreporting or fluctuations, so a potential hypothesis is that the BMI PGS may be a more precise predictor of health risk than actual BMI. However, because actual BMI may oscillate over the life course, a competing hypothesis is that the dynamic nature of BMI over several waves of data may carry more information and thus be a better predictor of diabetes.

Materials and Methods

Data

The HRS is a nationally representative survey of more than 30,000 US adults over the age of 50 and their spouses of any age. Since 1992, the HRS has biennially assessed the economic, physical, and mental health of its respondents to better understand the life circumstances of

older adults. The original sample has been refreshed with new birth cohorts every six years to maintain population representativeness. The HRS is sponsored by the National Institute on Aging (NIA U01AG009740) and is conducted by the University of Michigan. Detailed information on the HRS sample, design, and variables has been previously published.¹¹ We used the RAND HRS Longitudinal File 2016 (Version 2), a dataset based on the core HRS surveys.¹²

The HRS collected genetic data from a sub-sample of respondents who consented and provided salivary deoxyribonucleic acid (DNA) over the survey years 2006–2012. Information on sample selection, consent procedures, and assay processes have been described elsewhere.¹³ We linked the RAND HRS Longitudinal File with the third release of the HRS polygenic score genetic data containing the BMI PGS and restricted our analysis to non-Hispanic white respondents with available genetic information. After excluding four respondents who were missing their baseline measure of BMI, this resulted in an analytic sample of 12,086 respondents over the study period 1992 to 2016.

Measures

Outcome—We examined prevalent diabetes at baseline which was assessed at each respondent’s first HRS interview. We also examined incident diabetes defined as the age at which respondents first reported diabetes. Both prevalent and incident diabetes were determined by a respondent’s affirmative self-report to the question: “Since we last talked to you, that is since [last interview date], has a doctor ever told you that you have diabetes or high blood sugar?” Respondents who did not report diabetes or who died over the study period were censored.

Exposure—Our main covariate of interest was BMI, which was conceptualized in two different ways. The first was a continuous measure of BMI defined by a person’s mass in kilograms divided by their squared height in meters. Extreme values (BMI < 10; BMI > 75) were recoded as missing. We then standardized actual BMI to a normal distribution with a mean of zero and standard deviation of one. This was done to be comparable with our second conceptualization of BMI – a BMI PGS. Specifically, we used the third release of the HRS BMI PGS.¹³ The third release of the BMI PGS was constructed by HRS investigators based on the results from a joint genome-wide association study and Metachip meta-analysis comprised of 332,154 individuals across more than 2.5 million SNPs. From this analysis, 97 SNPs were identified as genome-wide significant, accounting for nearly 2.7% of BMI variation and suggesting that as much as 21% of BMI variation may be attributable to genetic variation.^{13,14} Among HRS respondents with European ancestry, 761,985 SNPs overlapped between the HRS genetic database and the GWAS meta-analysis. The HRS-provided PGSs were standardized within ethnicity to a normal distribution with a mean of zero and standard deviation of one. We residualized the BMI PGS by regressing it on the first ten genetic ancestry-specific principal components to account for population stratification, and then standardizing the residual values to a normal distribution with a mean of zero and standard deviation of one.

Covariates—Models were adjusted for covariates based on their documented associations with BMI and/or diabetes. Sociodemographic covariates included sex (male, female), whether the respondent was foreign born (yes, no), level of education (less than high school or GED, high school or some college, college and above), and partnership status (married/partnered, not married/partnered). Measures of economic well-being included labor force participation (employed, unemployed, retired, disabled, other), income (log-transformed), wealth (log-transformed), and whether the respondent had Medicare (yes, no), Medicaid (yes, no) or another form of health insurance (yes, no). We assessed behavioral and lifestyle characteristics by including respondent's self-report of physical activity (completed vigorous activity at least once per week), smoking status (never smoked, former smoker, active smoker), and alcohol consumption. US guidelines for alcohol consumption were used to classify respondents as heavy drinkers (more than eight drinks per week for women; 15+ drinks per week for men).¹⁵ We also included self-reported binary indicators of whether the respondent had been informed by a medical practitioner that they had high blood pressure, cardiovascular disease, stroke, lung problems, or arthritis in each survey wave. Medical comorbidities could impact a respondent's lifestyle and influence levels of actual BMI. Moreover, they are important health co-morbidities for diabetes. Additionally, we accounted for birth cohort to control for respondents entering the HRS at different survey waves. These estimates are not displayed for brevity but are available upon request from the authors. With the exception of BMI, missing covariate values at baseline were imputed using a form of multiple imputation by chained equations.¹⁶ Missing values over follow-up were imputed using a respondent's last known measure.

Statistical analysis

Structural equation modeling (SEM) was used to disentangle the different components of actual BMI.¹⁷ Covariates were grouped into latent variables for sociodemographic characteristics, economic well-being, behavioral and lifestyle characteristics, medical comorbidities, and genetic propensity for high BMI to understand how each of these domains were associated with BMI. For the purpose of interpretation, standardized estimates are presented to facilitate comparison of the effect sizes of latent variables across domains.

We then used multivariate logistic and Cox regression models to estimate the associations of BMI and the BMI PGS with diabetes. The logistic models were used to study prevalent diabetes by pooling data for all respondents' first survey wave (i.e., wave in which respondents entered the HRS). The Cox regression models were used to study incident diabetes; thus, we restricted the sample to respondents who did not have prevalent diabetes (n=11,129). Additionally, the Cox models included time-varying covariates for partnership status, economic well-being, behavioral and lifestyle characteristics, and medical comorbidities. For both the prevalent and incident analyses, we first fit separate models for actual BMI and the BMI PGS, along with the covariates. Then, we fit saturated models that included actual BMI and the BMI PGS concurrently, along with the covariates. McFadden's Adjusted R^2 was used to compare model performance for the prevalence analysis; concordance values were used to determine model performance relative to the saturated model obtained from Cox regression. We included cluster-robust standard errors to

account for household stratification in the HRS and to address potential within-household spillover effects.¹⁸

All statistical analyses were performed in R version 3.5.0 with the ‘lavaan’ package for SEM and the ‘survival’ package for Cox regression models.^{19–21} In all cases, significance was reported at the five-percent level.

Results

Age trends in BMI for the analytic sample are shown in Figure 1. The trends are stratified by BMI PGS tertiles. Age-specific BMI provides indication as to how dynamic of a measure BMI truly is, and the stratification allows for a glimpse into how BMI varies by genetic propensity for high BMI.

Figure 1 depicts a clear relationship between actual BMI and the BMI PGS over much of the lifespan. From about age 35, it appears that BMI trends formed a concave pattern; BMI increased before decreasing at older ages. The mean age-specific BMI levels were highest for those with the highest genetic propensity for high BMI, as expected, although there was a crossover at the oldest ages. It should be noted though that at the beginning of this age interval, the mean BMI for all three BMI PGS categories was above 25, which would be considered overweight. The correlation between actual BMI and the BMI PGS was 0.2531; that is, the two measures were positively correlated, though the relationship was relatively weak.

We utilized SEM to disentangle the genetic propensity for high BMI from the sociodemographic, economic, behavioral and lifestyle, and medical comorbidities domains and found that all of these were embedded within BMI. Table 1 displays the results of SEM for the sample. Despite all of the latent variables being significant, the absolute magnitude of the genetic propensity for high BMI was the smallest. This was followed by the sociodemographic, medical comorbidities, behavioral and lifestyle, and economic domains.

Table 2 presents the results from three logistic regression models of diabetes at baseline. We fit a saturated model to examine associations between prevalent diabetes, actual BMI, and the BMI PGS after covariate-adjustment. In addition, we fit covariate-adjusted models for prevalent diabetes as a function of actual BMI at baseline and separately for the BMI PGS. Because of the relatively weak correlation between BMI and the BMI PGS, multicollinearity was not an issue. Table 3 presents the results from the three corresponding Cox regression models of diabetes incidence. Time-varying standardized BMI and the BMI PGS were included as covariates in separate regressions, and then used together in the saturated model.

In both Tables 2 and 3, although both standardized BMI and the BMI PGS had significant associations with diabetes, the models with actual BMI performed better than those with the BMI PGS in regards to both R^2 and concordance values. Furthermore, the coefficients for standardized BMI were higher than those for BMI PGS. The odds ratio for baseline standardized BMI were 1.3 times as large as the odds ratio for the BMI PGS, and the hazard ratio for time-varying standardized BMI was 1.5 times as large as the hazard ratio

for the BMI PGS. The saturated models improved R^2 and concordance values, but the increases were minimal. Both actual BMI and the BMI PGS were statistically significant in the saturated logistic regression model for prevalent diabetes, but only actual BMI (and not the BMI PGS) was statistically significant in the saturated Cox regression model for incident diabetes. The odds ratio/hazard ratio for actual BMI remained similar in the saturated model, i.e., the addition of the BMI PGS did little to change the association between actual BMI and prevalent/incident diabetes.

Discussion

BMI is commonly used as a proxy for adiposity in epidemiological studies of health. However, it is subject to reporting biases and fluctuations over time. Thus, a more stable measure, such as genetic propensity for high BMI, might be more reliable. In this study, we observed a positive association with actual BMI and the BMI PGS although the association was not particularly strong. Our SEM analysis revealed that genetic risk had the lowest standardized magnitude when compared to other latent variables associated with actual BMI, including sociodemographic, economic, behavioral and lifestyle, and medical characteristics.

In analyses with prevalent and incident diabetes as outcomes, models that used standardized BMI as a covariate performed better than those that used the BMI PGS, as determined by R^2 and concordance values. Additionally, the magnitudes of the standardized BMI odds and hazard ratios were larger than those of the BMI PGS. That is, a one-unit increase in standard deviation of BMI had a stronger association with diabetes than a one-unit increase in the PGS. In saturated models that included both actual BMI and the BMI PGS, the coefficients for actual BMI remained largely unchanged, as did our measures of model performance, revealing that the BMI PGS had little additional explanatory power. In fact, the BMI PGS was only statistically significant in the logistic regression model of prevalent diabetes at survey entry. This may indicate that actual BMI explains more variation in prevalent and incident diabetes than the BMI PGS. If the BMI PGS had resulted in superior model performance, it could be argued that genetic testing is necessary for truly understanding the relationship between body mass and diabetes (and by extension, other chronic diseases). However, given our results and the challenges of genetic testing, collection of genetic data may not be of highest priority in a clinical setting to address questions such as the one posed in this study. This observation has been noted in prior literature suggesting that, although genetic information enhances our understanding of epidemiological relationships, it does not downplay the significant impact of social and environmental characteristics.^{22–24} However, the utility of polygenic scores depends on the research question and hypothesis under study.

The results from our SEM analysis suggest that the stronger association between actual BMI and diabetes may be driven by latent relationships between BMI and socioenvironmental characteristics which, in turn, may be associated with prevalent and incident diabetes. The associated magnitude of the BMI PGS was small relative to the magnitudes of the other domains, most notably characteristics representing economic well-being, which may be indicative of BMI capturing more than just adiposity. However, this begs the question as to whether BMI over-estimates the relationship between adiposity and diabetes. BMI is the most commonly used proxy of adiposity, but if BMI is capturing more than just adiposity,

as we demonstrated with SEM, its role in morbidities could be over-stated. It should be noted that the BMI PGS is not a perfect stand-in for adiposity, either. Although the BMI PGS is stable over time and not subject to misreporting, it is a measure of risk—not actual adiposity. Further, genetic correlations across related and unrelated phenotypes could exist (i.e., pleiotropy), in which case the BMI PGS may be indicative of phenotypes other than BMI and/or adiposity.

Our study is not without limitations. The analytic sample was comprised of non-Hispanic white respondents who consented and provided DNA samples for genotyping; thus, this is the population to which our results are generalizable. Those who selected into the genetic sample could be different than the full HRS sample. Mortality selection is also a concern, as adults who survived to provide genetic information may be advantaged or more robust relative to decedents from their birth cohort.²⁵ Although this is an issue with all studies using the HRS, respondents in studies also using the polygenic data had to survive to 2006–2012 in order to be included for potential genetic sampling. This mortality selection could produce biased estimates, especially when examining multiple birth cohorts.²⁶ Domingue and colleagues²⁶ reported that the BMI PGS of older cohorts (i.e., respondents from earlier birth years) were lower than those of younger cohorts (i.e., respondents from later birth years). We attempted to address this concern by including all available survey waves and accounting for birth cohort.

Despite these limitations, our analysis revealed that the strength of the relationship between actual BMI and diabetes was stronger than the relationship between diabetes and genetic propensity for high BMI among non-Hispanic whites. Actual BMI could be a more “loaded” variable that informs us of more than just adiposity. For understanding the relationship between adiposity and health among non-Hispanic whites, it could be that BMI overstates the association. But actual BMI is a more important covariate (in terms of magnitude of association and variation explained) and would be more useful than its genetic counterpart in prediction of diabetes.

References

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity among adults and youth: United States, 2015–2016. US Department of Health and Human Services, Centers for Disease Control and Prevention.; 2017.
2. Apovian CM. Obesity: definition, comorbidities, causes, and burden. *Am J Manag Care*. 2016;22(7 Suppl):s176–185. [PubMed: 27356115]
3. Jalkanen L, Tuomilehto J, Tanskanen A, Puska P. Accuracy of Self-Reported Body-Weight Compared to Measured Body-Weight - a Population Survey. *Scand J Soc Med*. 1987;15(3):191–198. [PubMed: 3616534]
4. Krul AJ, Daanen HAM, Choi H. Self-reported and measured weight, height and body mass index (BMI) in Italy, the Netherlands and North America. *European Journal of Public Health*. 2011;21(4):414–419. [PubMed: 20089678]
5. Kuczmarski MF, Kuczmarski RJ, Najjar M. Effects of age on validity of self-reported height, weight, and body mass index: Findings from the third National Health and Nutrition Examination Survey, 1988–1994. *Journal of the American Dietetic Association*. 2001;101(1):28–34. [PubMed: 11209581]

6. Flegal KM, Graubard BI, Williamson DF, Cooper RS. Reverse causation and illness-related weight loss in observational studies of body weight and mortality. *American journal of epidemiology*. 2010;173(1):1–9. [PubMed: 21059807]
7. Must A, Spadano J, Coakley EH, Field AE, Colditz G, Dietz WH. The disease burden associated with overweight and obesity. *Jama-J Am Med Assoc*. 1999;282(16):1523–1529.
8. Pi-Sunyer FX. The obesity epidemic: Pathophysiology and consequences of obesity. *Obesity Research*. 2002;10:97s–104s. [PubMed: 12490658]
9. Wyatt SB, Winters KP, Dubbert PM. Overweight and obesity: Prevalence, consequences, and causes of a growing public health problem. *American Journal of the Medical Sciences*. 2006;331(4):166–174. [PubMed: 16617231]
10. Dudbridge F. Polygenic epidemiology. *Genetic epidemiology*. 2016;40(4):268–272. [PubMed: 27061411]
11. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort profile: The health and retirement study (HRS). *International journal of epidemiology*. 2014;43(2):576–585. [PubMed: 24671021]
12. RAND. RAND HRS Longitudinal File 2016 (V2) public use dataset. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. 2020.
13. Ware E, Schmitz L, Gard A, Faul J. HRS Polygenic Scores—Release 3: 2006–2012 Genetic Data. Ann Arbor: Survey Research Center, University of Michigan. 2018.
14. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518(7538):197–206. [PubMed: 25673413]
15. U.S. Department of Health and Human Services and U.S. Department of Agriculture. 2015 – 2020 Dietary Guidelines for Americans. 2015.
16. Sv Buuren, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *Journal of statistical software*. 2010:1–68.
17. Singh K. Latent variable models: An introduction to factor, path, and structural equation analysis, 4th edition. *Appl Psych Meas*. 2006;30(3):256–258.
18. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B (Methodological)*. 1972;34(2):187–220.
19. R: A Language and Environment for Statistical Computing [computer program]. 2018.
20. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*. 2012;48(2):1–36.
21. Therneau TM, Lumley T. Package ‘survival’. *R Top Doc*. 2015;128.
22. Ng CD, Weiss J. Association of genetic and behavioral characteristics with the onset of diabetes. *BMC Public Health*. 2019;19(1):1297. [PubMed: 31615468]
23. Meigs JB, Shrader P, Sullivan LM, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med*. 2008;359(21):2208–2219. [PubMed: 19020323]
24. Lyssenko V, Jonsson A, Almgren P, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med*. 2008;359(21):2220–2232. [PubMed: 19020324]
25. Zajacova A, Sarah AB. Healthier, Wealthier, and Wiser: A Demonstration of Compositional Changes in Aging Cohorts Due to Selective Mortality. *Population Research and Policy Review*. 2013;32(3).
26. Domingue BW, Belsky DW, Harrati A, Conley D, Weir DR, Boardman JD. Mortality selection in a genetic sample and implications for association studies. *Int J Epidemiol*. 2017;46(4):1285–1294. [PubMed: 28402496]

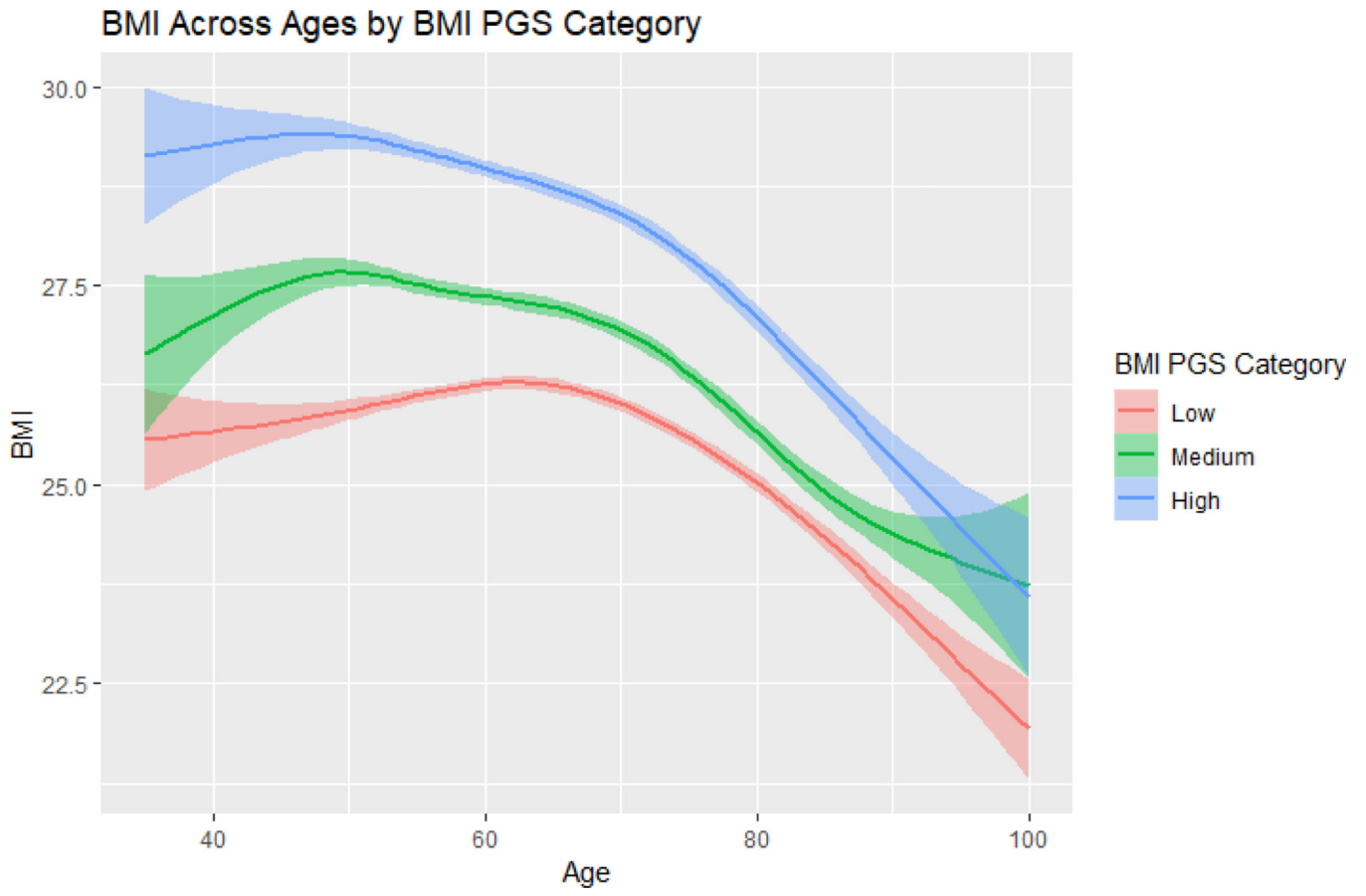


Figure 1. Mean age-specific BMI for the analytic sample by BMI PGS tertile
Notes. Abbreviations: BMI, body mass index; PGS, polygenic score

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Estimates of the relationship between BMI z-score and its latent variables (n=12,086)

Domain	Coefficient estimate (standard error)
Sociodemographic	-1.576 (0.047) ***
Economic	16.330 (1.044) ***
Behavioral and lifestyle	3.772 (0.229) ***
Medical comorbidities	2.135 (0.052) ***
BMI PGS	0.046 (0.007) ***

*Notes.** $p < 0.05$;** $p < 0.01$;*** $p < 0.001$.

Table 2.

Odds ratios and 95% CIs from multivariate logistic regression models of prevalent diabetes at survey entry (n=12,086)

Characteristic	Model with Actual BMI	Model with BMI PGS	Saturated Model
<i>Exposure</i>			
Actual BMI	1.60 ^{***} (1.50, 1.71)	—	1.57 ^{***} (1.46, 1.68)
BMI PGS	—	1.23 ^{***} (1.15, 1.32)	1.09 [*] (1.02, 1.18)
<i>Sociodemographic</i>			
Age	1.02 [*] (1.00, 1.03)	1.01 (0.99, 1.02)	1.02 [*] (1.00, 1.03)
Female	0.76 ^{***} (0.65, 0.89)	0.76 ^{***} (0.65, 0.89)	0.76 ^{***} (0.65, 0.89)
Foreign-born	1.26 (0.86, 1.84)	1.15 (0.78, 1.68)	1.26 (0.86, 1.84)
<i>Education</i>			
Less than high school/GED	1.23 [*] (1.02, 1.47)	1.23 [*] (1.02, 1.47)	1.22 [*] (1.02, 1.47)
High school/some college	1.00	1.00	1.00
College or above	1.05 (0.87, 1.26)	0.99 (0.83, 1.19)	1.05 (0.87, 1.26)
Married/partnered	1.11 (0.90, 1.37)	1.18 (0.96, 1.44)	1.11 (0.91, 1.37)
<i>Economic</i>			
<i>Employment status</i>			
Employed	1.00	1.00	1.00
Unemployed	1.11 (0.76, 1.62)	1.10 (0.76, 1.59)	1.10 (0.75, 1.62)
Retired	1.14 (0.92, 1.40)	1.15 (0.94, 1.42)	1.13 (0.92, 1.40)
Disabled	1.08 (0.69, 1.69)	1.12 (0.73, 1.73)	1.08 (0.69, 1.69)
Not in labor force	1.30 (0.99, 1.70)	1.26 (0.97, 1.64)	1.30 (1.00, 1.70)
Income	0.93 (0.84, 1.02)	0.92 (0.83, 1.01)	0.93 (0.84, 1.02)
Wealth	0.94 ^{***} (0.91, 0.98)	0.92 ^{***} (0.89, 0.95)	0.94 ^{***} (0.91, 0.98)
Medicare	1.42 [*] (1.07, 1.90)	1.44 [*] (1.09, 1.92)	1.43 [*] (1.07, 1.09)
Medicaid	0.99 (0.66, 1.48)	0.90 (0.60, 1.35)	0.97 (0.65, 1.47)
Health insurance	0.96 (0.79, 1.15)	0.96 (0.80, 1.16)	0.95 (0.79, 1.15)
Characteristic	Model with Actual BMI	Model with BMI PGS	Saturated Model
<i>Behavioral and lifestyle</i>			
Physically active	0.89 (0.76, 1.04)	0.78 ^{**} (0.67, 0.91)	0.89 (0.76, 1.04)
<i>Smoking status</i>			
Never smoked	1.00	1.00	1.00
Former smoker	1.06 (0.90, 1.24)	1.06 (0.91, 1.25)	1.05 (0.90, 1.24)
Active smoker	0.92 (0.74, 1.15)	0.74 ^{**} (0.59, 0.92)	0.90 (0.73, 1.13)
Heavy drinker	0.46 [*] (0.21, 0.99)	0.44 [*] (0.20, 0.96)	0.46 [*] (0.21, 0.99)
<i>Medical comorbidities</i>			
High blood pressure	2.40 ^{***} (2.08, 2.78)	2.86 ^{***} (2.48, 3.30)	2.40 ^{***} (2.07, 2.77)

Cardiovascular disease	1.64 ^{***} (1.37, 1.95)	1.64 ^{***} (1.38, 1.96)	1.63 ^{***} (1.37, 1.95)
Stroke	1.14 (0.86, 1.53)	1.15 (0.86, 1.53)	1.15 (0.87, 1.54)
Lung diseases	1.18 (0.91, 1.53)	1.21 (0.94, 1.56)	1.18 (0.91, 1.53)
Arthritis	1.09 (0.94, 1.26)	1.21 ^{**} (1.05, 1.40)	1.09 (0.94, 1.26)
McFadden's Adjusted R ²	0.1198	0.0958	0.1207

Notes. All models controlled for birth cohort.

Abbreviations: BMI, body mass index; GED, General Equivalency Diploma; PGS, polygenic score.

*
 $p < 0.05$;

**
 $p < 0.01$;

 $p < 0.001$.

Table 3.

Hazard ratios and 95% CIs from multivariate Cox regression models of incident diabetes for those who did not have diabetes at survey entry (n=11,129)

Characteristic	Model with Actual BMI	Model with BMI PGS	Saturated Model
<i>Exposure</i>			
Actual BMI	1.77 *** (1.66, 1.90)	—	1.76 *** (1.64, 1.88)
BMI PGS	—	1.16 *** (1.10, 1.23)	1.03 (0.97, 1.10)
<i>Sociodemographic</i>			
Female	0.88 * (0.77, 0.99)	0.85 ** (0.75, 0.96)	0.88 * (0.77, 0.99)
Foreign-born	1.01 (0.75, 1.37)	0.95 (0.71, 1.28)	1.02 (0.76, 1.38)
<i>Education</i>			
Less than high school/GED	1.15 (0.99, 1.33)	1.18 * (1.02, 1.36)	1.15 (0.99, 1.33)
High school/some college	1.00	1.00	1.00
College or above	0.98 (0.84, 1.15)	0.92 (0.79, 1.07)	0.99 (0.85, 1.15)
Married/partnered	1.29 * (1.04, 1.60)	1.29 * (1.05, 1.59)	1.29 * (1.04, 1.60)
<i>Economic</i>			
<i>Employment status</i>			
Employed	1.00	1.00	1.00
Unemployed	1.30 (0.91, 1.86)	1.41 * (1.00, 1.99)	1.30 (0.91, 1.86)
Retired	0.85 (0.72, 1.01)	0.91 (0.78, 1.06)	0.85 (0.72, 1.00)
Disabled	1.26 (0.80, 1.98)	1.26 (0.79, 2.02)	1.26 (0.80, 1.98)
Not in labor force	0.88 (0.70, 1.10)	0.89 (0.72, 1.11)	0.88 (0.70, 1.10)
Income	0.94 (0.86, 1.04)	0.95 (0.87, 1.04)	0.94 (0.86, 1.04)
Wealth	0.92 ** (0.87, 0.97)	0.89 *** (0.84, 0.94)	0.92 ** (0.87, 0.97)
Medicare	0.78 (0.59, 1.03)	0.72 * (0.54, 0.96)	0.78 (0.59, 1.03)
Medicaid	0.82 (0.46, 1.46)	0.69 (0.38, 1.25)	0.82 (0.46, 1.46)
Health insurance	0.96 (0.81, 1.12)	0.97 (0.83, 1.14)	0.96 (0.82, 1.12)
Characteristic	Model with Actual BMI	Model with BMI PGS	Saturated Model
<i>Behavioral and lifestyle</i>			
Physically active	0.91 (0.78, 1.06)	0.81 ** (0.70, 0.94)	0.91 (0.78, 1.06)
<i>Smoking status</i>			
Never smoked	1.00	1.00	1.00
Former smoker	1.00 (0.88, 1.14)	1.00 (0.88, 1.14)	1.00 (0.88, 1.15)
Active smoker	1.22 * (1.04, 1.44)	1.02 (0.87, 1.20)	1.21 * (1.03, 1.43)
Heavy drinker	0.64 ** (0.48, 0.86)	0.63 ** (0.47, 0.85)	0.65 ** (0.48, 0.87)
<i>Medical comorbidities</i>			
High blood pressure	1.46 *** (1.30, 1.65)	1.72 *** (1.53, 1.93)	1.46 *** (1.29, 1.65)

Cardiovascular disease	1.45 ^{***} (1.23, 1.72)	1.4 ^{***} (1.19, 1.65)	1.45 ^{***} (1.23, 1.72)
Stroke	1.00 (0.65, 1.53)	1.05 (0.69, 1.58)	1.00 (0.65, 1.52)
Lung diseases	0.98 (0.76, 1.25)	0.98 (0.78, 1.24)	0.98 (0.77, 1.25)
Arthritis	1.04 (0.93, 1.18)	1.14 [*] (1.02, 1.28)	1.04 (0.93, 1.18)
Concordance	0.7017	0.6402	0.7019

Notes. All models controlled for birth cohort.

Abbreviations: BMI, body mass index; GED, General Equivalency Diploma; PGS, polygenic score.

^{*}
 $p < 0.05$;

^{**}
 $p < 0.01$;

^{***}
 $p < 0.001$.