The causal effect of poor oral health on heart disease: a GWAS-by-subtraction genomic structural equation instrumental variables model

by

Chad Meyerhoefer, Lehigh University

Haiming Li, Lehigh University

Erin Ware, University of Michigan

John Moeller, University of Maryland School of Dentistry

Haiyan Chen, University of Maryland School of Dentistry

Jessica Faul, University of Michigan

Richard Manski, University of Maryland School of Dentistry

This version: December 18, 2025

Abstract:

We estimate the causal effect of edentulism, a common measure of poor oral health, on heart disease using genetic information in the 2006-2020 Health and Retirement Study. We use a GWAS-by-subtraction structural equation model to derive a polygenic score for poor oral health that is appropriate for instrumental analysis. This method improves upon earlier applications of Mendelian randomization by removing genetic variants from the instrument that have a direct effect on heart disease. Instrumental variables models indicate that edentulism increased the likelihood of a past heart disease diagnosis by 48 percentage points and the likelihood of a heart attack in the past two years by 3.9 – 6.9 percentage points. These treatment effects are significantly larger than OLS associations and represent up to a fivefold increase in average risk. The instruments exceed conventional thresholds for statistical power and pass falsification tests. Subpopulation estimated indicate that the risk of heart disease from poor oral health is significantly greater for men than women. Our analysis suggests that prior cost estimates of a proposed Medicare Part B dental benefit may be too high because they fail to account for cost offsets from lower heart disease rates.

Acknowledgement/Disclaimer: Research reported in this publication was supported by the National Institute of Dental & Craniofacial Research of the National Institutes of Health under Award Number R03DE034043. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Introduction

Following the nearly complete mapping of the human genome in 2003, statistical agencies have collected biological samples from participants in numerous large surveys.[1] Sequencing these samples permitted researchers to incorporate genetic information into studies of a broad range of biological, social and economics attributes of survey participants. Applications of genetic data to health economic studies are numerous and growing rapidly with the need to either control for genetic endowments or use such information to identify causal pathways through instrumental variables (IV) analysis. The latter is an extension of ongoing work involving the use of genetic IVs in the epidemiology literature, referred to as Mendelian randomization. In this paper, we review challenges associated with the use of genetic instruments. We then apply a recently developed method of mapping genes to traits (phenotypes) with the potential to improve the credibility of genetic IV studies to the estimation of the causal effect of poor oral health on heart disease.

Ideal instruments generate variation in the treatment that is as good as random in that it is not subject to feedback from the outcome variable and is orthogonal to unobservable factors in the error term (i.e., the exclusion restriction). The instrument(s) must also produce enough variation in the treatment to allow the detection of "small" treatment effects that can be generalized to the full study population. It is well known that instruments meeting these criteria are exceptionally difficult to construct, and most IV studies use instruments that balance statistical power against minor violations of the exclusion restriction. Genetic instruments have much promise because they are assigned at conception and therefore are not subject to feedback from the outcome variable. The extant literature focuses on the development of genetic instruments that have sufficient statistical power and are unlikely to be correlated with the error term, with varying levels of success.

Early studies used a small number of genes with known biological pathways as instruments. For example, Norton and Han (2008) used two to three genes that determine how the brain responds to the neurotransmitter dopamine as instruments to identify the effect of obesity on labor market outcomes. Similarly, Ding at el. (2009) used four genes that act on neurotransmitters or break down toxins in the liver as instruments for ADHD, depression and obesity in models of educational attainment. Fletcher and Lehrer (2009; 2011) also used a limited set of genes that affect neural response to the neurotransmitters

[1] Although the Human Genome Project provided an initial mapping of the human genome in 2003, it was not until 2022 that the last difficult-to-map 8% of the genome was sequenced by the Telomere-to-Telomere Consortium (Nurk et al. 2022).

dopamine and serotonin to investigate the effects of poor adolescent mental health and obesity on educational attainment, but incorporated family fixed effects on their models to overcome several threats to identification caused by dynastic effects and gene-environment interactions.

Dynastic effects refer to the process in which a parent's genes affect their behavior and the child's environment, which in turn influence the child's outcome. In a model where the child's genes are used instruments this could lead to a violation of the exclusion restriction due to a correlation between the instrument and observed environmental characteristics through the correlation with a parent's genes. Because siblings have the same parents, the inclusion of a family fixed effects removes the undesirable correlation. Likewise, family fixed effects may capture aspects of the environment shared across siblings that influence how the treatment is shaped by a gene (for example, how parenting style influences the relationship between the DRD4 gene's regulation of the dopamine D4 receptor and child mental health) (Koellinger and de Vlaming 2019). Nonetheless, Cawley, Han and Norton (2011) argued that genes influencing neurotransmitters could affect a wide range of behaviors associated with educational attainment, making it unlikely that they satisfy the exclusion restriction. Central to this critique is horizontal pleiotropy, the process whereby a gene may influence multiple traits through distinct biological pathways. For example, DRD4 affects attention and inhibition, which are part of the diagnostic criteria for ADHD, as well as curiosity and exploration (i.e. novelty seeking). While the former is negatively associated with academic performance, the latter is often rewarded in academic settings. This is why Menta et al. (2023) implemented a sensitivity test where they excluded genetic variants from their instrument that determine both the endogenous variable and factors associated with the outcome.

Aside from issues associated with dynastic effects, gene-environment interactions and pleiotropy, subsequent studies discuss other potential problems with the validity of genetic instruments.[2] For example, assortative mating in which individuals couple based on certain traits, such as height, education or alcohol use, creates a concentration of allele frequencies (the proportion of gene copies for a given variant) that determine certain traits within families across generations. Assortative mating can result in systematic associations between genes and the environment, parental behaviors or economic opportunities that lead to violations of the exclusion restriction (Hartwig, Davies and Smith 2018). On a more macro level, allele frequencies may differ systematically across racial and ethnic ancestral groups, which in turn have differential exposure to the treatment or similar values of the outcome. Such population stratification may cause spurious

[2] Pleiotropy is also commonly referred to as linkage disequilibria.

correlations in IV models (von Hinke 2016). A final issue is canalization, or biological processes that hide the effects of genes by compensating for their effects in development. In some cases, canalization may only reduce the statistical power of the instrument in the first stage, but if it creates a correlation between the gene and the outcome, IV estimates are biased (von Hinke Kessler Scholder et al. 2011).

Authors have used different strategies to address these threats to instrument validity. For example, in a genetic IV study of the effect of high body mass index on earnings and employment, Böckerman et al. (2019) included controls for the mother's and father's education to mitigate bias from assortative mating. Family fixed effects can also reduce some of the correlates between genetic composition and environmental attributes (Wang 2025). A common method to reduce bias from population stratification is to limit the analysis to a racially or ethnically homogenous sample, although other approaches involve the addition of control variables for stratification, such as principal components from genome-wide data on genetic variants (Cardon and Palmer 2003; Price et al. 2006).

Aside from instrument validity, another challenge in this literature is the statistical power of the IVs. Individual genetic variants are often weak instruments, which may lead to fragile estimates given that strong instruments are more robust to violations of the exclusion restriction (Small and Rosenbaum 2008). A common method of improving statistical power is the use of polygenic scores (PGS) as instruments rather than individual genetic variants. Researchers derive the components a PGS from the results of genome-wide association studies (GWAS), which are specific to a given trait or disease. GWAS are based on large samples of genotyped data processed using either single-nucleotide polymorphism (SNP) arrays or whole-genome sequencing. SNPs are the standard measure of genetic variability in both the epidemiological and health economics literatures on Mendelian randomization.[3] They indicate a position on the genome where individuals differ due to variation in the nucleotide. When conducting GWAS researchers first perform quality control and other adjustments on the sample by removing highly uncommon SNPs, imputing genotypes to a reference panel and correcting for population stratification. They then estimate millions of statistical correlations between the trait in question and each SNP, while controlling for basic individual information such as age, sex and principal components that account for population stratification.

[3] A nucleotide is a basic building block of DNA or RNA and a polymorphism is a genomic position where alleles (commonly) occur in a population. An allele is a particular version of a genetic sequence at a location (or locus) on the chromosome. Individuals inherit one allele from their biological mother and the other from their biological father.

A PGS is calculated as a weighted average of SNPs, where the weights are GWAS coefficients measuring the effect size of the SNP on the trait of interest.[4] Because millions of correlations are estimated to obtain GWAS coefficients, false positive rates are high at conventional levels of significance. Therefore, researchers often compare PGS based on a larger number of SNPs having lower levels of statistical significance to those computed using fewer SNPs having a significance threshold as low as $p < 5 \times 10^{-8}$.[5]

Most recent analyses use PGS as instruments or explanatory variables, or less sophisticated genetic risk scores constructed as unweighted sums of limited numbers of SNPs (Edwards, Bjørngaard and Mint Kinge 2021; Weng 2025; Fletcher 2023; Pehkonen et al. 2021; Menta et al. 2023; Brunello, Sanz-de-Galdeano and Terskaya 2020 Böckerman et al. 2018). Not only does the use of a PGS increase the statistical power of the instrument, but it reduces the risk of bias from horizontal pleiotropy (Palmer et al. 2012). As explained by Davey Smith (2011), if multiple SNPs serving as instruments (as those in a PGS) yield the same causal effect of the endogenous variable on the outcome by working through different pathways, it is less likely that a confounding relationship will persist jointly across independently operating SNPs. Nonetheless, pleiotropy can still significantly impact the magnitude of treatment effects and the bias from SNPs that have a direct effect on the outcome is difficult to overcome.

Genetic determinants of oral health and construction of polygenic scores using GWAS-by-substation genomic structural equation models

We use a method recently developed by Demange et al. (2021) that builds off prior work in genomic structural equation modelling (SEM) by Grotzinger et al. (2019) to derive PGS more suitable for IV analysis. The method involves the application of SEM to summary statistics from two GWAS to derive PGS weights that are correlated with the endogenous variable but uncorrelated with the outcome. Our objective is to estimate the causal effect of poor oral health on heart disease. We are motivated by a desire to determine whether improved oral health, through for example better access to dental insurance, may lower the risk of heart disease in a population of elderly Americans. Evidence for a causal effect is fundamental to future research on the costs and benefits of adding a dental benefit to Medicare Part B (Gangopadhyaya, Garret and Holahan 2023; Moeller et al. 2020).

Prior research documents a correlation between periodontitis and coronary heart disease that is thought to occur through chronic inflammation (Spahr et al. 2006;

[4] Prior to computing the PGS, some coefficients are excluded because they correspond to SNPs that are correlated with causal variants only because they are clustered at the same location on the genome due to linkage disequilibrium.

[5] This is referred to as "genome-wide significance" as it adjusts for multiple testing across the entire genome.

Humphrey et al 2008). In a comprehensive review, Sanz et al. (2020) identified bacteraemia (the presence of bacteria in the blood stream) and associated systemic inflammation evident in elevated C-reactive protein and oxidative stress as mechanisms through which periodontitis increases the risk of heart disease. Using models with individual fixed effects estimated using the 1992-2016 Health and Retirement Study, Meyerhoefer et al. (2021) found that edentulous individuals, a symptom of periodontitis and dental caries, are 41.7% more likely to be diagnosed with heart disease than those who retain their permanent teeth.

Because both poor oral health and heart disease are related to unobserved health behaviors and access to medical and dental care, we construct a PGS to serve as an identifying instrument for poor oral health in a model that predicts the incidence of heart disease. A series of earlier studies investigated the genetic determinants of periodontal disease, identifying over ten genetic regions and cytokines that affect susceptibility to periodontitis through variation in periodontal structural integrity and vulnerability of subgingival microbiota. The genetic regions and cytokines identified, including FcγRIIala, interleukin (IL)-1a, IL-1b, IL-6, IL-8, IL-10, IL-37, matrix metalloproteinase 2 (MMP2), MMP3, MMP8, MMP9 and VDR, regulate platelet activity, immunity, inflammation, tissue repair and tooth formation (ADA 2023). Estimated genetic heritability of periodontal disease ranges between 30 – 50 percent, with the level of heritability increasing with disease severity (ADA 2023; Nibali et al. 2019). However, several of these genetic variants, such as interleukin (IL)-1a, IL-1b, IL-6 and IL-8 have a direct effect on heart disease (Ridker 2017; Velásquez et al. 2022; Interleukin-6 Receptor Mendelian Randomisation Analysis Consortium 2012). Therefore, a PGS instrument based only on a single oral health GWAS would violate the IV exclusion restriction.

To implement the GWAS-by-subtraction method we start with a recent GWAS for heart disease by Aragam et al. (2022) and for oral health by Shungin et al. (2019) and create a new GWAS for a latent trait measuring the poor oral health net of genetic determinants of heart disease.[6] Figure 1 contains a path diagram for the model in which square boxes contain SNP data or GWAS summary statistics for the two phenotypes, circles contain latent factors and arrows denote linear regression associations pointing from the independent to dependent variables with accompanying estimable parameters/factor loadings.[7] The critical factor loading that we use to construct a PGS for poor oral health net

[6] Shungin et al. (2019) contains separate GWAS for decayed, missing filled tooth surfaces (DMFS) and dentures as well as for periodontitis (directly) and loose teeth. Although we estimate the GWAS-by-subtraction SEM with both oral health GWAS, the DMFS and dentures GWAS uses a significantly larger genetic sample, and as a result, produces much better results.

[7] For clarity of presentation, we omit variance and covariance relationships. Most important is the restriction of zero covariance between the two latent factors.

of the genetic determinants of heart disease is $\gamma_{Non-HD-OH}$. Below, we provide the specification for GWAS-by-subtraction SEM that we estimate on every SNP:

(1) $Heart\ disease\ (HD) = \gamma_{HD-HD} Latent\ HD$
(2) $Oral\ health\ (OH) = \gamma_{HD-OH} Latent\ HD + \gamma_{Non-HD-OH} Latent\ non-HD$
(3) $Latent\ HD = \beta_{HD} SNP + v_{HD}$
(4) $Latent\ non-HD = \beta_{Non-HD} SNP + v_{Non-}$

where v_{HD} and v_{Non-HD} are the residuals of the latent factors unaccounted for by the SNP. The covariance matrix of independent variables is:

(5) $cov(SNP, v_{HD}, v_{Non-H}\ ) = \begin{bmatrix} \sigma^2_{SNP} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Demange et al. (2021) show that a simplified genetic covariance matrix, with several very small parameter set to zero, can be expressed as:

(6) $\widehat{cov}(SNP, HD, OH) \approx$

$$\begin{bmatrix} 2pq & (2pq)\beta_{HD}\gamma_{HD-HD} & (2pq)\beta_{HD}\gamma_{HD-OH} + (2pq)\omega \\ (2pq)\beta_{HD}\gamma_{HD} & (\gamma_{HD-HD}\ )^2 & \gamma_{HD-HD}\gamma_{HD-OH} \\ (2pq)\beta_{HD}\gamma_{HD-OH} + (2pq)\omega & \gamma_{HD-HD}\gamma_{HD-OH} & (\gamma_{HD-OH}\ )^2 + (\gamma_{Non-HD-OH}\ )^2 \end{bmatrix},$$

where p is the minor allele frequency of the given SNP, $q = 1 - p$ and $\omega = \beta_{Non-HD}\gamma_{Non-HD-O}$. The empirical matrix is of the form:

(7) $cov(SNP, HD, OH) = \begin{bmatrix} 2pq & (2pq)\widetilde{\beta_{HD}} & (2pq)\widetilde{\beta_{OH}} \\ (2pq)\widetilde{\beta_{HD}} & h^2_{HD} & \sigma_{HD,OH} \\ (2pq)\widetilde{\beta_{OH}} & \sigma_{HD,OH} & h^2_{OH} \end{bmatrix},$

where $\widetilde{\beta_{HD}}$ and $\widetilde{\beta_{OH}}$ are the coefficients for the SNP reported in the GWAS summary statistics for heart disease and oral health, respectively. The quantities $\sigma_{HD,OH}$, h^2_{HD} and h^2_{OH} are genetic covariance and SNP heritabilities for heart disease and oral health, respectively, estimate from a linkage disequilibrium score regression (LDSC; Finucane et al. 2015). Following Grotzinger et al. (2019), the genomic SEM is estimated by minimizing the weighted difference between covariance functions (6) and (7) using numerical optimization.

We report in appendix Table A1 the factor loading estimates from a simple estimation of equations (1) and (2) without regressing the latent factors on a SNP. To obtain to full set of beta parameters and factor loadings for each SNP, we estimate equations (1) –

(4) on all of the 7,678,811 SNPs in our HapMap3 reference file.[8] We report the five largest SNP associations, all on chromosome 5, with our latent phenotype (poor oral health net of heart disease) in Table A2. Figure 2 shows a Manhattan plot for the new latent phenotype that depicts SNP correlations across the entire chromosome with standard thresholds for genome-wide significance and suggestive associations. There are 9,459 SNPs in Figure 2 that achieve genome-wide significance.

Health and Retirement Study and empirical model

We drew an individual-level longitudinal sample from the Health and Retirement Study (HRS), which is nationally representative of the U.S. population older than age 50. The HRS surveyed multiple study cohorts every two years, starting in 1992, although genotyping of saliva and venous blood samples did not begin until 2006. We used the 2006-2020 RAND HRS (waves 8 – 15) containing 31,455 individuals with observed over a period of 7.6 years, on average. After removing those with non-zero person weights, the sample drops to 29,593 and further to 17,127 after excluding individual without genetic information or with missing data for key dependent or independent variables. This translates to a final estimation sample for most of our models of 92,046 person-years.

We used the information on self-reported medical conditions to construct several measures of heart disease to serve as outcome variables in our analysis. First, we created an indicator variable for whether a doctor ever told the respondent that they have a heart problem. This variable is coded as 1 in the wave when the respondent initially reported the condition and in all subsequent waves or is equal to 1 in all waves if the condition was diagnosed prior to survey enrollment. We also constructed variables to indicate whether the respondent developed cardiovascular disease (CVD) or CVD and stroke since their first HRS wave, which are coded as 1 in all survey waves if the respondent acquired the condition after entering the survey. Finally, we created indicator variables to measure whether the respond developed a heart problem (broadly defined), angina, congestive heart failure or had a heart attack in the past two years (i.e. since the prior survey wave). These variables are only set equal to 1 in the first wave when the condition is reported.

Our measure of oral health is whether the respondent lost all of their permanent teeth, known in the dental literature as edentulism. Both dental caries and periodontal disease cause edentulism, with secondary determinants including smoking and alcohol use, poor oral hygiene, lack of access to dental care and low socioeconomic status (Roberto et al. 2019). The HRS contains a significant amount of information on the demographic and economic conditions for participants that we use to create control

[8] We arrive at this number after munging the data to remove SNPs not contained in the reference file or with low minor allele frequency.

variables for our models. These variables are listed in Table 1 and include measures of age, race/ethnicity, marital status, employment, region of residence, household income level, type(s) of health insurance, disability status, educational attainment and level of education of the respondent's mother and father.

We constructed our PGS instrument from the factor loadings, $\hat{\gamma}_{Non-HD-OH}$, estimated from the GWAS-by-subtraction genomic SEM and the genetic information on HRS participants. The form of the PGS for individual i is as follows:

$$PGS_i = \sum_{j=1}^{J} \hat{\gamma}_{Non-HD-OH,j} G_{ij}, \quad (8)$$

where $j = 1, \ldots, J$ indexes SNPs and $G_{ij} \in \{0,1,2\}$ is the number of effect alleles at SNP j. We then normalized the PGS by subtracting its sample mean and dividing by the standard deviation. We computed PGS separately for different ancestral groups (European, Hispanic and African) to capture different frequencies of effect alleles across SNPs. It is common in computational genetics literature to compute PGS based on multiple sets of SNPs that differ in their level of GWAS statistical significance. We computed four PGS by defining sets of SNPs based on the following p-value thresholds: $p < 0.001$; $p < 0.01$; $p < 0.05$; $p < 0.1$; $p < 0.3$; $p \leq 1$. The PGS where $p \leq 1$ is the genome-wide score. Since first stage statistical power of the PGS instruments and the second stage treatment effects for the various PGS generally lie within bounds defined by the endpoints ($p < 0.001$ and $p < 1$), we only report estimates from IV models based on the genome-wide PGS and the PGS with the strictest inclusion criteria.[9]

The specification of our IV model for a given PGS instrument is as follows:

$$CVD\ outcome_{it} = \alpha_1 + \beta_1 \cdot Edentulous_{it} + \boldsymbol{\gamma_1'} \boldsymbol{X_{it}} + \omega_{1t} + \varepsilon_{1it}, \quad (9)$$

$$Edentulous_{it} = \alpha_2 + \beta_2 \cdot \widetilde{PGS}_i + \boldsymbol{\gamma_2'} \boldsymbol{X_{it}} + \omega_{2t} + \varepsilon_{2it}, \quad (10)$$

where $\boldsymbol{X_{it}}$ is a vector of control variable for person i in year t, ω_{1t} and ω_{2t} are wave fixed effects and $\widetilde{PGS}_i$ is the normalized PGS instrument. We estimated equations (9) and (10) using two-stage least squares while incorporating HRS sampling weights to make the estimates nationally representative and adjusted the standard errors stratification and clustering.

Results

We report OLS and IV estimates of the treatment effect from losing all of one's permanent teeth in Table 2. There is a statistically significant and positive correlation between edentulism and all measures of heart disease from the OLS regressions in column (1). The

[9] Results for other PGS instruments are available upon request.

magnitudes of the treatment effects are larger are the IV model in column (2), but only statistically significant for two outcomes: a doctor every told the respondent they have a heart problem and having a heart attack in the past two years. The OLS association of edentulism with the former is 6.8 percentage points (27% relative to the sample mean), while the treatment effect from the IV model is 48 percentage points (189%). Likewise, edentulism is associated with a 0.7 percentage point (51%) increase in having a heart attack over the past two years in the OLS model, but a 6.9 percentage point (529%) increase in the IV model. We report full OLS and IV estimates for selected outcomes in appendix Tables A3 and A4.

The first stage F-statistic for the PGS instrument based on the threshold of $p < 0.001$ is 24.5, which exceeds conventional standards for statistical power (Stock and Yogo 2005). When we use the genome-wide PGS as the instrument, the first stage F-statistic increases significantly to 68.4. In both cases, the PGS instruments predict a higher likelihood of poor oral health. For example, a one standard deviation increase in the PGS IV with $p < 0.001$ is associated with a 1.6 percentage point (10.7%) increase in the likelihood of edentulism (Table A3), while a one standard deviation increase in genome-wide PGS is associated with a 2.4 percentage point (15.9%) increase in the likelihood of edentulism (Table A4).

The treatment effect of edentulism on having ever been told by a doctor of a heart problem reported in column (3) for the model with the more powerful genome-wide instrument is very similar in magnitude to the estimate in column (2). However, the treatment effect on a heart attack within the past two years, at 3.9 percentage points (296%), is significantly smaller than the estimate in column (2) using the less powerful instrument. Moreover, the effects of edentulism on acquiring CVD or CVD and stroke since entering the survey become statistically significant when the IV model is estimated using the genome-wide instrument. The treatment effect on acquiring CVD is 32.7 percentage points (119%) and 37 percentage points (110%) on acquiring CVD or having a stroke.

Prior studies in Mendelian randomization computed PGS based on a GWAS for the relevant phenotype without ensuring the instruments do not have a direct effect on the outcome. To gauge the bias frp, the conventional approach, we calculated PGS directly from the GWAS statistics of Shungin et al. (2019) and re-estimated the IV models. Columns (1) and (2) compare the two estimates for the PGS based on the strict p-value threshold and columns (3) and (4) compare the estimates when using the genome-wide PGS. In nearly all cases, the estimates from the conventional approach are larger than the GWAS-by-subtraction method; the one exception being the treatment effect on heart attack when using the genome-wide instrument. For example, the treatment effect of edentulism on being told by a doctor of heart problem is 22.6% larger using the conventional GWAS-based

PGS with $p < 0.001$ (0.59 versus 0.48) and 40.5% larger when using the genome-wide instrument (0.69 versus 0.49). Moreover, the conventional GWAS-based instrument yields statistically significant effects on angina and congestive heart failure that are not present when using GWAS-by-subtraction.

Next, we test the sensitivity of our estimates to access to dental care. The HRS asks respondents whether they visited a provider within the past two years, and we re-estimated our model after including this information using a control variable. Doing so increases the magnitudes of the treatment effects (Table 4). This suggests that the effect of poor oral health on heart disease that operates through an increase in the demand for dental care reduces the incidence of heart disease. Given that dental care provides both oral and non-oral health benefits, the results of this mediation analysis are not surprising (Meyerhoefer et al. 2021).

Falsification tests and heterogeneous treatment effects

Although the IV exclusion restriction is not directly testable, we can perform several empirical checks to investigate the instrument's validity. The first is a series of falsification tests where we estimate the IV models on health outcomes that should not be affected by the absence of permanent teeth. The HRS contains information on patient-reported cancer diagnoses, psychiatric problems and sleep disorders that we use for this purpose. Although edentulism is associated with cancers of the oral cavity and throat, these represent only 3% of all cancer diagnosis, so we do not expect edentulism to be significantly correlated with a measure of any cancer (Ellington et al. 2020). Moreover, although edentulism can be associated with self-rated mental health, it is not a major determinant of psychiatric illness, and it does not have an independent effect on the quality of sleep (Saman et al. 2014; Emami et al. 2012).

Table 5 contains OLS and IV estimates of the effect of edentulism on unrelated health conditions. We find statistically significant associations between being told by a doctor of psychiatric problems and sleep disorders in the OLS models, but these associations are not statistically significant in either of the IV models. There are no associations with cancer in any of the models. These treatment effect estimates are consistent with the validity of the PGS instrument.

Following McClellan, McNeil and Newhouse (1994) and Böckerman et al. (2019), we calculate the means of our control variables for individuals with a below and above average value of the instrument and test for statistical differences. Table A5 contains the balance test for the PGS instrument with $p<0.001$ and Table A6 contains the balance test for the genome-wide instrument. Although several characteristics have statistically significant

differences between the sub-samples, the magnitudes of most differences are small. The exceptions include the portion of the sample which is black, the number of ADLs and the number of IADLs, which are all larger for those with above average values of the instrument. In addition, the percentage of the sample with a college degree or higher is smaller for those with above average values of the instrument. Differences in characteristics for those with high or low values of the instrument are smaller for the genome-wide instrument than the instrument with p-value threshold of $p<0.001$. The only notable difference between the subsamples is for the proportion with a college degree or higher, which is lower for those with above average values of the instrument.

Finally, we investigate whether there are heterogenous treatment effects, by estimating our OLS and IV models for subsamples defined across certain demographic characteristics and report the results in Table 6. Edentulism is more consistently associated with heart problems for those over 65 than adults under 65. Nonetheless, there are a few large treatment effects for the under-65 population, possibly because edentulism at a younger age indicates more severe periodontitis. When we separate the sample by gender, we only find that edentulism increases the risk of CVD, heart attack and other cardiac issues for men in the IV models, but not for women. Finally, we fail to find any significant IV treatment effects for the non-white population, although this could reflected lower statistical power in the smaller sample or the calculation of PGS weights from GWAS of predominantly European ancestry (Martin et al. 2017).

Discussion

We use a GWAS-by-subtraction genomic structural equation model to create instruments for poor oral health that remove any direct genetic effect on heart disease outcomes. These instruments are polygenic scores (PGS) that make use of patterns of genetic inheritance across either thousands or millions of SNPs, depending on our inclusion criteria. The instruments are powerful by conventional standards with first stage F-statistics ranging between 25 – 68.

When we estimate OLS models of edentulism on heart disease outcomes, we find statistically significant associations with all outcomes, even after controlling for a large set of demographic and socioeconomic factors and disability status. Associations between edentulism and the development of angina, congestive heart failure and new heart problem (broadly defined) in the past two years are statistically insignificant in the IV models. However, IV treatment effects of edentulism on whether the respondent was ever told by a doctor of a heart problem and on having a heart attack in the past two years are statistically significant, irrespective of the instrument. Both IV treatment effects are

positive and much larger in magnitude than in the OLS models, suggesting attenuation bias.

Although it is possible that omitted variables could negatively bias OLS treatment effects, it is not clear what the omitted determinants could be. If unobserved poor SES were a significant confounder, we would predict a positive rather than negative bias but given the large set of control variables in the model, we expect that our models account for the primary determinants of SES (race/ethnicity, income, employment, education, insurance, household composition, geographic location and disability). Another possibility is that attenuation bias in the OLS models is largely driven by measurement error in self-reported edentulism. 16.7% of individuals in our sample do not report whether they have lost teeth, so we include an indicator variable to control for non-reporting behavior. Multiple studies have documented stigma from lost teeth, but evidence on the quality of self-reports of missing teeth is context-specific (Yan et al. 2025; Doughty et al. 2023). For example, Høvik et al. (2022) found that self-reports of tooth loss and edentulism among older Norwegian adults are close to clinical measures, unless respondents have cognitive impairments, while Shimazaki et al (2024) found that the self-reported number of teeth matched the observed number of teeth only 57.5% of the time among older Japanese adults.

Our results have important implications for health care policy. The largest group of Americans with inconsistent access to dental care are those enrolled in Medicare. This is because there is no dental benefit in Medicare Part B and many seniors find paying fully out-of-pocket for dental services a financial burden (Tranby et al. 2020). For example. Jacobson et al. (2021) found that one-fifth of older Americans spend in excess of $2,000 per year on dental care. Although a growing number of Medicare Part C plans offer dental plans, the coverage is often limited, and some seniors find the network restrictions associated with Part C plans unattractive. Overall, Tranby et al. (2020) found that Medicare enrollees (including those with Medicare C) paid 75% of dental care costs out-of-pocket. The expense to taxpayers of adding a dental benefit to Medicare Part B has limited the political viability of coverage proposals, but our results suggest that cost estimates are inflated because they fail to account for reductions in Medicare payments associated with the heart disease treatment.[10] Given that annual Medicare expenditures to treat heart disease exceed $55 billion, the failure to account for cost offsets through reductions in heart disease could significantly bias evaluations of Medicare dental benefit costs (Muhuri 2023).

[10] Gangopadhyaya, Garret and Holahan (2023) estimate the cost of a Medicare Part B dental benefit at $60.1 billion, annually.

Our study has limitations. When we estimate models that do not use GWAS-by-subtraction to report direct genetic effects, we find that the IV treatment effects are 23%-40% larger than in our main IV models. This suggests that failure to account for common genetic factors that determine both heart disease and edentulism results in positive bias. One limitation of our approach is that we only subtract genetic variants that have some role in determining heart disease. If, for example, a variant was a genetic determinant to poor oral health as well as another condition with no genetic link to heart disease, this variant would remain on our PGS. Nonetheless, it is difficult to think of phenotypes that are associated with heart disease but have no shared genetic determinants with heart disease and do determine poor oral health.[11] Nonetheless, we cannot completely rule out bias from pleiotropy using our approach and future studies should seek to investigate other latent pathways that could improve the credibility of instruments.

[11] For example, diabetes and dementia, both of which are linked to periodontal disease share genetic variants with heart disease (Sousa et al. 2011; Broce at al. 2019).

References

American Dental Association (ADA). (2023). Genetics and oral health. ADA Science and Research Institute (Chicago, IL). Accessed on December 12. 2025 at: https://www.ada.org/resources/ada-library/oral-health-topics/genetics-and-oral-health

Aragam, K. G., Jiang, T., Goel, A., Kanoni, S., Wolford, B. N., Atri, D. S., Weeks, E. M., Wang, M., Hindy, G., Zhou, W., Grace, C., Roselli, C., Marston, N. A., Kamanu, F. K., Surakka, I., Venegas, L. M., Sherliker, P., Koyama, S., Ishigaki, K., Åsvold, B. O., ... CARDIoGRAMplusC4D Consortium (2022). Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nature genetics*, *54*(12), 1803–1815.

Böckerman, P., Cawley, J., Viinikainen, J., Lehtimäki, T., Rovio, S., Seppälä, I., Pehkonen, J., & Raitakari, O. (2019). The effect of weight on labor market outcomes: An application of genetic instrumental variables. *Health economics*, *28*(1), 65–77.

Broce, I. J., Tan, C. H., Fan, C. C., Jansen, I., Savage, J. E., Witoelar, A., Wen, N., Hess, C. P., Dillon, W. P., Glastonbury, C. M., Glymour, M., Yokoyama, J. S., Elahi, F. M., Rabinovici, G. D., Miller, B. L., Mormino, E. C., Sperling, R. A., Bennett, D. A., McEvoy, L. K., Brewer, J. B., ... Desikan, R. S. (2019). Dissecting the genetic relationship between cardiovascular risk factors and Alzheimer's disease. *Acta neuropathologica*, *137*(2), 209–226.

Brunello, G., Sanz-de-Galdeano, A., & Terskaya, A. (2020). Not only in my genes: The effects of peers' genotype on obesity. *Journal of health economics*, *72*, 102349.

Cardon, L. R., & Palmer, L. J. (2003). Population stratification and spurious allelic association. *Lancet*, *361*(9357), 598–604.

Cawley, J., Han, E., & Norton, E. C. (2011). The validity of genes related to neurotransmitters as instrumental variables. *Health economics*, *20*(8), 884–888.

Davey Smith G. (2011). Random allocation in observational data: how small but robust effects could facilitate hypothesis-free causal inference. *Epidemiology*, *22*(4), 460–468.

Ding, W., Lehrer, S. F., Rosenquist, J. N., & Audrain-McGovern, J. (2009). The impact of poor health on academic performance: New evidence using genetic markers. *Journal of health economics*, *28*(3), 578–597.

Doughty, J., Macdonald, M. E., Muirhead, V., & Freeman, R. (2023). Oral health-related stigma: Describing and defining a ubiquitous phenomenon. *Community dentistry and oral epidemiology*, *51*(6), 1078–1083.

Edwards, C. H., Bjørngaard, J. H., & Minet Kinge, J. (2021). The relationship between body mass index and income: Using genetic variants from HUNT as instrumental variables. *Health economics*, *30*(8), 1933–1949.

Ellington, T. D., Henley, S. J., Senkomago, V., O'Neil, M. E., Wilson, R. J., Singh, S., Thomas, C. C., Wu, M., & Richardson, L. C. (2020). Trends in Incidence of Cancers of the Oral Cavity and Pharynx - United States 2007-2016. *MMWR. Morbidity and mortality weekly report*, *69*(15), 433–438.

Emami, E., Lavigne, G., de Grandmont, P., Rompré, P. H., & Feine, J. S. (2012). Perceived sleep quality among edentulous elders. *Gerodontology*, *29*(2), e128–e134.

Finucane, H. K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P. R., Anttila, V., Xu, H., Zang, C., Farh, K., Ripke, S., Day, F. R., ReproGen Consortium, Schizophrenia Working Group of the Psychiatric Genomics Consortium, RACI Consortium, Purcell, S., Stahl, E., Lindstrom, S., Perry, J. R., Okada, Y., … Price, A. L. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature genetics*, *47*(11), 1228–1235.

Fletcher J. (2023). Decoupling genetics from attainments: The role of social environments. *Economics and human biology*, *50*, 101259.

Fletcher, J. M., & Lehrer, S. F. (2009). The effects of adolescent health on educational outcomes: Causal evidence using genetic lotteries between siblings. *Forum for health economics & policy, 12*(2), Article 8.

_____. (2011). Genetic lotteries within families. *Journal of health economics*, *30*(4), 647–659.

Gangopadhyaya, A., Garret, B., & Holahan, J. (2023). Estimating the cost and effects of adding a dental benefit to Medicare Part B. Urban Institute Health Policy Center. September. Accessed on December 12, 2025 at: https://www.urban.org/research/publication/estimating-cost-and-effects-adding-dental-benefit-medicare-part-b

Hartwig, F. P., Davies, N. M., & Davey Smith, G. (2018). Bias in Mendelian randomization due to assortative mating. *Genetic epidemiology*, *42*(7), 608–620.

Høvik, H., Kolberg, M., Gjøra, L., Nymoen, L. C., Skudutyte-Rysstad, R., Hove, L. H., Sun, Y. Q., & Fagerhaug, T. N. (2022). The validity of self-reported number of teeth and edentulousness among Norwegian older adults, the HUNT Study. *BMC oral health*, *22*(1), 82.

Humphrey, L. L., Fu, R., Buckley, D. I., Freeman, M., & Helfand, M. (2008). Periodontal disease and coronary heart disease incidence: a systematic review and meta-analysis. *Journal of general internal medicine*, *23*(12), 2079–2086

Interleukin-6 Receptor Mendelian Randomisation Analysis Consortium, Swerdlow, D. I., Holmes, M. V., Kuchenbaecker, K. B., Engmann, J. E., Shah, T., Sofat, R., Guo, Y., Chung, C., Peasey, A., Pfister, R., Mooijaart, S. P., Ireland, H. A., Leusink, M., Langenberg, C., Li, K. W., Palmen, J., Howard, P., Cooper, J. A., Drenos, F., ... Casas, J. P. (2012). The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet*, *379*(9822), 1214–1224.

Jacobson, G., Cicchiello, A., Shah, A., Doty, M. M., & Williams II, R. D. (2021, October). *When costs are a barrier to getting health care: Reports from older adults in the United States and other high-income countries*. Commonwealth Fund. https://doi.org/10.26099/m7jm-2n91

Koellinger, P. D., & de Vlaming, R. (2019). Mendelian randomization: the challenge of unobserved environmental confounds. *International journal of epidemiology*, *48*(3), 665–671.

Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., Daly, M. J., Bustamante, C. D., & Kenny, E. E. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *American journal of human genetics*, *100*(4), 635–649.

McClellan, M., McNeil, B. J., & Newhouse, J. P. (1994). Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA*, *272*(11), 859–866.

Menta, G., Lepinteur, A., Clark, A. E., Ghislandi, S., & D'Ambrosio, C. (2023). Maternal genetic risk for depression and child human capital. *Journal of health economics*, *87*, 102718.

Meyerhoefer, C. D., Pepper, J. V., Manski, R. J., & Moeller, J. F. (2021). Dental Care Use, Edentulism, and Systemic Health among Older Adults. *Journal of dental research*, *100*(13), 1468–1474.

Moeller, J., Manski, R., Chen, H., Zuvekas, S., & Meyerhoefer, C. (2020). Does covering routine dental care for the Medicare population produce cost savings in Medicare? A preliminary 2-year analysis. *Journal of Public Health Dentistry, 80*(1), 31-42.

Muhuri, P. K. (2023). Healthcare Expenditures for Heart Disease among Adults Aged 18 and Older in the U.S. Civilian Noninstitutionalized Population, 2020. In *Statistical Brief (Medical Expenditure Panel Survey (US))*. Agency for Healthcare Research and Quality (US).

Nibali, L., Bayliss-Chapman, J., Almofareh, S. A., Zhou, Y., Divaris, K., & Vieira, A. R. (2019). What Is the Heritability of Periodontitis? A Systematic Review. *Journal of dental research*, *98*(6), 632–641.

Norton, E. C., & Han, E. (2008). Genetic information, obesity, and labor market outcomes. *Health economics*, *17*(9), 1089–1104.

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S. J., Diekhans, M., Logsdon, G. A., Alonge, M., Antonarakis, S. E., Borchers, M., Bouffard, G. G., Brooks, S. Y., Caldas, G. V., … Phillippy, A. M. (2022). The complete sequence of a human genome. *Science, 376*(6588), 44–53.

Palmer, T. M., Lawlor, D. A., Harbord, R. M., Sheehan, N. A., Tobias, J. H., Timpson, N. J., Davey Smith, G., & Sterne, J. A. (2012). Using multiple genetic variants as instrumental variables for modifiable risk factors. *Statistical methods in medical research*, *21*(3), 223–242.

Pehkonen, J., Viinikainen, J., Kari, J. T., Böckerman, P., Lehtimäki, T., & Raitakari, O. (2021). Birth weight and adult income: An examination of mediation through adult height and body mass. *Health economics*, *30*(10), 2383–2398.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, *38*(8), 904–909.

Ridker, P. M., Everett, B. M., Thuren, T., MacFadyen, J. G., Chang, W. H., Ballantyne, C., Fonseca, F., Nicolau, J., Koenig, W., Anker, S. D., Kastelein, J. J. P., Cornel, J. H., Pais, P., Pella, D., Genest, J., Cifkova, R., Lorenzatti, A., Forster, T., Kobalava, Z., Vida-Simiti, L., … CANTOS Trial Group (2017). Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *The New England journal of medicine*, *377*(12), 1119–1131.

Roberto, L. L., Crespo, T. S., Monteiro-Junior, R. S., Martins, A. M. E. B. L., De Paula, A. M. B., Ferreira, E. F., & Haikal, D. S. (2019). Sociodemographic determinants of edentulism in the elderly population: A systematic review and meta-analysis. *Gerodontology*, *36*(4), 325–337.

Saman, D. M., Lemieux, A., Arevalo, O., & Lutfiyya, M. N. (2014). A population-based study of edentulism in the US: does depression and rural residency matter after controlling for potential confounders?. *BMC public health*, *14*, 65.

Sanz, M., Marco Del Castillo, A., Jepsen, S., Gonzalez-Juanatey, J. R., D'Aiuto, F., Bouchard, P., Chapple, I., Dietrich, T., Gotsman, I., Graziani, F., Herrera, D., Loos, B., Madianos, P., Michel, J. B., Perel, P., Pieske, B., Shapira, L., Shechter, M., Tonetti, M., Vlachopoulos, C., ... Wimmer, G. (2020). Periodontitis and cardiovascular diseases: Consensus report. *Journal of clinical periodontology*, *47*(3), 268–288.

Shimazaki, Y., Saito, M., Nonoyama, T., & Inamoto, Y. (2024). Validity of the self-reported number of teeth in independent older people in Japan. *BMC geriatrics*, *24*(1), 900.

Shungin, D., Haworth, S., Divaris, K., Agler, C. S., Kamatani, Y., Keun Lee, M., Grinde, K., Hindy, G., Alaraudanjoki, V., Pesonen, P., Teumer, A., Holtfreter, B., Sakaue, S., Hirata, J., Yu, Y. H., Ridker, P. M., Giulianini, F., Chasman, D. I., Magnusson, P. K. E., Sudo, T., ... Johansson, I. (2019). Genome-wide analysis of dental caries and periodontitis combining clinical and self-reported data. *Nature communications*, *10*(1), 2773.

Small, D. S., Rosenbaum, P. R. (2008). War and wages. *Journal of the American Statistical Association, 103*(483), 924-933.

Sousa, A. G., Selvatici, L., Krieger, J. E., & Pereira, A. C. (2011). Association between genetics of diabetes, coronary artery disease, and macrovascular complications: exploring a common ground hypothesis. *The review of diabetic studies : RDS*, *8*(2), 230–244.

Spahr, A., Klein, E., Khuseyinova, N., Boeckh, C., Muche, R., Kunze, M., Rothenbacher, D., Pezeshki, G., Hoffmeister, A., & Koenig, W. (2006). Periodontal infections and coronary heart disease: role of periodontal bacteria and importance of total pathogen burden in the Coronary Event and Periodontal Disease (CORODONT) study. *Archives of internal medicine*, *166*(5), 554–559.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* (pp. 80–108). Cambridge University Press.

Tranby, E. P., Halasa-Rappel, Y. A., Brow, A. R., Jacob, M., & Frantsve-Hawley, J. (2020, February). *The burden of out-of-pocket expenditures for dental care on Medicare-enrolled elderly and disabled: Part 3 of 3*. CareQuest Institute for Oral Health. https://doi.org/10.35565/CQI.2020.2002.

Velásquez, I. M., Malarstig, A., Baldassarre, D., Borne, Y., de Faire, U., Engström, G., Eriksson, P., Giral, P., Humphries, S. E., Kurl, S., Leander, K., Lind, L., Lindén, A., Orsini, N., Pirro, M., Silveira, A., Smit, A. J., Tremoli, E., Veglia, F., Strawbridge, R. J., ... Gigante, B. (2023). Causal analysis of plasma IL-8 on carotid intima media thickness, a measure of subclinical atherosclerosis. *Current research in translational medicine*, *71*(1), 103374.

von Hinke, S., Davey Smith, G., Lawlor, D. A., Propper, C., & Windmeijer, F. (2016). Genetic markers as instrumental variables. *Journal of health economics*, *45*, 131–148.

von Hinke Kessler Scholder, S., Smith, G. D., Lawlor, D. A., Propper, C., & Windmeijer, F. (2011). Mendelian randomization: the use of genes in instrumental variable analyses. *Health economics*, *20*(8), 893–896.

Yan, Z. B., Zhu, R., Yi, X., Pei, Y. P., Zhan, M. J., Zhu, J. L., & Yu, H. Y. (2025). Is stigma correlated with oral health-related quality of life in prosthodontic patients? a cross-sectional study. *BMC oral health*, *25*(1), 1614.

Table 1. Descriptive statistics.

Variable	N·T	Mean	S.D.	Min	Max
Outcomes					
Heart problem in past 2 yrs.	88484	0.034	0.181	0	1
Dr. ever told of heart problem	92046	0.255	0.436	0	1
Acquired CVD since first wave	92046	0.275	0.446	0	1
Acquired CVD or stroke since first wave	92046	0.335	0.472	0	1
Heart attack in past 2 yrs.	92046	0.013	0.111	0	1
Angina in past 2 yrs.	92046	0.035	0.185	0	1
Cong. heart failure in past 2 yrs.	92046	0.030	0.172	0	1
Control variables					
Lost all permanent teeth	92046	0.149	0.356	0	1
Did not lose all permanent teeth	92046	0.684	0.465	0	1
Dentate status missing	92046	0.167	0.373	0	1
Age < 65	92046	0.370	0.483	0	1
65 ≤ Age < 70	92046	0.164	0.370	0	1
70 ≤ Age < 75	92046	0.154	0.361	0	1
75 ≤ Age < 80	92046	0.136	0.343	0	1
80 ≤ Age < 85	92046	0.098	0.297	0	1
85 ≤ Age	92046	0.077	0.267	0	1
Hispanic	92046	0.128	0.335	0	1
White non-Hispanic	92046	0.701	0.458	0	1
Black non-Hispanic	92046	0.169	0.375	0	1
Race/ethnicity missing	92046	0.002	0.040	0	1
Male	92046	0.415	0.493	0	1
Female	92046	0.585	0.493	0	1
Married or partnered	92046	0.635	0.482	0	1
Single	92046	0.365	0.481	0	1
Marital status missing	92046	0.001	0.027	0	1
No. of HH residents	92046	2.233	1.214	1	15
No. of living children	92046	3.170	2.041	0	11
Lives w/ spouse/partner	92046	0.635	0.482	0	1
Not in labor force & retired	92046	0.552	0.497	0	1
Not fully retired	92046	0.104	0.305	0	1
Working full time	92046	0.249	0.432	0	1
Unemployed	92046	0.011	0.105	0	1
Other work status	92046	0.083	0.276	0	1
Not in labor force & not retired	92046	0.062	0.242	0	1
New England resident	92046	0.037	0.188	0	1
Mid-Atlantic resident	92046	0.113	0.317	0	1
East-North-Central resident	92046	0.159	0.366	0	1
West-North-Central resident	92046	0.081	0.273	0	1
South-Atlantic resident	92046	0.238	0.426	0	1
East-South-Central resident	92046	0.064	0.245	0	1
West-South-Central resident	92046	0.109	0.311	0	1

Table 1, continued.

Variable	N·T	Mean	S.D.	Min	Max
Control variables					
Mountain resident	92046	0.064	0.245	0	1
Pacific resident	92046	0.134	0.341	0	1
Other non-US resident	92046	0.000	0.021	0	1
Census division missing	92046	0.000	0.017	0	1
HH income is ≤ 100% FPL	92046	0.106	0.308	0	1
HH income is 101% - 200% FPL	92046	0.184	0.388	0	1
HH income is 201% - 400% FPL	92046	0.306	0.461	0	1
HH income > 400% FPL	92046	0.403	0.491	0	1
Medicare	92046	0.648	0.477	0	1
Medicaid	92046	0.092	0.289	0	1
Private health insurance	92046	0.438	0.496	0	1
Other health insurance	92046	0.155	0.362	0	1
No ADLs	92046	0.927	0.261	0	1
No IADLs	92046	0.886	0.318	0	1
No. of ADLs	92046	0.318	0.869	0	5
No. of IADLs	92046	0.280	0.818	0	5
Less than high school education	92046	0.181	0.385	0	1
GED or high school graduate	92046	0.343	0.475	0	1
Some college	92046	0.246	0.431	0	1
College degree or higher	92046	0.230	0.421	0	1
Education is missing	92046	0.000	0.012	0	1
Mother had at least high school education	92046	0.136	0.343	0	1
Father had at least high school education	92046	0.140	0.347	0	1
GWAS-by-subtraction polygenic score instrumental variables					
PGS IV ($p < 0.001$)	92046	0.0001	0.9995	-3.3736	6.2627
PGS IV ($p < 1$)	92046	-0.0071	0.9983	-3.9612	4.9461
GWAS polygenic score instrumental variables for sensitivity test					
GWAS PGS IV ($p < 0.001$)	92046	-0.0019	0.9987	-3.8940	3.9610
GWAS PGS IV ($p < 1$)	92046	-0.0094	0.9963	-4.7896	3.8047
Additional control for sensitivity test					
Visited dental provider in last 2 yrs.	91959	0.643	0.479	0	1
Outcomes for falsification test					
Dr. ever told of cancer (except skin)	92046	0.159	0.366	0	1
Dr. ever told of psychiatric problem	92046	0.180	0.384	0	1
Dr. ever told of sleep disorder	28344	0.178	0.383	0	1
Had cancer in past 2 yrs.	88484	0.022	0.147	0	1
Had psychiatric problem in past 2 yrs.	88484	0.017	0.129	0	1

Notes: Means are weighted to be nationally representative.

Table 2. OLS and IV treatment effect of losing all permanent teeth on measures of heart disease.

Outcome variable	**OLS (1)**	**PGS IV (p < 0.001) (2)**	**PGS IV (p ≤ 1) (3)**
Heart problem in past 2 yrs.	0.0059**	0.0543	0.0339
	(0.0026)	(0.0440)	(0.0286)
Dr. ever told of heart problem	0.0685***	0.4832***	0.4879***
	(0.0115)	(0.2474)	(0.1745)
Acquired CVD since first wave	0.0254**	0.2728	0.3274***
	(0.0124)	(0.2613)	(0.1616)
Acquired CVD or stroke since first wave	0.0349***	0.4818	0.3699***
	(0.0116)	(0.3173)	(0.2024)
Heart attack in past 2 yrs.	0.0066***	0.0688***	0.0385***
	(0.0016)	(0.0375)	(0.0204)
Angina in past 2 yrs.	0.0118***	0.1177	0.0760
	(0.0035)	(0.0825)	(0.0499)
Cong. heart failure in past 2 yrs.	0.0163***	0.1209	0.0672
	(0.0042)	(0.0753)	(0.0486)
F-statistic		24.5	68.4

Notes: Models contain wave fixed effects and the set of control variables indicated in Table 1. Estimates are weighted and standard errors adjusted for the complex design of the Health and Retirement Study. ***,**,* indicates statistical significance at the 1%, 5% and 10% level, respectively.

Table 3. GWAS versus GWAS-by-Subtraction IV treatment effect of losing all permanent teeth on measures of heart disease.

Outcome var.	**GWAS PGS IV ($p < 0.001$) (1)**	**GWAS-by-subtr. PGS IV ($p < 0.001$) (2)**	**GWAS PGS IV ($p \leq 1$) (3)**	**GWAS-by-subtr. PGS IV ($p \leq 1$) (4)**
Heart problem in past 2 yrs.	0.0128	0.0543	0.0332	0.0339
	(0.0403)	(0.0440)	(0.0282)	(0.0286)
Dr. ever told of heart problem	0.5923***	0.4832***	0.6856***	0.4879***
	(0.2070)	(0.2474)	(0.1656)	(0.1745)
Acquired CVD since first wave	0.3242	0.2728	0.4814***	0.3274***
	(0.2452)	(0.2613)	(0.1612)	(0.1616)
Acquired CVD or stroke since first wave	0.2687	0.4818	0.4355***	0.3699***
	(0.2353)	(0.3173)	(0.1767)	(0.2024)
Heart attack in past 2 yrs.	0.0196	0.0688***	0.0335***	0.0385***
	(0.0276)	(0.0375)	(0.0169)	(0.0204)
Angina in past 2 yrs.	0.1540***	0.1177	0.1034***	0.0760
	(0.0703)	(0.0825)	(0.0469)	(0.0499)
Cong. heart failure in past 2 yrs.	0.1344***	0.1209	0.0769***	0.0672
	(0.0655)	(0.0753)	(0.0401)	(0.0486)
F-statistic	39.6	24.5	83.7	68.4

Notes: Models contain wave fixed effects and the set of control variables indicated in Table 1. Estimates are weighted and standard errors adjusted for the complex design of the Health and Retirement Study. ***, **,* indicates statistical significance at the 1%, 5% and 10% level, respectively.

Table 4. IV treatment effect of losing all permanent teeth on measures of heart disease, controlling for dental care use in the past two years.

Outcome var.	**PGS IV ($p < 0.001$) (1)**	**PGS IV ($p < 0.001$) w/ dental care (2)**	**PGS IV ($p \leq 1$) (3)**	**PGS IV ($p \leq 1$) w/ dental care (4)**
Heart problem in past 2 yrs.	0.0543	0.0626	0.0339	0.0410
	(0.0440)	(0.0518)	(0.0286)	(0.0365)
Dr. ever told of heart problem	0.4832***	0.5477***	0.4879***	0.5870***
	(0.2474)	(0.2920)	(0.1745)	(0.2257)
Acquired CVD since first wave	0.2728	0.3094	0.3274***	0.3976***
	(0.2613)	(0.3069)	(0.1616)	(0.2054)
Acquired CVD or stroke since first wave	0.4818	0.5489	0.3699***	0.4420***
	(0.3173)	(0.3725)	(0.2024)	(0.2566)
Heart attack in past 2 yrs.	0.0688***	0.0776***	0.0385***	0.0438***
	(0.0375)	(0.0445)	(0.0204)	(0.0261)
Angina in past 2 yrs.	0.1177	0.1364	0.0760	0.0925
	(0.0825)	(0.0979)	(0.0499)	(0.0633)
Cong. heart failure in past 2 yrs.	0.1209	0.1344	0.0672	0.0729
	(0.0753)	(0.0887)	(0.0486)	(0.0621)
F-statistic	24.5	24.7	68.4	57.9

Notes: Models contain wave fixed effects and the set of control variables indicated in Table 1. Estimates are weighted and standard errors adjusted for the complex design of the Health and Retirement Study. ***,**,* indicates statistical significance at the 1%, 5% and 10% level, respectively.

Table 5. OLS and IV treatment effect of losing all permanent teeth on unrelated outcomes.

Outcome variable	**OLS** **(1)**	**PGS IV (p < 0.001)** **(2)**	**PGS IV (p ≤ 1)** **(3)**
Dr. ever told of cancer (except skin)	0.0110	-0.0644	-0.1429
	(0.0107)	(0.1804)	(0.1332)
Dr. ever told of psychiatric problem	0.0283***	0.0976	0.0460
	(0.0105)	(0.2160)	(0.1536)
Dr. ever told of sleep disorder	-0.0255*	-0.5792	-0.4371
	(0.0148)	(0.3942)	(0.2877)
Had cancer in past 2 yrs.	0.0026	0.0062	0.0099
	(0.0018)	(0.0356)	(0.0258)
Had psychiatric problem in past 2 yrs.	-0.0001	-0.0157	0.0098
	(0.0016)	(0.0359)	(0.0226)
F-statistic		24.5	68.4

Notes: Models contain wave fixed effects and the set of control variables indicated in Table 1. Estimates are weighted and standard errors adjusted for the complex design of the Health and Retirement Study. ***, **,* indicates statistical significance at the 1%, 5% and 10% level, respectively.

Table 6. OLS and IV treatment effect of losing all permanent teeth on heart disease outcomes for subpopulations.

Outcome variable	OLS (1)	PGS IV (p < 0.001) (2)	PGS IV (p ≤ 1) (3)
Age ≥ 65 (N·T = 54,806)			
Heart problem in past 2 yrs.	0.0054**	0.0149	-0.0018
	(0.0027)	(0.0440)	(0.0273)
Dr. ever told of heart problem	0.0562***	0.3111	0.3812***
	(0.0112)	(0.2262)	(0.1399)
Acquired CVD since first wave	0.0248*	0.2267	0.2404***
	(0.0126)	(0.2444)	(0.1280)
Acquired CVD or stroke since first wave	0.0297**	0.2748	0.1996
	(0.0130)	(0.2823)	(0.1466)
Heart attack in past 2 yrs.	0.0067***	0.0877***	0.0464***
	(0.0018)	(0.0408)	(0.0189)
Angina in past 2 yrs.	0.0111**	0.1034	0.0744***
	(0.0047)	(0.0812)	(0.0434)
Cong. heart failure in past 2 yrs.	0.0182***	0.1497***	0.0986***
	(0.0049)	(0.0838)	(0.0458)
Age < 65 (N·T = 37,240)			
Heart problem in past 2 yrs.	0.0088*	0.1251	0.1253
	(0.0051)	(0.0876)	(0.0778)
Dr. ever told of heart problem	0.0937***	0.8164	0.7878***
	(0.0214)	(0.4883)	(0.4058)
Acquired CVD since first wave	0.0312	0.3329	0.5323
	(0.0198)	(0.4203)	(0.3934)
Acquired CVD or stroke since first wave	0.0490**	0.8434***	0.7793
	(0.0186)	(0.4985)	(0.4838)
Heart attack in past 2 yrs.	0.0061**	0.0340	0.0229
	(0.0029)	(0.0570)	(0.0451)
Angina in past 2 yrs.	0.0134**	0.1331	0.0752
	(0.0057)	(0.1341)	(0.1058)
Cong. heart failure in past 2 yrs.	0.0123**	0.0712	-0.0098
	(0.0056)	(0.0947)	(0.0838)
Men (N·T = 38,234)			
Heart problem in past 2 yrs.	0.0062*	0.1659	0.1049
	(0.0037)	(0.1134)	(0.0680)
Dr. ever told of heart problem	0.0817***	0.9628	1.1440***
	(0.0216)	(0.7340)	(0.4835)
Acquired CVD since first wave	0.0243	0.7576	0.9721***
	(0.0191)	(0.6830)	(0.4770)
Acquired CVD or stroke since first wave	0.0356*	0.7698	0.7076
	(0.0197)	(0.7359)	(0.4363)
Heart attack in past 2 yrs.	0.0097***	0.2422	0.1089***
	(0.0028)	(0.1457)	(0.0630)

Table 6, continued.

Outcome variable	OLS (1)	PGS IV ($p < 0.001$) (2)	PGS IV ($p \leq 1$) (3)
Men (N·T = 38,234)			
Angina in past 2 yrs.	0.0097*	0.1623	0.0466
	(0.0050)	(0.1548)	(0.1014)
Cong. heart failure in past 2 yrs.	0.0115*	0.2328	0.1741
	(0.0064)	(0.2051)	(0.1196)
Women (N·T = 53,812)			
Heart problem in past 2 yrs.	0.0060**	0.0106	0.0031
	(0.0029)	(0.0373)	(0.0255)
Dr. ever told of heart problem	0.0594***	0.3078	0.2293
	(0.0137)	(0.2133)	(0.1384)
Acquired CVD since first wave	0.0282*	0.0953	0.0675
	(0.0156)	(0.2012)	(0.1243)
Acquired CVD or stroke since first wave	0.0360**	0.3857	0.2363
	(0.0162)	(0.2650)	(0.1775)
Heart attack in past 2 yrs.	0.0041**	0.0054	0.0132
	(0.0018)	(0.0219)	(0.0169)
Angina in past 2 yrs.	0.0139**	0.0903	0.0811
	(0.0055)	(0.0871)	(0.0503)
Cong. heart failure in past 2 yrs.	0.0203***	0.0733	0.0204
	(0.0054)	(0.0630)	(0.0466)
Non-white (N·T = 20,052)			
Heart problem in past 2 yrs.	0.0017	0.0608	-0.0443
	(0.0043)	(0.0970)	(0.0782)
Dr. ever told of heart problem	0.0517**	0.7164	0.7309
	(0.0211)	(0.6225)	(0.6020)
Acquired CVD since first wave	0.0133	0.3500	-0.1516
	(0.0220)	(0.6253)	(0.4646)
Acquired CVD or stroke since first wave	0.0497**	0.4824	0.0554
	(0.0247)	(0.7872)	(0.5606)
Heart attack in past 2 yrs.	0.0039	0.0069	0.0579
	(0.0033)	(0.0553)	(0.0629)
Angina in past 2 yrs.	0.0079	0.1373	0.0215
	(0.0072)	(0.1437)	(0.1145)
Cong. heart failure in past 2 yrs.	0.0232**	0.2060	0.1175
	(0.0090)	(0.2430)	(0.1993)
White (N·T = 71,850)			
Heart problem in past 2 yrs.	0.0067**	0.0555	0.0427
	(0.0029)	(0.0483)	(0.0317)
Dr. ever told of heart problem	0.0705***	0.4464	0.4551***
	(0.0139)	(0.2705)	(0.1811)
Acquired CVD since first wave	0.0275*	0.2682	0.3747***
	(0.0138)	(0.2757)	(0.1800)

Table 6, continued.

Outcome variable	OLS (1)	PGS IV (p < 0.001) (2)	PGS IV (p ≤ 1) (3)
	White (N·T = 71,850)		
Acquired CVD or stroke since first wave	0.0309**	0.4907	0.4025***
	(0.0138)	(0.3222)	(0.2083)
Heart attack in past 2 yrs.	0.0070***	0.0775***	0.0379
	(0.0019)	(0.0433)	(0.0232)
Angina in past 2 yrs.	0.0123***	0.1087	0.0755
	(0.0037)	(0.0917)	(0.0536)
Cong. heart failure in past 2 yrs.	0.0148***	0.1053	0.0577
	(0.0045)	(0.0816)	(0.0496)

Notes: Models contain wave fixed effects and the set of control variables indicated in Table 1. Estimates are weighted and standard errors adjusted for the complex design of the Health and Retirement Study. ***,**,* indicates statistical significance at the 1%, 5% and 10% level, respectively.

Figure 1. Pathways in the GWAS-by-subtraction genomic SEM.

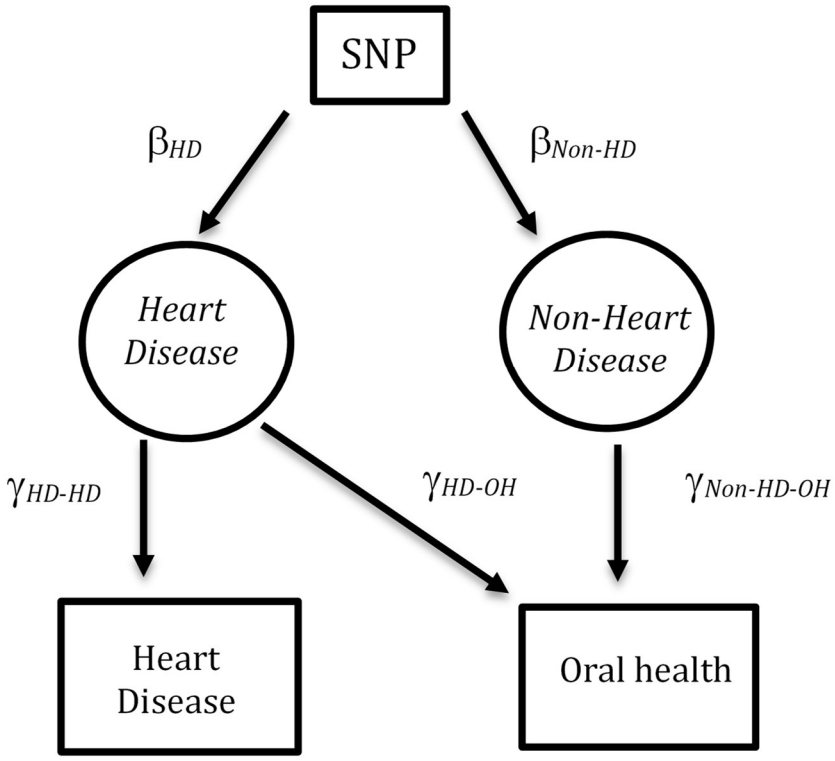
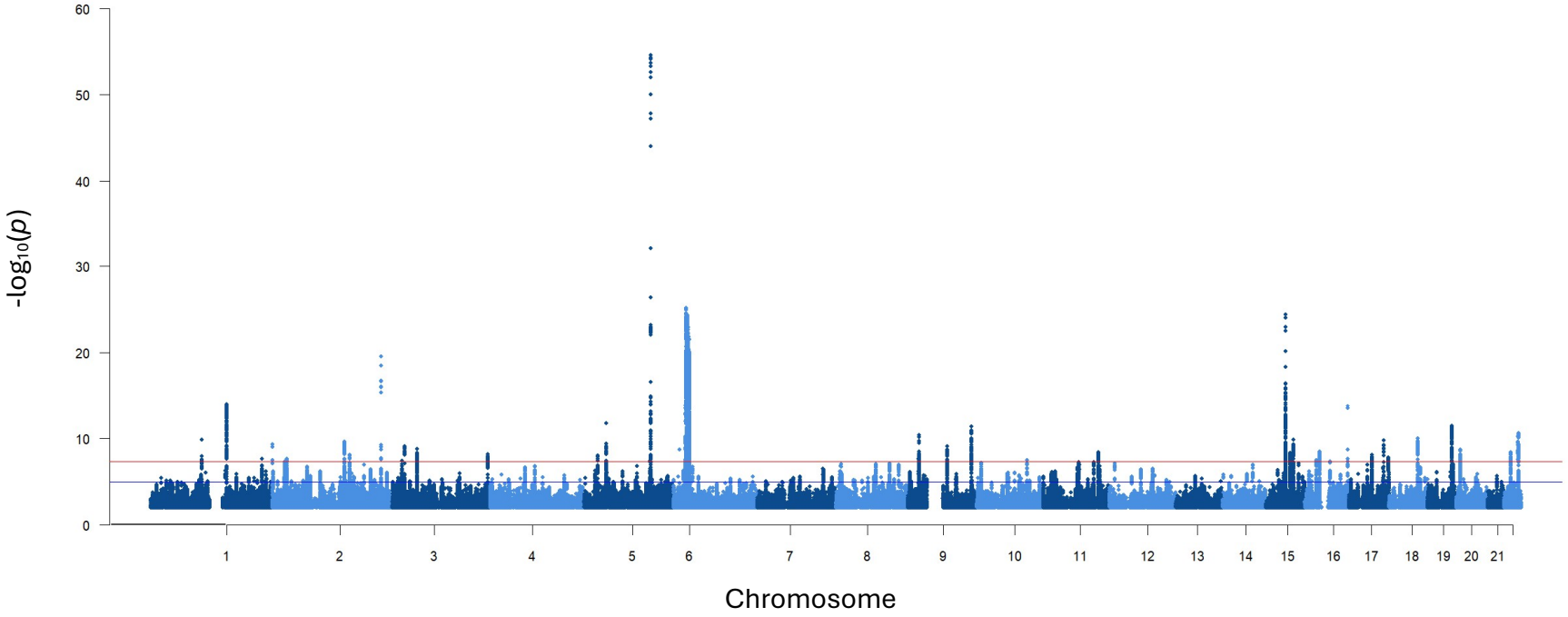


Figure 2. Manhattan plot of SNP associations with dental caries, net of associations with heart disease.



Notes: The red line denotes the threshold for genome-wide significance ($p = 5 \times 10^{-8}$) while the black line indicates suggestive threshold for association ($p = 1 \times 10^{-5}$).

Appendix tables

Table A1. Genomic SEM path loading estimates from Non-SNP model.

Path loading	Unstandardized estimate	Unstandardized estimate S.E.	Standardized estimate	Standardized estimate S.E.	p-value
$\gamma_{HD\text{-}HD}$	0.203	0.006	1.000	0.031	7.476×10^{-228}
$\gamma_{HD\text{-}OH}$	0.052	0.006	0.225	0.025	1.176×10^{-18}
$\gamma_{Non\text{-}HD\text{-}OD}$	0.224	0.004	0.974	0.019	$< 5 \times 10^{-300}$

Table A2. Genetic path correlations for SNPs with the largest associations with oral health (N=7,678,811).

SNP	CHR	BP	MAF	A1	A2	$\gamma_{Non\text{-}HD\text{-}OH}$	S.E.($\gamma_{Non\text{-}HD\text{-}OH}$)	Z-est.	p-value
rs1122171	5	134509987	0.429	C	T	-0.1605	0.0102	-15.66	2.875×10^{-55}
rs3749751	5	134509677	0.431	C	T	-0.1599	0.0102	-15.62	5.514×10^{-55}
rs10479093	5	134508559	0.431	G	A	-0.1597	0.0102	-15.60	7.647×10^{-55}
rs7729156	5	134507139	0.420	T	C	-0.1603	0.0103	-15.59	8.271×10^{-55}
rs10038879	5	134507859	0.427	C	T	-0.1591	0.0102	-15.53	2.028×10^{-54}

Notes: CHR = chromosome; BP = base pair; A1 = effect allele; A2 = non-effect allele; MAF = minor allele frequency.

Table A3. Complete OLS, first stage and second stage IV regression coefficients for "Dr. ever told you of heart problem [heart attack, coronary heart disease, angina, congestive heart failure or other heart problems]".

Control variables	OLS	First stage of IV	Second stage of IV
PGS IV (p< 0.001)		0.0160***	
		(0.0032)	
Lost all permanent teeth	0.0685***		0.4832***
	(0.0115)		(0.2474)
Dentate status missing	0.0378***	-0.2871***	0.1571***
	0.0685***	(0.0049)	(0.0710)
Age < 65	-0.1423***	-0.1072***	-0.0982***
	(0.0171)	(0.0167)	(0.0345)
65 ≤ Age < 70	-0.1450***	-0.0671***	-0.1177***
	(0.0128)	(0.0113)	(0.0240)
70 ≤ Age < 75	-0.1030***	-0.0460***	-0.0845***
	(0.0119)	(0.0106)	(0.0180)
75 ≤ Age < 80	-0.0685***	-0.0374***	-0.0533***
	(0.0108)	(0.0089)	(0.0164)
80 ≤ Age < 85	-0.0271***	-0.0205***	-0.0188
	(0.0097)	(0.0076)	(0.0124)
Hispanic	-0.0821***	-0.0857***	-0.0464***
	(0.0114)	(0.0136)	(0.0215)
Black non-Hispanic	-0.0380***	0.0021	-0.0389***
	(0.0109)	(0.0118)	(0.0105)
Race/ethnicity missing	-0.1155***	-0.0271	-0.1021***
	(0.0290)	(0.0331)	(0.0335)
Female	-0.0716***	-0.0162***	-0.0649***
	(0.0069)	(0.0054)	(0.0082)
Married or partnered	0.1549	-0.0646	0.1815
	(0.1528)	(0.1237)	(0.1651)
Marital status missing	0.0211	0.0465	0.0011
	(0.0782)	(0.0587)	(0.0894)
No. of HH residents	0.0049*	0.0042	0.0032
	(0.0028)	(0.0029)	(0.0030)
No. of living children	0.0008	0.0084***	-0.0028
	(0.0015)	(0.0012)	(0.0026)
Lives w/ spouse/partner	-0.1474	0.0403	-0.1639
	(0.1540)	(0.1235)	(0.1663)
Not working & retired	0.0904***	0.0237	0.0807***
	(0.0140)	(0.0165)	(0.0167)
Not fully retired	0.0477***	0.0139	0.0420***
	(0.0166)	(0.0168)	(0.0195)
Working full time	0.0199*	0.0145	0.0139
	(0.0113)	(0.0171)	(0.0147)
Other work status	0.0326	0.0321	0.0196
	(0.0305)	(0.0296)	(0.0312)

Table A3, continued.

Control variables	OLS	First stage of IV	Second stage of IV
Not in labor force & not retired	0.0202	-0.0161	0.0265
	(0.0302)	(0.0262)	(0.0302)
New England resident	-0.0153	0.0037	-0.0165
	(0.0138)	(0.0141)	(0.0148)
Mid-Atlantic resident	-0.0041	-0.0093	-0.0000
	(0.0139)	(0.0125)	(0.0141)
West North-Central resident	-0.0132	-0.0129	-0.0076
	(0.0155)	(0.0171)	(0.0180)
South-Atlantic resident	-0.0044	0.0182*	-0.0119
	(0.0112)	(0.0100)	(0.0115)
East South-Central resident	0.0081	0.0366**	-0.0079
	(0.0150)	(0.0160)	(0.0188)
West South-Central resident	-0.0132	0.0062	-0.0158
	(0.0121)	(0.0151)	(0.0120)
Mountain resident	-0.0302*	0.0088	-0.0340***
	(0.0167)	(0.0120)	(0.0158)
Pacific resident	-0.0237*	-0.0017	-0.0226***
	(0.0120)	(0.0109)	(0.0131)
Other non-US resident	-0.0809	-0.0195	-0.0741
	(0.0513)	(0.0759)	(0.0587)
Census division missing	-0.1631***	-0.1421***	-0.1045***
	(0.0326)	(0.0351)	(0.0444)
HH income is 101% - 200% FPL	0.0048	0.0118	-0.0002
	(0.0104)	(0.0075)	(0.0111)
HH income is 201% - 400% FPL	0.0125	-0.0194**	0.0207
	(0.0105)	(0.0088)	(0.0129)
HH income > 400% FPL	-0.0085	-0.0490***	0.0120
	(0.0103)	(0.0086)	(0.0177)
Medicare	0.0602***	-0.0112	0.0649***
	(0.0125)	(0.0103)	(0.0121)
Medicaid	0.0824***	0.0605***	0.0572***
	(0.0102)	(0.0136)	(0.0181)
Private health insurance	0.0152*	-0.0085*	0.0187***
	(0.0083)	(0.0049)	(0.0092)
Other health insurance	0.0205**	-0.0135**	0.0260***
	(0.0090)	(0.0064)	(0.0105)
No ADLs	-0.0215	-0.0253**	-0.0110
	(0.0179)	(0.0109)	(0.0202)
No IADLs	-0.0649***	-0.0334***	-0.0508***
	(0.0115)	(0.0080)	(0.0156)
No. of ADLs	0.0364***	0.0078*	0.0330***
	(0.0063)	(0.0041)	(0.0071)
No. of IADLs	-0.0023	0.0115**	-0.0070
	(0.0054)	(0.0044)	(0.0060)
Less than high school education	-0.0197	0.0914***	-0.0572***
	(0.0125)	(0.0112)	(0.0231)

Table A3, continued.

Control variables	OLS	First stage of IV	Second stage of IV
Some college	-0.0100	-0.0544***	0.0128
	(0.0087)	(0.0067)	(0.0135)
College degree or higher	-0.0225**	-0.0958***	0.0177
	(0.0106)	(0.0065)	(0.0244)
Education is missing	-0.2857***	0.2771	-0.4022***
	(0.0322)	(0.3543)	(0.1306)
Mother had at least high school education	-0.0199*	-0.0070	-0.0173
	(0.0116)	(0.0066)	(0.0118)
Father had at least high school education	-0.0180	-0.0127**	-0.0127
	(0.0143)	(0.0062)	(0.0143)
Constant	0.3378***	0.3027***	0.1245
	(0.0299)	(0.0350)	(0.1454)
F-statistic		24.5	

Notes: Models contain wave fixed effects. Estimates are weighted and standard errors adjusted for the complex design of the Health and Retirement Study. ***, **,* indicates statistical significance at the 1%, 5% and 10% level, respectively.

Table A4. Complete OLS, first stage and second stage IV regression coefficients for “Heart attack in the past 2 years”.

Control variables	OLS	First stage of IV	Second stage of IV
PGS IV (p≤ 1)		0.0237***	0.0385***
		(0.0029)	(0.0204)
Lost all permanent teeth	0.0066***		
	(0.0016)		
Dentate status missing	0.0055***	-0.2867***	0.0147***
	(0.0017)	(0.0049)	(0.0062)
Age < 65	0.0027	-0.1106***	0.0061
	(0.0032)	(0.0168)	(0.0040)
65 ≤ Age < 70	-0.0004	-0.0702***	0.0017
	(0.0024)	(0.0113)	(0.0028)
70 ≤ Age < 75	-0.0014	-0.0487***	0.0001
	(0.0025)	(0.0106)	(0.0027)
75 ≤ Age < 80	-0.0015	-0.0395***	-0.0003
	(0.0023)	(0.0089)	(0.0025)
80 ≤ Age < 85	-0.0019	-0.0215***	-0.0012
	(0.0026)	(0.0076)	(0.0026)
Hispanic	-0.0045**	-0.0823***	-0.0018
	(0.0018)	(0.0132)	(0.0025)
Black non-Hispanic	-0.0029**	0.0044	-0.0030***
	(0.0013)	(0.0118)	(0.0014)
Race/ethnicity missing	-0.0030	-0.0252	-0.0019
	(0.0057)	(0.0346)	(0.0059)
Female	-0.0079***	-0.0161***	-0.0073***
	(0.0010)	(0.0053)	(0.0010)
Married or partnered	0.0174*	-0.0663	0.0195***
	(0.0088)	(0.1221)	(0.0098)
Marital status missing	0.0014	0.0467	-0.0001
	(0.0115)	(0.0589)	(0.0117)
No. of HH residents	-0.0000	0.0039	-0.0001
	(0.0004)	(0.0028)	(0.0005)
No. of living children	0.0001	0.0081***	-0.0001
	(0.0002)	(0.0012)	(0.0003)
Lives w/ spouse/partner	-0.0160*	0.0425	-0.0172***
	(0.0086)	(0.1220)	(0.0096)
Not working & retired	0.0080***	0.0232	0.0072***
	(0.0015)	(0.0165)	(0.0015)
Not fully retired	0.0038**	0.0141	0.0034***
	(0.0016)	(0.0167)	(0.0017)
Working full time	0.0039***	0.0151	0.0034***
	(0.0013)	(0.0170)	(0.0014)
Other work status	0.0062	0.0326	0.0052
	(0.0053)	(0.0301)	(0.0053)

Table A4, continued.

Control variables	OLS	First stage of IV	Second stage of IV
Not in labor force & not retired	-0.0014	-0.0169	-0.0009
	(0.0054)	(0.0266)	(0.0053)
New England resident	-0.0008	0.0035	-0.0009
	(0.0027)	(0.0142)	(0.0026)
Mid-Atlantic resident	-0.0016	-0.0091	-0.0013
	(0.0018)	(0.0123)	(0.0017)
West North-Central resident	0.0010	-0.0124	0.0014
	(0.0015)	(0.0167)	(0.0015)
South-Atlantic resident	-0.0015	0.0168*	-0.0021
	(0.0013)	(0.0100)	(0.0015)
East South-Central resident	0.0033*	0.0328**	0.0021
	(0.0017)	(0.0159)	(0.0020)
West South-Central resident	-0.0010	0.0031	-0.0012
	(0.0017)	(0.0150)	(0.0019)
Mountain resident	0.0002	0.0069	-0.0001
	(0.0017)	(0.0116)	(0.0017)
Pacific resident	0.0015	-0.0017	0.0016
	(0.0021)	(0.0108)	(0.0022)
Other non-US resident	-0.0049**	-0.0224	-0.0044
	(0.0023)	(0.0794)	(0.0034)
Census division missing	-0.0081**	-0.1353***	-0.0036
	(0.0039)	(0.0356)	(0.0039)
HH income is 101% - 200% FPL	-0.0007	0.0118	-0.0011
	(0.0019)	(0.0076)	(0.0019)
HH income is 201% - 400% FPL	-0.0021	-0.0190**	-0.0015
	(0.0018)	(0.0088)	(0.0019)
HH income > 400% FPL	-0.0042**	-0.0485***	-0.0026
	(0.0017)	(0.0087)	(0.0020)
Medicare	0.0047	-0.0107	0.0051***
	(0.0029)	(0.0104)	(0.0028)
Medicaid	0.0069**	0.0603***	0.0050
	(0.0029)	(0.0137)	(0.0033)
Private health insurance	-0.0000	-0.0088*	0.0002
	(0.0011)	(0.0049)	(0.0011)
Other health insurance	-0.0020*	-0.0133**	-0.0016
	(0.0010)	(0.0063)	(0.0011)
No ADLs	-0.0085*	-0.0250**	-0.0077***
	(0.0043)	(0.0110)	(0.0045)
No IADLs	-0.0067*	-0.0332***	-0.0056
	(0.0035)	(0.0079)	(0.0035)
No. of ADLs	0.0024	0.0079*	0.0021
	(0.0015)	(0.0042)	(0.0015)
No. of IADLs	0.0009	0.0115**	0.0006
	(0.0017)	(0.0044)	(0.0016)
Less than high school education	-0.0011	0.0891***	-0.0040
	(0.0020)	(0.0111)	(0.0024)

Table A4, continued.

Control variables	OLS	First stage of IV	Second stage of IV
Some college	-0.0022*	-0.0534***	-0.0005
	(0.0012)	(0.0067)	(0.0017)
College degree or higher	-0.0037***	-0.0925***	-0.0006
	(0.0012)	(0.0065)	(0.0026)
Education is missing	-0.0156***	0.2890	-0.0245
	(0.0032)	(0.3431)	(0.0159)
Mother had at least high school education	-0.0004	-0.0070	-0.0002
	(0.0012)	(0.0066)	(0.0013)
Father had at least high school education	-0.0009	-0.0123*	-0.0005
	(0.0014)	(0.0063)	(0.0013)
Constant	0.0282***	0.3057***	0.0035
	(0.0067)	(0.0351)	(0.0136)
F-statistic		64.8	

Notes: Models contain wave fixed effects. Estimates are weighted and standard errors adjusted for the complex design of the Health and Retirement Study. ***,**,* indicates statistical significance at the 1%, 5% and 10% level, respectively.

Table A5. Balance test of PGS instrument with $p < 0.001$.

Variable	IV < $\bar{x}$	IV > $\bar{x}$	Diff.	p-value
Age < 65	0.456	0.455	-0.001	0.940
65 ≤ Age < 70	0.173	0.173	0.000	0.959
70 ≤ Age < 75	0.134	0.140	0.006	0.061
75 ≤ Age < 80	0.101	0.101	0.001	0.747
80 ≤ Age < 85	0.072	0.072	0.000	0.879
Hispanic	0.077	0.104	0.026	0.000
Black non-Hispanic	0.069	0.125	0.057	0.000
Race/ethnicity missing	0.001	0.001	0.000	0.645
Female	0.547	0.566	0.019	0.072
Married or partnered	0.675	0.646	-0.029	0.007
Marital status missing	0.001	0.001	0.000	0.217
No. of HH residents	2.182	2.217	0.035	0.041
No. of living children	2.800	2.971	0.171	0.000
Lives w/ spouse/partner	2.800	2.971	0.171	0.000
Not in labor force & retired	0.495	0.499	0.004	0.611
Not fully retired	0.110	0.110	0.000	0.983
Working full time	0.317	0.306	-0.012	0.109
Other work status	0.067	0.075	0.008	0.015
Not in labor force & not retired	0.051	0.059	0.008	0.008
New England resident	0.049	0.044	-0.005	0.515
Mid-Atlantic resident	0.126	0.117	-0.009	0.242
West-North-Central resident	0.093	0.076	-0.017	0.069
South-Atlantic resident	0.207	0.219	0.013	0.191
East-South-Central resident	0.057	0.074	0.017	0.000
West-South-Central resident	0.086	0.105	0.019	0.001
Mountain resident	0.070	0.069	-0.002	0.759
Pacific resident	0.135	0.124	-0.010	0.053
Other non-US resident	0.000	0.000	0.000	0.676
Census division missing	0.000	0.000	0.000	0.594
HH income is 101% - 200% FPL	0.144	0.164	0.021	0.000
HH income is 201% - 400% FPL	0.283	0.284	0.002	0.790
HH income > 400% FPL	0.501	0.458	-0.042	0.000
Medicare	0.560	0.567	0.007	0.363
Medicaid	0.059	0.082	0.022	0.000
Private health insurance	0.517	0.488	-0.028	0.000
Other health insurance	0.155	0.148	-0.007	0.248
No ADLs	0.945	0.935	-0.010	0.001
No IADLs	0.911	0.896	-0.016	0.000
No. of ADLs	0.247	0.299	0.052	0.000
No. of IADLs	0.213	0.256	0.043	0.000
Less than high school education	0.118	0.142	0.024	0.000
Some college	0.260	0.263	0.004	0.674
College degree or higher	0.303	0.260	-0.043	0.000
Education is missing	0.000	0.000	0.000	0.161

Table A5, continued.

Variable	IV < $\bar{x}$	IV > $\bar{x}$	Diff.	p-value
Mother had at least high school education	0.175	0.161	-0.013	0.079
Father had at least high school education	0.191	0.171	-0.020	0.021

Notes: Means are weighted and standard errors adjusted for the complex design of the Health and Retirement Study.

Table A6. Balance test of PGS instrument with $p \leq 1$.

Variable	IV < $\bar{x}$	IV > $\bar{x}$	Diff.	p-value
Age < 65	0.451	0.460	0.009	0.258
65 ≤ Age < 70	0.171	0.175	0.004	0.341
70 ≤ Age < 75	0.134	0.139	0.005	0.049
75 ≤ Age < 80	0.102	0.100	-0.002	0.333
80 ≤ Age < 85	0.075	0.069	-0.006	0.023
Hispanic	0.086	0.089	0.002	0.632
Black non-Hispanic	0.082	0.100	0.018	0.000
Race/ethnicity missing	0.001	0.001	0.000	0.831
Female	0.550	0.559	0.009	0.317
Married or partnered	0.671	0.656	-0.015	0.164
Marital status missing	0.001	0.001	0.000	0.771
No. of HH residents	2.174	2.218	0.044	0.027
No. of living children	2.774	2.968	0.194	0.000
Lives w/ spouse/partner	0.671	0.656	-0.015	0.163
Not in labor force & retired	0.491	0.502	0.010	0.194
Not fully retired	0.113	0.107	-0.006	0.180
Working full time	0.319	0.306	-0.013	0.163
Other work status	0.065	0.075	0.009	0.016
Not in labor force & not retired	0.050	0.059	0.009	0.003
New England resident	0.048	0.046	-0.002	0.646
Mid-Atlantic resident	0.125	0.120	-0.005	0.341
West-North-Central resident	0.096	0.076	-0.020	0.027
South-Atlantic resident	0.208	0.215	0.007	0.454
East-South-Central resident	0.054	0.075	0.021	0.003
West-South-Central resident	0.085	0.103	0.018	0.022
Mountain resident	0.068	0.072	0.004	0.492
Pacific resident	0.137	0.124	-0.014	0.061
Other non-US resident	0.000	0.000	0.000	0.913
Census division missing	0.000	0.000	0.000	0.759
HH income is 101% - 200% FPL	0.140	0.164	0.024	0.000
HH income is 201% - 400% FPL	0.285	0.282	-0.003	0.555
HH income > 400% FPL	0.502	0.465	-0.038	0.000
Medicare	0.564	0.562	-0.003	0.766
Medicaid	0.059	0.077	0.018	0.000
Private health insurance	0.515	0.496	-0.019	0.022
Other health insurance	0.156	0.149	-0.007	0.133
No ADLs	0.945	0.937	-0.008	0.011
No IADLs	0.911	0.899	-0.012	0.006
No. of ADLs	0.251	0.284	0.033	0.005
No. of IADLs	0.215	0.245	0.030	0.004
Less than high school education	0.110	0.146	0.036	0.000
Some college	0.264	0.258	-0.007	0.430
College degree or higher	0.317	0.253	-0.064	0.000
Education is missing	0.000	0.000	0.000	0.862

Table A6, continued.

Variable	IV < $\bar{x}$	IV > $\bar{x}$	Diff.	p-value
Mother had at least high school education	0.183	0.154	-0.029	0.001
Father had at least high school education	0.194	0.171	-0.023	0.008

Notes: Means are weighted and standard errors adjusted for the complex design of the Health and Retirement Study.