DIETARY PATTERNS, METABOLIC MEDIATORS AND RACIAL DISPARITIES IN MORTALITY IN THE UNITED STATES: POOLED INDIVIDUAL-LEVEL ANALYSES OF SEVEN PROSPECTIVE COHORTS, 1970-2016

Suman Chakrabarti

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2022

Reading Committee

Joseph Dieleman, Chair

Ashkan Afshin

Jessica Jones-Smith

Program Authorized to Offer Degree:

Global Health

UNIVERSITY OF WASHINGTON

ABSTRACT

Dietary patterns, metabolic mediators and racial disparities in all-cause mortality in the United States: pooled individual-level analyses of seven prospective cohorts, 1970-2016

Suman Chakrabarti

Chair of the Supervisory Committee:

Associate Professor Joseph Dieleman, PhD

Health Metrics Sciences

Suboptimal diets are important preventable risk factors for non-communicable diseases and mortality. Broadly, evidence from a plethora of systematic reviews suggests that food groups including whole grains, fibrous vegetables, fruits, nuts, legumes, red-meat, and processed meat, are linked to all-cause mortality. Despite the high quality evidence that has emerged from meta-analyses, high levels of heterogeneity reported across most systematic reviews presents challenges for rigorous estimation of average effects. It is also well established that suboptimal diets are important risk factors for high body mass index and elevated systolic blood pressure, which are leading metabolic risk factors for mortality. Moreover, evidence indicates that dietary patterns and mortality rates differ systematically across racial/ethnic groups in the US. Making progress towards reducing risks of mortality that stem from suboptimal diets requires a systematic and sequential approach that leverages learnings in one area as inputs to shed light on other emerging areas. Unpacking the interlinkages between dietary patterns, biological mediators and social factors will enable a holistic line of attack that will empower clinicians, researchers, and policy makers to take effective steps to address alarming emerging trends in morbidity and mortality in this century.

To address these research priorities, this study sought to pool secondary individual-level data from seven US based cohorts including Atherosclerosis Risk in Communities Study, Cardiovascular Health Study, Framingham Heart Study-Offspring, Multi-Ethnic Study of Atherosclerosis, Women's Health Initiative Study, NHANES I Epidemiologic Follow-up Study, Reasons for Geographic and Racial Differences in Stroke. Pooling several cohorts supported a systematic and coherent analysis that consistently accounted for potential confounders and addressed differences in exposure measurement. Across the cohorts, interviews were conducted in-person at regular examinations, follow-up telephone calls to participants or next of kin, state records, linkage to the National Death Index, surveillance of medical records, death certificates, obituaries, Centers for Medicare and Medicaid Services. A total of 127,279 individuals were followed-up for 1,579,299 person-years in the seven cohorts combined.

Principal Component Analysis of observed intake of food groups was used to estimate dietary patterns. Five dietary patterns were estimated. The ‘Western’ pattern was high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The ‘Prudent’ pattern was high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The ‘Fatty-grainy’ pattern was high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The ‘Lacto’ pattern was high in milk and moderate to low on all other food groups. The ‘Carnivore’ pattern was high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Dietary patterns were subsequently employed in estimating associations of dietary patterns with all-cause mortality, unpacking direct and indirect relationships with metabolic mediators, and explaining differences in mortality between racial/ethnic groups. First, the risk to all-cause mortality associated with commonly occurring dietary patterns in the US was estimated using Cox proportional hazard models. Second, the direct and indirect effects of dietary patterns on all-cause mortality through body mass index and systolic blood pressure were disentangled using methods including the percentage of excess risk mediated and percent mediated measure. Third, the roles of dietary patterns and other behavioral risk factors to close racial/ethnic differences in all-cause mortality in the US were decomposed using the Oaxaca-Blinder decomposition technique and population attributable fractions.

In seven large prospective cohorts with 8.5 to 36.7 median years of follow-up, greater adherence to Western and Carnivore patterns was consistently associated with higher risk of early mortality, while moderation eating patterns and those rich in plant foods lowered this risk. These findings support the recommendations of the Dietary Guidelines for Americans that multiple healthy eating patterns can be adapted to individual food traditions and preferences. Further, while interventions that reduce BMI and blood pressure, might address up to a third of the mortality risk associated attributed to suboptimal diets, maintenance of optimal diets is needed for complete health benefits. Finally, replacement of unhealthy diets with healthy ones could yield significant equitable improvements in mortality rates across all racial/ethnic groups in the US. Income disparities currently explain half the mortality differences between White and Black individuals in the US. Holistic improvements in income, physical activity, smoking and drinking are needed across all race/ethnic groups for further equitable reductions in mortality rates.

This study received ethical approval from the University of Washington IRB under Mod 8, HSD #46665.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I have many to thank for their support throughout my time in the PhD program. First, I thank IHME for endowing me with a scholarship that provided the financial support I needed to pursue higher education in the US. Thank you to the UW community and my fellow PhD cohort members for providing a welcoming environment to international students.

I would like to extend my heartfelt gratitude to my dissertation committee members, Joseph Dieleman, Ashkan Afshin, Jessica-Jones Smith, Aleksandr Aravkin, Peng Zheng and Sara Curran, of all whom provided tremendous support for my dissertation along the way. As the chair of my dissertation committee and my first supervisor, Joe supported me, gave me guidance and helped me during the most challenging times and I am beyond grateful for his support and kindness. Ashkan, my second supervisor at IHME, provided me with this research opportunity, and mentored me through the process of identifying my dissertation aims early in my PhD journey, for all of which I am extremely grateful. Jesse agreed to be on my committee during a very crucial period and provided in-depth knowledge and guidance for my dissertation research, for which I am very thankful. Thank you to Sasha and Peng for providing valuable inputs for this dissertation and for bringing positive energy in all our interactions. And thank you to Sara for teaching a wonderful class in demography and also for serving as the GSR on my committee.

Thank you to my friends and family, in India and the US. Thank you to my in-laws for always supporting me and never failing to cheer me on. Thank you, Sam, Colin, Pravati, Andrew, Vanessa and Leigh for being good friends and always being there to talk academia and more. Thank you Manu, Ankur, Aziz, Varun, Harsh, Apurv, Roopali, among others. I hope to enjoy many more hikes, board games, movies, dinners, vacations and happy hours in the future.

I thank my father for instilling a love for science in me that has shaped my career and curiosity. I thank Max, my departed best friend with whom I shared the fondest memories of my life. I have many women to thank. First, I thank my mother, who supported two children by herself after my father passed, loved and cared for us, and provided sustenance in endless ways. I thank my aunt who has been a second mother to me, being a lighthouse in a storm that kept us going in many difficult times. I thank my sister, who shares my love for science and has led by example, being an excellent student, sister, daughter, wife and mother. I thank my mentor, Purnima, who paved the way for me to pursue a career in public health. I am eternally grateful to you all.

Last but not least, I thank Anushree, my wife. None of this would be possible without her. We moved to Seattle together when I started this journey. She has been a pillar of unrelenting support and endless joy. All good things have come to me since I met her. Her smile, warmth, positivity and love keep me going.

GLOSSARY

Acronym	Full form
ACM	All-cause mortality
ARIC	Atherosclerosis Risk in Communities Study
BMI	Body mass index
CHD	Coronary heart disease
CHS	Cardiovascular Health Study
CVD	Cardiovascular disease
DP	Dietary pattern
FHSO	Framingham Heart Study-Offspring
MESA	Multi-Ethnic Study of Atherosclerosis
NFS	National Health and Nutrition Examination Follow-up Survey
OBD	Oaxaca-Blinder decomposition
PCA	Principal component analysis
PERM	Percentage of excess risk mediated
PMM	Percent mediated measure
REGARDS	Reasons for Geographic and Racial Differences in Stroke
SBP	Systolic Blood pressure
SEM	Structural equation modeling
WHI	Women's Health Initiative Study
US	United States

CHAPTER 1:

INTRODUCTION

Over half of all adults in the Unites States (US), approximately 120 million individuals, suffer from one or more chronic non-communicable diseases[1], for which suboptimal diets are important preventable risk factors[2]. In the past decade, numerous meta-analytical studies have systematically compiled and quantified summative evidence from a plethora of cohort studies relating dietary risk to all-cause mortality[3–22]. It is also well established that suboptimal diets are important risk factors for high body mass index (BMI) and elevated blood pressure (BP), which are leading metabolic risk factors for mortality[2,23]. The COVID-19 pandemic has further highlighted systemic inequity as a condition of social disadvantage and a public health issue[24], where evidence indicates that dietary patterns and mortality rates differ systematically across racial/ethnic groups in the US[25]. Despite the high quality evidence that has emerged from meta-analyses, high levels of heterogeneity reported across most systematic reviews presents challenges for rigorous estimation of pooled average effects[26] (**Table 11**). A review of systematic-reviews emphasizes the difficulty of detecting definitive associations in food group based nutrition, as demonstrated by a high degree of heterogeneity and lack of additivity of risk estimates[27]. Additionally, little is known about the risk from dietary patterns that is mediated by metabolic risk factors; let alone differences in mortality between racial/ethnic groups that can be explained by dietary differences.

Table 1 1 Summary of risk ratios and heterogeneity observed in systematic reviews of food groups

Dietary factor	No of studies	RR	95% CI	Heterogeneity I^2 (%)
Whole grains	11	0.92	0.89, 0.95	80
Refined grains	4	0.99	0.97, 1.01	7
Vegetables	17	0.96	0.95, 0.98	67
Fruits	17	0.94	0.92, 0.97	82
Nuts	16	0.76	0.69, 0.84	82
Legumes	6	0.96	0.90, 1.01	48
Eggs	5	1.15	0.99, 1.34	87
Dairy	16	0.98	0.93, 1.03	96
Fish	19	0.93	0.88, 0.98	53
Red meat	10	1.1	1.04, 1.18	92
Processed meat	7	1.23	1.12, 1.36	94

Notes: Adapted from Schwingshackl et al (2017)[26].

To address these research priorities, this study sought to pool individual-level data from several US based cohorts which allows for a systematic and coherent analysis that consistently accounts for potential confounders and addresses differences in exposure measurement[28,29]. Using these pooled data, commonly occurring dietary patterns were to be identified, and subsequently employed in estimating associations of diets with all-cause mortality, unpacking direct and indirect relationships with metabolic mediators, and explaining differences in mortality between racial/ethnic groups. The practicability of this study was supported by the fact that long-term randomized trials of mortality are not feasible for dietary risk factors; consequently supporting evidence for potential causal relationships must rely on prospective observational

cohort studies. Moreover, data-driven approaches have become increasingly popular in nutritional science to classify diets into patterns that have to been shown to have high stability[30].

I hypothesized that dietary patterns have significant explanatory power to advance our understanding of biological and social mechanisms that determine mortality. The pooled cohort approach addressed crucial evidence gaps and offered high quality evidence by addressing the risk of bias, precision, heterogeneity, and dose-response relationships[31]. Examining the role of healthy dietary patterns for improving life expectancy, reducing risk of chronic metabolic disease and addressing social disparities, are high-priority subject areas for researchers and practitioners[32].

Specifically, I pursued three specific aims:

Aim 1: Constructed *a posteriori* dietary patterns and examined associations with all-cause mortality

Aim 2: Disentangled the direct and indirect effects of dietary patterns on all-cause mortality through metabolic mediators

Aim 3: Decomposed the role of dietary patterns to close racial/ethnic differences in all-cause mortality

Making progress towards reducing risks of mortality that stem from suboptimal diets requires a systematic and sequential approach that leverages learnings in one area as inputs to shed light on other emerging areas. Unpacking the interlinkages between dietary patterns, biological mediators and social factors will enable a holistic line of attack that will empower clinicians, researchers, and policy makers to take effective steps to address alarming emerging trends in morbidity and mortality in this century[33].

CHAPTER 2:
THE EFFECTS OF DIETARY PATTERNS ON MORTALITY AND MORBIDITY IN THE UNITED STATES: A POOLED ANALYSIS OF SEVEN PROSPECTIVE COHORTS

ABSTRACT

Objectives: To quantify the risk to all-cause mortality associated with commonly occurring diet patterns in the US.

Participants and setting: Secondary analysis of pooled datasets from seven prospective cohorts including Atherosclerosis Risk in Communities Study, Cardiovascular Health Study, Framingham Heart Study-Offspring, Multi-Ethnic Study of Atherosclerosis, Women's Health Initiative Study, NHANES I Epidemiologic Follow-up Study, Reasons for Geographic and Racial Differences in Stroke. Across the cohorts, interviews were conducted in-person at regular examinations, follow-up telephone calls to participants or next of kin, state records, linkage to the National Death Index, surveillance of medical records, death certificates, obituaries, Centers for Medicare and Medicaid Services. 127,279 individuals were followed-up for 1,579,299 person-years.

Outcomes: Outcomes were 18388 deaths (binary) and the survival time until death/censoring (interval between start of the study and death or loss to follow-up).

Exposures: Principal Component Analysis of observed intake of food groups was used to estimate DPs coded as continuous variables ranging from 0 (low) to 1 (high). Five DPs were estimated. The 'Western' pattern was high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' pattern was high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' pattern was high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto' pattern was high in milk and moderate to low on all other food groups. The 'Carnivore' pattern was high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Design: Cox-Proportional Hazard models were used to examine associations between the outcome and exposures. Regressions were adjusted for age, sex, race, education, income, smoking, drinking, physical activity, birth-cohort fixed effects and survey (cohort) fixed effects. Spline models were used to investigate non-linear dose-response relationships. Sex specific regressions were used to test effect modification.

Results: Pooled fully adjusted HRs in the linear dose-response models were Western 1.12 [95%CI 1.05, 1.19], Prudent 0.89 [95%CI 0.85, 0.95], Fatty-grainy 0.95 [95%CI 0.90, 1.01], Lacto-mod 0.89 [95%CI 0.83, 0.95] and Carnivore 1.21 [95%CI 1.10, 1.33]. For Western and Lacto-mod, the linear model was a good approximation for the risk curve. For Prudent, Fatty-Grainy and Carnivore, there was a clear non-linear dose-response function. Females were more sensitive to the effects of Western and Prudent patterns.

Conclusions: In seven large prospective cohorts with 8.5 to 36.7 median years of follow-up, greater adherence to Western and Carnivore patterns high in meat, was consistently associated with higher risk of early mortality, while moderate eating patterns and those rich in plant foods lowered this risk. These findings support the recommendations of the Dietary Guidelines for Americans that multiple healthy eating patterns can be adapted to individual food traditions and preferences.

INTRODUCTION

Suboptimal diets are important preventable risk factors for non-communicable diseases and mortality[2]. Broadly, evidence from a plethora of systematic reviews suggests that food groups such as whole grains[7,16,17], refined grains[26], starchy vegetables, non-starchy vegetables[3,34], fruits[3,34], nuts[4], legumes[26], eggs[26], milk and dairy[18,21], fish[13], red-meat[14], processed meat[14], and white meat[35], are linked to all-cause mortality. In general, plant-based food groups appear to offer greater protection against early mortality[36–45]. Despite the high quality evidence that has emerged from meta-analyses, high levels of heterogeneity reported across most systematic reviews presents challenges for rigorous estimation of average effects[26]. Specifically, inconsistent findings across studies may stem from differences in follow-up periods, heterogeneity in control for confounding factors, inconsistency in precision due to a divergence in the number of mortality cases, age of the participants, and length of the study. A review of systematic-reviews emphasizes the difficulty of detecting definitive associations in food group based nutrition, as demonstrated by a high degree of heterogeneity[27].

A major criticism of examining food groups or nutrients in isolation is that dietary components are often consumed together and are interactive, adding dimensionality and bias to estimates for single food groups[46]. Quantitative estimates of average causal effects of food groups are suspect because in the majority of studies, each food group is studied in isolation from other food groups. This likely stems from statistical considerations such as a high degree of collinearity between food groups that are often consumed together. If all the meta-estimates of individual food groups were causal, then consuming all of them in the recommended dietary levels would effectively eliminate all the risk of premature mortality from specific causes[27].

Gradually, focus has moved away food groups towards overall dietary patterns (DPs)[47]. The DGAC of the United States Department of Agriculture (USDA) defines DPs as the quantities, proportions, variety or combinations of different foods and beverages in diets, and the frequency with which they are habitually consumed[1]. DPs usually connote diet quality in terms of balance, variety and moderation[48]. For example, a priori scales, such as Healthy Eating Index (HEI), the Diet Quality Index (DQI), the Healthy Diet Indicator (HDI) and the Mediterranean Diet Score (MDS)[12,20,49], place individuals into risk categories by calculating a single score generated from different predetermined components[50].

Over the past decade, data-driven approaches have recurrently been used in nutritional science to classify diets into patterns. These *a posteriori* methods generally employ statistical techniques such as exploratory factor analysis, principal components analysis (PCA), or cluster analysis to examine to approximate diet types[32] from data on individual foods or food groups. This approach offers the advantage of identifying commonly occurring dietary behavior in a population, which reflect realistic representations of eating choices[51].

The objective of this study was to quantify the risk to all-cause mortality associated with commonly occurring (*a posterior* defined) DPs in the US.

METHODS

Data

The cohorts included in the analysis are Atherosclerosis Risk in Communities Study (ARIC), Cardiovascular Health Study (CHS), Framingham Heart Study-Offspring (FHSO), Multi-Ethnic Study of Atherosclerosis (MESA), Women's Health Initiative Study (WHI), NHANES I Epidemiologic Follow-up Study, Reasons for Geographic and Racial Differences in Stroke (REGARDS) (Table A1).

Outcome Assessment

The primary outcome of interest for this study is all-cause mortality (deaths that occur in a population, regardless of the cause). Secondary outcomes of interest include diabetes, cardiovascular disease and cancer incidence, depending on cohort specific data availability. Participants were followed-up from baseline to the last round of examination, event of death or when they could be contacted for follow-up. A participant's time since baseline to death or last follow-up was a calculated in days. The final outcomes are event of a death (binary) and the survival time until death occurs (interval between start of the study and death or loss to follow-up). For those patients that are alive at the end of the observation period, the actual time-to-event is unknown, and will be treated as a case of right censored data. In ARIC, all-cause mortality was identified through annual (semi-annual since 2012) follow-up telephone calls to participants or their proxies, state records and linkage to the National Death Index (NDI)[52,53]. In CHS, all-cause mortality was identified by surveillance of medical records, death certificates, obituaries, Centers for Medicare and Medicaid Services databases and from next-of-kin interviews[54]. The FHSO participants were followed for the event of death through regular examinations at the FHS clinic, surveillance of hospital admissions, and death registries[55]. MESA included all deaths identified by follow-up call-ins, medical record abstractions, obituaries or through the National Death Index[56,57]. In NHEFS, all-cause mortality was identified by death certificates, and next-of-kin interviews[58]. In REGARDS, family members of participants were interviewed by telephone every 6 months to log all death events[58]. Following methodology adopted in other research, to minimize the risks of reverse-causation bias, the analyses for all-cause mortality removes participants who died within the first 3 years of baseline[59]. Secondary outcomes include any occurrence of cardio-vascular disease (CVD) (including coronary heart disease), stroke, cancer, and diabetes events. Individuals who were diagnosed with these events at baseline are removed from the analysis on secondary outcomes.

Dietary pattern estimation

Details of the methods used for estimating intake of food groups from Food Frequency Questionnaires (FFQs) are in the methods appendix for this chapter. A posteriori DPs are approximated using principal component analysis (PCA) where data on estimated daily intake (g/ml) of several food groups are compressed into dietary pattern scores that reflect various diet types that are commonly consumed in the US. PCA derived dietary patterns have to been shown to have high stability, suggesting that they are well suited for a pooled cohort design[30]. The primary motivation behind this method is to de-correlate data, i.e. remove second-order dependencies. PCA achieves this via dimensionality reduction, where it estimates a set of correlated food groups with a new set of comprehensive indexes (principal components) that are uncorrelated and retain the largest amount variance in the original sample as possible[60]. In other words, principal components explain variance/covariance between a set of observed variables in terms of a set of fewer unobserved components and weightings. When deriving DPs, it is common practice to pre-group food items (say into food groups) before calculating principal components through the optimal weighted linear combination of food groups based on their correlation. The pre-group step aids in enhancing interpretability of the components. Food groups used in the analyses are fruits, starchy vegetables, fibrous vegetables, whole grain, refined grain, legumes, oils/fats, nuts and seeds, milk, cheese, yogurt, fish/shellfish, poultry, processed meats, eggs, and red meat.

PCA is represented in the vector of equations formulation as:

$$\boldsymbol{y} = \Lambda \boldsymbol{H} + \boldsymbol{E} \qquad \ldots (1)$$

Where $\mathbf{y}$, $\mathbf{H}$ and $\mathbf{E}$ are vectors and Λ is a matrix of weights. Here, $\mathbf{y}$ represents the dietary intake variables, $\mathbf{H}$ the underlying and unobserved DPs, Λ the weights that relate dietary patterns to intake variables and $\mathbf{E}$ the unique residual variances of $\mathbf{y}$. While computational aspects of the PCA algorithm are quite involved, they rely on three key statistical measures. First, the covariance matrix that measures how each variable is related to another. Positive/negative values indicate a positive/negative association, and zero values indicate independence. Second, Eigenvectors indicate the directions in which the data are distributed and third, Eigenvalues, indicate the relative importance of these different directions. We drop Eigenvectors with Eigenvalues less than 1. The Varimax rotation is used to achieve sparsity. Since PCA is sensitive to outliers, we Winsorize the input variables which involves replacing data values for food groups where individuals that have consumption levels greater than a say 95th percentile with the level of the 95th percentile. After predicting component scores, the scores are re-scaled ranging from 0 to 1, with 0 and 1 representing lowest and highest consumption levels of a specific pattern, respectively. While each participant consumes all food groups, going for low to high levels on specific DPs indicates higher consumption of food groups that are highly correlated with that DP. The scaling enables us to perform a low versus high analysis by using all the data points, and specifying a linear dose-response function. We perform PCA analysis separately for each cohort to check for consistencies in patterns identified. We later pool the diet scores for each pattern across cohorts for the regression analysis.

Covariates

Following the Scientific Report of the 2015 Dietary Guidelines Advisory Committee (DGAC), a comprehensive, evidence-based set of important confounders are included[1]. Demographic controls include age at baseline in years, sex (dummy for male), and race/ethnicity specified dummy variables for Non-Hispanic White and Black (with other race/ethnicities grouped into one reference group due to small sample sizes for these groups). Socio-economic covariates include baseline income in USD (adjusted for inflation using Consumer Price Index deflators from the Bureau of Labor Statistics) and educational attainment in years. Behavioral covariates include smoking, alcohol consumption and physical activity. Smoking is captured by a dummy variable (current smoker) and a continuous variable for pack-years smoked to account for residual confounding. Alcohol consumption is measured as a categorical variable classified as no/low, moderate and high intake. Physical activity is measured as a categorical variable classified as mild, moderate and high activity levels[61,62]. Details of the specific methods used for constructing the alcohol and physical activity variables are in the methods appendix for this chapter. Systolic blood pressure (mmHg) and BMI (kg/m^2) are primary metabolic mediators of interest and are objectively measured in each cohort. We use the average of multiple baseline BMI and SBP measurements available for the participants.

Statistical analyses

The relationship between dietary patterns and all-cause mortality is the primary association of interest. We use Poisson (piecewise exponential) and Cox Proportional Hazards (PH) models to assess the effect of dietary scores as well as multiple confounders and covariates[63]. These models assume that censoring of observations is non-informative: individuals are not censored due to the influence of a known or unknown risk factor that could be linked to the future risk of death.

Formally, the proposed model will take the form

$$h(t) = h_0(t) \times exp(b_1x_1 + b_2x_2 + \cdots + b_px_p) \ldots (2)$$

In the equation, t represents survival time, $h(t)$ is the hazard function determined by a set of p covariates $(x_1 + x_2 + \cdots + x_p)$ with corresponding regression coefficients$(b_1 + b_2 + \cdots + b_p)$. h_0 is the baseline hazard. The Cox PH regression directly models the hazard function assuming that all individuals have a common baseline hazard that depends on time[64]. The exponentiated coefficients from the Cox PH model are interpreted as Hazard ratios for a unit difference in the levels of the outcome. The Poisson model is similar to the Cox PH, but is semi-parametric in that it assumes death rates are constant within specified intervals of time. Specifically, in the Cox model, the reference population's death rate over a specified interval of time is left unspecified, while in the Poisson model it is assumed constant[65]. Both models provide very similar results for relative risks, but the Poisson more conservative and can provide cluster robust standard errors.

The set of x variables includes the diet scores estimated from PCA and the full set of confounders mentioned above, excluding BMI and SBP, which are only used in sensitivity analyses. In addition to the covariates listed, the analysis controls for 5-year interval birth-cohort fixed effects that purge the estimates from biases resulting from differential aggregate early childhood exposures[66]. Adjustments are made for study-cohort fixed effects to account for unobserved time invariant differences across the different study-cohorts[67].

Sensitivity analyses

In sensitivity analyses, first, we run regressions on a subsample of non-smokers to further reduce threats from reverse-causation because non-smokers are more likely to represent healthy individuals who did not adopt specific DPs after affliction with a disease[23]. Second, we relax our assumption of a linear dose-response for the DPs by fitting restricted cubic splines with 4 knots to obtain a continuous smooth function that is linear before the first knot, a piecewise cubic polynomial between adjacent knots, and linear again after the last knot[68]. Third, we run regressions on sub-samples for males and females; and categories of BMI (normal, overweight and obese), to test for effect modification. Fourth, we add BMI and SBP as covariates to gauge sensitivity of the estimates to control for variables on the causal pathway[69]. Fifth, we run logit models to test if DPs are associated with specific disease events including CVD, stroke, cancer and diabetes. Associations with specific diseases reveal pathways though which DPs may be associated with ACM. Sixth, to estimate the strength that an unmeasured confounder would need to have to explain away DPs associations, we calculate the E-value for all dietary patterns[70].

Results

Across the seven cohorts, individuals were followed-up for 1.579 million person-years with 18388 deaths events, at an average mortality rate of 0.0116 per year [95%CI 0.0115; 0.0118] (**Table2 1**). Deaths rates among males 0.0191 [95%CI 0.0187, 0.0197] were significantly higher than females 0.01 [95%CI 0.0099, 0.0102]. Sex-specific survival curves are shown in **Figure A2 1**. Mean age at baseline was 62.1 years, 12.8% of the participants were Black and 80.5% were non-Hispanic White. Mean income was 78.4 thousand US$ (inflation adjusted) and participants had 14.2 years of education, on average. Smoking prevalence was higher among males (16.8%) compared to females (8.4%). Estimated mean energy intake was approximately 1820kcal among females and 2350kcal among males. Males (175.7cm) were taller than females (161.8cm),

but both had similar BMI (27.3 kg/m^2) and SBP (127.4mmhg), on average. Cohort specific summary tables are reported in the appendix (**Table A2 2**). Associations of covariates with ACM were: hazard ratios (HR) age 1.10 [95%CI 1.09, 1.11] per year, male 1.54 [95%CI 1.47, 1.61], Black-individuals 1.12 [95%CI 1.07, 1.17], other race/ethnicities 0.81 [95%CI 0.74, 0.89], height 1.02 [95%CI 1.01, 1.04] per cm, log real income 0.77 [95%CI 0.74, 0.79] per unit, education 0.99 [95%CI 0.98, 0.99] per year, current smoker 1.73 [95%CI 1.66, 1.82], pack-year 1.09 [95%CI 1.08, 1.10] per year, high alcohol intake 1.13 [95%CI 1.09, 1.17], and high physical activity 0.64 [95%CI 0.62, 0.66] (**Figure A2 2**).

Table 2 1: Mortality events and baseline summary statistics for covariates across seven prospective cohorts

	Females		Males		Total	
	Mean/%	95%CI/n	Mean/%	95%CI/n	Mean/%	95%CI/n
Person- years	1,305,925		273,374.1		1,579,299.1	
Deaths	13141		5247		18388	
Mortality rate	0.0100	[0.0099, 0.0102]	0.0191	[0.0187, 0.0197]	0.0116	[0.0115, 0.0118]
Age, years	62.5	[62.4,62.5]	60.2	[60.0,60.4]	62.1	[62.1,62.2]
African American (Black), %	11.6	12,504	19.0	3,770	12.8	16,274
Non-Hispanic White, %	81.3	87,351	75.9	15,056	80.5	102,407
Inflation adjusted income, 1000 $	77.1	[76.9,77.4]	85.0	[84.1,85.9]	78.4	[78.1,78.6]
Education, years	14.3	[14.3,14.3]	13.8	[13.7,13.8]	14.2	[14.2,14.2]
Current smoker, %	8.4	8,993	16.8	3,338	9.7	12,331
Pack-years, 10 years	0.2	[0.2,0.2]	1.8	[1.8,1.8]	0.5	[0.5,0.5]
Moderate/high alcohol consumption, %	29.4	31,587	26.0	5,162	28.9	36,749
Highly active, %	31.1	33,376	32.6	6,462	31.3	39,838
Energy intake, 100 kcal	18.2	[18.1,18.2]	23.5	[23.3,23.6]	19.0	[19.0,19.0]
Height, cm	161.8	[161.7,161.8]	175.7	[175.5,175.8]	163.9	[163.9,164.0]
Body mass index, kg/m^2	27.3	[27.2,27.3]	27.7	[27.7,27.8]	27.3	[27.3,27.4]
Systolic blood pressure, mmhg	127.0	[126.9,127.1]	129.6	[129.3,129.9]	127.4	[127.3,127.5]
N	107,431		19,848		127,279	

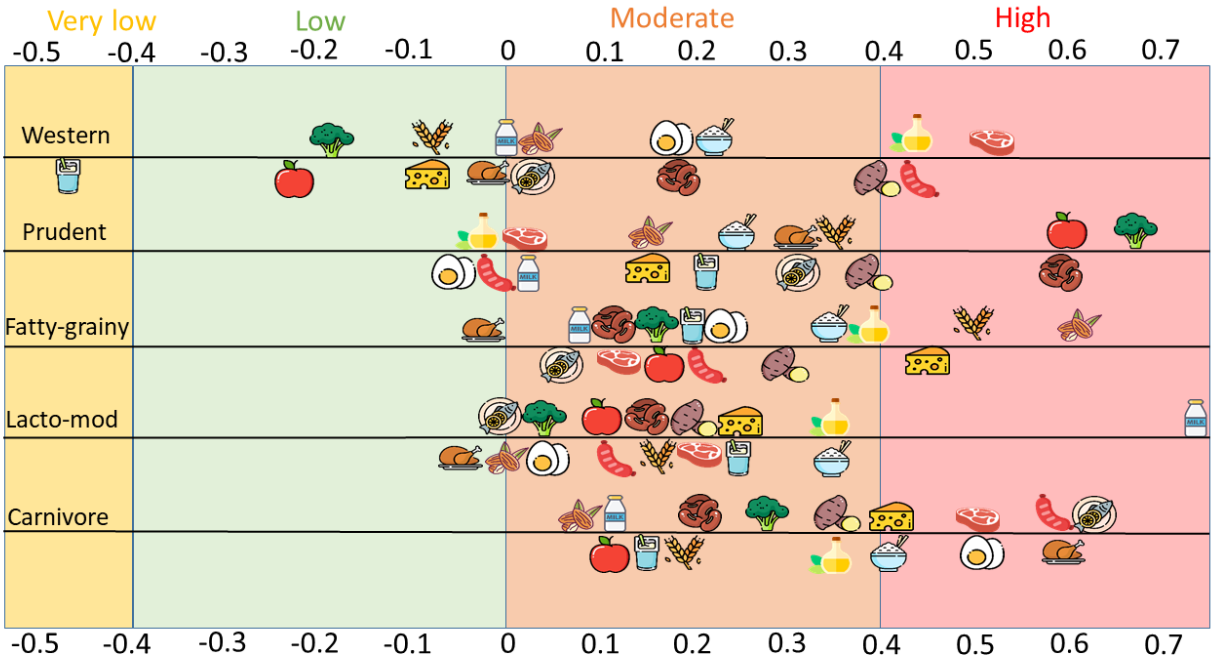
Notes: Numbers reported are sample means or percentages with 95% confidence intervals in parenthesis or counts for binary variables. Atherosclerosis Risk in Communities Study (ARIC), Cardiovascular Health Study (CHS), Framingham Heart Study-Offspring (FRAM), Multi-Ethnic Study of Atherosclerosis (MESA), NHANES I Epidemiologic Follow-up Study (NFS1), Reasons for Geographic and Racial Differences in Stroke (REGARDS), Women's Health Initiative Study (WHI).

PCA performed individually on each cohort revealed high consistency in the components present across the samples (**Tables A2 3 to A2 9**). Across most cohorts four components had Eigenvalues greater than 1. Cohort specific mean dietary intake of food groups used in PCA are reported in **Table A2 11.** **Figure 2 1** shows that five primary dietary patterns were retained for regression analysis. The horizontal lines depict how food groups are correlated (Pearson's Rho) with specific DPs and aims in classifying and naming the DPs based on these correlations in the pooled data. **Table A10** shows a heat map and the exact correlation coefficients depicted in the figure.

The 'Western' DP (not observed in the REGARDS cohort) was high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP was high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP was high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto' DP (not observed in the REGARDS and NHANES cohorts) was high in milk and moderate to low on all other food

groups. The 'Carnivore' DP was high in red and processed meats, poultry, fish, eggs and moderate on other food groups. This DP was only found in the WHI cohort. Individuals with higher scores on a specific DP are likely to consume more foods correlated with that DP.

Figure 2 1: Dietary patterns identified through principal component analyses across seven prospective cohorts

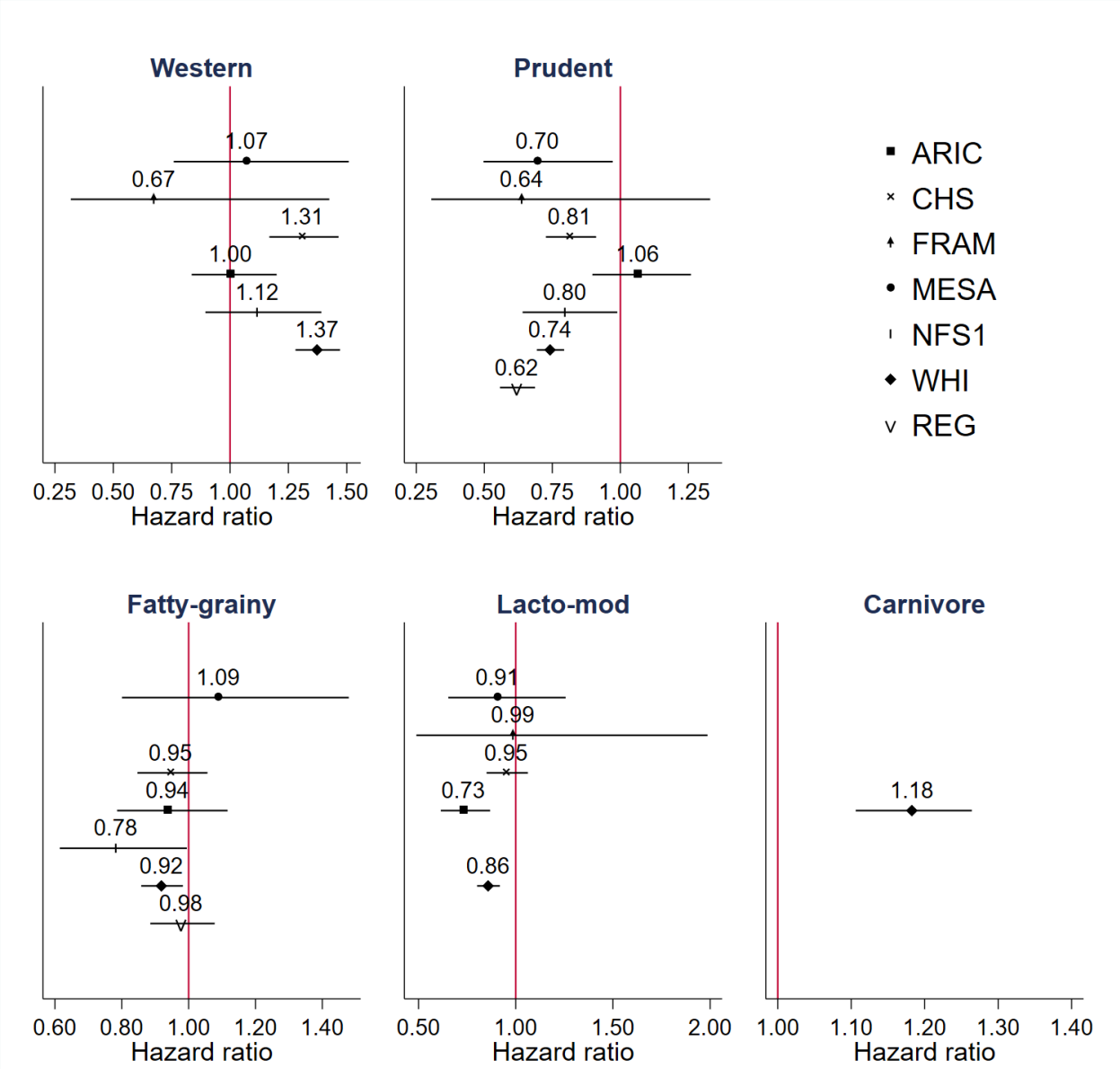


Notes: Classification is based on supplementary table A2.10. The Carnivore diet is only estimated in the Women's Health Initiative (WHI). Horizontal lines represent Pearson correlation coefficient scales. Correlations represent bivariate associations between dietary input variables and dietary patterns.

Age and sex adjusted cohort specific Cox-PH regressions revealed directionally similar results across most cohorts although the magnitude of the effects varied (**Figure 2 2**). Broadly, individuals in the highest quintiles (versus the lowest) on the Western and Carnivore DPs had higher rates of ACM, whereas the Prudent, Fatty-grainy and Lacto-mod DPs were associated with lower ACM. **Figure 2 3** shows pooled age and sex adjusted HR: Western 1.29 [95%CI 1.22, 1.36], Prudent 0.75 [95%CI 0.72, 0.79], Fatty-grainy 0.88 [95%CI 0.84, 0.93], Lacto-mod 0.87 [95%CI 0.82, 0.92] and Carnivore 1.19 [95%CI 1.10, 1.26].

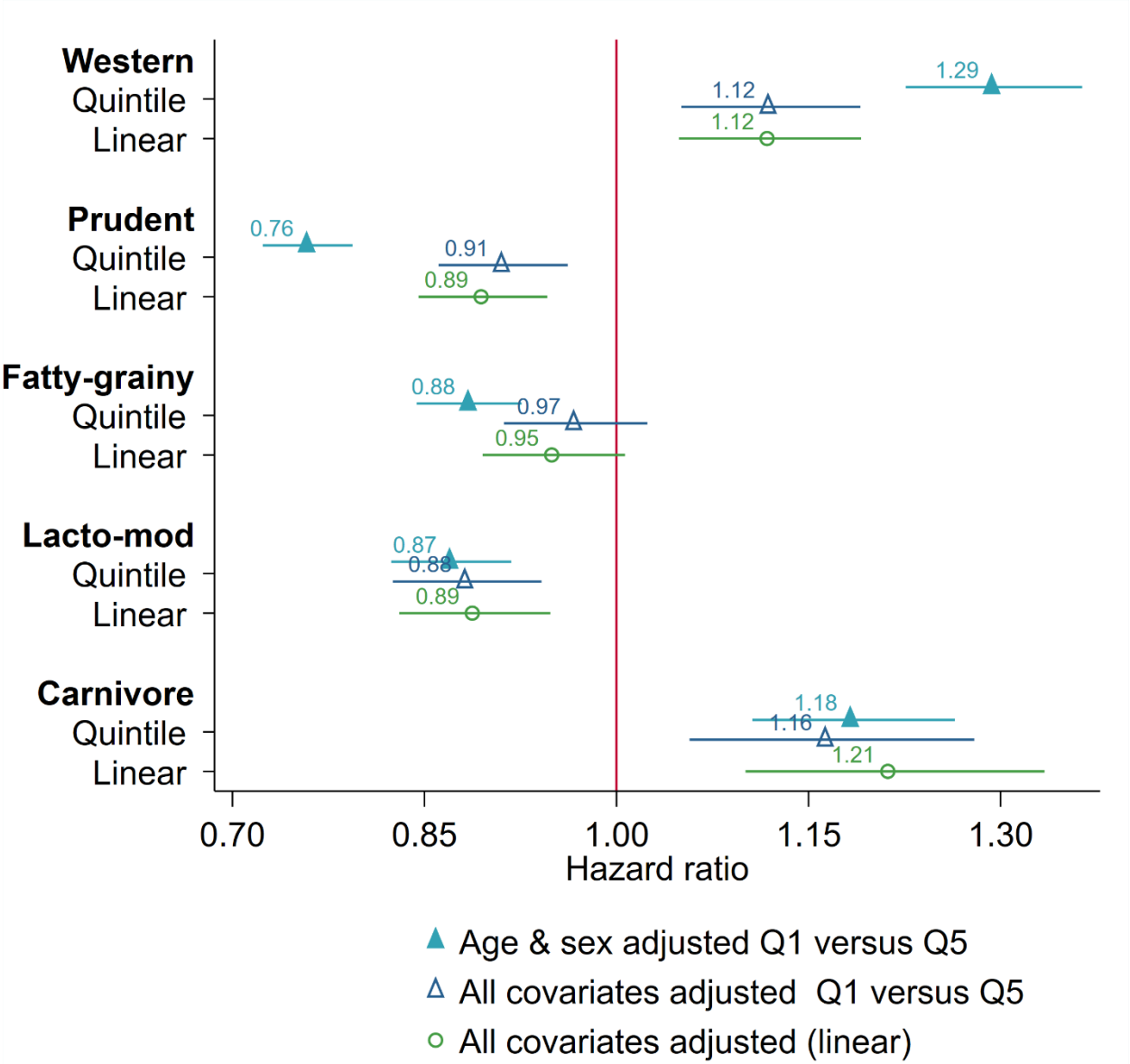
Figure 2 3 shows the fully adjusted Cox-PH models, which are directionally consistent with the age and sex adjusted models. Pooled fully adjusted HRs for highest versus lowest quintiles were Western 1.12 [95%CI 1.05, 1.19], Prudent 0.91 [95%CI 0.86, 0.96], Fatty-grainy 0.97 [95%CI 0.91, 1.02], Lacto-mod 0.88 [95%CI 0.83, 0.94] and Carnivore 1.16 [95%CI 1.06, 1.28]. The estimates from the linear dose-response specification were very similar to the quintile comparisons (**Figure 2 3)**. However, they are our preferred estimates since they utilize data on all individuals in the sample. Pooled fully adjusted HRs in the linear dose-response models were Western 1.12 [95%CI 1.05, 1.19], Prudent 0.89 [95%CI 0.85, 0.95], Fatty-grainy 0.95 [95%CI 0.90, 1.01], Lacto-mod 0.89 [95%CI 0.83, 0.95] and Carnivore 1.21 [95%CI 1.10, 1.33].

Figure 2 2: Cohort specific age and sex adjusted Cox-PH models for the association between all-cause mortality and dietary patterns: lowest versus highest quintile



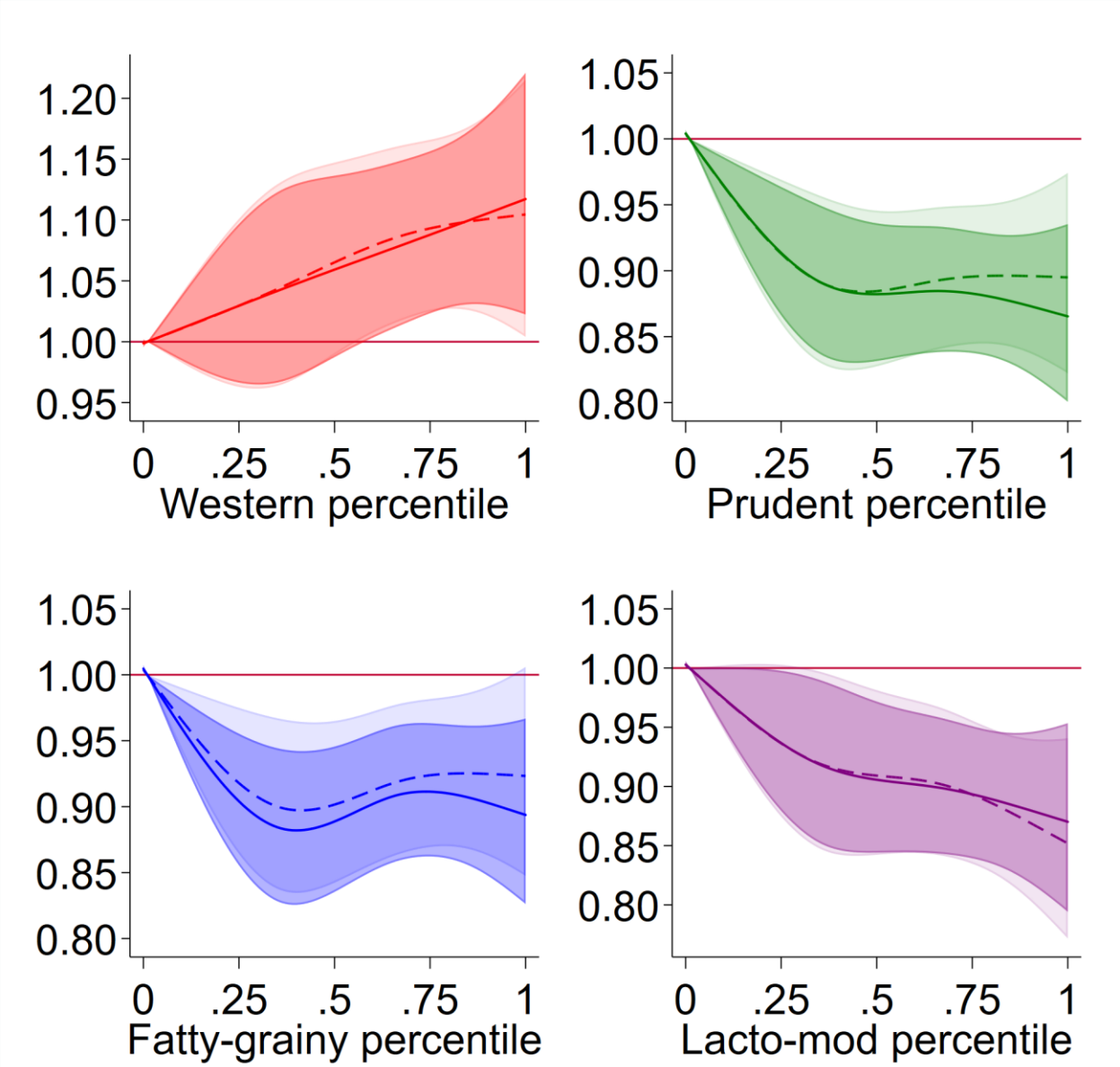
Notes: Coefficients are results run on individual cohorts adjusted for age and sex. Bands are 95% confidence intervals. The coefficient is interpreted as difference in hazard comparing the lowest quintile of the diet to the highest quintile. Atherosclerosis Risk in Communities Study (ARIC), Cardiovascular Health Study (CHS), Framingham Heart Study-Offspring (FRAM), Multi-Ethnic Study of Atherosclerosis (MESA), NHANES I Epidemiologic Follow-up Study (NFS1), Reasons for Geographic and Racial Differences in Stroke (REG), Women's Health Initiative Study (WHI). The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Figure 2 3: Pooled association between dietary patterns and all-cause mortality across seven prospective cohorts, low versus high intake



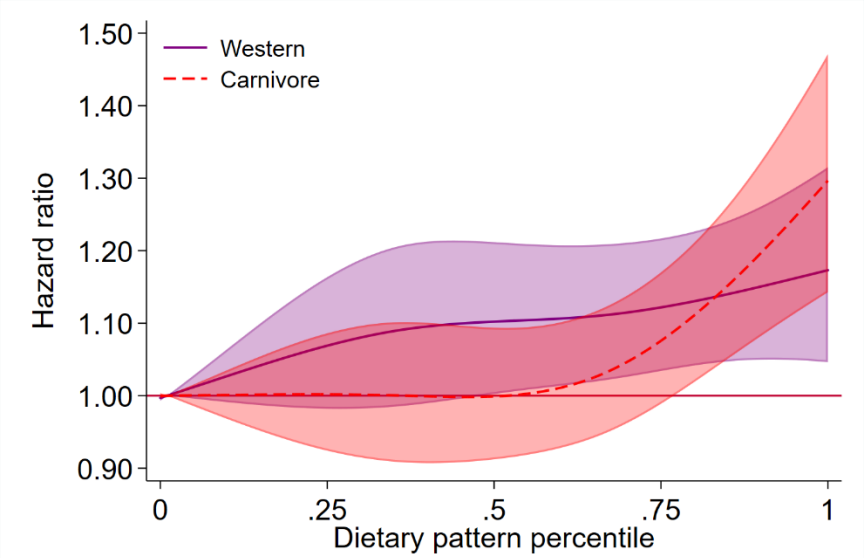
Notes: Q1=quintile 1 lowest, Q5=quintile 5 highest. All covariates include age, sex, race, income, education, smoking, alcohol, physical activity, calorie intake, height, study cohort fixed effects and birth cohort fixed effects. Standard error estimates are robust. Bands represent 95% confidence intervals. Diet scores range from 0 to 1. The quintile coefficient is interpreted as difference in hazard comparing the lowest quintile of the diet to the highest quintile. The linear coefficient is interpreted as difference in hazard comparing the lowest percentile of the diet to the highest percentile, assuming a linear risk curve. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Figure 2 4: Parametric non-linear dose-response association between dietary patterns and all-cause mortality across seven prospective cohorts



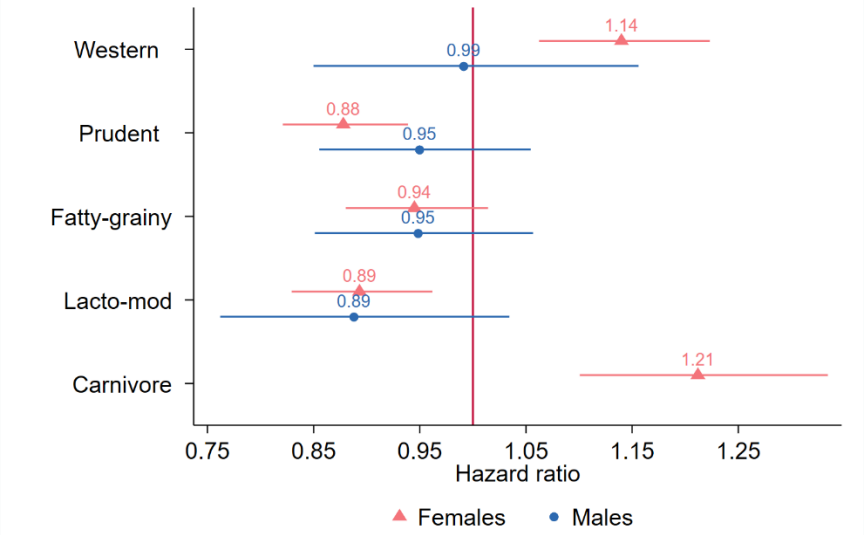
Notes: Solid lines represent the full sample. Dashed lines represent a sub-sample of non-smokers. Models are adjusted for age, sex, race, income, education, smoking, alcohol, physical activity, calorie intake, height, study cohort fixed effects and birth cohort fixed effects. Bands represent 95% confidence intervals. Non-linear curves are specified using splines with 4 knots. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups.

Figure 2 5: Comparing 'Western' and 'Carnivore' dietary patterns in the Women's Health Initiative (WHI) cohort



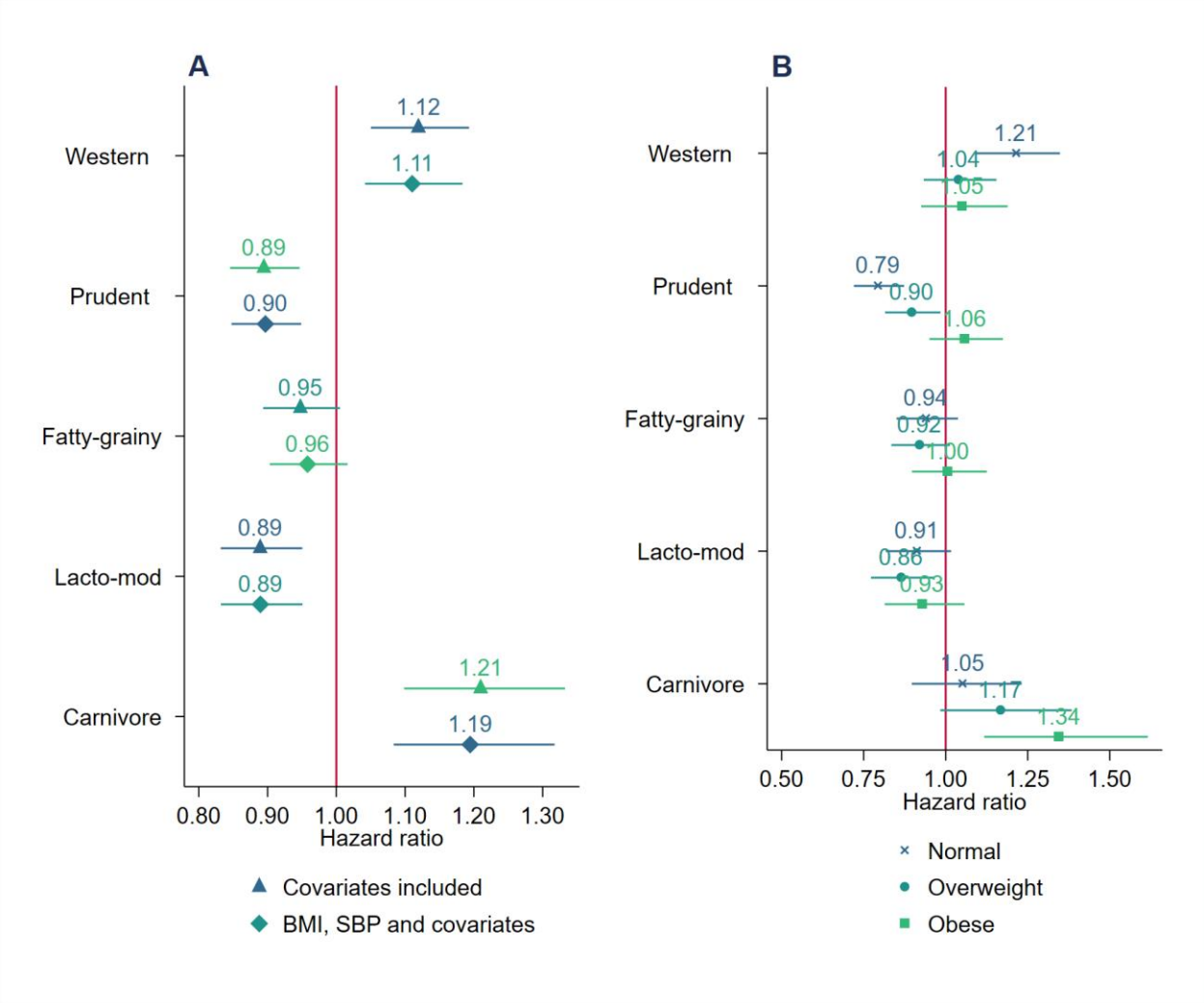
Notes: Non-linear curves are specified using splines with 4 knots. Models are adjusted for age, sex, race, income, education, smoking, alcohol, physical activity, calorie intake, height, study cohort fixed effects and birth cohort fixed effects. Bands represent 95% confidence intervals. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Figure 2 6: Sex specific associations between dietary patterns and all-cause mortality



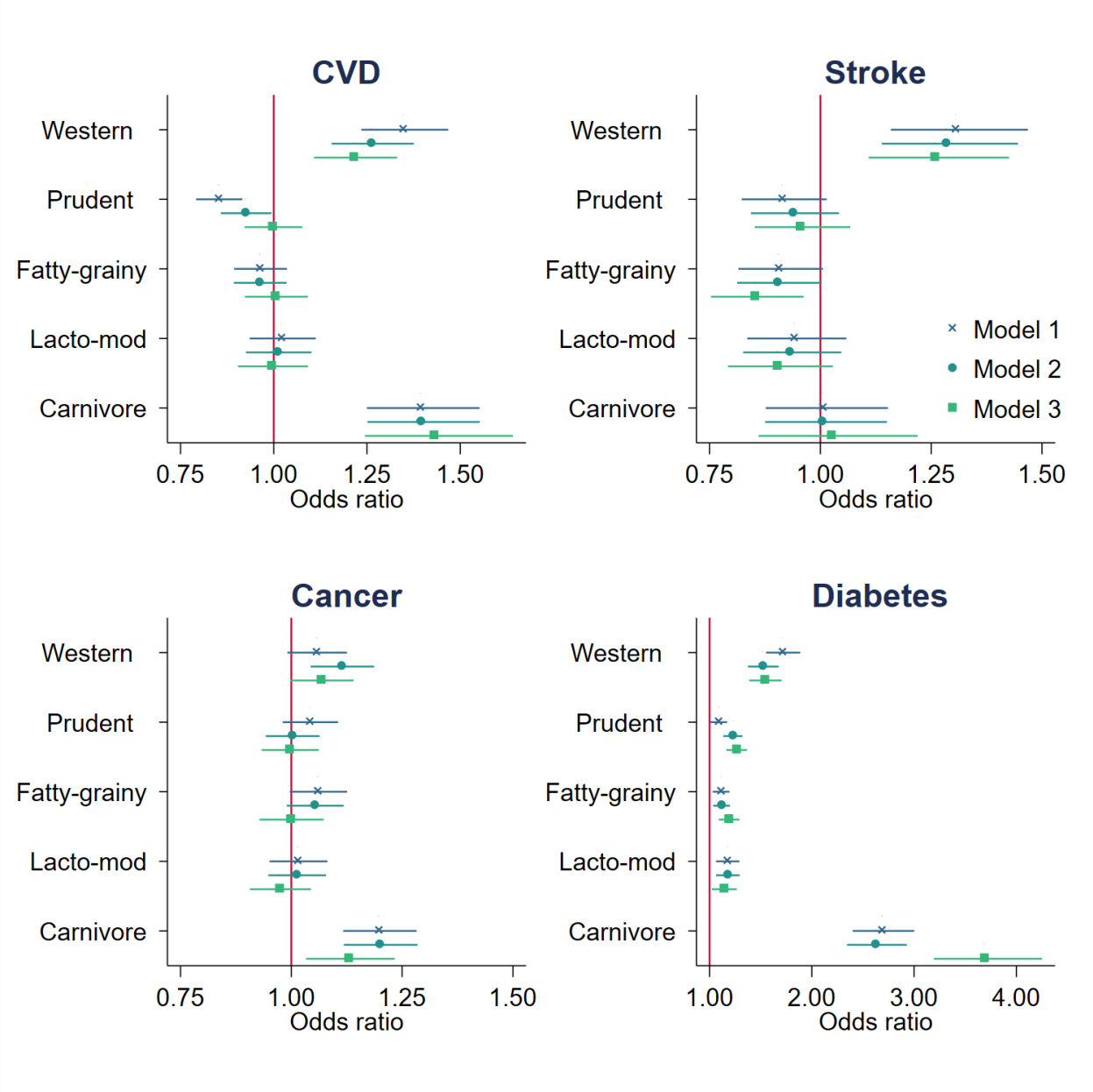
Notes: Regressions are run on sub-samples of males and females. Diet scores range from 0 to 1. The coefficient is interpreted as difference in hazard comparing the lowest percentile of the diet to the highest percentile, assuming a linear risk curve. Models are adjusted for age, race, income, education, smoking, alcohol, physical activity, calorie intake, height, study cohort fixed effects and birth cohort fixed effects. Bands represent 95% confidence intervals. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Figure 2 7: Sensitivity analyses with body mass index and systolic blood pressure



Notes: BMI= Body mass index, Normal BMI=18.5-25, Overweight=25-29.99, Obese≥30
SBP=Systolic blood pressure.
Diet scores range from 0 to 1. The coefficient is interpreted as difference in hazard comparing the lowest percentile of the diet to the highest percentile, assuming a linear risk curve. Bands represent 95% confidence intervals.
Covariates included in the models are age, sex, race, income, education, smoking, alcohol, physical activity, calorie intake, height, study cohort fixed effects and birth cohort fixed effects.
Panel A= tests for attenuation of effect sizes after including BMI and SBP.
Panel B= Runs regressions on sub-samples of BMI categories.
The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables.
The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry.
The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups.
The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Figure 2 8: Association between dietary patterns and disease occurrence



Notes: Model 1 include age, sex, race, height cohort fixed effects and birth cohort fixed effects.
Model2=Model 1+ education and income.
Model3=Model2 + smoking, and calorie intake.
Standard error estimates are robust. Bands represent 95% confidence intervals.
Diet scores range from 0 to 1. The linear coefficient is interpreted as difference in odds comparing the lowest percentile of the diet to the highest percentile, assuming a linear risk curve.
The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Parametric non-linear dose-response curves revealed further insights (**Figure 2 4**). For Western and Lacto-mod DPs, the linear model was a good approximation of the risk curve indicated by monotonically increasing risk at higher levels of intake. These curves were nearly identical in the sub-sample of non-smokers. For the Prudent and Fatty-grainy DPs, there was reduced risk until the 50th percentile followed by a steady level of low risk. These curves were only slightly attenuated in the sub-sample of non-smokers. Overall, the curves suggest a low degree of reverse-causation bias. **Figure 2 5** shows a comparison of the Western and Carnivore DPs in the WHI sample. The Carnivore DP has a non-linear dose-response function and has a distinct inflection towards high risk after the 50th percentile. The HRs at the 100th percentile in the WHI were 1.17 [95%CI 1.04, 1.32] and 1.30 [95%CI 1.14, 1.47] for the Western and Carnivore DPs, respectively in the WHI.

Sensitivity analyses revealed that Western and Prudent DPs had larger effect sizes for women than men, whereas Fatty-grainy and Lacto-mod DPs had similar associations among both sexes (**Figure 2 6**). Similarly, the Prudent DPs had larger effect sizes in sub-samples of individuals with normal BMI and the CIs did not overlap with other BMI categories indicating statistical significance (**Figure 2 7, panel B**). Adjusting for BMI and SBP slightly attenuated coefficients, but results remained statistically significant for all DPs excluding Fatty-grainy (**Figure 2 7, panel A**). Logit models showed that the Western DP was associated with significantly higher risk of CVD, stroke and diabetes, whereas the Carnivore DP was associated with higher risk of cancer and a very high risk of diabetes (**Figure 2 8**). The Prudent diet was negatively associated with CVD whereas the Fatty-mod was negatively associated with stroke. All DPs were associated with marginally elevated risks of diabetes. E-value calculations indicated that unmeasured confounders would need to be associated with ACM and DPs with HRs of 1.5, 1.5, 1.3, 1.5, and 1.7 for the Western, Prudent, Fatty-grainy, Lacto-mod and Carnivore DPs, respectively, to completely explain away dietary associations.

DISCUSSION

In this study we aimed to quantify the risk to ACM associated with commonly occurring DPs in the US. We identified four DPs that consistently explained variation across seven cohorts and an additional DP in the WHI. Since our analyses adjusted for calories, our findings speak to diet quality at the same levels of energy intake. Concurring with research from previous studies, diets high in meat were found to be high risk for ACM, while those comprised of nutrient dense plant foods were consistently and robustly associated with lower ACM risk[71]. Risk estimates for females with respect to the Western and Prudent DPs were larger compared to males. Further, results suggest that individuals with healthy BMI may benefit most from the Prudent DP.

Our study is novel in identifying that the widely researched Western DP has a linear dose-response relationship with ACM, while the Prudent DP has an L-shaped dose-response curve in US cohorts. More research is needed in other contexts to gauge the external validity of these relationships. L-shaped curves for the Prudent DP may suggest that moderation diets that are sufficiently high in fruits, legumes and vegetables may offer similar benefits as diets saturated with these foods groups. Similarly, the Lacto-mod DP, also a moderation diet, was associated with lower ACM risk. Since systematic reviews show that high intake of milk is neutral for ACM, is it is likely that a moderate to low intake of all other food groups in the Lacto-mod DP may offer ACM benefits[26]. However, there is a possibility that the lower ACM risk associated with Lacto-mod may be an artefact of residual confounding from race and income, since high levels of this DP are mostly prevalent among higher income White individuals (**Table A2 12**). There may also be

unobserved confounding from wealth. Further research is needed on the Lacto-mod DP to investigate its robustness in other cohorts and if it can offer an alternative to the 'Prudent' diet.

A second novel finding of this study, is the comparison of the Western DP with the Carnivore DP in the WHI. The Carnivore DP has recently gained popularity, but it's long term consequences are unknown[72]. Our study indicates that DPs that are saturated in meats are strongly associated with higher ACM risk and significantly higher risks of cancer and diabetes. This was clear at the sharp inflection in the ACM risk curve at high levels of the Carnivore DP. This may also suggest, though more research is needed, that the presence of refined grains may offset some of the ill effects of meat in the 'Western' diet[14,26,35]. Another conjectural explanation is that women adhering to the Carnivore DP in WHI are dieters following a high protein, low carbohydrate diet; and the association might reflect detriments of chronic dieting.

Evidence from this study concurs with a body of research suggesting that DPs characterized by higher consumption of vegetables, fruits, legumes, nuts, whole grains; and relatively low in red and processed meat; are associated with a decreased risk of ACM[71]. These food groups are also common across recommended protective diets like the HEI and MDS. Our estimates for the Prudent diet qualitatively agree with meta-analyses that estimated the risks for adherence to HEI/MDS[12,73,74]. Our comprehensive set of controls reduces large biases that limited previous US based studies. Importantly, our analyses suggest that studies using similar methods that do not adjust for income, likely provide estimates that are greatly biased away from the null, for the Western and Prudent DPs[75–78]. Thus, whereas our adjusted Prudent estimate is smaller than combined meta-estimates for individual food groups[3,7,16,17,26,34], it likely presents higher plausibility for being closer to the true protection that healthy foods provide in combination.

Limitations of the study include observational design and measurement error. DP analyses rely on observational prospective cohort studies because it is not feasible to randomize diets for long periods[79]. This comes at the cost of susceptibility to bias from residual and unobserved confounding, and reverse causation, so ultimately evidence on DPs must be evaluated with caution[80]. We minimized the presence of confounding by using a large set of controls, but confounders might have been measured with error, which leads to residual confounding (e.g., smoking and income). We assessed threats from unobserved confounders with the E-value. For major DPs, E-values indicated that an unobserved confounder would need to have relative risks of 1.5-1.8 to nullify our results. We know of no such confounder, given that only smoking had such a high relative risk in our sample. We mitigated threats from reverse causation by excluding participants who died within 3 years of observation and also analyzed a subset of nonsmokers for whom results were attenuated in some diets but still significant. Additionally, since this study relies on FFQs, the measured intake of foods is prone to random error due to non-differential misclassification, and can lead to biased estimates, generally towards the null[81]. Evidence indicates that a high degree of attenuation can be attributed to measurement error in FFQ[82]. To this end, the pooled cohort design provided greater statistical power thereby endowing the study with the ability to detect significant assoications[81]. Finally, DPs are not time invariant, thus, a person's lifetime risk to any DP depends on the lagged effects of the cumulative exposure to different DPs consumed throughout life. Our estimates may be biased due to the inability to account for cumulative exposure or lagged effects.

Strengths of this study include its large sample size, pooled cohort design, use of rigorous regression methods, low risk of confounding bias and generalizability[71]. The pooled cohort design supports statistical precision and enabled us to systematically compare and check for consistencies in dietary patterns across

multiple cohorts using the same food groups and selection criteria in PCA. Our study achieved consistency in deriving DPs and surmounted challenges faced by meta-analyses, which rely on estimates from varied studies employing different methods to identify DPs, adding high heterogeneity to the estimates[26]. The identification of DPs that are commonly consumed in the US support generalizability of our findings, making them readily translatable into eating behaviors, such as replacing a Western DP with a Prudent DP.

CONCLUSION

Greater adherence to several healthy eating patterns was consistently associated with lower risk of ACM. Our findings support the Dietary Guidelines for Americans and other studies[26,36–45], which recommend eating patterns rich in plant foods, and low in meat and processed foods, to adapt according to personal food traditions and preferences. In terms of guidance for public policy, our study reinforces the emerging focus on DPs as opposed to individual food groups[26,36–45]. A focus on the way people actually eat is more readily translatable into food choices at different life stages. Healthy DPs should be incorporated for individuals of all ages in institutional settings such as school meals, office cafeteria's and care homes. There is also a need to create more awareness about healthy eating patterns that support longevity instead of short term goals like weight loss. Indeed, our analyses indicate that diets rich in meat that are commonly suggested for weight loss may be highly damaging in the long run[83,84].

CHAPTER 3:
METABOLIC MEDIATORS OF THE EFFECTS OF DIETARY PATTERNS ON MORTALITY: A POOLED ANALYSIS OF SEVEN PROSPECTIVE COHORTS

ABSTRACT

Objectives: To disentangle the direct and indirect effects of dietary patterns (DP) on all-cause mortality (ACM) through body mass index (BMI) and systolic blood pressure (SBP).

Participants and setting: Secondary analysis of pooled datasets from seven prospective cohorts including Atherosclerosis Risk in Communities Study, Cardiovascular Health Study, Framingham Heart Study-Offspring, Multi-Ethnic Study of Atherosclerosis, Women's Health Initiative Study, NHANES I Epidemiologic Follow-up Study, Reasons for Geographic and Racial Differences in Stroke. 127,279 individuals were followed-up for 1,579,299 person-years.

Outcomes: Outcomes were 18,388 deaths (binary) and the survival time until death/censoring (interval between start of the study and death or loss to follow-up).

Exposures: Principal Component Analysis of observed intake of food groups was used to estimate DPs coded as continuous variables ranging from 0 (low) to 1 (high). Five dietary patterns were estimated. The 'Western' pattern was high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' pattern was high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' pattern was high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto' pattern was high in milk and moderate to low on all other food groups. The 'Carnivore' pattern was high in red and processed meats, poultry, fish, eggs and moderate on other food groups. Mediators BMI (kg/m^2) and SBP (mmhg) were objectively measured at baseline and averaged over follow-up visits.

Design: Seemingly unrelated Poisson and structural equation models (SEM) were used estimate direct, indirect and total effects. Regressions were adjusted for age, sex, race, education, income, smoking, drinking, physical activity, birth-cohort fixed effects and survey (cohort) fixed effects. Mediation was quantified using the percent of excess risk mediated (PERM) and proportion mediated measure (PMM).

Results: In the mediator unadjusted model, ACM and Western DP were associated with a hazard ratio (HR) of 1.09 [95%CI 1.02, 1.16], Prudent 0.91 [95%CI 0.86, 0.96], Fatty-grainy 0.97 [95%CI 0.92, 1.03], Lacto-mod 0.91 [95%CI 0.85, 0.97], and Carnivore 1.20 [95%CI 1.09, 1.32]. PERM for BMI and SBP combined was estimated to be 27.1%, 9.7%, 49.9%, 7.3%, 26.4%, for the Western, Prudent, Fatty-grainy, Lacto-mod and Carnivore DPs, respectively. Holding energy intake constant, SEMS models revealed that Western and Carnivore DPs were positively associated with BMI and SBP; and higher intakes of the other DPs were associated with lower BMI and SBP. PMM estimates were marginally larger than PERM for the Wester, Prudent, and Lacto-mod DPs. Mediation estimates were highly sensitive to not adjusting for energy intake due to the effect of energy on BMI.

Conclusions: Interventions that reduce BMI and blood pressure, might address up to a third of the risk of ACM associated attributed to suboptimal diets. Maintenance of optimal diets is needed for complete benefits.

INTRODUCTION

In the past two decades, the prevalence of overweight and obesity in the US has continuously risen to alarming levels[85], with 35-38% percent of the population currently classified as obese[86]. Similarly, in 2017, 32.5-36% of individuals had hypertension (elevated blood pressure) with 70% receiving treatment. During the same period, nationally representative estimates indicate that the levels of consumption of healthy diets (measured by the international diet-health index) in US population have declined between 2004 and 2014[87], a period when healthy life expectancy in the US also declined[88]. Continuous effective intervention efforts are needed to fight the ongoing chronic disease epidemic that is partly fueled by suboptimal dietary practices.

A recent comprehensive analysis from four continents demonstrated that both overweight and obesity were associated with increased all-cause mortality(ACM) [89]. Similarly, studies have demonstrated a strong causal link between elevated blood pressure (BP) and increased risk of mortality[90]. While substantial evidence links suboptimal diets to mortality[2], evidence also suggests that diets rich in fruits, vegetables, wholegrains, legumes, seeds, nuts, fish, and dairy can significantly lower BMI and BP[91,92]. Changes in weight are thought to be causally related to changes in dietary patterns via two predominant mechanims[79]. First, a positive energy balance (where calorie intake is higher than calorie expenditure) leads to accumulation of adipose tissue via overconsumption of energy[93]. Second, high-glycemic load diets may produce hormonal changes that cause adipose tissue accumulation, which in-turn may lead to increased hunger and lower energy expenditure[94]. Further, while diets may influence BP in complex ways, literature suggests that changes in the relative levels of sodium, potassium and calcium are important factors[95].

Taken together, albeit indirectly, the evidence indicates plausible mediation channels (A to C through B relationships) for protective effects of diets that work through lower BMI and BP. However, the direct and indirect effects of specific dietary patterns have not been rigorously studied[92]. These effects are important for clinicians and public health practitioners because they advance our understanding of the relative importance of treatments that directly address BMI (e.g. bariatric surgery)[96] and BP (e.g. antihypertensive drugs)[97] versus those of dietary practices (e.g. adopting the Mediterranean or DASH diet). Such mediation analyses can also shed light on whether protective diets can prevent mortality even among those in whom a protective diet fails to reduce weight or BP and if these effects differ by individual characteristics.

Methods for mediation analysis are being used frequently in epidemiology, but to the best of our knowledge these methods have not been applied to the diet-mortality association[92,98]. In this study, we applied analytical methods[99–103] to individual level data from seven pooled prospective cohort studies that have measurements of diet, body mass and blood pressure. While accounting for a comprehensive set of confounders, we quantified how much of the effects of diet on mortality were mediated through BMI and BP.

METHODS

Data, outcome and mediators

We use data from seven prospective cohorts for the analysis: Atherosclerosis Risk in Communities Study (ARIC), Cardiovascular Health Study (CHS), Framingham Heart Study-Offspring (FHSO), Multi-Ethnic Study of Atherosclerosis (MESA), Women's Health Initiative Study (WHI), NHANES I Epidemiologic Follow-up Study, Reasons for Geographic and Racial Differences in Stroke (REGARDS).

The outcome of interest for this study is all-cause mortality or ACM (deaths that occur in a population, regardless of the cause). Participants were followed-up from baseline to the last round of examination, event of death or when they could be contacted for follow-up. A participant's time since baseline to death or last follow-up was a calculated in days. The final outcomes are event of a death (binary) and the survival time until death occurs (interval between start of the study and death or loss to follow-up). For those patients that are alive at the end of the observation period, the actual time-to-event is unknown, and will be treated as a case of right censored data. Details of data sets and outcome assessment in the individual cohorts are described in chapter 2. Mediators include BMI (kg/m^2) and systolic BP (SBP mmhg), both measured objectively in all cohorts. We use two BMI and SBP measurements in the study. First, we use BMI and SBP taken at baseline in the primary analyses. Second, we use the average of measurements from all follow-up visits for participants who had more than one visit.

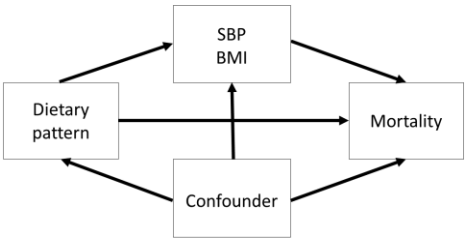
Dietary patterns

Dietary patterns (DPs) are approximated using principal component analysis (PCA) where data on intake of several food groups are compressed into DP scores that reflect individual diet types that are commonly consumed in the US. Food groups used in the analyses are fruits, starchy vegetables, fibrous vegetables, whole grain, refined grain, legumes, oils/fats, nuts and seeds, milk, cheese, yogurt, fish/shellfish, poultry, processed meats, eggs, and red meat. PCA derived DPs have been shown to have high stability, suggesting that they are well suited for a pooled cohort design[30]. PCA explains a set of correlated food groups with a new set of comprehensive indexes (principal components) that are uncorrelated and retain the largest amount variance in the original sample[60]. After predicting component scores (DPs), the scores are re-scaled ranging from 0 to 1, with 0 and 1 representing lowest and highest consumption levels of a specific pattern, respectively. PCA is conducted separately for each cohort to ensure that DPs represent each cohort. DP scores are pooled for the analysis.

Confounders

Conceptually, mediators differ from confounders because mediators fall on the causal pathway from an exposure to an outcome, whereas confounders do not. Confounders simultaneously predict both the exposure and the outcome. Mediators predict the outcomes but are predicted by the exposure[104] (**Figure A3 1**). In mediation analyses we make a set of strong assumptions about confounding, sequential occurrence and endogeneity on several variables. Since mediation requires that there is no confounding for (1) exposure-outcome, (2) mediator-outcome and (3) exposure-mediator relationships[103], we adjust for the set of covariates available in the data that satisfy these assumptions[105,106]. The full set of covariates included in the analysis are age, sex, race/ethnicity, height, income, education, physical activity, smoking, alcohol and calorie intake. We also include cohort (survey) fixed effects to account for measurement differences across cohorts; and birth-cohort fixed effects to account for unmeasured confounding from early childhood factors. Further, we assume there is no factor in the set of covariates that is a consequence of DPs[103]. While this assumption cannot be verified from the data, the known mortality related confounders for hypertension, obesity and diets are presented in **Table A3 1** to anticipate the direction of the relationships. **Figure 3 1** shows mediation framework with arrows depicting the direction of hypothesized relationships.

Figure 3 1 Directed acyclic graph showing mediation from diet to outcome



Notes: Given our observational study, we are unable to estimate the natural effects without making further assumptions. In the directed acyclic graph (DAG) shown, given the adjustment set of confounders, sufficient conditions for identifying the Natural Direct Effects with observational data are (1) No member of the set of confounders is a descendant of dietary patterns; (2) the set of confounders blocks all backdoor paths from mediators (BMI.SBP) to mortality; (3) there are no backdoor paths from dietary patterns to mortality that we cannot block; and (4) there are no backdoor paths from dietary patters to mortality that we cannot block. With these assumptions, we can estimate the natural direct effects with our observational data[69].

Statistical analysis

The definitions of direct and indirect effects within a counterfactual framework are provided in the appendix for chapter 3. Briefly, direct effects are interpreted as effects that flow directly from the exposure to the outcome, whereas indirect effects are those that flow from the exposure to the outcome through the mediator. We use Poisson regressions to estimate the rate ratios because they are computationally amenable to a mediation framework and produce results similar to the Cox proportional hazards model[65]. We analyze diet, BMI and SBP as a continuous variables in relation to ACM. We use (1) the difference and (2) the product methods to estimate direct and indirect effects[99,103,107]. Since underweight individuals comprised less than 1% of the sample and had significantly higher ACM risk (**Figure 3 2**), they were excluded from the regression analysis to enforce linearity. We mitigated threats from reverse causation by excluding participants who died within 3 years of enrollment[23].

Difference method

The difference method entails running a set of seemingly unrelated nested Poisson regressions[65,108], where we first estimate the effect of DPs on ACM with adjustment for confounders. Coefficients on the DPs in these models are interpreted as total effects. We then add mediators to the model, separately, and together. The coefficients on DPs in the mediator adjusted regressions are interpreted as direct effects. We compare the differences (attenuation) in the coefficients on the DPs in models, with and without mediators, to visualize the indirect effects. Next, the proportion of excess risk mediated (PERM)[23] is calculated as

$$PERM = \frac{Mediator\ unadjusted\ RR - Mediator\ adjusted\ RR}{Mediator\ unadjusted\ RR - 1} \quad (3.1)$$

Here the numerator is the indirect effect (the difference in the exposure coefficients of the two regressions) and the denominator is the total effect minus 1. The subtraction of 1 in the denominator reflects that on the ratio scale, 1 is the reference point. Running the regressions simultaneously allows us to estimate 95% confidence intervals for the PERM. We use percentile based bootstrap confidence interval estimates for

indirect effects because they are usually positively skewed and kurtotic; thus preclude the use of standard errors derived from the delta method[109].

Product method

The product method entails fitting structural equation models (SEM) that estimate confounder adjusted coefficients for DPs and mediators on ACM as well as coefficients of DPs on the mediators[65]. Specifically, for outcome t, exposure a, mediator m, and confounder c, we fit the following SEMs separately for each dietary pattern.

$$\lambda_t(t|a,m,c) = \lambda_0 + \lambda_1 a + \lambda_2' c + \lambda_3 m \quad (3.2)$$

$$E[M|a,c] = \beta_0 + \beta_1 a + \beta_2' c \quad (3.3)$$

Here the direct effect is λ_1, and the indirect effect is $\beta_1\lambda_3$.

We next estimate the proportion mediated measure (PMM)[104] as

$$PMM = \frac{\beta_1\lambda_3}{\beta_1\lambda_3 + \lambda_1} \quad (3.4)$$

All calculations for PERM and PMM are made on the log scale[99]. For rare outcomes (such as ACM), PMM and PERM should provide similar estimates.

Sensitivity analyses

We perform four sets of sensitivity analyses. First, we use BMI and SBP measurements averaged over all follow-up visits to check the sensitivity of mediation estimates to mediator measurement timing. Second, we gauge changes in mediation after excluding energy intake as a covariate. This model tests for quantity and quality aspects of DPs together. Third, we test for effect modification, by running models that include interactions of BMI/SBP with DPs. Fourth, threats from unmeasured confounding are assessed using bias correction factors[100].

RESULTS

Across the seven cohorts, 127,279 individuals were followed-up for 1.579 million person-years with 18,388 deaths events, at an average mortality rate of 0.012 per year [95%CI 0.0114; 0.0118] (**Table3 1**). The mean starting age of individuals across cohorts was 62.1 years, and 15.6% were male. Participants had 14.2 years of education and inflation adjusted income of $78,400, on average. At baseline, 9.7% were current smokers, and 28.9% had moderate to high alcohol consumption. Mean energy intake was 1900 kcal. Participants were 163.9 cm tall, on average. A large proportion of individuals were overweight (35.7%) and obese (26.0%), whereas 43.5% had 'at-risk' SBP and 21.8% were hypertensive at baseline. Kaplan-Meier curves stratified by BMI categories showed significantly higher ACM risk among individuals who were overweight and obese, with being underweight as the highest risk category. ACM risk curves were even starker for SBP, with hypertensive individuals at substantially higher risk than those with normal SBP levels (**Figure3 2**).

PCA performed individually on each cohort revealed high consistency in the components present across the samples (**Tables A2 3 to A2 9**). Across most cohorts four components had Eigenvalues greater than 1. Cohort specific mean dietary intake of food groups used in PCA are reported in **Table A2 11.** **Table A10** shows a heat map depicting how food groups are correlated (Pearson's Rho) with specific DPs and aims in classifying and naming the DPs based on these correlations in the pooled data. Broadly, the 'Western' DP (not observed in the REGARDS cohort) was high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP was high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP was high in nuts and seeds,

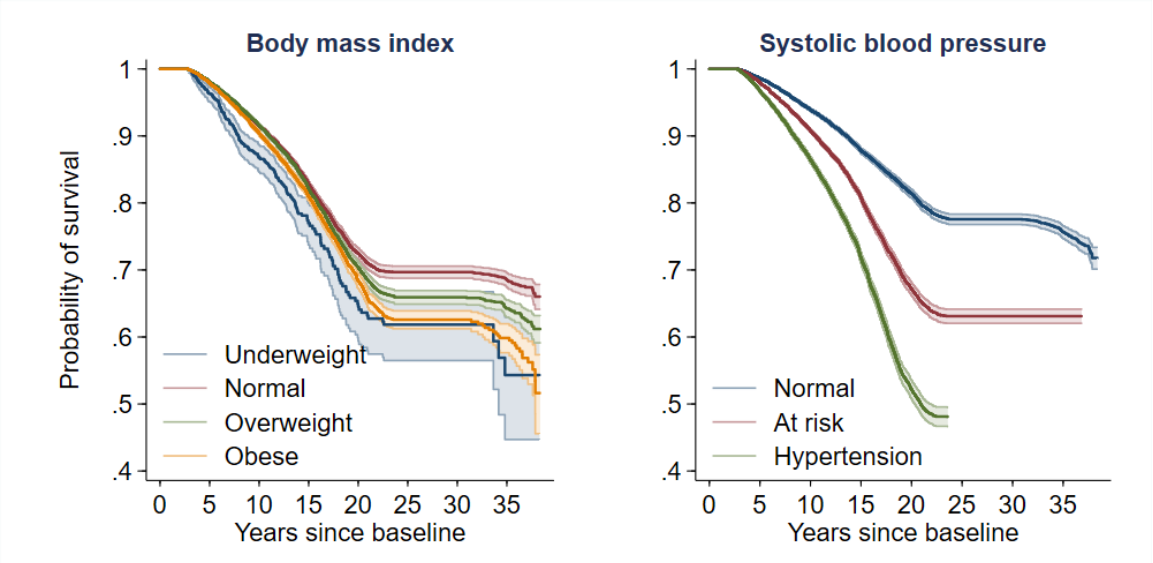
cheese, whole grains, fats/oils and low in poultry. The 'Lacto' DP (not observed in the REGARDS and NHANES cohorts) was high in milk and moderate to low on all other food groups. The 'Carnivore' DP was high in red and processed meats, poultry, fish, eggs and moderate on other food groups. This DP was only found in the WHI cohort. Individuals with higher scores on a specific DP are likely to consume more foods correlated with that DP.

Table 3 1: Baseline summary statistics

Denominator (N)= 127,279	**Mean/%**	**95% CI/ Numerator**
Person- years	1,579,299	
All-cause deaths	18,388	
Rate	0.012	[0.0114; 0.0118]
Body mass index, kg/m^2		
<18.5	0.9	N=1,184
18.5-25	37.4	N=45,329
25-30	35.7	N=45,738
>30	26.0	N=35,028
Systolic blood pressure, mmhg		
<120	34.8	N=48,133
120-140	43.5	N=55,868
>140	21.8	N=23,278
Covariates		
Demographic		
Age, years	62.1	[62.1,62.2]
Male, %	15.6	N=19,848
Height, cm	163.9	[163.9,164.0]
Socio-economic		
Inflation adjusted income, $	78.4	[78.1,78.6]
Education, years	14.2	[14.1,14.2]
Behavioral		
Current smoker, %	9.7	N=12,331
Pack-years, 10 years	0.5	[0.5,0.5]
High alcohol consumption, %	28.9	N=36,749
Highly active, %	31.3	N=39,838
Energy intake, 100 kcal	19.0	[18.9,19.0]

Notes: Atherosclerosis Risk in Communities Study (ARIC), Cardiovascular Health Study (CHS), Framingham Heart Study-Offspring (FRAM), Multi-Ethnic Study of Atherosclerosis (MESA), NHANES I Epidemiologic Follow-up Study (NFS1), Reasons for Geographic and Racial Differences in Stroke (REGARDS), Women's Health Initiative Study (WHI).

Figure 3 2: Kaplan-Meier survival curves for all-cause mortality stratified by risk categories of body mass and blood pressure



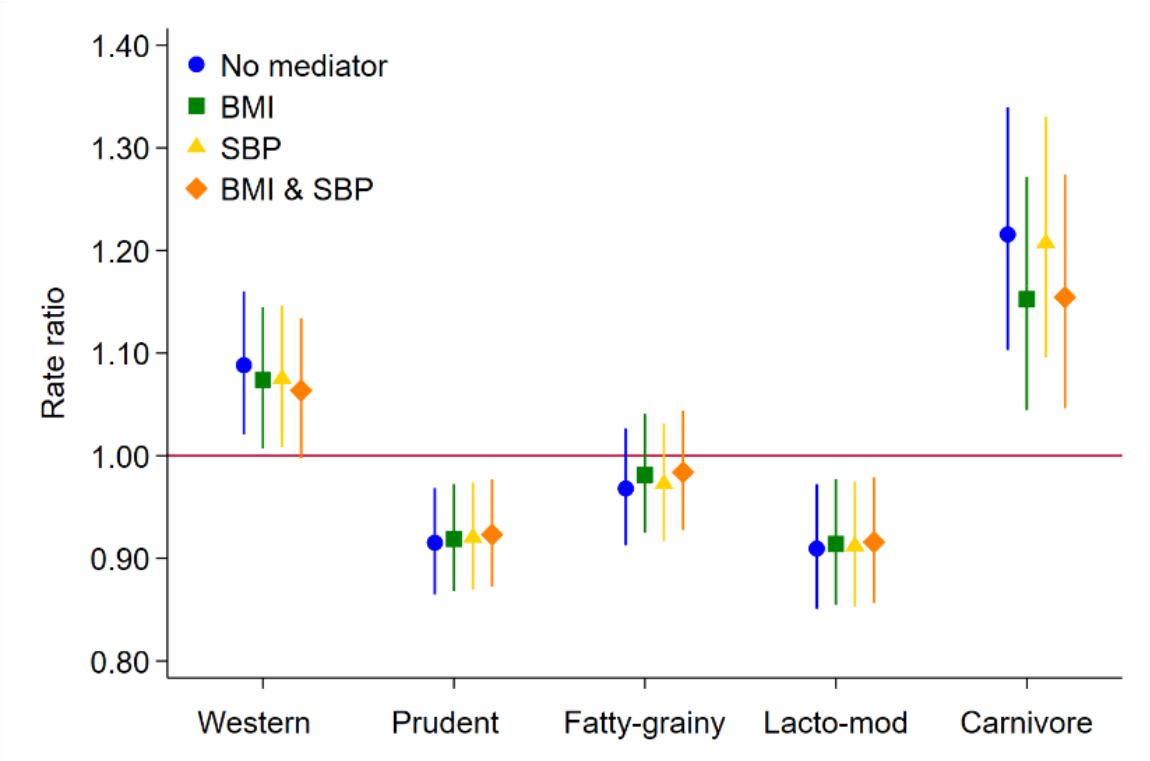
Notes: Width of the curves represent 95% confidence intervals. BMI: Underweight=<18.5, Normal=18.5-25, Overweight=25-30, Obese=>30, SBP: Normal=<120, at risk=120-140, Hypertension=>140. BMI and SBP are measured at baseline.

Seemingly unrelated nested Poisson regression models showed that the Western and Carnivore DPs were positively associated with ACM, whereas the Prudent, Fatty-grainy and Lacto-mod DPs were inversely association with ACM (**Figure 3 3**). In the mediator unadjusted model, ACM and Western DP were associated with a hazard ratio (HR) of 1.09 [95%CI 1.02, 1.16], Prudent 0.91 [95%CI 0.86, 0.96], Fatty-grainy 0.97 [95%CI 0.92, 1.03], Lacto-mod 0.91 [95%CI 0.85, 0.97], and Carnivore 1.20 [95%CI 1.09, 1.32]. After adding mediators separately, coefficients for all DPs were attenuated suggesting the presence of mediation. In the BMI and SBP adjusted models, ACM and Western DP were associated with a hazard ratio (HR) of 1.06 [95%CI 1.00, 1.13], Prudent 0.92 [95%CI 0.87, 0.98], Fatty-grainy 0.98 [95%CI 0.93, 1.04], Lacto-mod 0.92 [95%CI 0.86, 0.98], and Carnivore 1.15 [95%CI 1.05, 1.27]. Fully adjusted SEM models revealed that a 5 unit difference in BMI was associated with a 1.05 [95%CI 1.04, 1.07] higher risk of ACM, on average (**TableA3 3, Panel A**). Similarly, a 10 unit difference in SBP was associated with a 1.08 [95%CI 1.06, 1.10] higher risk of ACM. The Western and Carnivore DPs were positively associated with BMI and SBP (**TableA3 3, Panel B**). At the same level of energy intake, among all DPs, high intake of the Carnivore DP was associated with the highest BMI level, whereas the Western DP was associated with the highest SBP level. Conversely, higher intakes of other DPs were associated with lower BMI and SBP. Coefficients on DPs in the SEMs were highly consistent with the nested Poisson regression models (**TableA3 3, Panel A**).

Figure 3 4 shows mediation estimates from the nested Poisson and SEM models using BMI and SBP measurements taken at baseline. For the Western DP, PMM was estimated to be 15.8% through BMI, 14.1% through SBP and 27.1% together (Panel A). For the Prudent DP, PMM was 4.4% through BMI, 6.3% through SBP and 9.7% together. For the Lacto-mod DP, PMM was 5.2% through BMI, 2.9% through SBP and 7.3% together. For the Carnivore DP, PMM was 27.3% through BMI, 3.6% through SBP and 26.4% together. Although, mediation coefficients for the Fatty-grainy DP were large 49.9%, they were very imprecise with CIs crossing 0 and 100%. Panel B shows PMM estimates from SEM models. PMMs were larger for the

Western (34.1%) and Prudent (12.8%) DPs, but were qualitatively similar for other DPs. In sensitivity analyses, we gauge differences in PERM and PMM when using BMI and SBP at follow-up visits instead of baseline (**Figure 3 5**). Across all DPs, we see significant attenuation in the PMM and PERM estimates. PMMs were smaller for the Western (19.3%), Prudent (5.9%), and Carnivore (12.9%) DPs, but were directionally similar to results with measurements taken at baseline. Next, PMM and PERM without adjustment for energy intake using baseline BMI and SBP showed high sensitivity for the Prudent, Fatty-grainy and Lacto-mod DPs (**Table A3 5**). Mediation though BMI and SBP was absent for these DPs without accounting for energy intake.

Figure 3 3: Changes in dietary effects on all-cause mortality with and without adjustment for body mass index and systolic blood pressure

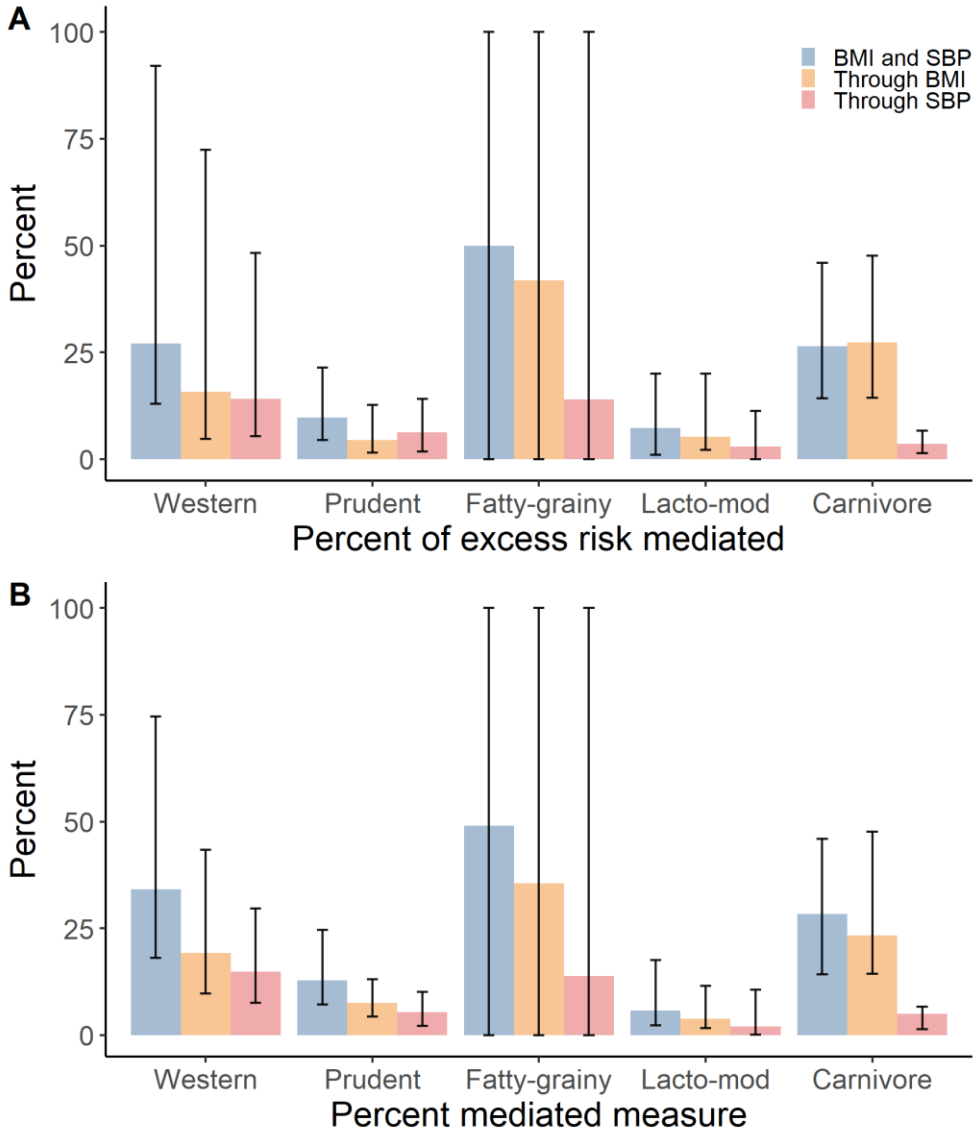


Notes: Models are adjusted for age, sex, race, income, education, smoking, alcohol, physical activity, calorie intake, study cohort fixed effects and birth cohort fixed effects. Bands represent 95% confidence intervals. Each marker represents a separate regression.
Diet scores range from 0 to 1. The coefficient is interpreted as difference in risk comparing the lowest percentile of the diet to the highest percentile, assuming a linear risk curve.
BMI=Baseline body mass index, SBP=Baseline systolic blood pressure.
The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry.
The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups.
The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Effect modification analyses revealed that for individuals with normal BMI and SBP, the Western DP was strongly positively associated with ACM, HR 1.76[95%CI 1.13, 2.73]; and the Prudent DP strongly negatively associated, 0.50 [95% CI 0.33, 0.75] (**Table 3 2**). Effect modification was only marginally statistically significant for the Carnivore DP for BMI, and only marginally significant for the Prudent DP for SBP. Bias

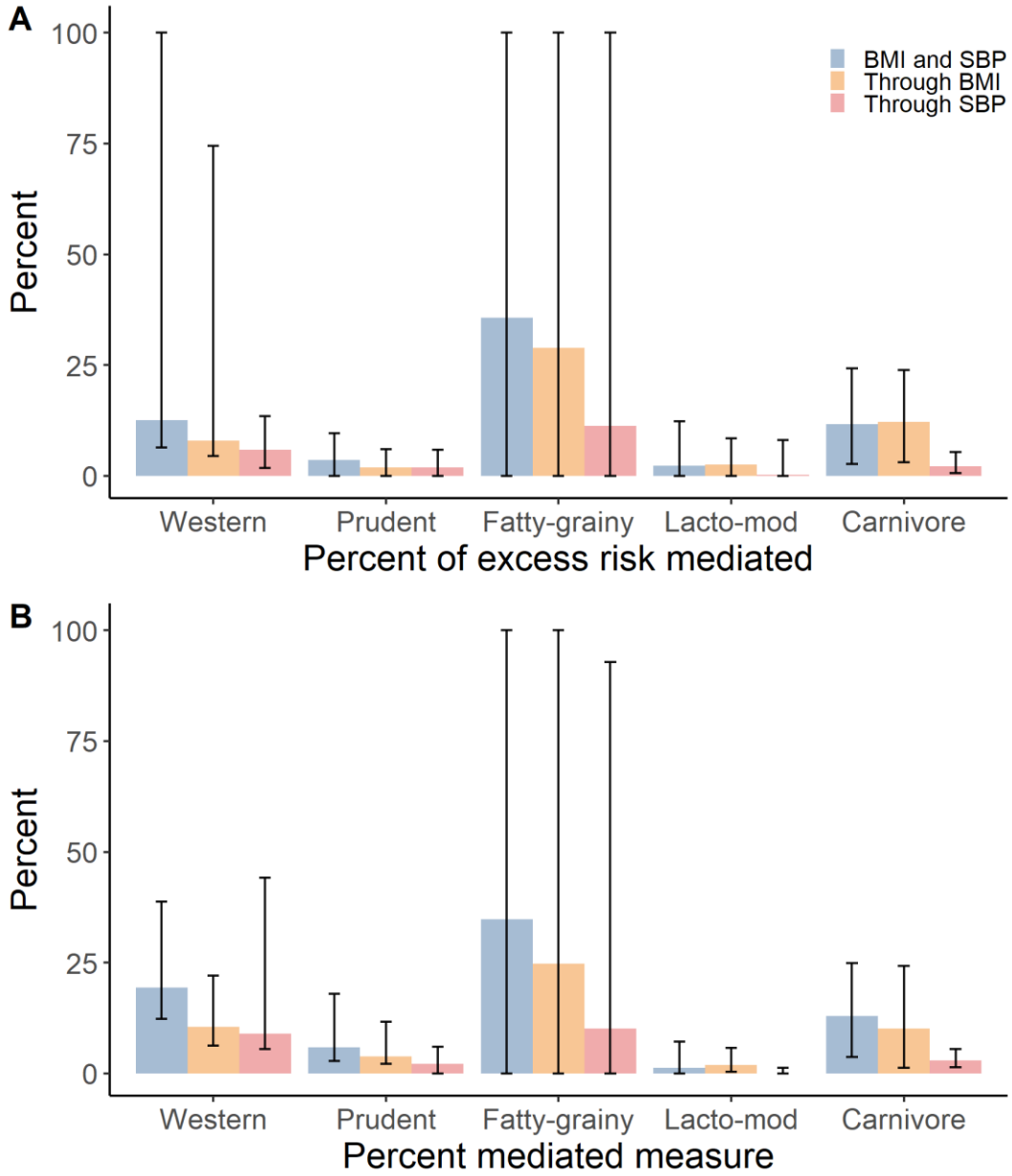
correction analyses showed that an unmeasured confounder would need to have an RR of 1.35, 0.5, and 1.95 to nullify the direct effects of the Western, Prudent and Carnivore DPs, respectively.

Figure 3 4 Percent of excess risk mediated and percent mediated measures for body mass index (BMI) and systolic blood pressure (SBP) measured at baseline



Percent of excess risk mediated calculated post-hoc from seemingly unrelated nested Poisson models. Percent mediated measure calculated post-hoc from structural equation Poisson models. Percentile based CIs are calculated using the bootstrap method. CIs are truncated at 0 and 100%. Models are adjusted for age, sex, race, income, education, smoking, alcohol, physical activity, calorie intake, study cohort fixed effects and birth cohort fixed effects. . Diet scores range from 0 to 1. The coefficient is interpreted as difference in hazard comparing the lowest percentile of the diet to the highest percentile, assuming a linear risk curve. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Figure 3 5 Percent of excess risk mediated and percent mediated measures for body mass index (BMI) and systolic blood pressure (SBP) measured at follow-up visits



Percent of excess risk mediated calculated post-hoc from seemingly unrelated nested Poisson models. Percent mediated measure calculated post-hoc from structural equation Poisson models. Percentile based CIs are calculated using the bootstrap method. CIs are truncated at 0 and 100%. Models are adjusted for age, sex, race, income, education, smoking, alcohol, physical activity, calorie intake, study cohort fixed effects and birth cohort fixed effects. . Diet scores range from 0 to 1. The coefficient is interpreted as difference in hazard comparing the lowest percentile of the diet to the highest percentile, assuming a linear risk curve. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Table 3 2: Effect modification by body mass index and systolic blood pressure for the association between dietary patterns and all-cause mortality

	Western		Prudent		Fatty-grainy		Lacto-mod		Carnivore	
	HR	**95%CI**	**HR**	**95%CI**	**HR**	**95%CI**	**HR**	**95%CI**	**HR**	**95%CI**
Diet	1.76	[1.13,2.73]	0.50	[0.33,0.75]	1.19	[0.80,1.77]	0.99	[0.63,1.56]	0.89	[0.48,1.63]
BMI, 5 units	1.11	[1.07,1.14]	1.06	[1.03,1.09]	1.07	[1.04,1.10]	1.09	[1.06,1.13]	1.02	[0.98,1.07]
SBP, 10 units	1.07	[1.05,1.08]	1.04	[1.03,1.06]	1.07	[1.05,1.08]	1.05	[1.04,1.07]	1.06	[1.04,1.09]
Diet x BMI	0.96	[0.91,1.02]	1.04	[0.99,1.09]	1.01	[0.96,1.05]	0.98	[0.93,1.04]	1.08	[1.00,1.15]
Diet x SBP	0.98	[0.95,1.00]	1.03	[1.01,1.06]	0.98	[0.96,1.01]	1.00	[0.97,1.03]	0.99	[0.95,1.03]

Diet scores range from 0 to 1. The coefficient is interpreted as difference in risk comparing the lowest percentile of the diet to the highest percentile, assuming a linear risk curve. BMI=Body mass index per 5 unit increment. SBP=Systolic blood pressure per 10 mmhg. Models are adjusted for age, sex, race, income, education, smoking, alcohol, physical activity, calorie intake, height, cohort fixed effects and birth cohort fixed effects. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

DISCUSSION

In this study we aimed to quantify the direct and indirect effects of DPs on ACM associated with BMI and SBP. We identified four DPs that consistently explained variation across six cohorts and an additional DP in the WHI. Since our analyses adjusted for calories, our findings speak to diet quality at the same levels of energy intake. Concurring with research from previous studies, diets high in meat were found to be high risk for ACM, while those comprised of nutrient dense plant foods were consistently and robustly associated with lower ACM risk[71]. The main finding of this study is that BMI and SBP significantly mediate the effects of the widely researched Western and Prudent DPs. The most important mediator was BMI, especially for the Western and Carnivore DPs, accounting for about one fourth of the risk. However, the direct effects of DPs remained large and significant, indicating that interventions targeting BMI and SBP can only provide a fraction of the benefits associated with adopting healthy DPs[110,111].

Mediation analysis is a developing field in nutritional epidemiology and relatively few studies have investigated the role of metabolic factors in mediating dietary risk for fatal endpoints. For example, a multi-cohort study found that type-2 diabetes (T2D) mediated risks from inflammatory diets to pancreatic cancer development[112]. Another recent cohort study found that the Mediterranean diet reduced the risk of T2D, and BMI mediated two-thirds of this effect[113]. Further, a seminal study that focused only on BMI and metabolic mediators found evidence of significant mediation for coronary heart disease and stroke[23]. To our knowledge, this is the first study to consider mediation of dietary risk through BMI and SBP to ACM.

Our mediation results have strong empirical support and biological plausibility. One of the largest individual level analysis of the BMI-mortality association estimated that the hazard for ACM increases by 30% for every 5 point increase in BMI[114]. A meta-analysis further estimated that obesity is associated with an 18% higher risk of ACM[115]. Similarly, numerous studies have demonstrated causal links between high BP and fatal endpoints[116–119]. For biological mechanisms, it is established that unhealthy diets linked to adipose tissue

deposition can lead to increased sympathetic nervous system activity, impaired compliance of large conduit arteries, activation of local angiotensin II in heart and arteries, increased cardiac output, abnormal pressure natriuresis, and increased tissue angiotensin II in kidneys and adrenal glands[120]. Yet, our study demonstrates that unhealthy diets have adverse effects that go well beyond the pathways described here, indicated by substantial direct effects on ACM. Further research is needed to unpack the multiple pathways through DPs influence fatal endpoints.

Encouragingly, our SEM analyses showed that, at the same levels of energy intake, the Prudent, Fatty-grainy and Lacto-mod DPs were associated with significantly lower BMI and SBP at baseline and follow-up. Of these, Fatty-grainy had the largest indirect effects through BMI, but wide confidence intervals on the PERM and PMM for this DP limit the precision of these results. Qualitatively, the Prudent DP is similar to the Mediterranean diet, which is associated with significant weight loss in long-term randomized trials[121]. Importantly, the Western and Carnivore DPs had the largest and significant associations with higher BMI and SBP, suggesting that replacement with the Prudent DP could lead to significant benefits in terms of lower BMI, SBP and ACM risk. However, a caveat based on our sensitivity analyses is that calories must be held at optimal levels because healthy DPs may be associated with weight gain at high levels of energy intake. Measurement timing also plays a significant role for BMI and SBP. Our results indicated that follow-up measurements showed lower mediation compared to baseline measurements of BMI and SBP. These results may be due to DPs changing overtime, participants changing their behavior, or that BMI and SBP are naturally time varying. More research is needed on identifying causal effects of DPs on BMI and SBP over several time points.

Our study has potentially important policy implications. The large direct effects of DPs, regardless of BMI or SBP, suggest that healthy diets can prevent early mortality even if they do not lead to improvements in weight or BP. Evidence suggests that interventions targeting SBP have proved more successful than those targeting BMI[122–124]. Dietary recommendations towards adoption of healthy DPs are likely to achieve reductions in ACM and be supportive of BMI and SBP in healthier ranges. In fact, our study shows that the Western DP is associated with a 70% higher risk of early mortality in participants with BMI and SBP in the healthy range. Conversely, the Prudent DP was associated with a 50% lower ACM risk in such metabolically healthy individuals. For the Carnivore DP, there was significant interaction with BMI, suggesting that this DP is associated with higher mortality risks at higher BMI levels. These results are cautionary because the Carnivore DP has recently gained popularity as a weight loss diet, but its long term consequences are unknown[72].

Our study is the largest pooling mediation analysis of diet and ACM, with 120 thousand participants, and over 18000 mortality events. This large sample size allowed us to study the extent of mediation, and if DPs interact with mediators. Further, similarities in the results obtained from the difference and product methods lent confidence to the accuracy of our mediation estimates. Moreover, we accounted for a large range of sociodemographic, economic and behavioral factors to minimize the possibility of potential confounding.

Limitations of the study include observational design and measurement error. Mediation analyses are susceptible to confounding on the exposure and mediators. An unbiased study would need to randomize individuals to both diets and levels of BMI and SBP, which is unrealistic in most settings. Thus, we relied on observational prospective cohort studies[79]. This comes at the cost of susceptibility to bias from residual and

unobserved confounding, so ultimately evidence on DPs about mediation must be evaluated with caution[80]. We minimized the presence of confounding by using a large set of controls, but confounders might have been measured with error, which leads to residual confounding (e.g., smoking and income). Further, unobserved confounders pose threats to the validity of mediation estimates. We assessed the magnitude of such threats with bias correction values. For major DPs, these values indicated that an unobserved confounder would need to have relative risks of 1.5 to 2 to nullify our results. We know of no such confounder, given that only smoking had such high risks in our sample. Additionally, since this study relies on Food Frequency Questionnaires (FFQs), the measured intake of food groups is prone to random error due to non-differential misclassification, and can lead to biased estimates, generally towards the null[81]. Evidence indicates that a high degree of attenuation can be attributed to measurement error in FFQs[82]. To this end, the pooled cohort design provided greater statistical power thereby endowing the study with the ability to detect significant assoications[81]. Another concern is that DPs are not time invariant, thus, a person's lifetime risk to any DP depends on the lagged effects of the cumulative exposure to different DPs consumed throughout life. Our estimates may be biased due to the inability to account for cumulative exposure or lagged effects.

CONCLUSION

To conclude, we found that BMI and SBP may have a small but significant mediating effect in the relationship between widely researched Western and Prudent DPs and risk of ACM. Interestingly, we found a large direct effect of the several DPs, regardless of BMI or SBP, suggesting that healthy diets can prevent early mortality even if they do not lead to improvements in weight or BP. These results are clinically relevant for public health, in particular for encouraging a Prudent-like dietary pattern adoption for primary prevention of fatal endpoints and improving mediators. Our results are also cautionary about fad diets like the Carnivore DP.

CHAPTER 4:
CAN DIETARY PATTERNS EXPLAIN MORTALITY DIFFERENCES ACROSS RACIAL/ETHNIC GROUPS IN THE UNITED STATES? A POOLED ANALYSIS OF SIX PROSPECTIVE COHORTS

ABSTRACT

Objectives: To decompose the role of dietary patterns (DPs) and other behavioral risk factors to close racial/ethnic differences in all-cause mortality (ACM) in the United States.

Participants and setting: Secondary analysis of pooled datasets from six prospective cohorts including Atherosclerosis Risk in Communities Study, Cardiovascular Health Study, Multi-Ethnic Study of Atherosclerosis, Women's Health Initiative Study, NHANES I Epidemiologic Follow-up Study, Reasons for Geographic and Racial Differences in Stroke. Non-Hispanic White (NHW) (N=100667), Black (N=16255), and Hispanic/Asian/Other (N=8595) participants were followed-up for 1250631, 178901, 85655 person-years, respectively.

Outcomes: Outcomes were all-cause deaths (binary) and the survival time until death/censoring (interval between start of the study and death or loss to follow-up).

Exposures: Principal Component Analysis of observed intake of food groups was used to estimate DPs coded as continuous variables ranging from 0 (low) to 1 (high). Five dietary patterns were estimated. The Western pattern was high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' pattern was high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' pattern was high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto' pattern was high in milk and moderate to low on all other food groups. The 'Carnivore' pattern was high in red and processed meats, poultry, fish, eggs and moderate on other food groups. Other exposures were age, sex, race, education, income, smoking, drinking, physical activity, body mass index and systolic blood pressure.

Design: Poisson models were used estimate race/ethnicity specific risks for DPs. The Oaxaca-Blinder decomposition was used quantify the proportion of exposure specific risk that explains the mortality gap due to diet, income, smoking, physical activity and smoking. Population attributable fractions were used calculate race/ethnicity specific ACM reductions based on hypothetical scenarios.

Results: DPs were significantly associated with ACM across all racial/ethnic groups. The Western DP had high ACM risks among Hispanic/Asian/Others RR 2.0 [1.95, 2.1]. The Carnivore DP had high ACM risks among Black individuals RR 1.7 [1.2, 2.3]. The Prudent DP was robustly associated with lower risk of mortality among all racial/ethnic groups. Dietary patterns explained a small part of the White-Black mortality rate difference, though income was the primary determinant for the White versus Black mortality advantage. Eliminating the Western DP could reduce the mortality burden among Hispanic/Asian/Others by 30%. Eliminating the Carnivore DP could reduce the mortality burden among NHW and Black individuals by 7.4% and 18.3% respectively. Adoption of the Prudent diet by Black individuals would reduce their mortality burden further by 7.2%. Exposing all groups to high physical activity would produce the most reductions in mortality (33.7-41.7%), on average.

Conclusions: Replacement of unhealthy DPs with healthy ones could yield significant equitable improvements in mortality rates across all racial/ethnic groups in the US. Income disparities currently explain half the mortality differences between White and Black individuals in the US. Holistic improvements in physical activity, smoking and drinking are needed across all race/ethnic groups for further equitable reductions in mortality rates.

Introduction

Racism is an organized system based on the categorization and ranking of social groups into races that undervalues, disempowers, and differentially allocates societal opportunities and resources to racial groups regarded as inferior[125]. Racial and ethnic minorities in the US experience disparities in their health and health care that stem from interrelated factors, including discrimination, social drivers of health, health care access and quality, and individual behavior[126]. Health related racial/ethnic disparities begin with childhood health conditions such as prematurity, obesity, and asthma[127]. In later life, these racial/ethnic disparities manifest into differential rates of morbidity, mortality, and overall well-being[128]. These disparities persist, despite decades of focused research to ameliorate them[129].

Differences in life expectancy at birth (LE) between non-Hispanic White and non-Hispanic Black individuals (henceforth referred to as White and Black, respectively) have decreased considerably in the US since 1950[130]. In 2020, average LE was 78.5 and 74.9 years, among White and Black individuals, respectively[131]. The literature suggests that pervasive racial/ethnic differences in the major social and behavioral risk factors for morbidity and premature mortality[132], such as diets[133–135], income[136], education[137–139], substance abuse[140,141], smoking[142,143], and physical activity[144], among others, may contribute to explaining mortality differences between racial/ethnic groups. Notably, these causes may share the common underlying cause of various forms of structural and interpersonal racism[125]. Moreover, two problematic mortality trends have emerged since 2000. First, Black individuals have experienced a more rapid increase in obesity rates[145], suggesting that future mortality trends for Black individuals may not be as favorable as they have been in the past. Second, White individuals have witnessed an increase in mortality in midlife, a trend not seen in other Western countries[146]. Thus, it is timely to examine the relative importance of modifiable risk factors for mortality across racial/ethnic groups in the US, and evaluate reductions in mortality that could be achieved if gaps in exposure to preventable risk factors were reduced. Indeed, racism itself is an underlying preventable risk factor that has multiple downstream effects.

While the cultural appropriateness of a uniform dietary recommendation has recently faced criticism[147], it is well established that multicultural healthy and high quality diets share common traits[147]. Studies have documented mixed trends in diet quality in the US, and Black individuals have persistently shown lower diet quality than other groups[87,148–151]. However, beyond a handful of such studies, to the best of our knowledge, the role of dietary differences across racial/ethnic groups in relation to all-cause mortality (ACM) have not been systematically evaluated with multi-cohort data[25,152,153]. The nexus between diet-related public health strategies, race/ethnicity and mortality is an urgent area for scientific research. To address this research gap, in this study, we decompose the role of dietary patterns (DPs) and other behavioral risk factors to close racial/ethnic disparities in ACM in the United States. The study focuses on five DPs which includes the widely researched Western and Prudent diets[71]. We assess whether common DPs currently explain mortality differences in race/ethnicity. We further predict mortality gains within racial/ethnic groups that may be associated with improvements in DPs and other risk factors.

METHODS

Data and outcome

We use data from six prospective cohorts for the analysis: Atherosclerosis Risk in Communities Study (ARIC), Cardiovascular Health Study (CHS), Multi-Ethnic Study of Atherosclerosis (MESA), Women's Health Initiative Study (WHI), NHANES I Epidemiologic Follow-up Study, Reasons for Geographic and Racial Differences in

Stroke (REGARDS). We exclude the Framingham cohort for this analysis because it only sampled White individuals.

The outcome of interest for this study is ACM (deaths that occur in a population, regardless of the cause). Participants were followed-up from baseline to the last round of examination, event of death or when they could be contacted for follow-up. A participant's time since baseline to death or last follow-up was a calculated in days. The final outcomes are event of a death (binary) and the survival time until death occurs (interval between start of the study and death or loss to follow-up). For those patients that are alive at the end of the observation period, the actual time-to-event is unknown, and will be treated as a case of right censored data. Details of data sets and outcome assessment in the individual cohorts are described in chapter 2. We mitigated threats from reverse causation by excluding participants who died within 3 years of enrollment[23].

Racial and ethnic groups

We group racial/ethnic groups into three categories: (1) Non-Hispanic White, (2) Black and (3) Hispanic/Asian/other. Hispanics, Asians and other ethnicities had smaller sample sizes relative to White and Black participants. Since Hispanics and Asians live longer than White and Black individuals in the US, on average, it is analytically practical to bin the ethnicities into one group in our analyses[154]. Moreover, some cohorts do not collected data on all racial/ethnicities thereby precluding further sub-categorization.

Dietary patterns

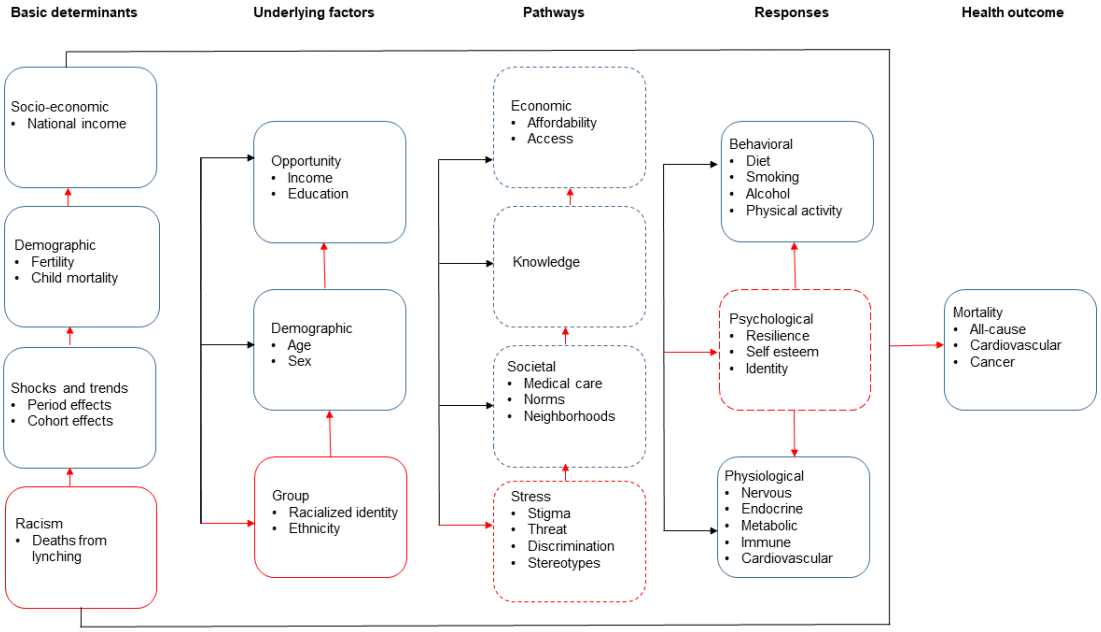
DPs are approximated using principal component analysis (PCA) where data on intake of several food groups are compressed into DP scores that reflect individual diet types that are commonly consumed in the US. Food groups used in the analyses are fruits, starchy vegetables, fibrous vegetables, whole grain, refined grain, legumes, oils/fats, nuts and seeds, milk, cheese, yogurt, fish/shellfish, poultry, processed meats, eggs, and red meat. PCA derived DPs have to been shown to have high stability, suggesting that they are well suited for a pooled cohort design[30]. PCA explains a set of correlated food groups with a new set of comprehensive indexes (principal components) that are uncorrelated and retain the largest amount variance in the original sample as possible[60]. After predicting component scores (DPs), the scores are re-scaled ranging from 0 to 1, with 0 and 1 representing lowest and highest consumption levels of a specific pattern, respectively. PCA is conducted separately for each cohort and the scores are pooled for the analysis. See chapter 2 for details.

Covariates

This study anchors itself on the framework proposed by Williams and Mohammed in their seminal article on racism and health[125]; starting with the notion that racial discrimination persists in contemporary society, and it is pervasive across institutional mechanisms and processes. **Figure 4 1** adapts the framework for the current study, organizing risk factors that go from the most upstream towards more proximal determinants that eventually lead to mortality outcomes (See methods appendix for details). Per the framework, we adjust for a set of covariates including age (years), sex (male/female), income (inflation adjusted US$), education (years), physical activity (low/medium/high), smoking (current and pack-years), drinking (mild/moderate/heavy) and calorie intake (kcal). We also include cohort fixed effects to account for measurement differences across cohorts (survey). Prior research on racism suggests that further adjustment should be made for early childhood factors that may confound the relationship between race and mortality in adulthood[129,155–157] (see methods appendix for details). To this end we adjust for birth-cohort fixed effects,

birth year specific Gross Domestic Product, Infant Mortality Rate, Total Fertility Rate, and total cases of lynching to account for ecological level factors that capture aggregate economic, developmental and racial climate that the participant was exposed to in the past. Falling on the causal pathway from diet to mortality, body mass index (BMI) and systolic blood pressure (SBP) are used in sensitivity analyses but not in primary regressions.

Figure 4 1: A framework for the study of racism, diet quality and mortality



Notes: Red arrows indicate pathways through which race affects health. Red boxes indicate the presence of an important aspect of racism. Dotted outline indicates factors that are not adjusted for in the study. The framework asserts that racism is a dominant force that shapes the relationships between multiple variables at various levels of effect. The effects of race cannot be easily disentangled from the effects of causal pathways or mediating variables. See methods appendix for details.

Statistical analysis

We utilize the Oaxaca-Blinder decomposition (OBD) technique to detect the magnitude of differences in ACM that would be reduced if individuals from minority race/ethnicities had the same average levels of protective diets and other factors compared with White individuals[158]. Originally developed by economists, OBD is now being applied in health research to examine racial/ethnic differences across various outcomes[159–162]. OBD enables us to decompose racial/ethnic disparities in mortality into differences in the underlying characteristics (i.e., means of the independent variables) and differences that are not explained by underlying characteristics (i.e., differences across groups in regression coefficients)[161].

The OBD is carried out in two steps. First, we run separate regressions for each racial/ethnic group in the sample. For an analysis comparing White versus Black individuals, these are represented as,

$$\bar{Y}_b = \hat{B}_{b0} + \sum_{j=1}^{N} \bar{X}'_{bj}\hat{B}_{bj} \dots (1)$$

$$\bar{Y}_w = \hat{B}_{w0} + \sum_{j=1}^{N} \bar{X}'_{wj}\hat{B}_{wj} \dots (2)$$

Where the subscript *b* corresponds to Black individuals, *w* to White individuals, $\bar{Y}$ is the mean mortality rate, $\hat{B}_0$ is the intercept, and $\hat{B}_j$ is a column vector of estimated slope coefficients for the array of *j* explanatory variables, $\bar{X}_j'$. The primary exposure of interest in set $\bar{X}_j'$ is the diet score along with other important explanatory variables. We use semi-parametric Poisson models assuming the reference population's death rate over a specified interval of time is assumed constant[65]. The Poisson estimates rate ratios, and is more amenable to OBD compared to the conventional Cox-PH model for incorporating cluster robust standard errors.

In the second step, the difference in mortality between White and Black individuals, $\bar{Y}_b - \bar{Y}_w$ is decomposed as,

$$\bar{Y}_b - \bar{Y}_w = \left(\hat{B}_{b0} - \hat{B}_{w0}\right) + \sum_{j=1}^{N} \bar{X}_{wj}'(\hat{B}_{bj} - \hat{B}_{wj}) + \sum_{j=1}^{N} (\bar{X}_{bj}' - \bar{X}_{wj}')\,\hat{B}_{bj} \ldots (3)$$

The term $\sum_{j=1}^{N} (\bar{X}_{bj}' - \bar{X}_{wj}')\,\hat{B}_{bj}$ represents the 'explained part' that is caused by differences in the means of the risk factors in the model. The extent to which specific risk factors contribute to differences in mortality across groups can be calculated by evaluating this part. This is component that is most policy relevant and the focus of the analysis.

The terms $\left(\hat{B}_{b0} - \hat{B}_{w0}\right) + \sum_{j=1}^{N} \bar{X}_{wj}'(\hat{B}_{bj} - \hat{B}_{wj})$ represent the 'unexplained part', which is caused by differences in the intercepts and coefficient estimates across the two regression equations.

Population attributable fractions

We extend the OBD analyses by calculating the population attributable fractions (PAFs) at hypothetical scenarios for each group to gauge the reduction in mortality that would be achieved if all racial/ethnic groups had the same hypothetical improved levels of exposure to risk factors[163]. The scenarios include (1) the highest level consumption of DPs negatively associated with ACM; (2) the lowest level consumption of DPs positively associated with ACM; (3) smoking at 5% prevalence (based on the lowest level observed across the cohorts); (4) physical activity set at the highest tertile; (5) very low consumption or abstaining from alcohol (bottom tertile); (6) real income at $100,000 per head, on average (threshold based on the 2020 income distribution of the US where 33% of households earned more than $100,000 per year[164]); (7) mean BMI at 22.5 kg/m^2 and (8) mean SBP at 100 mmhg. Unlike OBD, which compares two groups at a time, this analysis is performed with all racial groups together. The estimates are made with all race/ethnicities in the same the model and interaction terms between exposures and race/ethnicities.

RESULTS

Across the cohorts, the majority of participants were White, and were followed-up for 1.26 million person-years with 15271 deaths events, at an average mortality rate of 0.012 per year [95% CI 0.0119; 0.0123] (**Table4 1**). Compared to Whites, death rates among Black individuals (0.015 [95%CI 0.0143, 0.0154]) were significantly higher, but were lower among other race/ethnicities (0.006 [95%CI 0.0059, 0.0070]). Compared to individuals of other races/ethnicities, a higher percentage of Black individuals in the sample were male (23.2%) and smokers (16.7%). Compared to individuals of other races/ethnicities, fewer Black individuals were physically active (22.8%) and had high alcohol intake (22.8%). Black individuals further had the highest mean BMI (30.3), SBP (132.6) and the lowest income ($57,700). Hispanic/Asian/Others were similar Whites

on most covariates, except for income ($68200 v $79300), education (13.2 v 14.4 years) and height (158.3 v 164 cm), where White individuals had an advantage.

PCA performed individually on each cohort revealed high consistency in the components present across the samples (**Tables A2 3 to A2 9**). Across most cohorts four components had Eigenvalues greater than 1. Cohort specific mean dietary intake of food groups used in PCA are reported in **Table A2 11.** **Table A10** shows a heat map depicting how food groups are correlated (Pearson's Rho) with specific DPs and aims in classifying and naming the DPs based on these correlations in the pooled data. Broadly, the 'Western' DP (not observed in the REGARDS cohort) was high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP was high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP was high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto' DP (not observed in the REGARDS and NHANES cohorts) was high in milk and moderate to low on all other food groups. The 'Carnivore' DP was high in red and processed meats, poultry, fish, eggs and moderate on other food groups (only observed in the WHI cohort). Individuals with higher scores on a specific DP are likely to consume more foods correlated with that DP.

There were differences across racial/ethnic groups on the consumption of the highest levels DPs measured by highest quintile of the DP score (**Table4 1**). Hispanic/Asian/others (14.9%) had a smaller proportion of participants consuming the top quintile of the Western DP, relative to White (19.5%) and Black (18.7%) participants. This distribution was also similar for the Fatty-grainy DP. White individuals had the highest share of participants consuming the top quintile of the Lacto-mod DP (20.7) compared to Black (10.3%) and Hispanic/Asian/other (13.5%) participants. Black participants had the highest share of the top quintile of the Carnivore DP (26.1%). Prudent was the only DP that had a similar racial/ethnic distribution in the top consumption quintile.

Table 4 1 Mortality events and baseline summary statistics for covariates

	NHW (N=100667)		**Black (N=16255)**		**Hispanic/Asian/Other (N=8595)**	
	Mean/%	**95% CI**	**Mean/%**	**95% CI**	**Mean/%**	**95% CI**
All-cause mortality						
Person- years	1250630.5		178900.9		85655.1	
Deaths	15096		2658		550	
Rate	0.012	[0.0119; 0.0123]	0.015	[0.0143; 0.0154]	0.006	[0.0059; 0.0070]
Participants consuming highest quintile of dietary patterns						
Western (N=104,903), %	19.5	n=16,924	18.7	n=1,819	14.9	n=1,278
Prudent (N=124,599), %	19.2	n=19,166	18.9	n=3,047	19.1	n=1,643
Fatty-grainy (N=124,599), %	19.6	n=19,554	18.5	n=2,993	14.1	n=1,210
Lacto-mod (N=97,505), %	20.7	n=16,545	10.3	n=915	13.5	n=1,154
Carnivore (N=82,206), %	19.7	n=13,751	26.1	n=1,507	18.3	n=1,181
Covariates		**N=100667**		**N=16255**		**N=8595**
Age, years	62.7	[62.7,62.8]	61.3	[61.2,61.5]	62.1	[61.9,62.3]
Male, %	14.1	n=14,205	23.2	n=3,764	11.8	n=1,013
Inflation adjusted income, 1000 $	79.3	[79.0, 79.5]	57.7	[57.2,58.3]	68.2	[67.4,69.1]
Education, years	14.4	[14.4,14.4]	13.6	[13.5,13.6]	13.2	[13.1,13.2]
Current smoker, %	8.8	n=8,855	16.7	n=2,721	6.6	n=570
Pack-years, 10 years	0.5	[0.5,0.5]	0.7	[0.7,0.7]	0.2	[0.2,0.2]
Moderate/high alcohol	29.1	n= 29,342	22.8	n= 3,703	27.2	n=2,338

consumption, %						
Highly active, %	32.5	n=32,669	25.4	n=4,133	28.7	n=2,470
Energy intake, 100 kcal	19.0	[19.0,19.1]	19.4	[19.3,19.6]	17.6	[17.5,17.8]
Height, cm	164.0	[164.0,164.1]	165.8	[165.7,166.0]	158.3	[158.2,158.5]
Body mass index, kg/m^2	26.9	[26.9,27.0]	30.3	[30.2,30.4]	26.8	[26.7,26.9]
Systolic blood pressure, mmhg	127.3	[127.2,127.4]	132.6	[132.3,132.9]	127.2	[126.8,127.6]

Notes: N=Denominator; n=numerator; Denominators for dietary patterns depend on the number of cohorts they were observed in. NHW=Non-Hispanic White. Other=Hispanic/Asian/Others. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables (not present in REGARDS cohort). The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups (not present in REGARDS and NHANES cohort). The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups (only observed in the WHI cohort).

Figure 4 2 shows the results from the first stage of the OBD. The Western DP was positively associated with ACM among Whites, Rate Ratio (RR) 1.08 [95%CI 1.03,1.13]; and strongly positively associated among Hispanics/Asians/Others RR 2.04 [95%CI 1.95,2.12]. The Prudent was negatively associated with ACM for all race/ethnicities but lowered risk most among Black individuals RR 0.88[95%CI 0.80,0.97]. The Fatty-grainy was also negatively associated with ACM risk in all groups, but only significant among Black individuals. Similarly, the Lacto-mod DP was significantly negatively associated with ACM among White individuals RR 0.91 [0.87,0.96]. The Carnivore DP has positively associated with ACM across all groups, but had the largest effect sizes among Black individuals RR 1.7 [95%CI 1.18,2.32] followed by Whites RR 1.16 [1.05,1.29]. Differences in regression coefficients did exist for other covariates, though they were not large and had overlapping CIs. Notably, males had much lower ACM risk among Hispanics/Asian/Others. Current smokers had the highest risks among all groups with RRs ranging between1.44 and 1.76. Moderate-high physical activity and income, were both strongly and negatively associated with ACM. High alcohol intake was positively associated with ACM and had the largest risks among Hispanics/Asian/others 1.29 [95%CI 1.13,1.48].

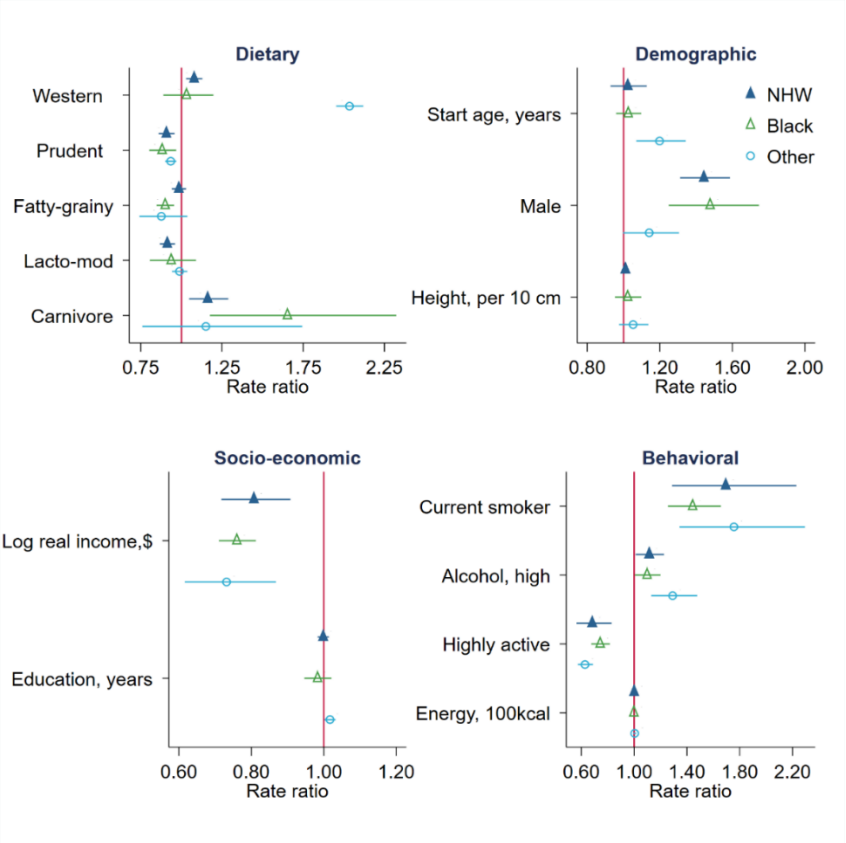
Figure 4 3 shows the results from the second stage of the OBD which quantifies the proportion of the racial/ethnic mortality gap that was attributable to differences in mean covariate levels. In the White versus Black decomposition, DPs could explain part of the higher Black mortality rates, though CIs were wide for most DPs. A small but significant proportion of the Black mortality disadvantage existed due to Black individuals consuming higher levels of the Carnivore DP (2.8%) and lower levels of the Fatty-grainy DP(2.8%) (**Panel A**). Apart from DPs, income was the primary determinant for the White versus Black mortality advantage, accounting for 50.9% of the difference (**Panel B**). Higher smoking rates among Black individuals explained 21.1% of the mortality gap. Together, income, smoking and physical activity explained more than 80% of the mortality difference. For other race/ethnicities, no DPs significantly explained the mortality differences between White and Hispanic/Asians/others. Among behavioral factors, significant in the White versus other race/ethnicity decomposition were income (10.9%), smoking (-6.9%) and physical activity (2.5%), but taken together differences covariate profiles did not explain much of the mortality difference.

Figure 4 4 shows the race/ethnicity specific PAFs for hypothetical scenarios of improvements in exposures. For White individuals, eliminating the Western and Carnivore DPs would reduce the mortality burden by 7.4% and 4.5%, respectively. Replacement with the Lacto-mod (4.5%) or Prudent (4.1%) DPs would reduce the burden further (**Panel A**). For Black individuals, eliminating the Carnivore DP would reduce the mortality burden by 18.3%. Replacement with the Fatty-grainy (4.5%) or Prudent (7.2%) DPs would reduce the burden further. For Hispanic/Asian/other individuals, eliminating the Western and Carnivore DPs would reduce the

mortality burden by 29.6% and 14.1%; and replacement with the Prudent DP would reduce the burden further by 3.7%. Among other risk factors, exposing all race/ethnicities to high levels of physical activity would produce the most reductions in mortality (33.7-41.7%) (**Panel B**). Of similar importance was high mean income (12.2-21.8%), followed by reducing smoking to 5% prevalence (1.1-11.9%) and then reducing alcohol consumption (0.6-9.9%).

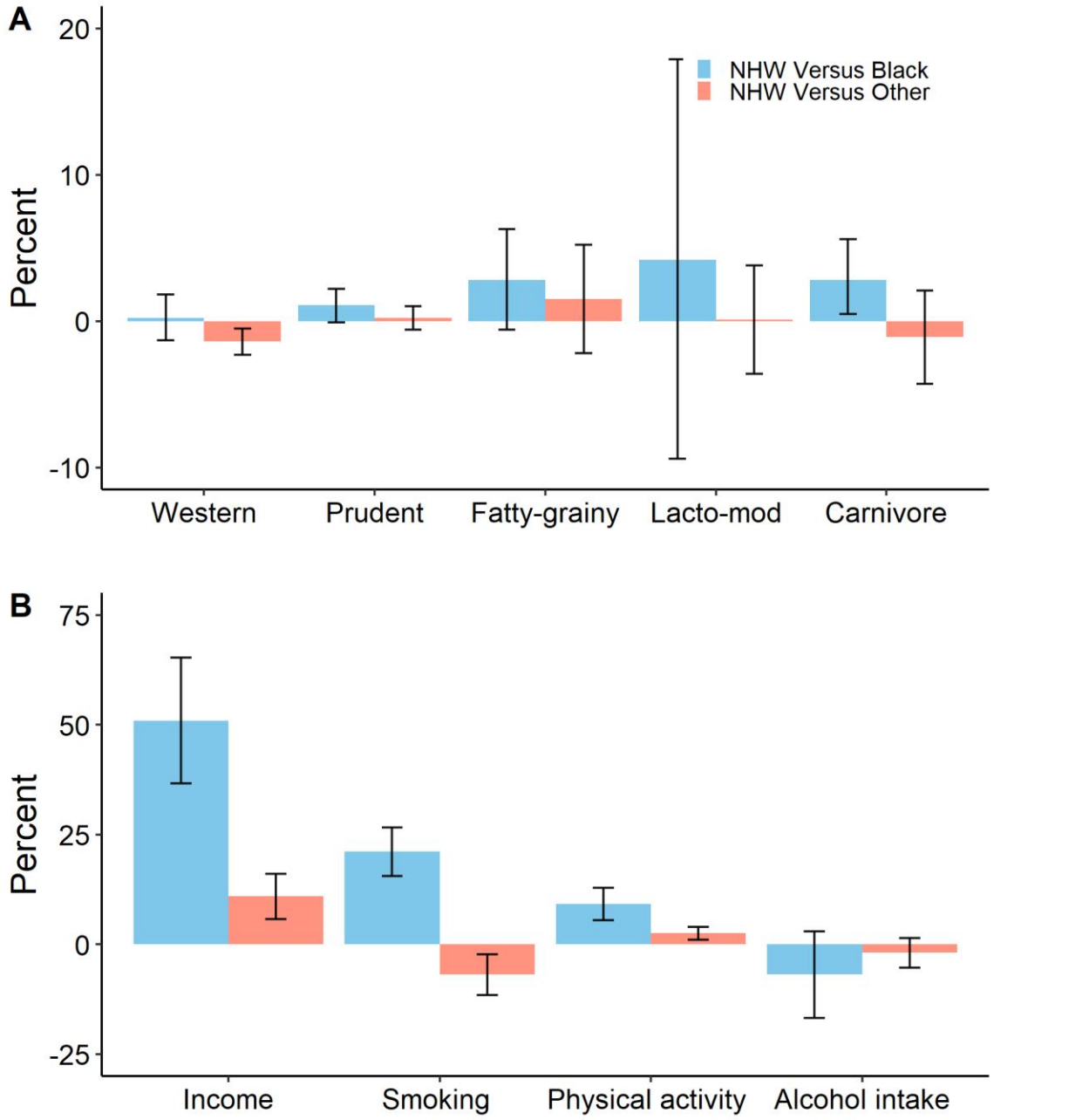
Figure 4 5 shows sensitivity analyses using BMI and SBP as substitutes of DPs. Using BMI and SBP instead of DPs revealed that higher levels of mean BMI and SBP among Black individuals can explain 13.1% and 15.5% of the White-Black mortality difference (**Panel A**). However, neither BMI nor SBP could explain mortality differences between Hispanics/Asian/others and White individuals. In hypothetical scenarios of improved BMI and SBP, the US population with mean BMI at 22.5 would translate to 5.7%, 10.1% and 12.3% reductions in the mortality burden among White, Black and Hispanic/Asian/other groups, respectively. Reducing mean SBP levels to 100 mmhg would produce even larger improvements among White (18.2%), Black (18.3%) and Hispanic/Asian/others (13.1%).

Figure 4 2 Association between dietary patterns, covariates and all-cause mortality within race/ethnicity groups



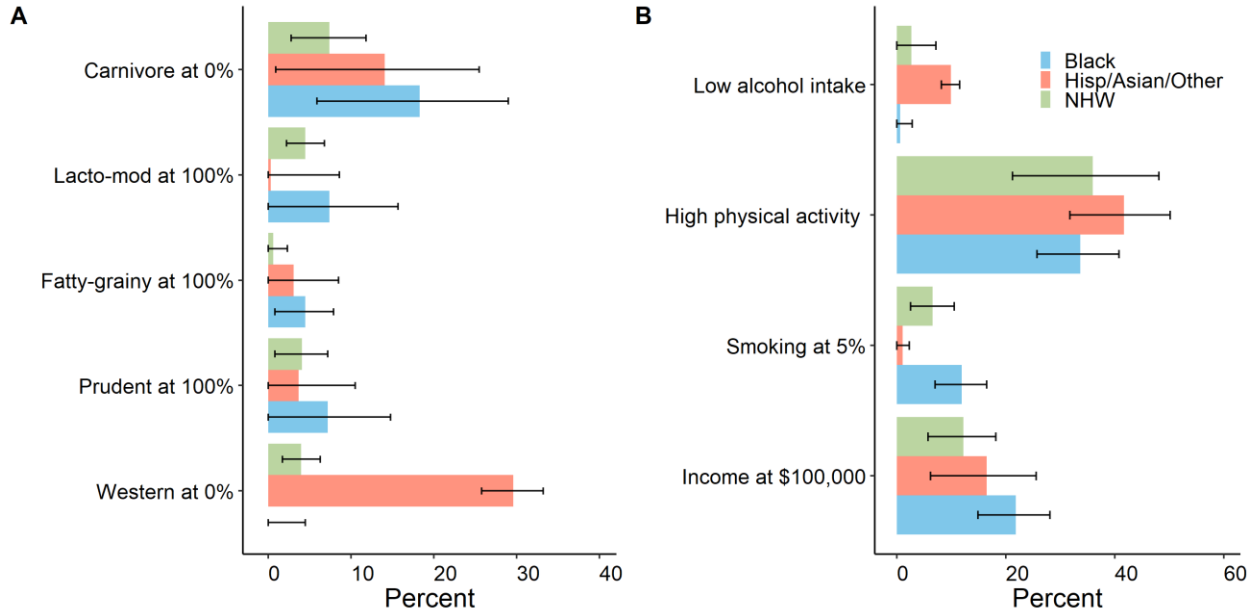
Notes: Models are adjusted for age, sex, income, education, smoking, alcohol, physical activity, calorie intake, height, cohort fixed effects; birth cohort fixed effects and birth-year specific GDP, fertility, infant mortality and lynching cases. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables (not present in REGARDS cohort). The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups (not present in REGARDS and NHANES cohort). The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups (only observed in the WHI cohort).

Figure 4 3 Differences in mortality rates between racial/ethnic groups explained by dietary patterns and other risk factors



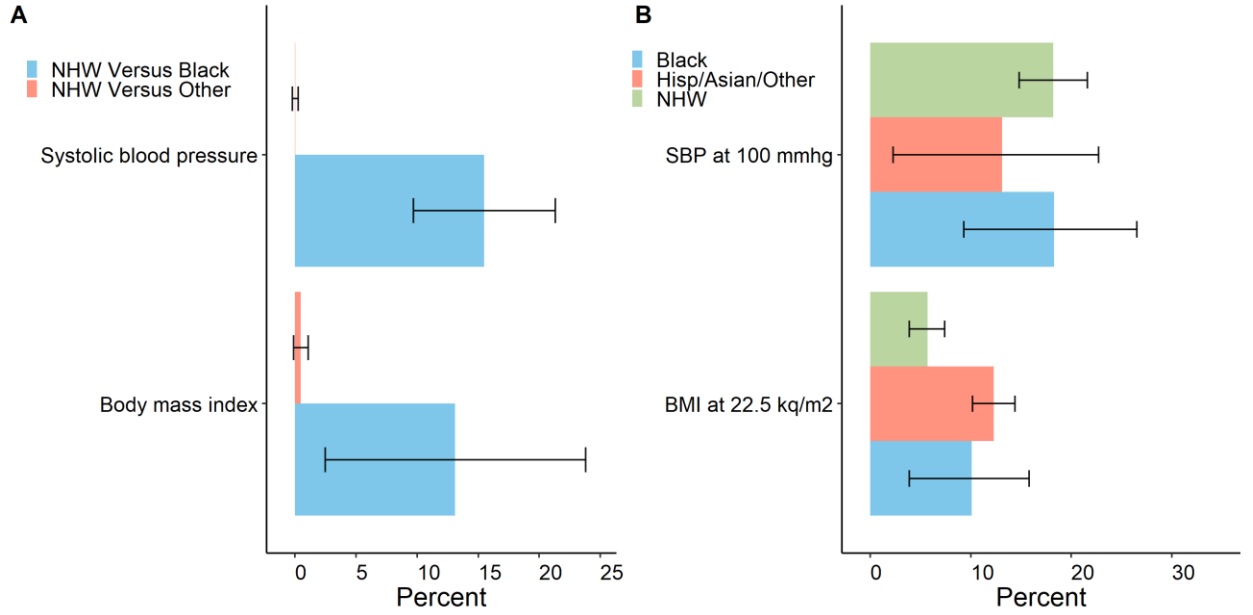
Notes: Models are adjusted for age, sex, income, education, smoking, alcohol, physical activity, calorie intake, height, cohort fixed effects. For factors other than dietary patterns, decomposition results are shown for risk factors that were significant in the regression models. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables (not present in REGARDS cohort). The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups (not present in REGARDS and NHANES cohort). The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups (only observed in the WHI cohort).

Figure 4 4 Reductions in mortality rates across racial/ethnic groups for hypothetical future scenarios of dietary patterns and other risk factors



Notes: Models are adjusted for age, sex, income, education, smoking, alcohol, physical activity, calorie intake, height, cohort fixed effects. Panel A shows scenarios for dietary patterns at optimal levels. Panel B shows scenarios covariates at improved levels.

Figure 4 5 Sensitivity analyses using body mass index and systolic blood pressure



Notes: Models are adjusted for age, sex, income, education, smoking, alcohol, physical activity, calorie intake, height, cohort fixed effects. Panel A shows Oaxaca-Blinder decomposition results for BMI and SBP. Panel B shows scenarios for BMI and SBP at improved levels.

DISCUSSION

In this study we found that the widely studied Western and Prudent DPs have different effects on racial/ethnic groups in the US. The Western DP was associated with very high ACM risk among Hispanics and Asians but was less so for Black individuals. The Prudent DP was beneficial for all groups, but most so for Black individuals. Further, the recently popularized Carnivore DP was associated with very high ACM risks, particularly among Black individuals[72]. Notably, average differences in DPs did not substantially explain racial/ethnic mortality gaps in the sample. However, in hypothetical future scenarios, replacing the Western or Carnivore DPs with the Prudent DP could provide significant mortality reductions across all demographic groups in the US, and particularly large ones for Black, Hispanic and Asian individuals. Importantly, this study shows that among the major risk factors for ACM, improvements in physical activity are likely to generate the largest population level reductions in ACM across all racial/ethnic groups[165].

Across several studies, differences in socio-economic status have explained differences in outcomes of White and Black individuals[160,161,166]. For example, Black individuals have a higher prevalence of coronary heart disease and cardiovascular disease (CVD)[25,167], with age, wealth, healthcare access, and health behaviors accounting for about half the differences in CVD rates between White and Black individuals[162]. Yet, little is known about how much ACM variation, socio-economic and behavioral factors can explain. Our study contributes to this literature by demonstrating that income is currently the primary driver of Black-White mortality disparities. Income differences could explain half the White-Black mortality gap in the sample, suggesting that income support may be the strongest determinant for achieving health equity in the US. Given that we did not control for differences in wealth, ours are likely lower bound estimates. This finding also raises questions about policy responses to ameliorate income disparities for improvements in health. It is well established that racial discrimination is one of the root causes of income disparities and tackling racism at the level of institutions and communities may be one policy lever for bridging the income gap[168]. Moreover, long run evidence from other contexts suggests that targeted cash transfers hold potential to deliver significant equity gains in prosperity and health[169,170]. In the same spirit, proposals for a Universal Basic Income and 'baby bonds' may be considered as candidates for a comprehensive social safety net[171,172].

Currently, there is scant research on DPs as they relate to Black individuals in the US. One study on Black women showed that diets tailored to reduce hypertension (DASH diet) was associated with a 25% reduction in the risk of ACM, while another showed significant effects on lowering the risk of obesity[152,153]. However, both studies did not adjust for income, suggesting that their results may be biased upwards. Our results for the Prudent diet (similar to DASH) are more conservative (12% lower risk for Black individuals) likely because we control for a more comprehensive set of confounders. Our PAF analysis shows that elimination of the Carnivore DP would reduce the Black mortality burden by 18.3% and adoption of the Prudent diet could reduce it further by 7.5%. Our race-specific regression coefficients for other risk factors revealed that Black individuals had larger effect sizes with respect to changes in income, and smaller effect sizes for smoking, than White individuals. However, disadvantage in their covariate profiles relative to White individuals, contributes significantly towards the current mortality disparity. Our sensitivity analyses reveal that currently highly levels of mean BMI also contributed significantly to the Black mortality burden. Further, PAF analyses showed that lower mean SBP scenarios could reduce the mortality burden by 20%. Similarly high PAFs were noted for smoking.

In the US, the Hispanic mortality advantage (20% lower ACM risk) despite an average covariate profile has been widely documented[154]. Asians live longer still, on average, with a 20% advantage over Hispanics[154]. These features, along with a smaller sample size for these groups, made it computationally practical to bin the ethnicities into one group in our analyses. Our results indicated that the Western DP more than doubled ACM risks for Hispanics and Asians. Currently, White, Hispanic and Asian individuals consume the Western DP at similar levels. However, the risk profile implies that switching from Western to Prudent diets may reduce mortality among Hispanics/Asians by 30%. It is also noteworthy that differences in covariate means, including BMI and SBP, did not explain a substantial part of the Hispanic/Asian-White mortality gap. In our regressions, Hispanic/Asian males had substantially lower ACM risks compared to White and Black males. Hispanics/Asians further had larger effect sizes for income and physical activity. In the OBD framework, such differences in coefficients would be attributed to unexplained ethnicity specific attributes. This may suggest that unknown pathways of resilience exist for these ethnic groups that buffer against stress[154]. More research is needed to detect risk and resilience pathways as well as to understand interaction patterns that may explain the lack of explanatory ability in frequently used covariates for these groups.

To our knowledge, this is first decomposition analysis of diet and ACM using large pooled cohort data, with over 120,000 participants, and 18,000 mortality events. This large sample size allowed us to precisely decompose the effects of DPs in relation to race/ethnicity. Moreover, we accounted for a large range of sociodemographic, economic and behavioral factors to minimize the possibility of potential confounding and also gauge their relative importance versus DPs.

Limitations of the study include observational design and measurement error. We relied on observational prospective cohort studies[79], which are prone to bias from residual and unobserved confounding, so ultimately evidence on DPs must be evaluated with caution[80]. We minimized the presence of confounding by using a large set of controls, but confounders might have been measured with error, which leads to residual confounding (e.g., smoking and income). Additionally, causal interpretations of results for an OBD assume that there is no time dependent confounding. While this assumption cannot be verified given the nature of the data, we used covariates measured at the birth-year of the participants to control for these effects. Further, we cannot rule out that our estimates are biased to some degree due to the influence of race-specific confounders or the constituents of food groups consumed by different racial/ethnic groups. Finally, since this study relies on FFQs, the measured intake of foods is prone to random error due to non-differential misclassification, and can lead to biased estimates, generally towards the null[81]. Evidence indicates that a high degree of attenuation can be attributed to measurement error in FFQ[82]. To this end, the pooled cohort design provided greater statistical power thereby endowing the study with the ability to detect significant associations[81].

CONCLUSION

In the US, research supported by historical precedents[173,174], indicates that, compared to White individuals, Black individuals experience dissimilar health risks from numerous causes[175–181]. Several adverse factors may contribute to widening the racial/ethnic mortality gap in the US, especially in a post-pandemic setting[24]. Addressing racial discrimination might be an efficient way to tackle multiple downstream factors. This study reaffirms what is already known, and adds that replacing Western and Carnivore diets with Prudent diets, along with income support and interventions to promote physical activity, should be holistically employed to tackle racial/ethnic disparities in mortality in the US[182,183].

CHAPTER 5:

CONCLUSION

SUMMARY

This dissertation is a compilation of three publishable manuscripts that present findings from analyses of pooled data from seven prospective cohorts. The overarching motivations for pursuing this line of research were to advance our understanding of biological and social mechanisms that determine mortality to improve healthy life expectancy, reduce the burden of chronic metabolic disease and address social disparities in health in the US. This research is of interest to researchers working across multiple country contexts and relevant for public health nutrition practitioners. Important findings have emerged from this study of the diet-mortality connection in the US.

In aim one, we quantified the risk to ACM associated with commonly occurring DPs in the US. We identified four DPs that consistently explained variation across six cohorts and an additional DP in the WHI. Since our analyses adjusted for calories, our findings speak to diet quality at the same levels of energy intake. Concurring with research from previous studies, diets high in meat were found to be high risk for ACM, while those comprised of nutrient dense plant foods were consistently and robustly associated with lower ACM rates. Risk estimates for females with respect to the Western and Prudent DPs were larger compared to males. Further, results suggest that individuals with healthy BMI may benefit most from the Prudent DP. These findings support the Dietary Guidelines for Americans and other studies, which recommend eating patterns rich in plant foods, and low in meat and processed foods, to adapt according to personal food traditions and preferences. In terms of guidance for public policy, our study reinforces the emerging focus on DPs as opposed to individual food groups. A focus on the way people actually eat is more readily translatable into food choices at different life stages. This study concludes that diets that founded on whole plant based foods and otherwise moderate to low in the consumption of other food groups are supportive of longer lifespans.

In aim two, we quantified the direct and indirect effects of DPs on ACM associated with BMI and SBP. The main finding here was that BMI and SBP significantly mediate the effects of the widely researched Western and Prudent DPs. The most important mediator was BMI, especially for the Western and Carnivore DPs, accounting for about one fourth of the risk. However, the direct effects of DPs remained large and significant, indicating that interventions targeting BMI and SBP can only provide a fraction of the benefits associated with adopting healthy DPs. In other words, healthy diets can prevent early mortality even if they do not lead to improvements in weight or BP. These results are clinically relevant for public health, in particular for encouraging a Prudent-like dietary pattern adoption for primary prevention of fatal endpoints and improving mediators. Our results are also cautionary about fad diets like the Carnivore DP. This study concludes that interventions address BMI and SBP including physical activity, antihypertensive drugs and bariatric surgery can only provide a fraction of the benefits that adopting healthy DPs can provide.

In aim three, we found that the widely studied Western and Prudent DPs have different effects on racial/ethnic groups in the US. The Western DP was associated with very high ACM risk among Hispanics and Asians but was less so for Black individuals. The Prudent DP was beneficial for all groups, but most so for Black individuals. Further, the Carnivore DP was associated with very high ACM risks, particularly among

Black individuals. Notably, average differences in DPs did not substantially explain racial/ethnic mortality gaps in the sample. However, in hypothetical future scenarios, replacing the Western or Carnivore DPs with the Prudent DP could provide significant mortality reductions across all demographic groups in the US, and particularly large ones for Black, Hispanic and Asian individuals. Importantly, this study shows that among the major risk factors for ACM, improvements in physical activity and income are likely to generate the largest population level reductions in ACM across all racial/ethnic groups. In the US, research supported by historical precedents, indicates that, compared to White individuals, Black individuals experience dissimilar health risks from numerous causes. Several adverse factors may contribute to widening the White-Black mortality gap in the US, especially in a post-pandemic setting. Addressing racial discrimination with a policy focus on addressing income disparities is likely the most efficient way to tackle multiple downstream factors. This study reaffirms what is already known, and adds that replacing Western and Carnivore diets with Prudent diets, along with income support programs and interventions to promote physical activity, should be holistically employed to tackle racial/ethnic disparities in mortality in the US.

FUTURE DIRECTIONS

A specific area for future research identified through aims 1 and 2 is the interaction between BMI and diet. Our analyses revealed that diet may influence health outcomes differently at varying levels of BMI. While our sample was largely overweight/obese and female, we found that individuals with BMI in the normal range were most protected by healthy diets. This finding needs to be corroborated in other studies because it might have implications for individual specific health recommendations. Another area left unresolved in this study was the Hispanic-Asian mortality advantage puzzle in the US. While the advantage is well documented, no studies have been able to definitively find risk factors that can explain it. This suggests that focused cohorts are needed to study ethnic groups in the US that cover aspects of cultural practices and other factors that are conducive to longer life. Finally, an important ancillary finding of this study was that income differences between White and Black individuals can explain half of their mortality rate differences. This finding has important implications for future social safety net programs that are designed to provide income support with a focus on groups that have been historically disadvantaged or are currently languishing.

Extensions of this work would require larger collaborative pooling studies that employ data from numerous cohorts across varying country contexts. However, it is likely that such studies would face challenges including lack of data on a standardized set of dietary input variables as well as demographic, spatial, socio-economic and behavioral variables. While pooling studies are data and time intensive, they provide superior evidence because they overcome the heterogeneity challenges faced by meta-analyses of dietary risk factors. However, there are abundant data access restrictions due to ethical concerns in proprietary cohorts. This suggests that donor investments need to be made in implementing nationally representative cross country cohorts that use standardized questionnaires and make de-identified data publically available. The Demographic Health Surveys are examples of such initiatives that make it possible to track chronic disease trends in low-and-middle-income countries. Such data are essential to pin down causal relationships and track dietary changes in countries as they move through the epidemiologic transition. Rigorously assessing the relative importance of different factors, such as diet, smoking, alcohol, stress, genetics, environment and physical activity, has tremendous implications in terms of policy-making and prioritization of health interventions.

BIBLIOGRAPHY

1. Anderson, C. A. M. & Campbell, W. W. 2015 Dietary Guidelines Advisory Committee Report. *Nutr. Today* **50**, 172–173 (2015).

2. Afshin, A. *et al.* Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* (2019) doi:10.1016/S0140-6736(19)30041-8.

3. Aune, D. *et al.* Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality-A systematic review and dose-response meta-analysis of prospective studies. *Int. J. Epidemiol.* **46**, 1029–1056 (2017).

4. Aune, D. *et al.* Nut consumption and risk of cardiovascular disease, total cancer, all-cause and cause-specific mortality: A systematic review and dose-response meta-analysis of prospective studies. *BMC Med.* **14**, 1–14 (2016).

5. Larsson, S. C., Crippa, A., Orsini, N., Wolk, A. & Michaëlsson, K. Milk consumption and mortality from all causes, cardiovascular disease, and cancer: A systematic review and meta-analysis. *Nutrients* **7**, 7749–7763 (2015).

6. Li, B. *et al.* Consumption of whole grains in relation to mortality from all causes, cardiovascular disease, and diabetes: Dose-response meta-analysis of prospective cohort studies. *Med. (United States)* **95**, (2016).

7. Ma, X. *et al.* Association between whole grain intake and all-cause mortality: A meta-analysis of cohort studies. *Oncotarget* **7**, 61996–62005 (2016).

8. Naghshi, S., Sadeghi, O., Willett, W. C. & Esmaillzadeh, A. Dietary intake of total, animal, and plant proteins and risk of all cause, cardiovascular, and cancer mortality: Systematic review and dose-response meta-analysis of prospective cohort studies. *BMJ* **370**, (2020).

9. Schwingshackl, L., Bogensberger, B. & Hoffmann, G. Diet Quality as Assessed by the Healthy Eating Index, Alternate Healthy Eating Index, Dietary Approaches to Stop Hypertension Score, and Health Outcomes: An Updated Systematic Review and Meta-Analysis of Cohort Studies. *J. Acad. Nutr. Diet.* **118**, 74-100.e11 (2018).

10. Seidelmann, S. B. *et al.* Dietary carbohydrate intake and mortality: a prospective cohort study and

meta-analysis. *Lancet Public Heal.* **3**, e419–e428 (2018).

11. Soltani, S., Arablou, T., Jayedi, A. & Salehi-Abargouei, A. Adherence to the dietary approaches to stop hypertension (DASH) diet in relation to all-cause and cause-specific mortality: A systematic review and dose-response meta-analysis of prospective cohort studies. *Nutr. J.* **19**, 1029–1039 (2020).

12. Soltani, S., Jayedi, A., Shab-Bidar, S., Becerra-Tomás, N. & Salas-Salvadó, J. Adherence to the Mediterranean Diet in Relation to All-Cause Mortality: A Systematic Review and Dose-Response Meta-Analysis of Prospective Cohort Studies. *Adv. Nutr.* **10**, 1029–1039 (2019).

13. Wan, Y., Zheng, J., Wang, F. & Li, D. Fish, long chain omega-3 polyunsaturated fatty acids consumption, and risk of all-cause mortality: A systematic review and dose-response meta-analysis from 23 independent prospective cohort studies. *Asia Pac. J. Clin. Nutr.* **26**, 939–956 (2017).

14. Zeraatkar, D. *et al.* Red and processed meat consumption and risk for all-cause mortality and cardiometabolic outcomes a systematic review and meta-analysis of cohort studies. *Ann. Intern. Med.* **171**, 703–710 (2019).

15. Aune, D. *et al.* Dietary intake and blood concentrations of antioxidants and the risk of cardiovascular disease, total cancer, and all-cause mortality: A systematic review and dose-response meta-analysis of prospective studies. *Am. J. Clin. Nutr.* **108**, 1069–1091 (2018).

16. Aune, D. *et al.* Whole grain consumption and risk of cardiovascular disease, cancer, and all cause and cause specific mortality: Systematic review and dose-response meta-analysis of prospective studies. *BMJ* **353**, 1–14 (2016).

17. Benisi-Kohansal, S., Saneei, P., Salehi-Marzijarani, M., Larijani, B. & Esmaillzadeh, A. Whole-grain intake and mortality from all causes, cardiovascular disease, and cancer: A systematic review and dose-response meta-analysis of prospective cohort studies. *Adv. Nutr.* **7**, 1052–1065 (2016).

18. Cavero-Redondo, I. *et al.* Milk and Dairy Product Consumption and Risk of Mortality: An Overview of Systematic Reviews and Meta-Analyses. *Adv. Nutr.* **10**, S97–S104 (2019).

19. Chen, Z. *et al.* Dietary protein intake and all-cause and cause-specific mortality: results from the Rotterdam Study and a meta-analysis of prospective cohort studies. *Eur. J. Epidemiol.* **35**, 411–429 (2020).

20. Eleftheriou, D., Benetou, V., Trichopoulou, A., La Vecchia, C. & Bamia, C. Mediterranean diet and its components in relation to all-cause mortality: Meta-analysis. *Br. J. Nutr.* **120**, 1081–1097 (2018).

21. Guo, J. *et al.* Milk and dairy consumption and risk of cardiovascular diseases and all-cause mortality: dose–response meta-analysis of prospective cohort studies. *Eur. J. Epidemiol.* **32**, 269–287 (2017).

22. Lacombe, J., Armstrong, M. E. G., Wright, F. L. & Foster, C. The impact of physical activity and an additional behavioural risk factor on cardiovascular disease, cancer and all-cause mortality: A systematic review. *BMC Public Health* **19**, (2019).

23. Lu, Y. *et al.* Metabolic mediators of the effects of body-mass index, overweight, and obesity on coronary heart disease and stroke: a pooled analysis of 97 prospective cohorts with 1·8 million participants. *Lancet* **383**, 970–983 (2014).

24. Chakrabarti, S., Hamlet, L. C., Kaminsky, J. & Subramanian, S. V. Association of Human Mobility Restrictions and Race/Ethnicity–Based, Sex-Based, and Income-Based Factors With Inequities in Well-being During the COVID-19 Pandemic in the United States. *JAMA Netw. Open* **4**, e217373–e217373 (2021).

25. Cuevas, A. G. *et al.* Assessing the Role of Health Behaviors, Socioeconomic Status, and Cumulative Stress for Racial/Ethnic Disparities in Obesity. *Obesity* **28**, 161–170 (2020).

26. Schwingshackl, L. *et al.* Food groups and risk of all-cause mortality: A systematic review and meta-analysis of prospective studies. *American Journal of Clinical Nutrition* vol. 105 1462–1473 (2017).

27. Jacobs, D. R. & Steffen, L. M. Nutrients, foods, and dietary patterns as exposures in research: A framework for food synergy. in *American Journal of Clinical Nutrition* vol. 78 (Am J Clin Nutr, 2003).

28. AM, S. *et al.* Relation of body fat mass and fat-free mass to total mortality: results from 7 prospective cohort studies. *Am. J. Clin. Nutr.* **113**, 639–646 (2021).

29. Shan, Z. *et al.* Association Between Healthy Eating Patterns and Risk of Cardiovascular Disease. *JAMA Intern. Med.* **180**, 1090–1100 (2020).

30. Fransen, H. P. *et al.* A posteriori dietary patterns: How many patterns to retain? *J. Nutr.* **144**, 1274–1282 (2014).

31. Jayedi, A., Soltani, S., Abdolshahi, A. & Shab-Bidar, S. Healthy and unhealthy dietary patterns and the

risk of chronic disease: An umbrella review of meta-analyses of prospective cohort studies. *British Journal of Nutrition* vol. 124 1133–1144 (2020).

32. Edefonti, V. *et al.* Reproducibility and Validity of A Posteriori Dietary Patterns: A Systematic Review. *Advances in Nutrition* vol. 11 293–326 (2020).

33. 2015-2020 Dietary Guidelines | health.gov. https://health.gov/our-work/nutrition-physical-activity/dietary-guidelines/previous-dietary-guidelines/2015.

34. Wang, X. *et al.* Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: Systematic review and dose-response meta-analysis of prospective cohort studies. *BMJ* **349**, 1–14 (2014).

35. Lupoli, R. *et al.* White meat consumption, all-cause mortality, and cardiovascular events: A meta-analysis of prospective cohort studies. *Nutrients* **13**, 1–16 (2021).

36. Lee, K. W., Loh, H. C., Ching, S. M., Devaraj, N. K. & Hoo, F. K. Effects of vegetarian diets on blood pressure lowering: A systematic review with meta-analysis and trial sequential analysis. *Nutrients* **12**, 1–17 (2020).

37. Glenn, A. J. *et al.* Relation of vegetarian dietary patterns with major cardiovascular outcomes: A systematic review and meta-analysis of prospective cohort studies. *Front. Nutr.* **6**, (2019).

38. Menzel, J. *et al.* Systematic review and meta-analysis of the associations of vegan and vegetarian diets with inflammatory biomarkers. *Sci. Rep.* **10**, 1–11 (2020).

39. Huang, R. Y., Huang, C. C., Hu, F. B. & Chavarro, J. E. Vegetarian Diets and Weight Reduction: a Meta-Analysis of Randomized Controlled Trials. *J. Gen. Intern. Med.* **31**, 109–116 (2016).

40. Craddock, J. C., Neale, E. P., Peoples, G. E. & Probst, Y. C. Vegetarian-Based Dietary Patterns and their Relation with Inflammatory and Immune Biomarkers: A Systematic Review and Meta-Analysis. *Adv. Nutr.* **10**, 433–451 (2019).

41. Lee, Y. & Park, K. Adherence to a vegetarian diet and diabetes risk: A systematic review and meta-analysis of observational studies. *Nutrients* **9**, (2017).

42. Yokoyama, Y., Levin, S. M. & Barnard, N. D. Association between plant-based diets and plasma lipids: A systematic review and meta-analysis. *Nutr. Rev.* **75**, 683–698 (2017).

43. Benatar, J. R. & Stewart, R. A. H. Cardiometabolic risk factors in vegans; A meta-analysis of observational studies. *PLoS One* **13**, 1–23 (2018).

44. Huang, T. *et al.* Cardiovascular disease mortality and cancer incidence in vegetarians: A meta-analysis and systematic review. *Ann. Nutr. Metab.* **60**, 233–240 (2012).

45. Wang, F. *et al.* Effects of vegetarian diets on blood lipids: A systematic review and meta-analysis of randomized controlled trials. *J. Am. Heart Assoc.* **4**, (2015).

46. Hu, F. B. Dietary pattern analysis: A new direction in nutritional epidemiology. *Current Opinion in Lipidology* vol. 13 3–9 (2002).

47. Dietary Guidelines for Americans, 2020-2025 and Online Materials | Dietary Guidelines for Americans. https://www.dietaryguidelines.gov/resources/2020-2025-dietary-guidelines-online-materials.

48. Wirt, A. & Collins, C. E. Diet quality - What is it and does it matter? *Public Health Nutr.* **12**, 2473–2492 (2009).

49. Gil, Á., de Victoria, E. M. & Olza, J. Indicadores de evaluación de la calidad de la dieta. *Nutr. Hosp.* **31**, 128–144 (2015).

50. Bach, A. *et al.* The use of indexes evaluating the adherence to the Mediterranean diet in epidemiological studies: a review. *Public Health Nutr.* **9**, 132–146 (2006).

51. Tucker, K. L. Dietary patterns, approaches, and multicultural perspective. *Appl. Physiol. Nutr. Metab.* **35**, 211–218 (2010).

52. Atherosclerosis Risk in Communities Study Description | Atherosclerosis Risk in Communities. https://sites.cscc.unc.edu/aric/description.

53. The atherosclerosis risk in communit (aric) stui)y: Design and objectwes. *Am. J. Epidemiol.* **129**, 687–702 (1989).

54. Ives, D. G. *et al.* Surveillance and ascertainment of cardiovascular events. The Cardiovascular Health Study. *Ann. Epidemiol.* **5**, 278–285 (1995).

55. Kannel, W. B., Feinleib, M., Mcnamara, P. M., Garrison, R. J. & Castelli, W. P. An investigation of

coronary heart disease in families: The framingham offspring study. *Am. J. Epidemiol.* **110**, 281–290 (1979).

56. Bild, D. E. Multi-Ethnic Study of Atherosclerosis: Objectives and Design. *Am. J. Epidemiol.* **156**, 871–881 (2002).

57. Nomura, S. O., Karger, A. B., Weir, N. L., Duprez, D. A. & Tsai, M. Y. Free fatty acids, cardiovascular disease, and mortality in the Multi-Ethnic Study of Atherosclerosis. *J. Clin. Lipidol.* **14**, 531–541 (2020).

58. Anderson, G. *et al.* Design of the Women's Health Initiative clinical trial and observational study. *Control. Clin. Trials* **19**, 61–109 (1998).

59. Angelantonio, E. Di *et al.* Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet* **388**, 776–786 (2016).

60. Zhao, J. *et al.* A review of statistical methods for dietary pattern analysis. *Nutrition Journal* vol. 20 1–18 (2021).

61. Geidl, W., Schlesinger, S., Mino, E., Miranda, L. & Pfeifer, K. Dose-response relationship between physical activity and mortality in adults with noncommunicable diseases: A systematic review and meta-analysis of prospective observational studies. *International Journal of Behavioral Nutrition and Physical Activity* vol. 17 (2020).

62. Ekelund, U. *et al.* Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: Systematic review and harmonised meta-analysis. *BMJ* **366**, (2019).

63. Bradburn, M. J., Clark, T. G., Love, S. B. & Altman, D. G. Survival Analysis Part II: Multivariate data analysis- An introduction to concepts and methods. *British Journal of Cancer* vol. 89 431–436 (2003).

64. Cox, D. R. Regression Models and Life-Tables. *J. R. Stat. Soc. Ser. B* **34**, 187–202 (1972).

65. Vonesh, E. F., Schaubel, D. E., Hao, W. & Collins, A. J. Statistical methods for comparing mortality among ESRD patients: Examples of regional/international variations. *Kidney Int.* **57**, S19–S27 (2000).

66. Verbeek, M. & Vella, F. Estimating dynamic models from repeated cross-sections. *J. Econom.* (2005) doi:10.1016/j.jeconom.2004.06.004.

67. Deaton, A. Panel data from time series of cross-sections. *J. Econom.* (1985) doi:10.1016/0304-4076(85)90134-4.

68. Harrell , F. E. Regression Modeling Strategies. (2015) doi:10.1007/978-3-319-19425-7.

69. Causal Inference in Statistics: A Primer - Judea Pearl, Madelyn Glymour, Nicholas P. Jewell - Google Books. https://books.google.com/books?hl=en&lr=&id=L3G-CgAAQBAJ&oi=fnd&pg=PR9&dq=judea+pearl+causality&ots=CsYiCKvOdn&sig=XKAyqFNu1NI6H0TFaraVZ4dj8OI#v=onepage&q=judea pearl causality&f=false.

70. Van Der Weele, T. J. & Ding, P. Sensitivity analysis in observational research: Introducing the E-Value. *Ann. Intern. Med.* **167**, 268–274 (2017).

71. English, L. K. *et al.* Evaluation of Dietary Patterns and All-Cause Mortality. *JAMA Netw. Open* **4**, e2122277 (2021).

72. O'Hearn, A. Can a carnivore diet provide all essential nutrients? *Curr. Opin. Endocrinol. Diabetes. Obes.* **27**, 312–316 (2020).

73. Li, F. *et al.* Associations of dietary patterns with the risk of all-cause, CVD and stroke mortality: A meta-Analysis of prospective cohort studies. *British Journal of Nutrition* vol. 113 16–24 (2015).

74. Onvani, S., Haghighatdoost, F., Surkan, P. J., Larijani, B. & Azadbakht, L. Adherence to the Healthy Eating Index and Alternative Healthy Eating Index dietary patterns and mortality from all causes, cardiovascular disease and cancer: a meta-analysis of observational studies. *J. Hum. Nutr. Diet.* **30**, 216–226 (2017).

75. Heidemann, C. *et al.* Dietary patterns and risk of mortality from cardiovascular disease, cancer, and all causes in a prospective cohort of women. *Circulation* **118**, 230–237 (2008).

76. Boggs, D. A., Ban, Y., Palmer, J. R. & Rosenberg, L. Higher Diet Quality Is Inversely Associated with Mortality in African-American Women. *J. Nutr.* **145**, 547 (2015).

77. Hsiao, P. Y. *et al.* Dietary Patterns and Relationship to Obesity-Related Health Outcomes and Mortality in Adults 75 Years of Age or Greater. *J. Nutr. Health Aging* **17**, 566 (2013).

78. Anderson, A. L. *et al.* Dietary Patterns and Survival of Older Adults. *J. Am. Diet. Assoc.* **111**, 84 (2011).

79. Hall, K. D. *et al.* Effect of a plant-based, low-fat diet versus an animal-based, ketogenic diet on ad libitum energy intake. *Nat. Med. 2021 272* **27**, 344–353 (2021).

80. Delgado-Rodríguez, M. & Llorca, J. Bias. *J. Epidemiol. Community Heal.* **58**, 635–641 (2004).

81. Freedman, L. S., Schatzkin, A., Midthune, D. & Kipnis, V. Dealing with dietary measurement error in nutritional cohort studies. *Journal of the National Cancer Institute* vol. 103 1086–1092 (2011).

82. Subar, A. F. *et al.* Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: The OPEN study. *Am. J. Epidemiol.* **158**, 1–13 (2003).

83. Joshi, S., Ostfeld, R. J. & McMacken, M. The Ketogenic Diet for Obesity and Diabetes—Enthusiasm Outpaces Evidence. *JAMA Intern. Med.* **179**, 1163–1164 (2019).

84. Hall, K. D. & Guo, J. Obesity Energetics: Body Weight Regulation and the Effects of Diet Composition. *Gastroenterology* **152**, 1718-1727.e3 (2017).

85. Liu, B. *et al.* Trends in obesity and adiposity measures by race or ethnicity among adults in the United States 2011-18: population based study. *BMJ* **372**, 2011–2029 (2021).

86. Wang, Y. *et al.* Has the prevalence of overweight, obesity and central obesity levelled off in the United States? Trends, patterns, disparities, and future projections for the obesity epidemic. *Int. J. Epidemiol.* **49**, 810–823 (2020).

87. Bai, Y. *et al.* A Novel Diet-Health Index Reveals Worsening Trends and Widening Disparities in the Health Impacts of Diet in the United States. *Curr. Dev. Nutr.* **5**, 101–101 (2021).

88. Healthy life expectancy (HALE) at birth (years). https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-ghe-hale-healthy-life-expectancy-at-birth.

89. Di Angelantonio, E. *et al.* Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet* **388**, 776–786 (2016).

90. Brunström, M. & Carlberg, B. Association of blood pressure lowering with mortality and cardiovascular disease across blood pressure levels a systematic review and meta-analysis. *JAMA Internal Medicine* vol. 178 28–36 (2018).

91. Ndanuko, R. N., Tapsell, L. C., Charlton, K. E., Neale, E. P. & Batterham, M. J. Dietary Patterns and Blood Pressure in Adults: A Systematic Review and Meta-Analysis of Randomized Controlled Trials 1,2. doi:10.3945/an.115.009753.

92. Stelmach-Mardas, M. & Walkowiak, J. Dietary interventions and changes in cardio-metabolic parameters in metabolically healthy obese subjects: A systematic review with meta-analysis. *Nutrients* vol. 8 (2016).

93. Blundell, J. E. & MacDiarmid, J. I. Passive overconsumption: Fat intake and short-term energy balance. *Ann. N. Y. Acad. Sci.* **827**, 392–407 (1997).

94. Ludwig, D. S. & Ebbeling, C. B. The carbohydrate-insulin model of obesity: Beyond "calories in, calories out". *JAMA Internal Medicine* vol. 178 1098–1103 (2018).

95. Filippou, C. D. *et al.* Dietary Approaches to Stop Hypertension (DASH) Diet and Blood Pressure Reduction in Adults with and without Hypertension: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Adv. Nutr.* **11**, 1150 (2020).

96. Buchwald, H. *et al.* Bariatric surgery: A systematic review and meta-analysis. *Journal of the American Medical Association* vol. 292 1724–1737 (2004).

97. Danaei, G. *et al.* National, regional, and global trends in systolic blood pressure since 1980: Systematic analysis of health examination surveys and epidemiological studies with 786 country-years and 5·4 million participants. *Lancet* **377**, 568–577 (2011).

98. Lu, Y., Hajifathalian, K., Rimm, E. B., Ezzati, M. & Danaei, G. Mediators of the effect of body mass index on coronary heart disease. *Epidemiology* **26**, 153–162 (2015).

99. VanderWeele, T. J. Causal mediation analysis with survival data. *Epidemiology* vol. 22 582–585 (2011).

100. VanderWeele, T. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology* **21**, 540–551 (2010).

101. VanderWeele, T. J. Invited Commentary: Structural Equation Models and Epidemiologic Analysis. *Am. J. Epidemiol.* **176**, 608 (2012).

102. VanderWeele, T. J. Unmeasured confounding and hazard scales: Sensitivity analysis for total, direct,

and indirect effects. *European Journal of Epidemiology* vol. 28 113–117 (2013).

103. VanderWeele, T. J. Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health* vol. 37 17–32 (2016).

104. Fairchild, A. J. & McDaniel, H. L. Best (but oft-forgotten) practices: Mediation analysis. *American Journal of Clinical Nutrition* vol. 105 1259–1271 (2017).

105. High blood pressure (hypertension) - Symptoms and causes - Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/high-blood-pressure/symptoms-causes/syc-20373410.

106. Obesity - Symptoms and causes - Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/obesity/symptoms-causes/syc-20375742.

107. Tein, J.-Y. & MacKinnon, D. P. Estimating Mediated Effects with Survival Data. in *New Developments in Psychometrics* 405–412 (Springer, Tokyo, 2003). doi:10.1007/978-4-431-66996-8_46.

108. Zellner, A. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *J. Am. Stat. Assoc.* **57**, 348–368 (1962).

109. Preacher, K. J. & Hayes, A. F. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav. Res. Methods* **40**, 879–891 (2008).

110. Law, M. R., Morris, J. K. & Wald, N. J. Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. *BMJ* **338**, (2009).

111. Sjöström, L. *et al. From the Institutes of Medicine*. vol. 23 www.nejm.org (2007).

112. Zheng, J. Inflammatory Potential of Diet And Pancreatic Cancer Risk: Interaction And Mediation Analysis In Two Prospective Cohorts. *Theses Diss.* (2017).

113. André, P. *et al.* The role of overweight in the association between the Mediterranean diet and the risk of type 2 diabetes mellitus: a mediation analysis among 21 585 UK biobank participants. *Int. J. Epidemiol.* **49**, 1582–1590 (2020).

114. MacMahon, S. *et al.* Body-mass index and cause-specific mortality in 900 000 adults: Collaborative

analyses of 57 prospective studies. *Lancet* **373**, 1083–1096 (2009).

115. Flegal, K. M., Kit, B. K., Orpana, H. & Graubard, B. I. *Association of All-Cause Mortality With Overweight and Obesity Using Standard Body Mass Index Categories A Systematic Review and Meta-analysis*. www.jama.com.

116. Murakami, Y., Hozawa, A., Okamura, T. & Ueshima, H. Relation of Blood Pressure and All-Cause Mortality in 180 000 Japanese Participants Pooled Analysis of 13 Cohort Studies. (2008) doi:10.1161/HYPERTENSIONAHA.107.102459.

117. Vlachopoulos, C. *et al.* Prediction of cardiovascular events and all-cause mortality with central haemodynamics: a systematic review and meta-analysis. doi:10.1093/eurheartj/ehq125.

118. Wang, J. *et al.* Blood pressure categories defined by the 2017 ACC/AHA guideline and all-cause mortality: a national cohort study in China and meta-analysis. *J. Hum. Hypertens.* **36**, 95–105 (2022).

119. Li, W. feng, Huang, Y. qing & Feng, Y. qing. Association between central haemodynamics and risk of all-cause mortality and cardiovascular disease: a systematic review and meta-analysis. *J. Hum. Hypertens.* **33**, 531–541 (2019).

120. Sacks, F. M. & Campos, H. Dietary Therapy in Hypertension. *N. Engl. J. Med.* **362**, 2102–2112 (2010).

121. Mancini, J. G., Filion, K. B., Atallah, R. & Eisenberg, M. J. Systematic Review of the Mediterranean Diet for Long-Term Weight Loss. *Am. J. Med.* **129**, 407-415.e4 (2016).

122. Law, M. R., Morris, J. K. & Wald, N. J. Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. *BMJ* **338**, 1245 (2009).

123. Douketis, J. D., Macie, C., Thabane, L. & Williamson, D. F. Systematic review of long-term weight loss studies in obese adults: clinical significance and applicability to clinical practice. *Int. J. Obes. (Lond).* **29**, 1153–1167 (2005).

124. Franz, M. J. *et al.* Weight-loss outcomes: a systematic review and meta-analysis of weight-loss clinical trials with a minimum 1-year follow-up. *J. Am. Diet. Assoc.* **107**, 1755–1767 (2007).

125. Williams, D. R. & Mohammed, S. A. Racism and Health I: Pathways and Scientific Evidence. *American Behavioral Scientist* vol. 57 1152–1173 (2013).

126. Serchen, J., Doherty, R., Atiq, O. & Hilden, D. A Comprehensive Policy Framework to Understand and Address Disparities and Discrimination in Health and Health Care: A Policy Paper From the American College of Physicians. *Ann. Intern. Med.* **174**, 529–532 (2021).

127. Matsui, E. C., Perry, T. T. & Adamson, A. S. An Antiracist Framework for Racial and Ethnic Health Disparities Research. *Pediatrics* vol. 146 (2020).

128. Ford, C. L. & Airhihenbuwa, C. O. Critical race theory, race equity, and public health: Toward antiracism praxis. *American Journal of Public Health* vol. 100 S30 (2010).

129. Yearby, R. Structural Racism and Health Disparities: Reconfiguring the Social Determinants of Health Framework to Include the Root Cause. *J. Law, Med. Ethics* **48**, 518–526 (2020).

130. Over the past century, African-American life expectancy and education levels have soared | The Economist. https://www.economist.com/graphic-detail/2021/05/20/over-the-past-century-african-american-life-expectancy-and-education-levels-have-soared.

131. Andrasfay, T. & Goldman, N. Reductions in 2020 US life expectancy due to COVID-19 and the disproportionate impact on the Black and Latino populations. *Proc. Natl. Acad. Sci.* **118**, (2021).

132. Virtanen, M. *et al.* Lifestyle factors and risk of sickness absence from work: a multicohort study. *Lancet Public Heal.* **3**, e545–e554 (2018).

133. Diaz, V. A., Mainous, A. G., Koopman, R. J., Carek, P. J. & Geesey, M. E. Race and diet in the overweight: association with cardiovascular risk in a nationally representative sample. *Nutrition* **21**, 718–725 (2005).

134. Sijtsma, F. P. *et al.* Longitudinal trends in diet and effects of sex, race, and education on dietary quality score change: the Coronary Artery Risk Development in Young Adults study. *Am. J. Clin. Nutr.* **95**, 580–586 (2012).

135. Raffensperger, S. *et al.* Effect of Race and Predictors of Socioeconomic Status on Diet Quality in the HANDLS Study Sample. *J. Natl. Med. Assoc.* **102**, 923–930 (2010).

136. Beydoun, M. A. *et al.* Racial disparities in adult all-cause and cause-specific mortality among us adults: mediating and moderating factors. *BMC Public Heal. 2016 161* **16**, 1–13 (2016).

137. Roy, B. *et al.* Education, race/ethnicity, and causes of premature mortality among middle-aged adults

in 4 US urban communities: Results from CARDIA, 1985-2017. *American Journal of Public Health* vol. 110 530–536 (2020).

138. Assari, S. & Lankarani, M. M. Race and Urbanity Alter the Protective Effect of Education but not Income on Mortality. *Front. Public Heal.* **0**, 100 (2016).

139. Leive, A. A. & Ruhm, C. J. Education Gradients in Mortality Trends by Gender and Race. (2021) doi:10.3386/W28419.

140. Evans, E. A., Grella, C. E., Washington, D. L. & Upchurch, D. M. Gender and race/ethnic differences in the persistence of alcohol, drug, and poly-substance use disorders. *Drug Alcohol Depend.* **174**, 128–136 (2017).

141. Graham, C. & Pinto, S. Unequal hopes and lives in the USA: optimism, race, place, and premature mortality. *J. Popul. Econ.* **32**, 665–733 (2019).

142. El-Toukhy, S., Sabado, M. & Choi, K. Trends in susceptibility to smoking. By race and Ethnicity. *Pediatrics* **138**, (2016).

143. White, H. R., Nagin, D., Replogle, E. & Stouthamer-Loeber, M. Racial differences in trajectories of cigarette use. *Drug Alcohol Depend.* **76**, 219–227 (2004).

144. Vásquez, E. *et al.* Differences in All-cause and Cardiovascular Mortality by Race, Ethnicity and Physical Activity. *Ethn. Health* **23**, 72 (2018).

145. Flegal, K. M., Kruszon-Moran, D., Carroll, M. D., Fryar, C. D. & Ogden, C. L. Trends in Obesity Among Adults in the United States, 2005 to 2014. *JAMA* **315**, 2284–2291 (2016).

146. Case, A. & Deaton, A. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proc. Natl. Acad. Sci.* **112**, 15078–15083 (2015).

147. Burt, K. The whiteness of the Mediterranean Diet: A historical, sociopolitical, and dietary analysis using Critical Race Theory. *Crit. Diet.* **5**, 41–52 (2021).

148. Kirkpatrick, S. I., Dodd, K. W., Reedy, J. & Krebs-Smith, S. M. Income and Race/Ethnicity Are Associated with Adherence to Food-Based Dietary Guidance among US Adults and Children. *J. Acad. Nutr. Diet.* **112**, (2012).

149. Kant, A. K., Graubard, B. I. & Kumanyika, S. K. Trends in Black-White Differentials in Dietary Intakes of U.S. Adults, 1971-2002. *Am. J. Prev. Med.* **32**, (2007).

150. Sijtsma, F. P. *et al.* Longitudinal trends in diet and effects of sex, race, and education on dietary quality score change: the Coronary Artery Risk Development in Young Adults study. *Am. J. Clin. Nutr.* **95**, 580–586 (2012).

151. Hiza, H. A. B., Casavale, K. O., Guenther, P. M. & Davis, C. A. Diet Quality of Americans Differs by Age, Sex, Race/Ethnicity, Income, and Education Level. *J. Acad. Nutr. Diet.* **113**, 297–306 (2013).

152. Boggs, D. A., Rosenberg, L., Rodríguez-Bernal, C. L. & Palmer, J. R. Long-Term Diet Quality Is Associated with Lower Obesity Risk in Young African American Women with Normal BMI at Baseline. *J. Nutr.* **143**, 1636–1641 (2013).

153. Boggs, D. A., Ban, Y., Palmer, J. R. & Rosenberg, L. Higher Diet Quality Is Inversely Associated with Mortality in African-American Women. *J. Nutr.* **145**, 547–554 (2015).

154. Ruiz, J. M., Steffen, P. & Smith, T. B. Hispanic mortality paradox: A systematic review and meta-analysis of the longitudinal literature. *Am. J. Public Health* **103**, (2013).

155. Ford, C. L. & Airhihenbuwa, C. O. The public health critical race methodology: Praxis for antiracism research. *Soc. Sci. Med.* **71**, 1390–1398 (2010).

156. Matsui, E. C., Perry, T. T. & Adamson, A. S. An Antiracist Framework for Racial and Ethnic Health Disparities Research. *Pediatrics* vol. 146 (2020).

157. Gee, G. C., Walsemann, K. M. & Brondolo, E. A life course perspective on how racism may be related to health inequities. *American Journal of Public Health* vol. 102 967–974 (2012).

158. Elder, T. E., Goddeeris, J. H. & Haider, S. J. Unexplained gaps and Oaxaca-Blinder decompositions. *Labour Econ.* **17**, 284–290 (2010).

159. Powell, L. M., Wada, R., Krauss, R. C. & Wang, Y. Ethnic disparities in adolescent body mass index in the United States: The role of parental socioeconomic status and economic contextual factors. *Soc. Sci. Med.* **75**, 469–476 (2012).

160. Lhila, A. & Long, S. What is driving the black-white difference in low birthweight in the US? *Health Econ.* **21**, 301–315 (2012).

161. Kirby, J. B., Taliaferro, G. & Zuvekas, S. H. Explaining racial and ethnic disparities in health care. *Med. Care* **44**, (2006).

162. Gaskin, D. J., Zare, H., Jackson, J. W., Ibe, C. & Slocum, J. Decomposing Race and Ethnic Differences in CVD Risk Factors for Mid-life Women. *J. Racial Ethn. Heal. Disparities* **8**, 174–185 (2021).

163. Newson, R. PUNAF : Stata module to compute population attributable fractions for cohort studies fractions and other scenario comparisons. (2016).

164. Shrider, E. A., Kollar, M., Chen, F. & Semega, J. Income and Poverty in the United States: 2020 Current Population Reports. (2020).

165. Ramakrishnan, R. *et al.* Objectively measured physical activity and all cause mortality: A systematic review and meta-analysis. *Prev. Med. (Baltim).* **143**, 106356 (2021).

166. Powell, L. M., Wada, R., Krauss, R. C. & Wang, Y. Ethnic disparities in adolescent body mass index in the United States: The role of parental socioeconomic status and economic contextual factors. *Soc. Sci. Med.* **75**, 469–476 (2012).

167. Hruby, A. & Hu, F. B. The Epidemiology of Obesity: A Big Picture. *Pharmacoeconomics* **33**, 673 (2015).

168. Woolf, S. H. Commentary: Progress in achieving health equity requires attention to root causes. *Health Aff.* **36**, 984–991 (2017).

169. Pescarini, J. M. *et al.* Evaluating the impact of the Bolsa Familia conditional cash transfer program on premature cardiovascular and all-cause mortality using the 100 million Brazilian cohort: a natural experiment study protocol. *BMJ Open* **10**, e039658 (2020).

170. Neves, J. A. *et al.* The Brazilian cash transfer program (Bolsa Família): A tool for reducing inequalities and achieving social rights in Brazil. *https://doi.org/10.1080/17441692.2020.1850828* **17**, 26–42 (2020).

171. Zewde, N. Universal Baby Bonds Reduce Black-White Wealth Inequality, Progressively Raise Net Worth of All Young Adults: *https://doi.org/10.1177/0034644619885321* **47**, 3–19 (2019).

172. Oliver, M. L. & Shapiro, T. M. Disrupting the Racial Wealth Gap: *https://doi.org/10.1177/1536504219830672* **18**, 16–21 (2019).

173. Mehra, R., Boyd, L. M. & Ickovics, J. R. Racial residential segregation and adverse birth outcomes: A systematic review and meta-analysis. *Soc. Sci. Med.* **191**, 237–250 (2017).

174. Watkins, S. L. & Gerrish, E. The relationship between urban forests and race: A meta-analysis. *J. Environ. Manage.* **209**, 152–168 (2018).

175. Gijsberts, C. M. *et al.* Race/Ethnic Differences in the Associations of the Framingham Risk Factors with Carotid IMT and Cardiovascular Events. *PLoS One* **10**, e0132321 (2015).

176. Lee, W. *et al.* Systematic review and meta-analysis of patient race/ethnicity, socioeconomics, and quality for adult type 2 diabetes. *Health Serv. Res.* **55**, 741–772 (2020).

177. Vart, P., van Zon, S. K. R., Gansevoort, R. T., Bültmann, U. & Reijneveld, S. A. SES, Chronic Kidney Disease, and Race in the U.S.: A Systematic Review and Meta-analysis. *American Journal of Preventive Medicine* vol. 53 730–739 (2017).

178. Rich, N. E. *et al.* Racial and Ethnic Disparities in Nonalcoholic Fatty Liver Disease Prevalence, Severity, and Outcomes in the United States: A Systematic Review and Meta-analysis. *Clin. Gastroenterol. Hepatol.* **16**, 198-210.e2 (2018).

179. Korous, K. M., Causadias, J. M. & Casper, D. M. Racial discrimination and cortisol output: A meta-analysis. *Social Science and Medicine* vol. 193 90–100 (2017).

180. Conway, B. N. *et al.* Cause-specific mortality by race in low-income Black and White people with Type 2 diabetes. *Diabet. Med.* **32**, 33–41 (2015).

181. Chi, G. C. *et al.* Trends in Acute Myocardial Infarction by Race and Ethnicity. *J. Am. Heart Assoc.* **9**, (2020).

182. The Lancet: advancing racial equality. https://www.thelancet.com/racial-equality?dgcid=etoc-edschoice_email_lancetbhm21.

183. Rivara, F. P. & Fihn, S. D. Call for Papers on Prevention and the Effects of Systemic Racism in Health. *JAMA Netw. Open* **3**, e2016825–e2016825 (2020).

184. USDA National Nutrient Database for Standard Reference, Legacy Release | Ag Data Commons. https://data.nal.usda.gov/dataset/usda-national-nutrient-database-standard-reference-legacy-release.

185. Griswold, M. G. *et al.* Alcohol use and burden for 195 countries and territories, 1990â€“2016: a systematic analysis for the Global Burden of Disease Study 2016. (2018) doi:10.1016/S0140-6736(18)31310-2.

186. BioLINCC: Atherosclerosis Risk in Communities Study (ARIC). https://biolincc.nhlbi.nih.gov/studies/aric/.

187. Cardiovascular Health Study (CHS) | NHLBI, NIH. https://www.nhlbi.nih.gov/science/cardiovascular-health-study-chs.

188. Feinleib, M., Kannel, W. B., Garrison, R. J., McNamara, P. M. & Castelli, W. P. The framingham offspring study. Design and preliminary data. *Prev. Med. (Baltim).* **4**, 518–525 (1975).

189. MESA - Multi-Ethnic Study of Atherosclerosis. https://www.mesa-nhlbi.org/.

190. Women's Health Initiative (WHI) | NHLBI, NIH. https://www.nhlbi.nih.gov/science/womens-health-initiative-whi.

191. NHANES I - Epidemiologic Followup Study. https://wwwn.cdc.gov/nchs/nhanes/nhefs/publicuse.aspx.

192. SOPH - REGARDS Study | UAB. https://www.uab.edu/soph/regardsstudy/.

193. Deaton, A. & Lubotsky, D. Mortality, inequality and race in American cities and states. *Soc. Sci. Med.* **56**, 1139–1153 (2003).

194. Murray, C. J. L. *et al.* Eight Americas: Investigating mortality disparities across races, counties, and race-counties in the United States. *PLoS Med.* **3**, 1513–1524 (2006).

195. Witbrodt, J., Mulia, N., Zemore, S. E. & Kerr, W. C. Racial/Ethnic disparities in alcohol-related problems: Differences by gender and level of heavy drinking. *Alcohol. Clin. Exp. Res.* **38**, 1662–1670 (2014).

196. Kwate, N. O. A. & Meyer, I. H. ON STICKS AND STONES AND BROKEN BONES. *Du Bois Rev. Soc. Sci. Res. Race* **8**, 191–198 (2011).

197. Factor, R., Williams, D. R. & Kawachi, I. Social resistance framework for understanding high-risk behavior among nondominant minorities: Preliminary evidence. *Am. J. Public Health* **103**, 2245–2251

(2013).

198. Poti, J. M., Mendez, M. A., Ng, S. W. & Popkin, B. M. Highly processed and ready-to-eat packaged food and beverage purchases differ by race/ethnicity among US households. *J. Nutr.* **146**, 1722–1730 (2016).

199. Brooks, C. J., Gortmaker, S. L., Long, M. W., Cradock, A. L. & Kenney, E. L. Racial/ethnic & socioeconomic disparities in hydration status among US adults & the role of tap water & other beverage intake. *Am. J. Public Health* **107**, 1387–1394 (2017).

200. Bahr, P. R. Race and nutrition: An investigation of Black-White differences in health-related nutritional behaviours. *Sociol. Heal. Illn.* **29**, 831–856 (2007).

201. Richman, L. S. & Jonassaint, C. The effects of race-related stress on cortisol reactivity in the laboratory: Implications of the duke lacrosse scandal. *Ann. Behav. Med.* **35**, 105–110 (2008).

202. Chowkwanyun, M. & Reed, A. L. Racial Health Disparities and Covid-19 — Caution and Context. *N. Engl. J. Med.* **383**, 201–203 (2020).

203. Pollack, C. E. *et al.* Should Health Studies Measure Wealth?. A Systematic Review. *American Journal of Preventive Medicine* vol. 33 250–264 (2007).

204. Orfield, G., Frankenberg, E. & Garces, L. M. Statement of American social scientists of research on school desegregation to the U.S. Supreme court in parents v. Seattle school district and Meredith v. Jefferson county. *Urban Rev.* **40**, 96–136 (2008).

205. Catalano, R. The effect of deviations from trends in national income on mortality: The Danish and USA data revisited. *Eur. J. Epidemiol.* **13**, 737–743 (1997).

206. Baird, S., Friedman, J. & Schady, N. Aggregate income shocks and infant mortality in the developing world. *Rev. Econ. Stat.* **93**, 847–856 (2011).

207. Ferreira, F. H. G. & Schady, N. Aggregate economic shocks, child schooling, and child health. *World Bank Res. Obs.* **24**, 147–181 (2009).

208. Despard, M., Grinstein-Weiss, M., Guo, S., Taylor, S. & Russell, B. Financial Shocks, Liquid Assets, and Material Hardship in Low- and Moderate-Income Households: Differences by Race. *J. Econ. Race, Policy* **1**, 205–216 (2018).

209. Seguin, C. & Rigby, D. National Crimes: A New National Data Set of Lynchings in the United States, 1883 to 1941. *Socius Sociol. Res. a Dyn. World* **5**, 237802311984178 (2019).

210. Deaton, A. Height, health, and development. *Proc. Natl. Acad. Sci. U. S. A.* (2007) doi:10.1073/pnas.0611500104.

211. Duggan, C. P., Kurpad, A., Stanford, F. C., Sunguya, B. & Wells, J. C. Race, ethnicity, and racism in the nutrition literature: an update for 2020. *Am. J. Clin. Nutr.* **112**, 1409–1414 (2020).

APPENDIX: CHAPTER 2

Methods

Food and nutrient intake estimation

Since each cohort used a different food frequency questionnaire (FFQ), it was necessary to harmonize dietary intake to a standard reference period. FFQs provided data on the frequency (number of times consumed) and/or quantities (number and size of servings) of different foods consumed within a specific recall period. Using this information we estimated dietary intake with the following steps:

1. Quantities of individual foods consumed in metric units (mg/gm/ml/lt) were calculated separately for each cohort using conversion factors made available by USDA's National Nutrient Database for Standard Reference[184]. For example if a period reported that they eat chicken breast 3 times a week and ate 2 servings in a meal then the estimated number of servings per week is 6. A typical serving of chicken breast is 42 g, thus the estimated weekly intake is 6 x 42g = 252g.
2. We also calculated the energy (kcal) intake for each food in this step[81]. For example 100g of chicken breast contain 79 kcal so 252g contain about 199 calories.
3. The individual food items were then grouped into broad food groups including fruits, starchy vegetables, fibrous vegetables, whole grain, refined grain, legumes, lipids, nuts and seeds, dairy (milk, cheese, and yogurt), fish/shellfish, poultry, processed meats, eggs, and red meat. For example chicken and turkey are assigned to the food group poultry.
4. We calculated the total intake for each specific food group by adding up the intakes of individual foods comprising the respective group. For example 252 g of chicken plus 48 g of Turkey is 300 g of poultry consumed per week.
5. Next, we harmonized the recalls periods into one uniform 24 hour reference period. For example, if the recall for a FFQ was one week, then the total estimate would be divided by 7 to estimate daily mean intake.

Construction of physical activity tertiles

Questions on physical activity differed substantially across the cohorts and thus precluded the construction of a standardized measure such as metabolic equivalent of task minutes. Depending on the cohort, these questions included questions on whether the individual engaged in, and the total amount of time spent on activities including various sports, exercise, walking, bike riding, jogging, household chores, sitting, and sleeping, among others.

We followed a pragmatic approach using all the physical activity questions available in each cohort and performed principal component analyses (PCA) to extract the first component (largest amount of shared variance among the questions). The first component provided a composite index of physical activity status. PCA was performed for each cohort separately and cohort specific tertiles for calculated for each cohort by binning the first component into three categories reflecting mild, moderate and heavy physical activity. The REGARDS cohort only had one question on physical activity (times per week exercise) which were used to classify individuals into three categories. The final indicator represents a ‘within-in’ cohort measure of physical activity intensity. The limitation is that tertiles across cohort may differ slightly on which individuals are categorized as highly active.

Alcohol intake tertiles

Alcohol intake was measured by combing data on past and current consumption in a variable with three categories: (1) no or low (2) moderate and (3) heavy. A classification that accounts for past consumption is necessary because it accurately estimates an appropriate cumulative dose response function for alcohol[185]. We took the average of the current and past reported consumption and created cohort specific tertiles as the final indicator. A uniform standard was applied for males and females.

Table A2 1: Details of prospective cohorts used in analyses

Cohort	Year established	Sample size	Description
Atherosclerosis Risk in Communities Study (ARIC)	1987	14950	Data were collected from four US communities (Winston-Salem, NC, Jackson, MS, Minneapolis, MN and Baltimore, MD) for men and women aged 45-64, at baseline.[186]
Cardiovascular Health Study (CHS)	1988	5888	Data were collected from four US communities (Sacramento, CA; Hagerstown, MD; Winston-Salem, NC; and Pittsburgh, PA), for men and women aged 65 or older, at baseline.[187]
Framingham Heart Study-Offspring (FHSO)	1971	5013	Data were collected from Framingham, Massachusetts, for men and women aged 13-59 years, at baseline.[188]
Multi-Ethnic Study of Atherosclerosis (MESA)	1999	6814	Data were collected from six US communities (Columbia University, New York; Johns Hopkins University, Baltimore; Northwestern University, Chicago; UCLA, Los Angeles; University of Minnesota, Twin Cities; Wake Forest University, Winston Salem) for men and women aged 44-84 years, at baseline.[189]
Women's Health Initiative Study (WHI)	1991	93676	Data were collected from 40 centers across the US for women aged 49-81 years, at baseline.[190]
NHANES I Epidemiologic Follow-up Study (NHEFS)	1982	9281	Data were collected from multiple centers across the US for men and women aged 25-75 years, at baseline.[191]
Reasons for Geographic and Racial Differences in Stroke (REGARDS)	2003	29716	Data were collected from multiple centers across the US for men and women aged 45-98 years, at baseline.[192]

Table A2 2: Mortality events and baseline summary statistics for covariates across seven prospective cohorts

	ARIC	CHS	FRAM	MESA	NFS1	REGARDS	WHI
All-cause mortality							
Person- years	90802.5	59,255.8	63496.5	50,043.6	178,255.1	192,575.9	944,869.4
Deaths	1388	3471	82	390	856	3937	8264
Rate	0.0152	0.0585	0.0013	0.0078	0.0048	0.0204	0.0087
95% CI	[0.014-0.016]	[0.056-0.061]	[0.001-0.002]	[0.007-0.009]	[0.004-0.005]	[0.019-0.021]	[0.008-0.009]
Covariates							
Age, years	55.5	72.1	33.9	62.0	45.1	64.9	63.6
	[55.3,55.6]	[72.0,72.3]	[33.5,34.3]	[61.8,62.3]	[44.8,45.4]	[64.7,65.0]	[63.6,63.7]
Male, %	52.6	42.0	48.9	47.0	36.0	43.9	0.0
	[51.3,53.9]	[40.5,43.4]	[46.6,51.3]	[45.7,48.3]	[34.9,37.1]	[43.2,44.6]	[0.0,0.0]
African American (Black), %	23.2	4.7	0.0	27.9	11.3	32.7	7.0
	[22.1,24.4]	[4.0,5.3]	[0.0,0.0]	[26.7,29.0]	[10.6,12.0]	[32.1,33.4]	[6.9,7.2]
Non-Hispanic White, %	76.8	94.8	98.8	38.7	88.0	67.3	85.1
	[75.6,77.9]	[94.2,95.5]	[98.3,99.4]	[37.5,39.9]	[87.3,88.8]	[66.6,67.9]	[84.9,85.3]
Inflation adjusted income, 1000 $	82.4	61.5	268.6	76.9	68.2	61.6	80.0
	[81.0,83.8]	[60.0,62.9]	[263.9,273.3]	[75.6,78.2]	[67.1,69.4]	[61.1,62.0]	[79.8,80.3]
Education, years	14.0	11.6	15.0	13.5	11.6	14.1	14.7
	[13.8,14.1]	[11.5,11.7]	[14.9,15.2]	[13.4,13.5]	[11.5,11.7]	[14.1,14.1]	[14.7,14.7]
Current smoker, %	26.6	11.4	10.1	12.8	28.1	13.0	5.7
	[25.4,27.7]	[10.5,12.4]	[8.7,11.6]	[12.0,13.7]	[27.1,29.1]	[12.6,13.5]	[5.5,5.8]
Pack-years, 10 years	2.1	1.6	0.1	1.1	1.3	1.2	0.0
	[2.1,2.2]	[1.5,1.6]	[0.1,0.1]	[1.0,1.1]	[1.2,1.3]	[1.2,1.2]	[0.0,0.0]
Moderate/high alcohol consumption, %	20.6	52.6	78.5	31.2	28.1	7.1	32.3
	[19.5,21.7]	[51.1,54.1]	[76.5,80.4]	[30.0,32.3]	[27.1,29.1]	[6.7,7.5]	[32.0,32.6]
Highly active, %	30.9	33.8	32.1	33.6	9.6	30.1	33.4
	[29.7,32.1]	[32.4,35.2]	[29.9,34.3]	[32.4,34.7]	[9.0,10.3]	[29.5,30.7]	[33.0,33.7]
Energy intake, 100 kcal	21.9	21.9	19.3	20.8	18.2	21.9	17.9
	[21.7,22.1]	[21.7,22.1]	[18.9,19.7]	[20.6,21.0]	[18.0,18.3]	[21.8,22.1]	[17.8,17.9]
Height, cm	167.4	164.8	168.3	166.4	166.8	169.5	161.8
	[167.1,167.7]	[164.6,165.1]	[167.8,168.7]	[166.2,166.7]	[166.4,167.1]	[169.4,169.7]	[161.7,161.8]
Body mass index, kg/m^2	28.4	26.4	25.3	28.3	26.2	29.0	27.0
	[28.3,28.6]	[26.3,26.5]	[25.1,25.5]	[28.1,28.4]	[26.1,26.3]	[28.9,29.0]	[27.0,27.0]
Systolic blood pressure, mmhg	143.5	140.8	84.8	126.8	128.5	126.6	126.6
	[143.1,144.0]	[140.2,141.4]	[84.3,85.2]	[126.3,127.4]	[128.1,128.9]	[126.4,126.8]	[126.5,126.8]
N	5500	4252	1736	6087	7802	19696	82206

Table A2 3: Principal Component Analysis of dietary intake in the Women's Health Initiative Study (WHI) cohort

Components	**1**	**2**	**3**	**4**	**5**
Mapped to diet	1	2	3	4	5
Eigenvalues	2.1	2.0	1.5	1.4	1.1
Eigenvectors					
Refined grains			0.20		0.20
Whole grains		0.17	0.37		
Nuts and seeds			0.68		-0.18
Starchy vegetables		0.33		0.42	
Fibrous vegetables		0.50		-0.16	
Fruits		0.45		-0.20	
Legumes		0.53		0.26	
Milk					0.88
Yogurt				-0.56	0.24
Cheese	0.25		0.36	-0.24	
Poultry	0.48		-0.26		
Fish	0.48				
Red meat	0.29			0.38	
Processed meat	0.41			0.20	
Fats/oils			0.29	0.29	0.18
Eggs	0.38	-0.20	0.16		

Notes: Blank cells indicate that the eigenvector was below 0.15

Table A2 4: Principal Component Analysis of dietary intake in the Atherosclerosis Risk in Communities Study (ARIC) cohort

Components	**1**	**2**	**3**	**4**	**5**
Mapped to dietary pattern in WHI	2	4	5		3
Eigenvalues	**2.1**	**1.8**	**1.4**	**1.2**	**1.1**
Eigenvectors					
Refined grains		0.42		-0.17	0.18
Whole grains	0.26	-0.29	0.29	-0.32	0.34
Nuts and seeds					0.86
Starchy vegetables	0.16	0.34	0.20		
Fibrous vegetables	0.44			0.20	
Fruits	0.30	-0.16	0.21	0.22	
Legumes	0.44	0.24			
Milk			0.59		
Yogurt			-0.15	0.65	0.21
Cheese			0.25	0.57	
Poultry	0.45				
Fish	0.47		-0.17		
Red meat		0.51			
Processed meat		0.48			
Fats/oils			0.59		
Eggs					

Notes: Blank cells indicate that the eigenvector was below 0.15

Table A2 5: Principal Component Analysis of dietary intake in the Multi-Ethnic Study of Atherosclerosis (MESA) cohort

Components	**1**	**2**	**3**	**4**	**5**
Mapped to dietary pattern in WHI	2	1	5	3	
Eigenvalues	**2.6**	**2.2**	**1.5**	**1.4**	**1.3**
Eigenvectors					
Refined grains		0.17	0.48		
Whole grains				0.37	0.36
Nuts and seeds				0.38	0.22
Starchy vegetables	0.18	0.29	0.18		
Fibrous vegetables	0.49				
Fruits	0.41	-0.20			
Legumes	0.43		0.16		
Milk			0.68		
Yogurt				0.78	-0.15
Cheese	0.37	0.36			-0.28
Poultry	0.43		-0.24		
Fish		0.58			
Red meat		0.55			0.20
Processed meat			0.42		0.28
Fats/oils				-0.15	0.74
Eggs		0.17	0.48		

Notes: Blank cells indicate that the eigenvector was below 0.15

Table A2 6: Principal Component Analysis of dietary intake in the Framingham Heart Study-Offspring (FHSO) cohort

Components	**1**	**2**	**3**	**4**	**5**
Mapped to dietary pattern in WHI	2	4			5
Eigenvalues	2.0	2.0	1.2	1.2	1.1
Eigenvectors					
Refined grains		0.47			
Whole grains	0.41				0.36
Nuts and seeds	0.37	0.34	-0.20	-0.31	
Starchy vegetables		0.51			
Fibrous vegetables	0.41		0.24		-0.25
Fruits	0.54				-0.15
Legumes					
Milk					0.85
Yogurt	0.33	-0.20		0.17	
Cheese			0.76		
Poultry	0.27		0.37		
Fish	-0.19	0.37		0.26	
Red meat		0.27	-0.27	0.48	
Processed meat		0.36	0.24	-0.21	
Fats/oils				0.71	
Eggs					0.85

Notes: Blank cells indicate that the eigenvector was below 0.15

Table A2 7: Principal Component Analysis of dietary intake in the NHANES I Epidemiologic Follow-up Study (NFS1) cohort

Components	**1**	**2**	**3**	**4**	**5**
Mapped to dietary pattern in WHI	4	2			3
Eigenvalues	**1.6**	**1.6**	**1.5**	**1.4**	**1.4**
Eigenvectors					
Refined grains				0.60	
Whole grains					0.59
Nuts and seeds		0.15		0.15	0.32
Starchy vegetables	0.37	0.29	-0.22		
Fibrous vegetables		0.61			
Fruits		0.55			
Legumes		0.19		0.50	
Milk					0.56
Yogurt			0.35	-0.30	
Cheese	0.35	0.16	0.29	-0.33	
Poultry	-0.16		0.53	0.18	
Fish			0.57		
Red meat	0.57				-0.19
Processed meat	0.33	-0.19		0.31	
Fats/oils	0.32				0.37
Eggs	0.34	-0.33	0.29		0.17

Notes: Blank cells indicate that the eigenvector was below 0.15

Table A2 8: Principal Component Analysis of dietary intake in the Cardiovascular Health Study (CHS) cohort

Components	**1**	**2**	**3**	**4**	**5**
Mapped to dietary pattern in WHI	2	4		3	5
Eigenvalues	**2.2**	**2.1**	**1.4**	**1.3**	**1.2**
Eigenvectors					
Refined grains		0.28		0.24	0.24
Whole grains	0.15	-0.22	0.20	0.23	0.37
Nuts and seeds				0.74	
Starchy vegetables		0.31	0.37		
Fibrous vegetables	0.40		0.35		
Fruits	0.33	-0.16	0.24		0.16
Legumes			0.64		
Milk					0.82
Yogurt	0.33	-0.21	-0.25	0.32	
Cheese	0.31		-0.31	0.33	
Poultry	0.49				
Fish	0.48				
Red meat		0.51		-0.17	
Processed meat		0.52			
Fats/oils		0.38	-0.22		0.21
Eggs		0.28		0.24	0.24

Notes: Blank cells indicate that the eigenvector was below 0.15

Table A2 9: Principal Component Analysis of dietary intake in the Reasons for Geographic and Racial Differences in Stroke (REGARDS) cohort

Components	**1**	**2**	**3**
Mapped to dietary pattern in WHI		2	3
Eigenvalues	1.9	1.4	1.4
Eigenvectors			
Refined grains		-0.12	0.73
Whole grains		0.23	0.62
Fibrous vegetables	0.40	0.40	
Fruits		0.69	
Legumes	0.56	0.13	
Milk		0.35	0.14
Poultry	0.63	-0.15	
Fats/oils	0.35	-0.36	0.22

Notes: Blank cells indicate that the eigenvector was below 0.15

Table A2 10: Pooled correlations of dietary patterns with food groups across seven prospective cohorts

	Western	Prudent	Fatty-grainy	Lacto-mod	Carnivore
Refined grains	0.21	0.25	0.35	0.34	0.41
Whole grains	-0.08	0.34	0.49	0.2	0.18
Nuts and seeds	0.06	0.16	0.62	0	0.09
Starchy vegetables	0.41	0.37	0.29	0.21	0.36
Fibrous vegetables	-0.19	0.68	0.18	0.06	0.28
Fruits	-0.24	0.6	0.18	0.11	0.11
Legumes	0.18	0.59	0.13	0.15	0.21
Milk	0	0.1	0.13	0.75	0.11
Yogurt	-0.48	0.22	0.19	0.23	0.14
Cheese	-0.09	0.16	0.45	0.22	0.43
Poultry	-0.02	0.33	-0.04	0.03	0.6
Fish	0.03	0.32	0.08	0	0.62
Red meat	0.51	0.03	0.15	0.2	0.53
Processed meat	0.44	-0.05	0.21	0.11	0.6
Fats/oils	0.44	-0.04	0.39	0.34	0.35
Eggs	0.19	-0.05	0.21	0.04	0.49
N	104,390	104,390	104,390	96,992	81670

Notes: Correlation between dietary input variables and final dietary patterns used in analysis.

Table A2 11: Baseline consumption levels of food groups across seven prospective cohorts

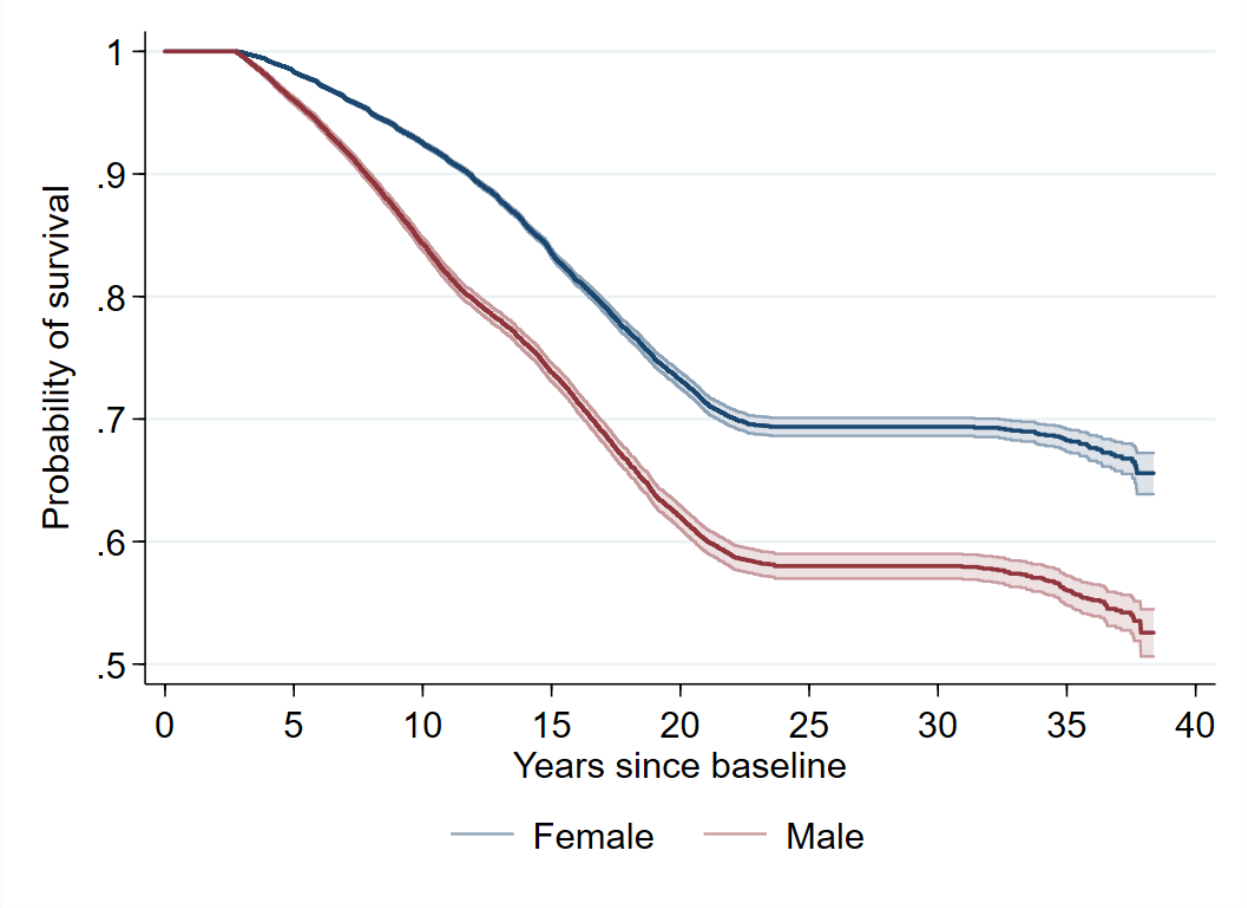
	ARIC	CHS	FRAM	MESA	NFS1	REGARDS	WHI
Fruits, grams per day	193.0	175.1	56.4	122.9	239.1	205.8	125.1
	[189.4,196.6]	[172.8,177.3]	[54.5,58.2]	[120.7,125.1]	[236.9,241.3]	[203.9,207.7]	[124.6,125.7]
Starchy vegetables, grams per day	118.5	56.7	77.4	37.2	49.8		43.0
	[116.6,120.4]	[55.9,57.5]	[75.3,79.4]	[36.5,37.9]	[49.3,50.2]		[42.8,43.2]
Vegetables, grams per day	76.5	212.0	47.8	139.7	184.0	210.5	256.6
	[75.2,77.8]	[209.5,214.4]	[46.7,48.8]	[137.5,142.0]	[182.6,185.3]	[208.8,212.2]	[255.6,257.5]
Whole grains, grams per day	41.2	32.8	32.3	35.4	32.1	24.4	50.4
	[40.2,42.2]	[32.3,33.3]	[31.1,33.5]	[34.6,36.2]	[31.7,32.4]	[24.2,24.6]	[50.1,50.7]
Refined grains, grams per day	108.3	44.0	52.8	91.7	90.7	103.2	129.5
	[106.5,110.1]	[43.4,44.6]	[51.6,54.1]	[90.1,93.4]	[90.0,91.4]	[102.2,104.3]	[129.1,130.0]
Legumes, grams per day	41.8	15.8		29.8	6.8	26.7	45.2
	[41.1,42.5]	[15.4,16.1]		[29.2,30.5]	[6.7,6.9]	[26.5,27.0]	[45.0,45.4]
Lipids, grams per day	72.2	74.9	68.9	74.5	58.2	71.8	53.4
	[70.6,73.7]	[73.6,76.1]	[66.7,71.1]	[73.1,75.9]	[57.7,58.7]	[71.3,72.3]	[53.1,53.8]
Nuts and seeds, grams per day	6.6	7.2	8.8	24.0	9.6		12.7
	[6.4,6.8]	[7.0,7.4]	[8.5,9.2]	[23.3,24.6]	[9.5,9.8]		[12.5,12.8]
Milk/cheese/ yogurt, ml/g per day	211.3	78.9	494.0	62.0	152.0	234.9	122.2
	[206.3,216.4]	[77.8,80.0]	[474.5,513.4]	[60.5,63.5]	[150.7,153.2]	[232.4,237.4]	[121.7,122.8]
Processed meat, grams per day	17.6	18.1	27.2	10.6	17.6		17.9
	[17.2,18.0]	[17.7,18.4]	[26.2,28.1]	[10.3,10.9]	[17.4,17.8]		[17.8,18.1]
Red meat, grams per day	58.8	16.3	60.3	14.6	60.4		26.1
	[57.7,59.9]	[16.0,16.7]	[58.8,61.8]	[14.2,15.0]	[59.9,60.9]		[25.9,26.2]
Poultry/fish, grams per day	58.9	34.9	97.2	41.5	53.5	44.0	34.9
	[58.0,59.9]	[34.4,35.3]	[95.1,99.4]	[40.8,42.2]	[53.2,53.8]	[43.7,44.4]	[34.7,35.0]
N	5500	4252	1736	6087	7802	19696	82206

Notes: Numbers reported are sample means with 95% confidence intervals in parenthesis. Atherosclerosis Risk in Communities Study (ARIC), Cardiovascular Health Study (CHS), Framingham Heart Study-Offspring (FRAM), Multi-Ethnic Study of Atherosclerosis (MESA), NHANES I Epidemiologic Follow-up Study (NFS1), Reasons for Geographic and Racial Differences in Stroke (REGARDS), Women's Health Initiative Study (WHI).

Table A2 12 Baseline covariate distribution within highest quintile of dietary patterns

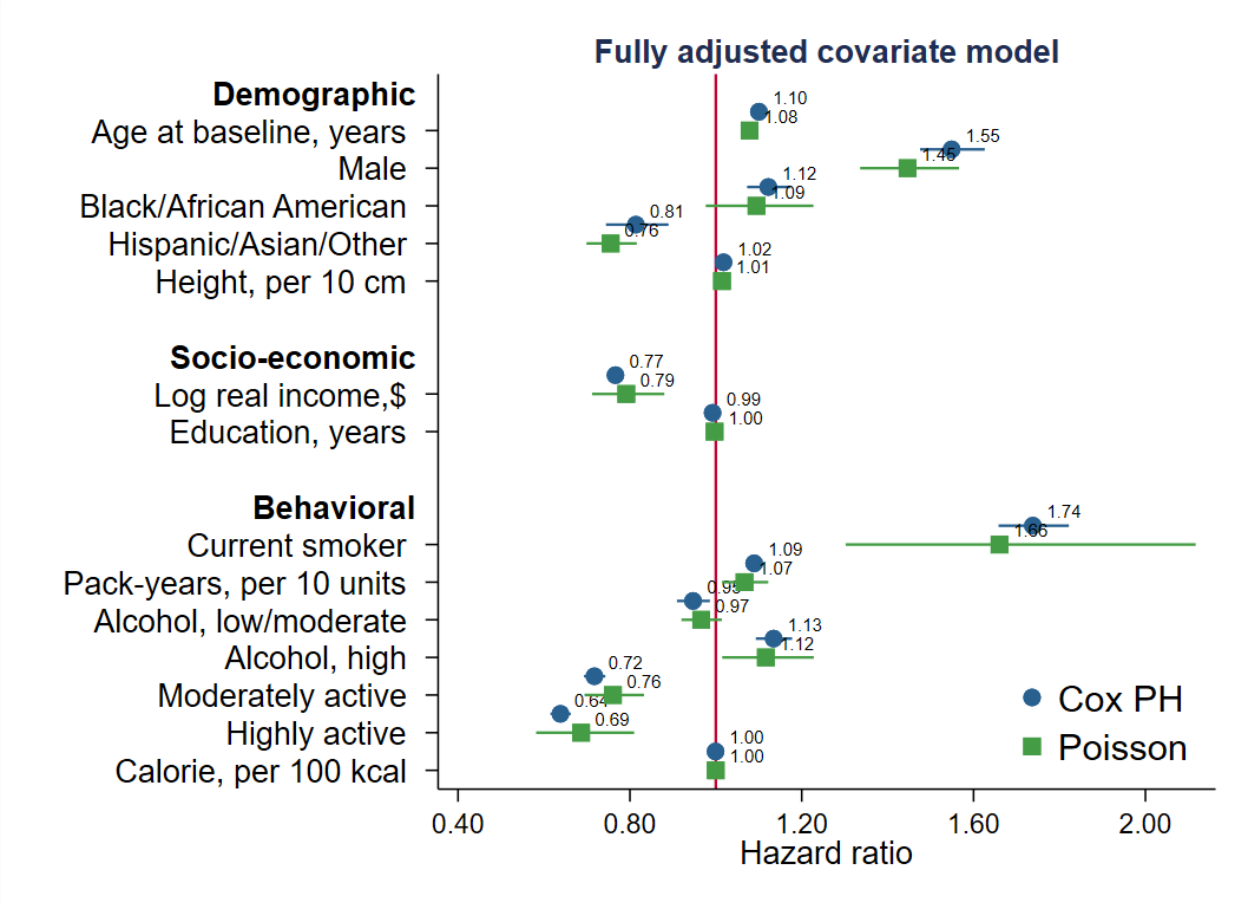
	Western	Prudent	Fatty-grainy	Lacto-mod	Carnivore
Age at baseline, years	61.3	62.6	63.2	62.7	62.7
	[61.2,61.5]	[62.5,62.8]	[63.1,63.3]	[62.6,62.9]	[62.6,62.8]
Male, %	12.9	13.6	16.1	8.5	0.0
	[12.4,13.4]	[13.2,14.0]	[15.6,16.6]	[8.1,8.9]	[0.0,0.0]
Black, %	8.9	12.6	12.6	4.8	9.2
	[8.6,9.3]	[12.2,13.0]	[12.2,13.0]	[4.5,5.1]	[8.7,9.6]
White, %	84.8	80.6	82.3	89.0	83.6
	[84.3,85.3]	[80.1,81.1]	[81.8,82.8]	[88.6,89.5]	[83.1,84.2]
Income, 1000 $	75.7	80.9	74.2	82.3	78.0
	[75.1,76.4]	[80.3,81.5]	[73.7,74.7]	[81.7,83.0]	[77.4,78.6]
Education, years	13.8	14.6	14.3	14.6	14.6
	[13.7,13.8]	[14.6,14.7]	[14.3,14.4]	[14.6,14.6]	[14.6,14.6]
Current smoker, %	12.6	6.0	8.8	7.5	7.3
	[12.1,13.0]	[5.7,6.3]	[8.5,9.2]	[7.1,7.9]	[6.9,7.7]
Pack-years, per 10 units	0.4	0.4	0.5	0.3	0.0
	[0.4,0.4]	[0.4,0.4]	[0.5,0.5]	[0.2,0.3]	[0.0,0.0]
Alcohol high intake, %	33.3	28.5	28.9	31.3	34.3
	[32.6,33.9]	[28.0,29.1]	[28.3,29.5]	[30.6,31.9]	[33.6,35.0]
Moderately active, %	32.5	32.4	32.8	33.3	32.3
	[31.8,33.1]	[31.8,33.0]	[32.2,33.4]	[32.7,34.0]	[31.6,33.0]
Highly active, %	23.2	40.4	33.2	35.4	30.5
	[22.6,23.8]	[39.8,41.0]	[32.6,33.8]	[34.7,36.1]	[29.8,31.2]
Calorie, per 100 kcal	22.4	22.2	24.1	22.7	24.5
	[22.3,22.5]	[22.2,22.3]	[24.0,24.1]	[22.6,22.8]	[24.4,24.6]
Height, cm	163.6	163.9	164.6	163.2	162.3
	[163.5,163.7]	[163.8,164.0]	[164.5,164.7]	[163.1,163.3]	[162.2,162.4]
BMI	28.2	27.1	27.5	27.2	29.2
	[28.1,28.3]	[27.1,27.2]	[27.4,27.6]	[27.1,27.3]	[29.1,29.3]
SBP	128.8	127.2	128.2	127.0	127.3
	[128.5,129.0]	[127.0,127.4]	[128.0,128.5]	[126.7,127.2]	[127.0,127.5]
N	20365	24183	23761	18940	16439

Figure A2 1: Sex-specific Kaplan-Meier survival curves for all-cause mortality



Notes: Width of the bands are 95% confidence intervals.

Figure A2 2: Covariate model for all-cause mortality



Notes: Cox PH= Cox proportional hazard model. Circles and squares represent Hazard ratios and bands represent 95% confidence intervals. Model include fixed effects for each prospective cohort and 5 year birth interval cohorts. The seven cohorts used are Atherosclerosis Risk in Communities Study (ARIC), Cardiovascular Health Study (CHS), Framingham Heart Study-Offspring (FRAM), Multi-Ethnic Study of Atherosclerosis (MESA), NHANES I Epidemiologic Follow-up Study, Reasons for Geographic and Racial Differences in Stroke (REGARDS), Women's Health Initiative Study (WHI).

Natural direct and indirect effects

The mediation model considers the impact of an intervening variable which is theorized to transmit the effect of an independent variable onto an outcome[104]. Consider a dietary pattern *T* (e.g., Mediterranean diet) and an individual-level mediator *M* (e.g., blood pressure or body mass). Here, *T* can take two values, 1 or 0, which represent the individual consuming and not consuming the foods that correspond to the dietary pattern, respectively. We define these dietary states with subscript *t*. Similarly, *M* can also take two values, 1 or 0, which represent the potential states of the mediator in which the individual conforms or does not conform to the dietary pattern, respectively. We define these individual mediation states with subscript *m*. In addition, we assume that *M* is a function of variable set *W* and *T*.

For an outcome $Y_{t,m}$, the natural direct effect (NDE) of *T* on *Y* can be expressed as $\boldsymbol{E}\left[Y_{1,M0} - Y_{0,M0}\right]$. The expression represents the difference between the potential outcomes that would be observed under *t*=1 versus *t*=0, holding the mediator constant at the value it would take under *t*=0 (e.g. its natural value under a no-Mediterranean diet scenario). More succinctly, the NDE is represented as:

$$\sum_{\mathrm{m}}\sum_{\mathrm{w}}\left(\mathbf{E}[\mathrm{Y} \mid T=1, M=m, W=w] - \mathbf{E}[\mathrm{Y} \mid T=0, M=m, W=w]\right)$$
$$\cdot\ P(M=m \mid T=0, W=w)P(W=w)$$

The natural indirect effect (NIE) can be expressed as $\mathbf{E}\left[\mathrm{Y}_{0,\mathrm{M1}} - \mathrm{Y}_{0,\mathrm{M0}}\right]$, which represents the difference in the potential outcomes that would be observed under *m*=1 versus *m*=0, holding the treatment constant at *t*=0. In the linear case, the total effect (TE) can be estimated by adding the NDE and NIE.

Figure A3 1: Conceptual difference between confounders and mediators

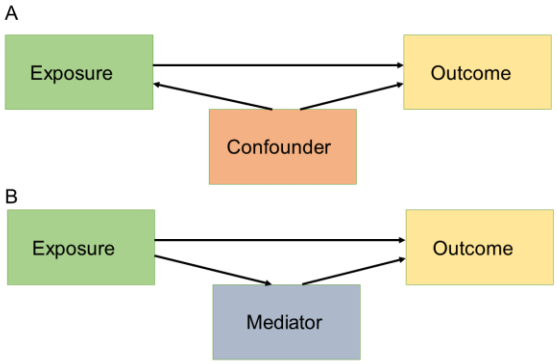


Table A3 1: Known confounders of the exposure-mediator-outcome relationship

Confounder	Justification and pathways
High blood pressure	
Age	Blood pressure tends to be higher in older individuals
Race/ethnicity	Individuals of African heritage tend to develop high blood pressure earlier than White individuals.
Family history	There may be a genetic pathway to inherit high blood pressure
Weight	Higher body weight requires more oxygen and nutrients for cells which increase the flow of blood through vessels. Consequently blood pressure increases.
Physical activity	Low levels of physical activity are associated with higher heart rates, blood pressure and weight.
Smoking	Smoking can immediately raise your blood pressure temporarily, and the chemicals in tobacco can damage the lining of your artery walls, raising blood pressure.
Sodium and potassium	Sodium in your diet can cause your body to retain fluid, which increases blood pressure. Potassium helps balance the amount of sodium in your cells, thus regulates blood pressure.
Alcohol	Heavy drinking can damage your heart thereby affecting blood pressure.
Stress	High stress levels are associated with elevated blood pressure.
Pre-existing conditions	Certain chronic conditions also may increase your risk of high blood pressure.
Overweight/obesity	
Age	Hormonal changes, lower muscle mass and a less active lifestyle associated with aging increase your risk of obesity.
Sex	Females often gain weight during pregnancy which may be retained.
Family history	Genes regulate body fat storage, distribution and conversion of food to energy.
Physical activity	Low levels of physical activity are associated with a positive energy balance i.e. higher calorie intake than calories metabolized leading to weight gain.
Smoking	Quitting smoking is often associated with weight gain.
Stress	Individuals often seek more high-calorie food when experiencing stressful situations.
Alcohol	Excess alcohol consumption may lead to positive energy balance via high calorie intake.
Diet	
Income	Higher income may provide access to healthier diets and modes to enable higher or lower physical activity.
Education	Higher education may be associated with a better knowledge of healthy diets, ways of cooking, perils of smoking and benefits of exercise.
Smoking	Smoking may suppress appetite.
Alcohol	Frequent drinking may be associated with poor dietary choices.
Physical activity	High physical activity may stimulate appetite.
Family history	Families tend to share dietary habits.

Table A3 2 Percent of excess risk mediated by body mass index and systolic blood pressure for the association between dietary patterns and all-cause mortality

	Western		Prudent		Fatty-grainy		Lacto-mod		Carnivore	
	%	95%CI	%	95%CI	%	95%CI	%	95%CI	%	95%CI
Through BMI	15.8	[15.8,72.4]	4.4	[1.5,12.7]	41.9	[0,100]	5.2	[2.1,20.0]	27.3	[14.3,47.6]
Through SBP	14.1	[5.4,48.3]	6.3	[1.8,14.1]	13.9	[0,100]	2.9	[0,11.3]	3.6	[1.4,6.7]
BMI and SBP	27.1	[12.9,92.1]	9.7	[4.5,21.4]	49.9	[0,100]	7.3	[1.0,20.0]	26.4	[14.2,46.0]

Percentages are calculated post-hoc from regressions shown in figure 2 using equation 3.1. Percentile based CIs are calculated using the bootstrap method. CIs are truncated at 0 and 100%. BMI=Body mass index, SBP=Systolic blood pressure. . The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Table A3 3 Structural equation models: Direct and indirect effects of dietary patterns on all-cause mortality through body mass and blood pressure

	Western		Prudent		Fatty-grainy		Lacto-mod		Carnivore	
	RR	**95%CI**	**RR**	**95%CI**	**RR**	**95%CI**	**RR**	**95%CI**	**RR**	**95%CI**
Panel A (SEM equation 3.2)										
Poisson model, outcome=death										
Diet coefficient, low v high	1.06	[1.00,1.13]	0.92	[0.88,0.97]	0.98	[0.93,1.04]	0.92	[0.86,0.97]	1.15	[1.05,1.27]
BMI coefficient, per 5 units	1.05	[1.05,1.06]	1.06	[1.05,1.07]	1.06	[1.05,1.07]	1.05	[1.04,1.06]	1.06	[1.04,1.07]
SBP coefficient, per 10 units	1.08	[1.07,1.10]	1.08	[1.06,1.09]	1.08	[1.06,1.09]	1.08	[1.07,1.10]	1.06	[1.04,1.09]
Panel B (SEM equation 3.3)										
Ordinary least squares model, exposure=diet, low v high										
Outcome= BMI, per 5 units	0.22	[0.20,0.24]	-0.09	[-0.11,-0.07]	-0.15	[-0.17,-0.13]	-0.04	[-0.07,-0.02]	0.75	[0.72,0.78]
Outcome= SBP, per 10 units	0.27	[0.23,0.30]	-0.09	[-0.12,-0.05]	-0.08	[-0.12,-0.04]	-0.04	[-0.08,0.01]	0.18	[0.13,0.23]
Panel C (mediation)										
Total effect of diet, RR	1.10	[1.03,1.16]	0.91	[0.86,0.96]	0.97	[0.92,1.02]	0.91	[0.86,0.96]	1.22	[1.11,1.34]
Indirect effect of diet, RR	1.03	[1.03,1.04]	0.99	[0.98,0.99]	0.98	[0.98,0.99]	0.99	[0.99,1.00]	1.06	[1.0,1.07]
Panel D (PMM)										
Indirect effect of diet, %	34.1	[18.1;74.6]	12.8	[7.2;24.6]	49.1	[0,100]	5.8	[2.3;17.6]	28.3	[14.2,46.0]
Indirect effect though BMI, %	19.2	[9.7;43.4]	7.5	[4.3;13.0]	35.6	[0,100]	3.8	[1.6;11.5]	23.3	[14.3,47.6]
Indirect effect through SBP, %	14.9	[7.5;29.6]	5.3	[2.2;10.1]	13.8	[0,100]	2.0	[0.1;10.6]	5.0	[1.4,6.7]

PMM= proportion mediated measure. Percentile based CIs are calculated using the bootstrap method for PMM. CIs are truncated at 0 and 100% for PMM. Diet scores range from 0 to 1. The coefficient is interpreted as difference in hazard comparing the lowest percentile of the diet to the highest percentile, assuming a linear risk curve. BMI=Body mass index per 5 unit increment. SBP=Systolic blood pressure per 10 mmhg. Models are adjusted for age, sex, race, income, education, smoking, alcohol, physical activity, calorie intake, study cohort fixed effects and birth cohort fixed effects. . The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Table A3 4 Mediation by body mass index and systolic blood pressure measured at follow-up

	Western		Prudent		Fatty-grainy		Lacto-mod		Carnivore	
	%	95%CI	%	95%CI	%	95%CI	%	95%CI	%	95%CI
PERM										
Through BMI	7.9	[4.5;74.5]	1.9	[0,6.0]	28.9	[0,100]	2.5	[0,8.5]	12.1	[3.0,23.8]
Through SBP	5.9	[1.8;13.5]	1.9	[0,5.9]	11.3	[0,100]	0.2	[0,8.1]	2.1	[0.6,5.4]
BMI and SBP	12.6	[6.4,100]	3.5	[0,9.6]	35.7	[0,100]	2.3	[0,12.3]	11.7	[2.6,24.2]
PMM										
Through BMI	10.5	[6.2,22.1]	3.8	[2.2,11.6]	24.7	[0,100]	1.9	[0.3,5.8]	10.1	[1.2,24.2]
Through SBP	8.9	[5.5,44.2]	2.1	[0,6.0]	10.1	[0,92.8]	0	[0,1.2]	2.9	[1.4,5.5]
BMI and SBP	19.3	[12.3,38.8]	5.9	[2.8,17.9]	34.8	[0,100]	1.2	[0,7.1]	12.9	[3.7,24.9]

Percentile based CIs are calculated using the bootstrap method. CIs are truncated at 0 and 100%. BMI=Body mass index at follow-up, SBP=Systolic blood pressure at follow-up. . The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

Table A3 5 Mediation by body mass index and systolic blood pressure measured not adjusting for energy intake

	Western	Prudent	Fatty-grainy	Lacto-mod	Carnivore
	%	%	%	%	%
PERM					
Through BMI	15.2	-3.9	3.7	-5.9	22.1
Through SBP	6.6	0.6	4.8	-0.3	5.6
BMI and SBP	19.6	-2.9	7.5	-5.6	23.2
PMM					
Through BMI	17.7	-4.9	-21.2	-11.9	18.6
Through SBP	11.3	-0.9	-6.2	-4.1	6.7
BMI and SBP	29.0	-5.9	-21.4	-15.9	25.4

Values are truncated at 0 and 100%. BMI=Body mass index at follow-up, SBP=Systolic blood pressure at follow-up. . The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables. The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups. The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups.

APPENDIX: CHAPTER 4

Methods

Theoretical concepts for analyzing the role of racism in public health

Several conceptual frameworks have been proposed to understand the complex relationships between race/ethnicity and health outcomes. These include Public Health Critical Race Theory (PHCRT), the Antiracism framework (ARF), the Social Determinants of Health (SODH) framework, and the Life Course framework (LCF), among others[129,155–157]. PHCRT has four core principles: (1) racialization – that socially constructed racial/ethnic categories are used to order groups in society, (2) race consciousness – the acknowledgement of racial dynamics in society and personal life, (3) social location – an individual's or group's position within a social hierarchy, and (4) inequity – lack of fairness or justice towards racial/ethnic groups[155]. ARF is grounded on the principles of opposing racism and promoting racial tolerance in all domains of socio-economic status (SES), environment, and access to quality health care[156]. SODH contends that the environments in which people are born, live, learn, work, lay, worship and age, all collectively determine health, and quality-of-life outcomes and risks[129]. LCF asserts that racism and health inequities should be viewed from a life course perspective of social pathways (the paths that individuals take through the life course), sensitive periods (periods of profound importance), linked lives (social networks), latency periods (the period between exposure and disease appearance), stress proliferation (one stressor leads to other stressors), historic period (historical events affect lives uniformly), and cohort effects (social change affects lives differentially over time)[157]. Elements found in each of these frameworks are important in determining ways in which race/ethnicity may influence both diet quality and mortality.

Outcomes: At the level of the outcome, there is substantial evidence to show that racial/ethnic differences in mortality and morbidity are large and persistent. Black individuals have higher mortality rates than White individuals, and aggregate regional mortality rates are robustly associated with racial composition[193]. Moreover, there are large differences across racial/ethnic lines between the best and worst off groups. For example, one study found that the life expectancy gap between urban Black males and Asian females was 20.7 years[194]. Black individuals have a higher prevalence of coronary heart disease (CHD) and cardiovascular disease (CVD)[25,167], with age, wealth, healthcare access, and health behaviors accounting for only about half the differences in CVD rates between White and Black individuals[162]. Evidently, differences in health outcomes between racial/ethnic groups are dependent on a large set of factors that may systematically drive convergent or divergent tendencies.

Responses/ immediate determinants: At the immediate level (labeled 'response' in figure 7), mortality is determined by behavioral, psychological and physiological factors. Behavioral factors are the most commonly adjusted factors in nutrition epidemiological studies; they include smoking, alcohol and physical activity[132]. This set of variables comprises the most important risk factors that explain mortality and are linked to diets in the population. These factors are also moderated by racial/ethnic differences. For example, compared with White men, Black and Hispanic men have higher rates of injuries, accidents, adverse health outcomes, social costs, and greater work/legal penalties related to alcohol consumption[195]. There are also documented differential rates of smoking across race/ethnicity. Trends in smoking differ by race and cigarette smoking is more prevalent and more frequent for White than Black males, at least through young adulthood[142,143]. Another study reported that Black individuals are significantly less physically active compared to White individuals[144]. Taken together, the evidence suggests that racial/ethnic dynamics related

to behavioral factors are far from straightforward. While some deleterious factors clearly increase the racial gap in mortality, others may close it.

After behavioral factors are psychological factors that include resilience, self-esteem and identity. By accepting innate deficiencies of one's self and one's group, internalized racism can lead to lower self-esteem and psychological well-being[196]. On resilience, research suggests that unhealthy behaviors of ethnic minorities may reflect resistance—an effort to express opposition, assert independence, and reject the dominant society's norms[197]. Such factors however are difficult to capture in secondary data, unless specifically designed questionnaires are implemented to ascertain them. Therefore, these factors are not directly accounted for. Instead, the study controls for macro-level factors that may be upstream indicators of a larger environment of racism. Next are physiological responses that represent objectively measurable markers of health status, including cardiovascular and metabolic responses. Physiological responses are usually adequately measured in cohort studies. For example, blood pressure and body-mass index are frequently measured, and have strong associations with racial/ethnic disparities in the US[145]. However, these factors may also be mediators of diet, so are not included in sensitivity analyses to prevent blocking variables on causal pathways.

Nutrition and race/ethnicity: Importantly, studies consistently report higher diet quality scores among White individuals compared to Black individuals[134,135]. However, evidence also suggests that compared with White households, Black and Hispanic households have lower purchases of highly processed and ready-to-eat foods[198]. Further, Black individuals are also more likely to be inadequately hydrated compared to White individuals; a difference that is largely dependent on tap water access across communities[199]. This is an example of how race-based differences in access to utilities can influence intake of essential nutrients. Another study found that Black individuals exhibit less healthy nutritional practice, relative to White individuals, and that this relationship independent of SES and cost[200]. This is partially attributable to the fact that Black individuals can be subjected to intense residential segregation characterized by limited nutritional infrastructure. It is also well established that stressful conditions are associated with making unhealthy food choices, thus, racism induced stress may even explain some of differences in diet quality across race/ethnicity. Moreover, some differences in diets are attributable to cultural aspects of race/ethnicity. Deep engrained dietary practices may persist for centuries and some foods that are eaten from racially dominant groups may be culturally inappropriate from some minority groups. Thus dietary differences across racial/ethnic groups are complex and require careful nuanced interpretations. Dietary recommendations in particular should adopt a race/ethnicity conscious position.

Underlying factors and related pathways

One step removed from the response variables are underlying factors. The effect of the underlying factors is theoretically mediated via 'pathways' listed in figure 7. For example, affordability of optimal diets may be a function of income and identifying the right foods might be a function of knowledge. Gender may also play a role in how race effects individuals in different situations. For instance, one study indicated that Black females had higher levels of cortisol after media coverage of a racially charged incident[201]. Racism can even work via geographic stigmatization, place-based risks and resource deficits. For example the uneven geographic distribution of preventive care services across the US or the concentration of respiratory exposures and toxic sites in low-income, minority-populated parts[202]. Studies further show that the median wealth of White individuals is an order of magnitude larger than that of Black individuals and that wealth affects health over and above education[203]. Moreover, Black individuals have the highest average mortality

rates in the US, at all education levels, over time[137–139]. Racism restricts socioeconomic mobility by limiting access to quality elementary and high school education[204]. Racism can further unfavorably affect health through stigma, stereotypes, and discrimination. These facets of racism can mediate effects of differential income and education levels and to wide-ranging societal resources[125]. While this study does not measure or control for all the pathways listed, it acknowledges their importance as plausible mediating mechanisms. The study adjusts for income, education, sex and age to account for these important sources of variation in diets and mortality.

Basic determinants

This study takes a life-course perspective with regards to the most distal level, and acknowledges that racism is a determinant of inequities in the (macro) basic-determinants of mortality in the population via mechanisms that operate through the lifecourse[157]. While the effects of aggregate current systematic geographic differences will be accounted for using study cohort fixed effects (period effects), there needs to be explicit control for sensitive and latency periods. Whereas these factors are difficult to measure, it is important to control for them because they proxy an unobservable environment of stress and discrimination in the population when individuals were growing up. From a statistical perspective, these latent factors need to be controlled because there may be time dependent confounding from racial disparities experienced during childhood, to diet and mortality relationships estimated using data collected in adulthood.

The influence of early childhood factors will be captured via aggregate variables (country level) that are matched with the individual's birth year to account for conditions in early childhood. At a fundamental level GDP is a determinant of mortality in the US[205] and recent analyses indicate significant economic losses that are associated with racial disparities[205]. Some studies also show that shocks (wars, recessions, epidemics) experienced in infancy are strongly correlated with mortality[206]; while others claim that in high income countries child health outcomes improve during aggregate income shocks[207]. It is also documented that the relationship between shocks and hardship is also moderated by race/ethnicity[208]. In the current study, effects of these aggregate shocks are controlled directly via GDP at birth year and residual aggregate effects by 5-year interval birth-cohort fixed effects. The study also tries to capture the presence of overt aggregate racism experienced at birth by controlling for total cases of lynching that were recorded in the birth year of the participants[209]. Other factors that are useful in capturing the effects of racial disparities at the aggregate level are infant mortality (infants born to Black women suffer the highest risks in the US) and fertility rate (Black women have higher fertility rates in the US). In addition, to account for cumulative nutritional effects experienced during childhood and adolescence as well as some genetic endowment effects, the study adjusts for attained adult height. Height is an indicator of health that is useful in historical work, because of its association with higher economic output, physical capability of work, and cognitive or/and genetic potential[210].

Finally, the study will borrows concepts from PHCRT and ARF for interpreting the results of the decomposition. This study acknowledges that race is not a viable variable to capture genetic differences, but rather represents a composite index of social constructs that may disadvantage one group versus another. After numerous variables are adjusted for and differences in means are accounted, the unexplained variation will be interpreted with race consciousness and with an objective of promoting equity along racial lines, in a manner that benefits all sections of the society[211].

Table A4 1 Association between dietary patterns and all-cause mortality within race/ethnicity groups

	RR	95% CI	RR	95% CI	RR	95% CI
	Non-Hispanic White		Black/African American		Hispanic/Asian/Other	
Western	1.08	[1.03,1.13]	1.03	[0.89,1.20]	2.04	[1.95,2.12]
Prudent	0.91	[0.86,0.96]	0.88	[0.80,0.97]	0.93	[0.90,0.97]
Fatty-grainy	0.98	[0.94,1.03]	0.90	[0.85,0.96]	0.88	[0.74,1.04]
Lacto-mod	0.91	[0.87,0.96]	0.94	[0.81,1.09]	0.99	[0.94,1.04]
Carnivore	1.16	[1.05,1.29]	1.65	[1.18,2.32]	1.15	[0.76,1.74]
Age, yr	1.02	[0.93,1.13]	1.03	[0.96,1.10]	1.20	[1.07,1.34]
Male, 0/1	1.44	[1.31,1.59]	1.48	[1.25,1.75]	1.14	[1.00,1.31]
Height, cm	1.01	[1.00,1.02]	1.02	[0.95,1.10]	1.05	[0.97,1.14]
Log real income,$	0.81	[0.72,0.91]	0.76	[0.71,0.81]	0.73	[0.62,0.87]
Education, years	1.00	[0.98,1.01]	0.98	[0.95,1.02]	1.02	[1.00,1.03]
Current smoker, 0/1	1.69	[1.29,2.23]	1.44	[1.26,1.66]	1.76	[1.34,2.29]
Pack-years, per 10 units	1.07	[1.01,1.13]	1.07	[1.02,1.12]	1.11	[1.05,1.18]
Alcohol moderate, 0/1	0.95	[0.88,1.01]	1.06	[0.92,1.22]	1.09	[1.03,1.14]
Alcohol high, 0/1	1.11	[1.01,1.23]	1.10	[1.00,1.20]	1.29	[1.13,1.48]
Moderately active, 0/1	0.77	[0.69,0.87]	0.72	[0.65,0.79]	0.71	[0.54,0.93]
Highly active, 0/1	0.68	[0.56,0.83]	0.74	[0.68,0.82]	0.63	[0.57,0.69]
N	100667		16255		8595	

Notes: Models are adjusted for age, sex, income, education, smoking, alcohol, physical activity, calorie intake, height, cohort fixed effects and birth cohort fixed effects. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables (not present in REGARDS cohort). The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups (not present in REGARDS and NHANES cohort). The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups (only observed in the WHI cohort).

Table A4 2 Oaxaca-Blinder decomposition for explained differences in mortality rates due to differences in the sample means of exposures by race/ethnicity

	%	95% CI	%	95% CI
	Black/African American versus Non-Hispanic White		Hispanic/Asian/Other versus Non-Hispanic White	
Western	0.2	[-1.3,1.8]	-1.4	[-2.3,-0.5]
Prudent	1.1	[-0.1,2.2]	0.2	[-0.6,1.0]
Fatty-grainy	2.8	[-0.6,6.3]	1.5	[-2.2,5.2]
Lacto-mod	4.2	[-9.4,17.9]	0.1	[-3.6,3.8]
Carnivore	2.8	[0.5,5.6]	-1.1	[-4.3,2.1]
Income	50.9	[36.6,65.3]	10.9	[5.8,16.1]
Smoking	21.1	[15.6,26.6]	-6.9	[-11.6,-2.3]
Physical activity	9.2	[5.5,12.9]	2.5	[1.1,3.9]
Alcohol intake	-6.9	[-16.8,2.9]	-1.9	[-5.3,1.4]

Notes: Models are adjusted for age, sex, income, education, smoking, alcohol, physical activity, calorie intake, height, cohort fixed effects. For factors other than dietary patterns, decomposition results are shown for risk factors that were significant in the regression models. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables (not present in REGARDS cohort). The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups (not present in REGARDS and NHANES cohort). The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups (only observed in the WHI cohort).

Table A4 3 Population attributable fractions for predicted changes in mortality rates at hypothetical scenarios for levels exposures by race/ethnicity

	%	Min,Max	%	Min,Max	%	Min,Max
	Non-Hispanic White		Black/African American		Hispanic/Asian/Other	
Western at 0%	4.0	[1.7,6.3]	-0.8	[-6.4,4.5]	29.6	[25.8,33.2]
Prudent at 100%	4.1	[0.8,7.2]	7.2	[-1.1,14.8]	3.7	[-3.5,10.5]
Fatty-grainy at 100%	0.6	[-1.1,2.3]	4.5	[0.8,7.9]	3.1	[-2.7,8.5]
Lacto-mod at 100%	4.5	[2.2,6.8]	7.4	[-1.8,15.7]	0.3	[-8.7,8.6]
Carnivore at 0%	7.4	[2.8,11.8]	18.3	[5.9,29.0]	14.1	[0.9,25.5]
Income at $100,000	12.2	[5.7,18.2]	21.8	[14.9,28.1]	16.5	[6.2,25.6]
Smoking at 5%	6.6	[2.5,10.5]	11.9	[7.0,16.5]	1.1	[-0.1,2.3]
High physical activity at 100%	36.0	[21.2,48.1]	33.7	[25.7,40.8]	41.7	[31.8,50.2]
Alcohol intake at 0%	2.7	[-2.5,7.2]	0.6	[-1.7,2.8]	9.9	[8.2,11.5]
BMI at 22.5 kq/m^2	5.7	[3.9,7.4]	10.1	[3.9,15.8]	12.3	[10.2,14.4]
SBP at 100 mmhg	18.2	[14.8,21.6]	18.3	[9.3,26.5]	13.1	[2.3,22.7]

Notes: Models are adjusted for age, sex, income, education, smoking, alcohol, physical activity, calorie intake, height, cohort fixed effects. The 'Western' DP: high in red and processed meats, fats/oils, starchy vegetables, and low in yogurt, fruits and vegetables (not present in REGARDS cohort). The 'Prudent' DP: high in fibrous vegetables, fruits, legumes, starchy vegetables, whole grains; and low in eggs, fats and processed meat. The 'Fatty-grainy' DP: high in nuts and seeds, cheese, whole grains, fats/oils and low in poultry. The 'Lacto-mod' DP: high in milk and moderate to low on all other food groups (not present in REGARDS and NHANES cohort). The 'Carnivore' DP: high in red and processed meats, poultry, fish, eggs and moderate on other food groups (only observed in the WHI cohort).

Table A4 4 Differences in all-cause mortality attributable to differences in body mass index and systolic blood pressure across racial/ethnic groups using Oaxaca-Blinder Decomposition

	% Black/African American versus Non-Hispanic White	95% CI	% Hispanic/Asian/Other versus Non-Hispanic White	95% CI
Body mass index	13.1	[2.5,23.8]	0.5	[-0.1,1.1]
Systolic blood pressure	15.5	[9.7,21.3]	0.04	[-0.2,0.3]

Notes: Models are adjusted for age, sex, income, education, smoking, alcohol, physical activity, calorie intake, height, cohort fixed effects.

Figure A4 1: Basic determinants of all-cause mortality in the US, 1898-1958

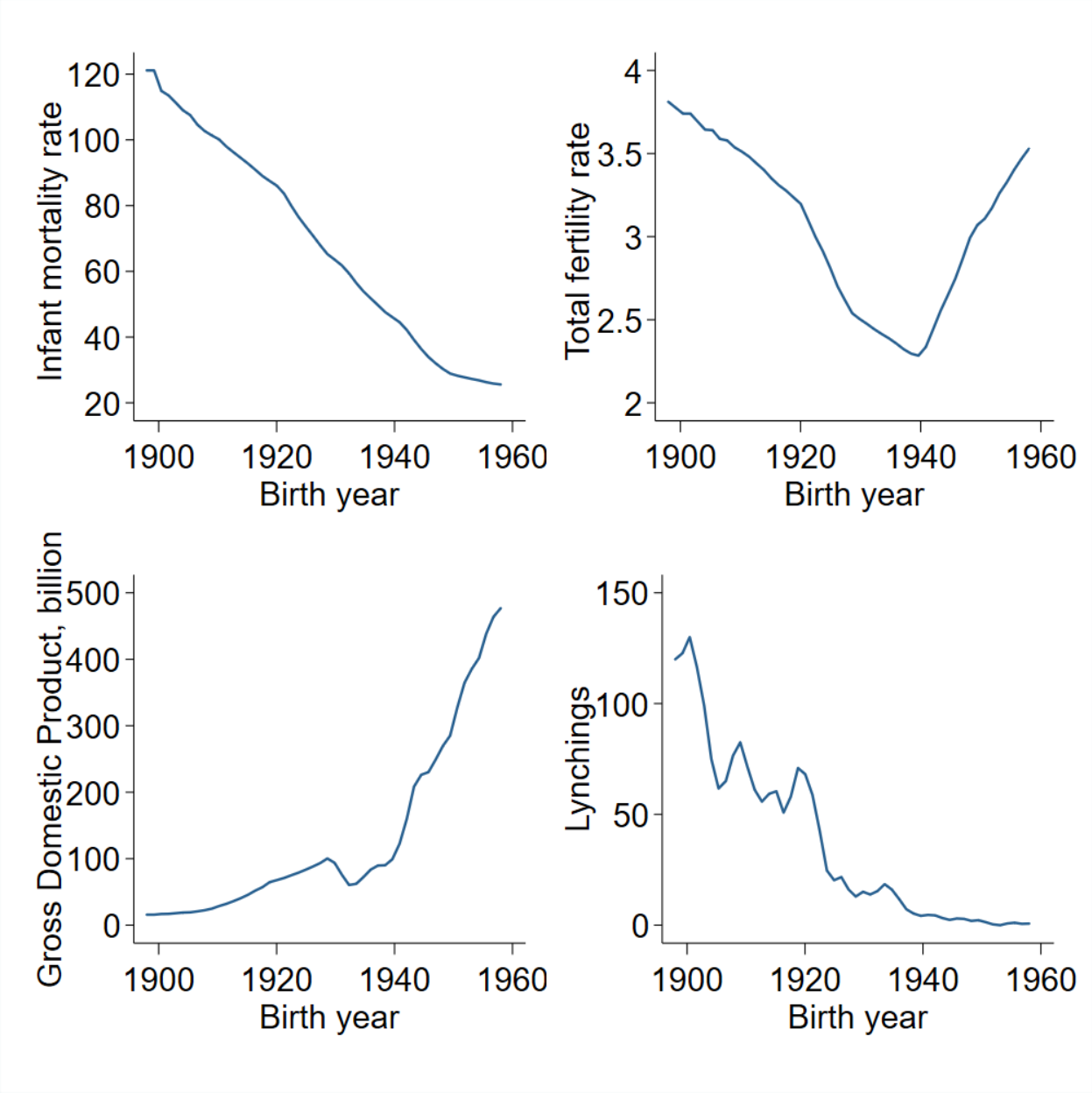
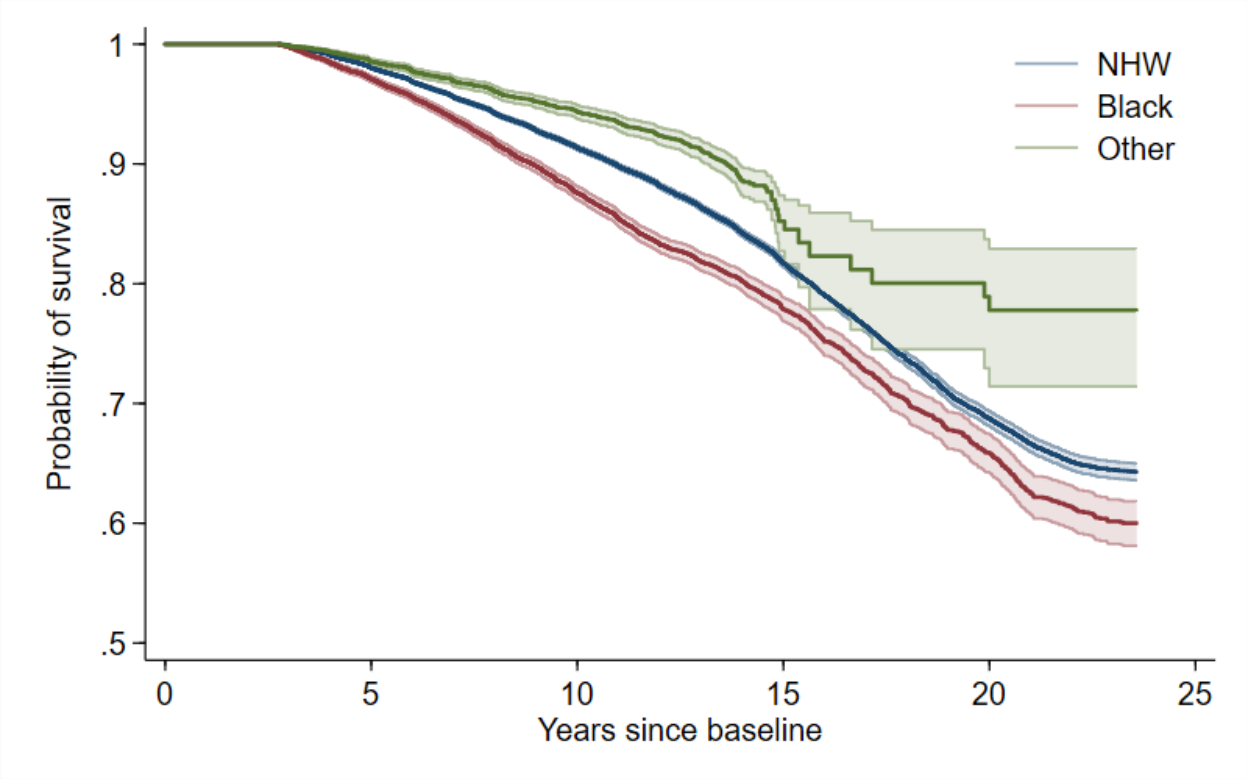


Figure A4 2: Kaplan-Meier survival curves for all-cause mortality stratified by race/ethnicity



Notes: NHW=Non-Hispanic White