Supplementary Online Content

Bhattacharya R, Zekavat SM, Uddin MM, et al. Association of diet quality with prevalence of clonal hematopoiesis and adverse cardiovascular outcomes. *JAMA Cardiol*. Published online June 9, 2021. doi:10.1001/jamacardio.2021.1678

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Diet Quality Categorization, Diet Quality Treated as a Continuous Variable, Dietary Survey Response Scores, and Definition of Cardiovascular Events

Diet Quality Categorization:

The primary exposure in this study is diet quality, rated in an ordinal categorical scale: unhealthy, intermediate, or healthy. In defining the quality of a participant's diet, we compared available questionnaire response fields to the Healthy Eating Index 2015 (HEI), which is derived from the U.S. Department of Health and Human Services and U.S. Department of Agriculture joint "2015-2020 Dietary Guidelines for Americans" Key Recommendations. The document recommends that a healthy diet should be enhanced for healthy foods including a variety of vegetables, whole fruits, grains (at least half whole grains), fat-free or low-fat dairy, oils and a variety of protein foods including seafood, lean meats, poultry, eggs, legumes, nuts, seeds and soy products.

In agreement with the HEI an individual's diet was categorized as unhealthy if a participant reported below-median intake of fruits and vegetables (field codes 1289, 1299, 1309, 1319), and above-median intake of red meat or processed meat intake (field codes 1369, 1379, 1389, 1349), or above-median frequency of adding salt to their diet (field code 1478). Frequency of lean meats such as fish and poultry (field codes 1329, 1339, 1359) were not included in diet classification. Conversely, diets were classified as healthy if intake of healthy elements was above the median, and unhealthy elements was below the median. The remainder of individuals were categorized as having intermediate diet quality.

Diet Quality Treated as a Continuous Variable:

As a means of replicating the primary analysis with an alternate method of scoring diet quality, diet quality was treated as a continuous variable by assigning point values to each of three dietary elements: (1) Fruits and Vegetables (field codes 1289, 1299, 1309, 1319); (2) Red and Processed Meats (field codes 1369, 1379, 1389, 1349); and (3) Added salt (field code 1478). For the fruits and vegetables category, and the red and processed meat category each individual's intake was ranked as a percentile. For fruits and vegetables, the highest quantile of intake (top 20%) was assigned 3 points, the lowest quantile (lowest 20%) was assigned 1 point, and intermediate values were assigned 2 points. For red and processed meats the lowest quantile of intake (lowest 20%) was assigned 3 points, the highest quantile (highest 20%) was assigned 1 point, and intermediate values were assigned 2 points. For added salt, those who indicated they added salt to their food "usually" or "always" were assigned 1 point, those who indicated "sometimes" were assigned 2 points, and those who indicated "never" were assigned 3 points. The dietary quality score generated ranges from 3 to 9 for each individual. Such a diet score assumes equal contribution of each dietary element to the total score, and as such was not used for the primary analysis.

The continuous diet quality score was then used to replicate the primary analysis examining the relationship between the presence of CHIP and diet quality. Multivariate logistic regression was used while adjusting for age, sex, smoking status, type 2 diabetes mellitus, Townsend deprivation index, and the first 10 principal components of genetic ancestry. The diet score was also binned into "Low Diet Score" (3-5), "Intermediate Diet Score" (6-7), and "High Diet Score" (8-9) to examine prevalence of CHIP (**Supplementary Figure 7).**

Dietary Survey Response Scores:

Healthy dietary elements included raw vegetable/salad intake, cooked vegetable intake, fresh fruit intake and dried fruit intake. Vegetable intake was measured in tablespoons per day, while fruit intake was measured in pieces per day.

Unhealthy dietary elements included processed meat intake, beef intake, pork intake, lamb intake and frequency of adding salt to one's food. Meat intake was measured on an ordinal categorial scale (0="Never", 1= "Less than once per week", 2="1 time per week", 3="2-4 times per week", 4="5-6 times per week”, or 5=“Once or more daily”). Added salt was also measured on an ordinal categorical scale (1=“Never/Rarely”, 2=“Sometimes”, 3=“Usually”, or 4=“Always”).

Definition of Cardiovascular Events:

Incident CAD and death were defined in the UK Biobank by means of enrollment interview and linking participant information to national electronic health records and death records. Participants’ electronic health records (EHRs), including inpatient International Classification of Disease (ICD-10) diagnosis codes and Office of Population and Censuses Surveys (OPCS-4) procedure codes, were integrated into UK Biobank. Individuals were defined as having CAD on the basis of at least one of the following criteria: (1) Myocardial infarction (MI), coronary artery bypass grafting, or coronary artery angioplasty documented in their medical history at the time of enrollment by a trained nurse; (2) Hospitalization for an ICD-10 code for acute myocardial infarction (I21.0, I21.1, I21.2, I21.4, I21.9); (3) Hospitalization for an OPCS-4 coded procedure: coronary artery bypass grafting (K40.1–40.4, K41.1–41.4, K45.1–45.5); (4) Hospitalization for an OPCS-4 coded procedure: coronary angioplasty with or without stenting (K49.1–49.2, K49.8–49.9, K50.2, K75.1–75.4, K75.8–75.9).

eTable 1. Distribution of Dietary Elements

	Overall		Unhealthy Diet		Intermediate Diet		Healthy Diet	
Dietary Element	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Raw Salad	2.33	2.28	1.3	1.14	2.26	2.16	3.71	3.23
Cooked Vegetables	2.78	1.85	2	1.08	2.74	1.79	3.69	2.53
Fresh Fruit	2.27	1.61	1.3	1.01	2.23	1.57	3.33	1.89
Dried Fruit	0.91	1.85	0.2	0.62	0.86	1.8	1.83	2.49
Oily Fish	1.64	0.92	1.4	0.85	1.65	0.89	1.75	1.16
Non-Oily Fish	1.78	0.79	1.7	0.73	1.79	0.77	1.69	1.04
Poultry	2.3	0.9	2.4	0.75	2.36	0.83	1.52	1.3
Beef	1.38	0.82	1.8	0.76	1.44	0.79	0.38	0.54
Pork	1.14	0.73	1.5	0.71	1.19	0.69	0.26	0.47
Lamb	1.11	0.71	1.4	0.7	1.16	0.68	0.31	0.51
Processed Meats	1.85	1.06	2.7	0.84	1.92	1	0.48	0.66
Added Salt	1.62	0.84	3.3	0.45	1.55	0.77	1.25	0.43

eTable 2. Frequency of CHIP Driver Genes

Gene	N	Proportion of CHIP
DNMT3A	1600	0.6382
TET2	391	0.156
ASXL1	156	0.0622
TP53	41	0.0164
PPM1D	39	0.0156
SRSF2	25	0.01
JAK2	24	0.0096
GNB1	23	0.0092
SF3B1	20	0.008
CBL	14	0.0056
GNAS	14	0.0056
NF1	13	0.0052
BCORL1	12	0.0048
CUX1	10	0.004
RUNX1	10	0.004
BCOR	8	0.0032
SMC1A	8	0.0032
BRCC3	7	0.0028
CALR	7	0.0028
ASXL2	6	0.0024
SETD2	6	0.0024
KDM6A	5	0.002
KRAS	5	0.002
ZRSR2	5	0.002
CTCF	4	0.0016
ETV6	4	0.0016
IDH2	4	0.0016
MPL	4	0.0016
NRAS	4	0.0016
PHF6	4	0.0016
CREBBP	3	0.0012
PRPF8	3	0.0012
RAD21	3	0.0012
SETBP1	3	0.0012
STAG2	3	0.0012
CSF3R	2	0.0008
EZH2	2	0.0008
FLT3	2	0.0008
IKZF1	2	0.0008
PDS5B	2	0.0008
SMC3	2	0.0008
BRAF	1	0.0004
CEBPA	1	0.0004
EP300	1	0.0004
ETNK1	1	0.0004
GATA2	1	0.0004
PRPF40B	1	0.0004
SUZ12	1	0.0004

CHIP counts by gene and as a proportion of total CHIP positive individuals in the study population. (CHIP, clonal hematopoiesis of indeterminate potential)

eTable 3. Association Between CHIP Prevalence and Diet Quality

Diet Quality	CHIP Count	Stratum Subtotal Count	CHIP Prevalence	P-value
VAF >2%				
Unhealthy	162	2271	7.13%	2.6×10^{-3}
Intermediate	2177	38552	5.65%	
Healthy	168	3288	5.11%	
VAF >10% (Large CHIP Clones only)				
Unhealthy	77	2271	3.39%	4.0×10^{-4}
Intermediate	897	38552	2.33%	
Healthy	61	3288	1.86%	

Raw counts and rates of CHIP (VAF>2%), and Large CHIP (VAF>10%) demonstrate a significant trend of lower prevalence of both CHIP and Large CHIP across improving dietary categories using the Chi-squared test of trend. (CHIP, clonal hematopoiesis of indeterminate potential; VAF, variant allele fraction)

eTable 4. Sensitivity Analysis Including Additional Covariates to Account for Potential Confounding in the Association Between CHIP Prevalence and Diet Quality

Variable	Odds Ratio	Confidence Interval	P-value
Unhealthy Diet	1.24	[1.02-1.51]]	0.03
Intermediate Diet	1	--	Ref
Healthy Diet	0.91	[0.74-1.11]	0.38
Age	1.08	[1.07-1.08]	<0.001
Male	0.95	[0.86-1.05]	0.34
Current Smoking	1.08	[0.98-1.20]	0.11
Hypertension	1.05	[0.94-1.17]	0.42
Systolic Blood Pressure	1.00	[1.00-1.00]	0.76
Hyperlipidemia	0.89	[0.78-1.02]	0.09
Total Cholesterol	0.97	[0.92-1.01]	0.16
Triglycerides	1.01	[0.96-1.06]	0.78
T2DM	0.92	[0.65-1.26]	0.62
Blood Sugar	0.99	[0.94-1.04]	0.75
Townsend Deprivation Index	0.01	[0.99-1.02]	0.51
PC1	1.00	[0.97-1.04]	0.79
PC2	1.01	[0.98-1.04]	0.47
PC3	1.00	[0.97-1.03]	0.86
PC4	0.98	[0.96-1.00]	0.07
PC5	1.01	[1.00-1.02]	0.02
PC6	0.98	[0.95-1.01]	0.15
PC7	1.00	[0.98-1.03]	0.88
PC8	0.99	[0.97-1.02]	0.68
PC9	1.00	[0.99-1.01]	0.84
PC10	1.00	[0.97-1.02]	0.79

In this model, CHIP prevalence is predicted by diet quality category with adjustment for age, sex, hypertension, hyperlipidemia, diabetes, smoking status, Townsend deprivation index, systolic blood pressure, blood glucose at enrollment visit, total cholesterol at enrollment, triglyceride level at enrollment, and the first 10 principal components of genetic ancestry. (CHIP, clonal hematopoiesis of indeterminate potential)

eTable 5. Sensitivity Analysis Showing Odds Ratios for the Presence of CHIP Predicted by Diet Quality After Excluding Patients With Prevalent Ischemic Stroke and Heart Failure From the Study Cohort

Variable	Odds Ratio	Confidence Interval	P-value
Unhealthy Diet	1.23	[1.02-1.48]	0.03
Intermediate Diet	1.00	--	Ref
Healthy Diet	0.87	[0.72-1.05]	0.16
Age	1.07	[1.07-1.08]	<0.001
Male	0.96	[0.88-1.05]	0.38
Current Smoking	1.08	[0.99-1.19]	0.09
T2DM	1.04	[0.78-1.37]	0.77
Townsend Deprivation Index	1.00	[0.98-1.02]	0.95
PC1	1.00	[0.97-1.03]	0.84
PC2	1.01	[0.98-1.04]	0.45
PC3	1.01	[0.98-1.04]	0.64
PC4	0.98	[0.96-1.00]	0.08
PC5	1.01	[1.00-1.02]	0.07
PC6	0.98	[0.95-1.01]	0.11
PC7	1.00	[0.97-1.02]	0.90
PC8	0.99	[0.97-1.02]	0.48
PC9	1.00	[0.99-1.01]	0.58
PC10	1.00	[0.98-1.02]	0.87

This analysis additionally controls for age, sex, smoking status, type 2 diabetes mellitus, Townsend deprivation index and the first 10 principal components of genetic ancestry. (CHIP, Clonal Hematopoiesis of Indeterminate Potential)

eTable 6. Odds Ratios for Presence of CHIP Predicted by a Continuous Dietary Score and Adjusted for Key Covariates

Variable	**Odds Ratio**	**Confidence Interval**	**P-value**
Diet Score	0.96	[0.92-0.99]	0.02
Age	1.07	[1.07-1.08]	<0.001
Male	0.96	[0.87-1.05]	0.34
Current Smoking	1.01	[0.99-1.18]	0.1
T2DM	1.00	[0.76-1.32]	0.95
Townsend Deprivation Index	1.08	[0.98-1.02]	0.98
PC1	1.00	[0.97-1.03]	0.86
PC2	1.01	[0.98-1.04]	0.45
PC3	1.01	[0.97-1.04]	0.73
PC4	0.98	[0.96-1.00]	0.09
PC5	1.01	[1.00-1.02]	0.07
PC6	0.98	[0.95-1.01]	0.14
PC7	1.00	[0.97-1.02]	0.86
PC8	0.99	[0.97-1.02]	0.46
PC9	1.00	[0.99-1.01]	0.53
PC10	1.00	[0.98-1.02]	0.93

Here a dietary score ranging in values from 3 to 9 was modeled as a continuous variable to predict presence of CHIP while adjusting for age, sex, current smoking, type 2 diabetes mellitus, Townsend deprivation index, and the first 10 principal components of genetic ancestry. For every 1 point increase in diet score, the odds of CHIP being present decreased by 4.5%. (CHIP, Clonal Hematopoiesis of Indeterminate Potential)

eTable 7. Frequency of CHIP Driver Gene Variations by Diet Quality Category

Gene	CHIP Count	CHIP Prevalence
Unhealthy	**162**	**7.1%**
DNMT3A	95	58.6%
TET2	26	16.0%
ASXL1	10	6.2%
SRSF2	5	3.1%
JAK2	3	1.9%
Intermediate	**2177**	**5.7%**
DNMT3A	1392	63.9%
TET2	339	15.6%
ASXL1	131	6.0%
PPM1D	37	1.7%
TP53	36	1.7%
Healthy	**168**	**5.1%**
DNMT3A	113	67.3%
TET2	26	15.5%
ASXL1	15	8.9%
ASXL2	2	1.2%
CBL	2	1.2%

The prevalence of CHIP was highest in the unhealthy diet category, but distribution of the most frequently mutated genes was similar among all diet classes. (CHIP, clonal hematopoiesis of indeterminate potential)

eTable 8. Incidence of Adverse Events by Diet Quality Category

Diet Quality	Incident CAD + Death	Stratum Subtotal Count	Cumulative Incidence of Events	P-value
Unhealthy	276	2271	12.15%	8.8×10^{-4}
Intermediate	3382	38552	8.77%	
Healthy	197	3288	5.99%	

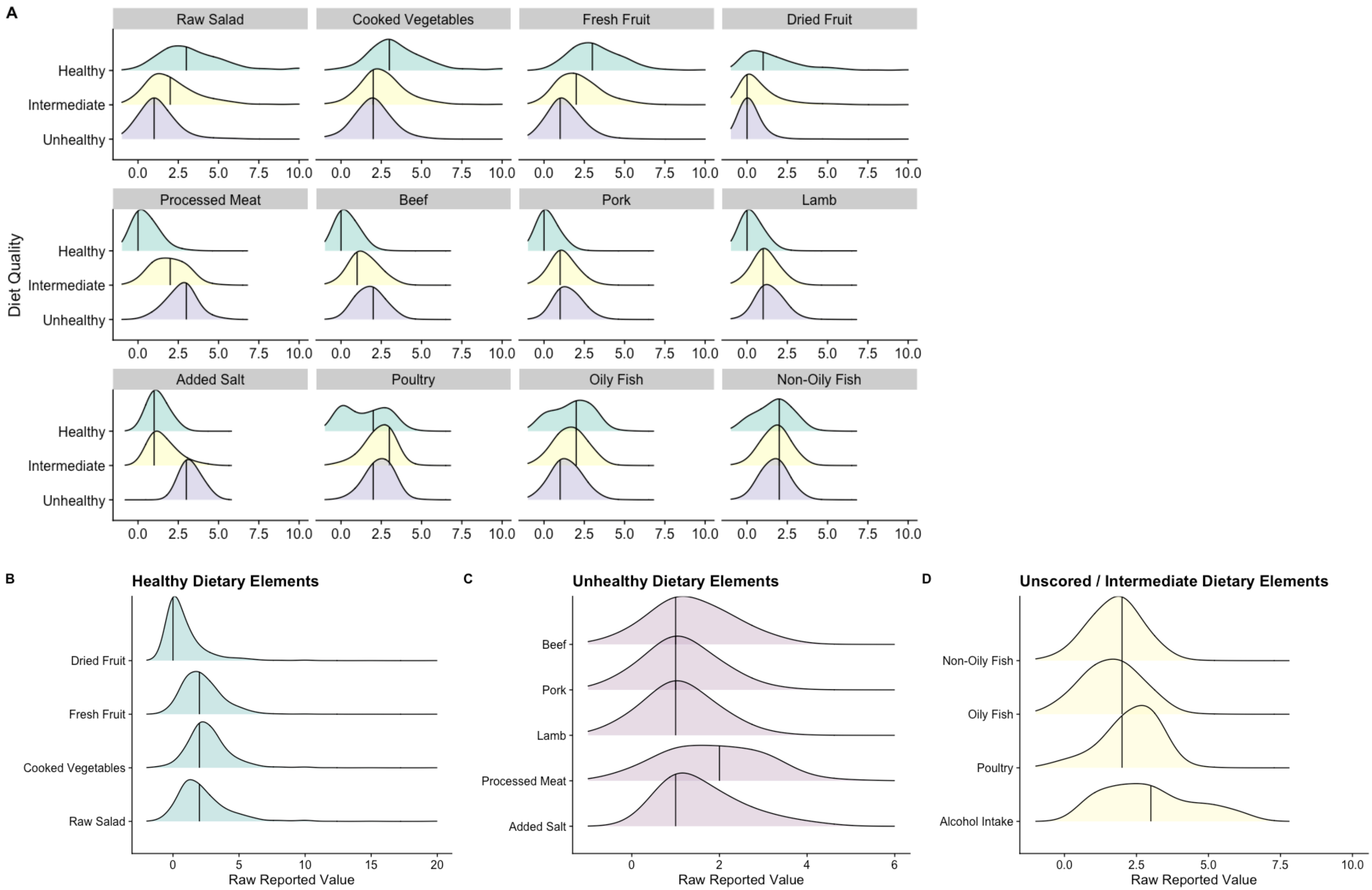
Incidence of adverse events was greatest in the unhealthy diet category and decreased in intermediate and healthy diet categories. (CAD, coronary artery disease)

eTable 9. Sensitivity Analysis of Incidence of Adverse Events by CHIP Presence and Diet Quality Category After Exclusion of Individuals With Cardiac MRI Data Available

Variable	Hazard Ratio	Confidence Interval	P-value
CHIP Present	1.21	[1.08-1.37]	0.001
Unhealthy Diet	1.26	[1.11-1.44]	<0.001
Intermediate Diet	1	--	Ref
Healthy Diet	0.78	[0.67-0.91]	0.002
Age	1.08	[1.07-1.08]	<0.001
Current Smoking	1.05	[1.0-1.14]	0.16
T2DM	2.19	[1.82-2.39]	<0.001
Male	1.78	[0.54-0.62]	<0.001
PC1	1.00	[0.99-1.00]	0.25
PC2	1.00	[0.99-1.00]	0.54
PC3	1.00	[1.00-1.00]	0.05
PC4	1.00	[1.00-1.00]	0.03
PC5	1.00	[0.99-1.00]	0.31
PC6	1.00	[0.99-1.00]	0.65
PC7	0.99	[0.99-1.00]	0.14
PC8	1.01	[1.00-1.01]	0.002
PC9	0.99	[0.99-1.00]	0.15
PC10	1.01	[0.99-1.02]	0.07

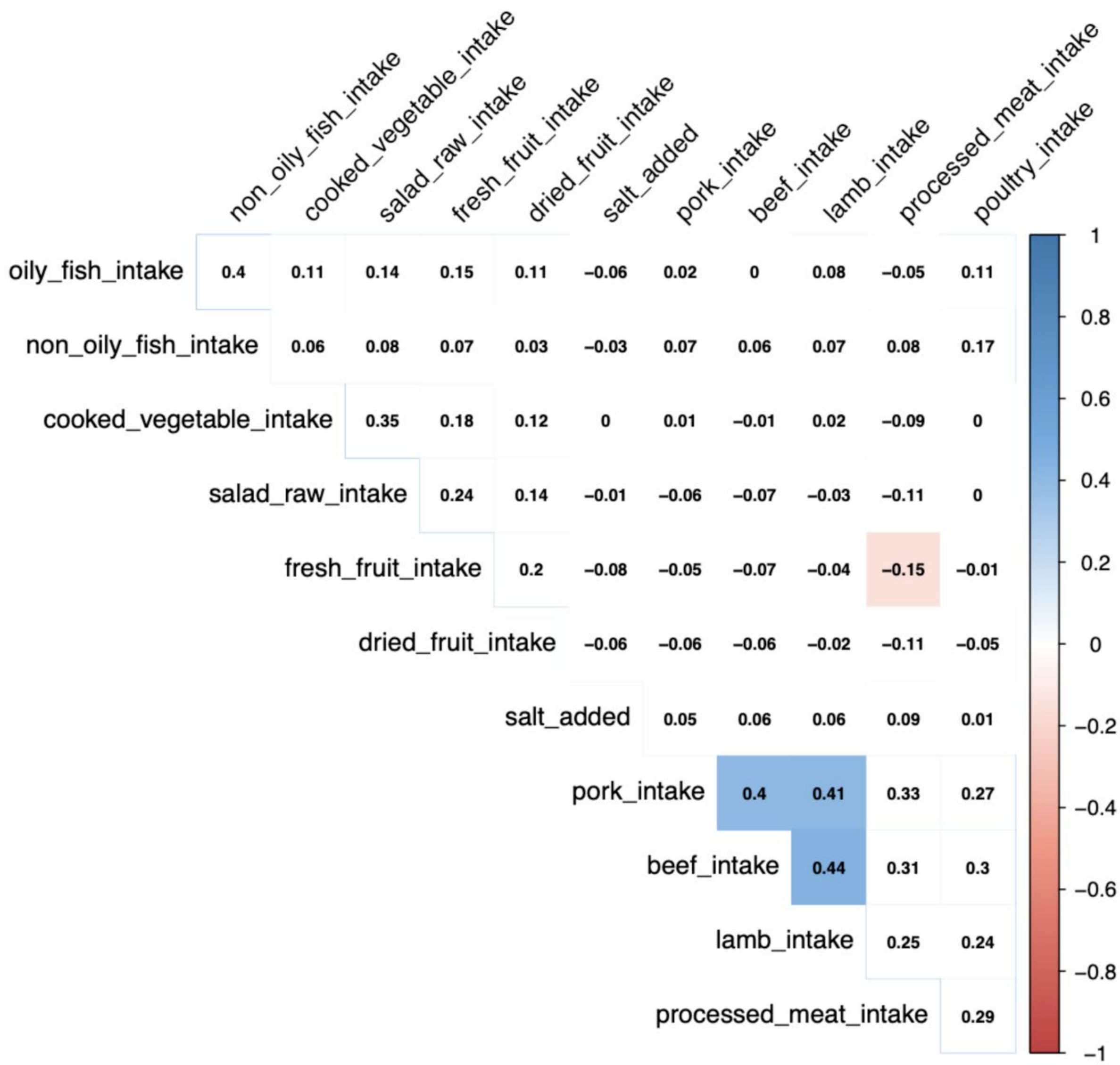
Because some individuals in the UK Biobank were selected for WES based on availability of cardiac MRI imaging data, previous analyses have demonstrated immortal time bias associated with these individuals in survival analyses. To assess for the presence of confounding due to immortal time bias, sensitivity analysis was conducted excluding those individuals who were selected due to cardiac MRI data presence. CHIP and diet quality category both independently associate with adverse cardiovascular events and death after exclusion of these individuals. (CHIP, clonal hematopoiesis of indeterminate potential; MRI, magnetic resonance imaging; WES, whole exome sequencing)

eFigure 1. Distribution of Dietary Elements by Diet Quality Category and Overall



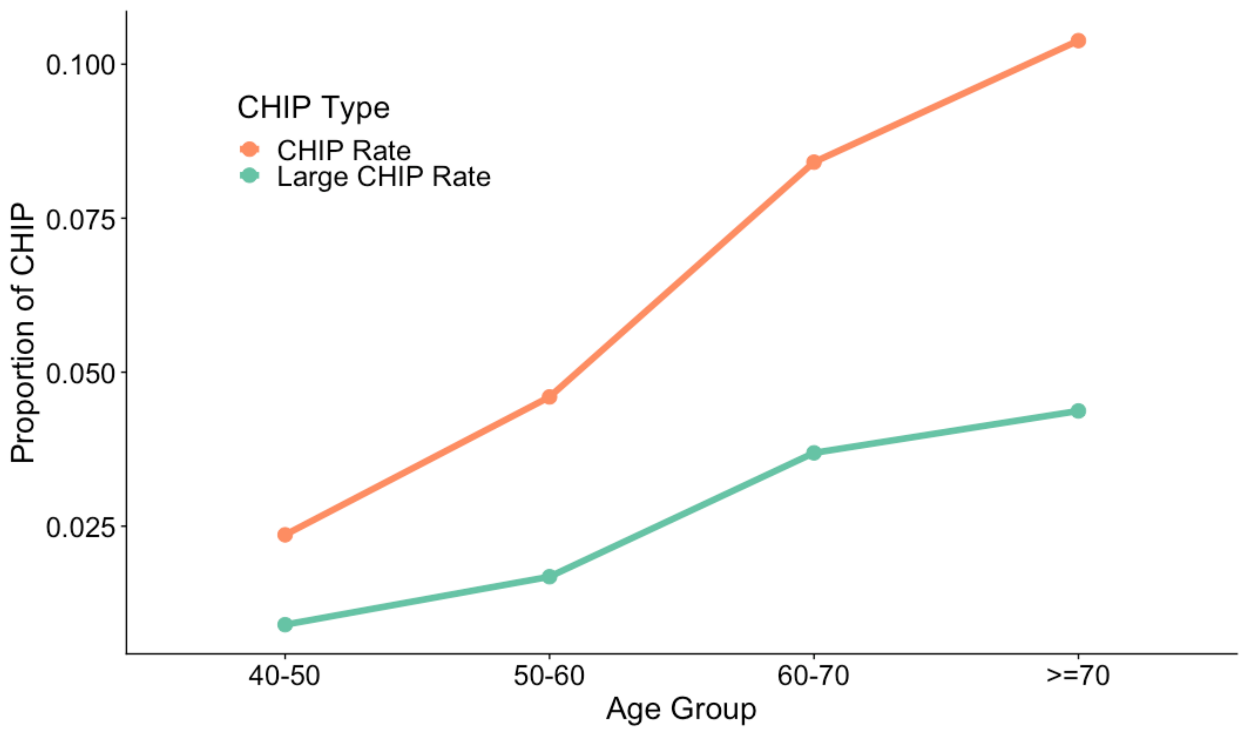
(A) Healthy dietary elements had higher median scores among those with a healthy diet, and long right tails (x axis curtailed here for ease of viewing), while those with unhealthy diet had lower values for healthy dietary elements and higher values for unhealthy dietary elements. (B) In the overall study population healthy dietary elements had generally low values, but long right tails. (C) In the overall study population, unhealthy dietary elements were homogenous with a median value of eating red meat once weekly. (D) Dietary elements that were left unscored included non-oily fish, oily fish, poultry and alcohol intake.

eFigure 2. Correlation Matrix of Dietary Element Intake Values



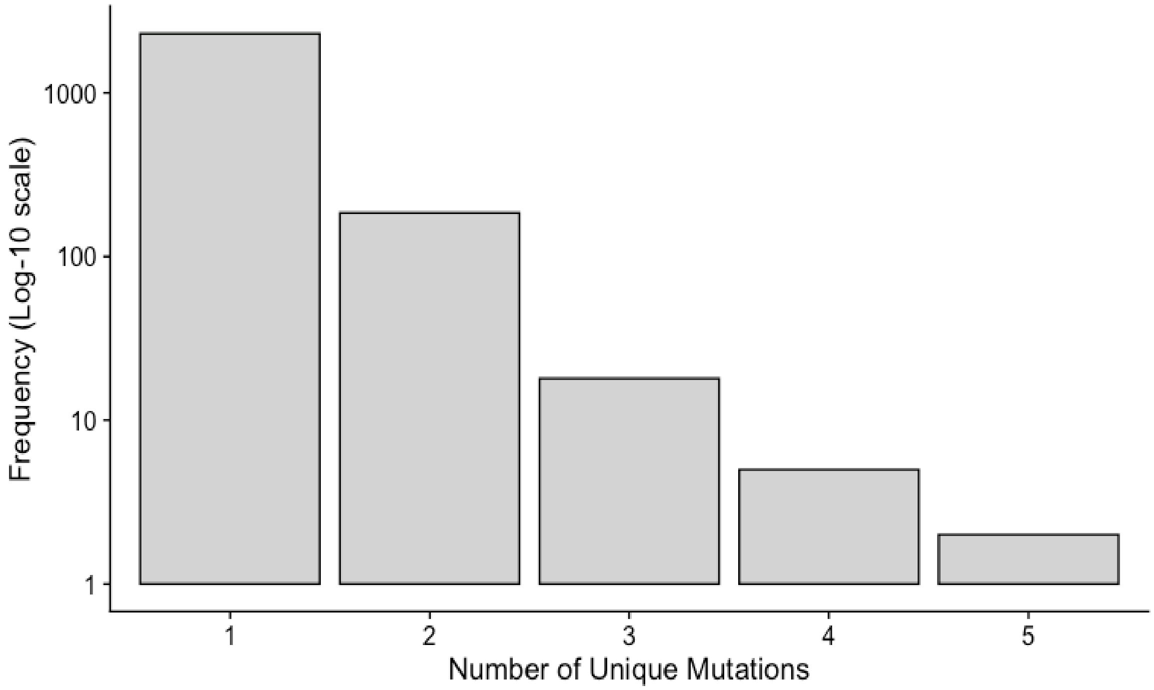
Colored squares represent correlation coefficients below the pre-specified significance level of p<0.05. Pork beef and lamb intake were correlated. Fresh fruit intake and processed meat intake were inversely correlated. The greatest correlation existed between beef and lamb intake (r=0.44). The greatest negative correlation existed between fresh fruit intake and processed meat intake (r=-0.15).

eFigure 3. CHIP Prevalence by Age Group



Both CHIP and large CHIP clones with a VAF ≥ 10% increase with age, such that CHIP prevalence is >0.10 in the group ≥ 70 years of age. (CHIP, clonal hematopoiesis of indeterminate potential; VAF, variant allele fraction)

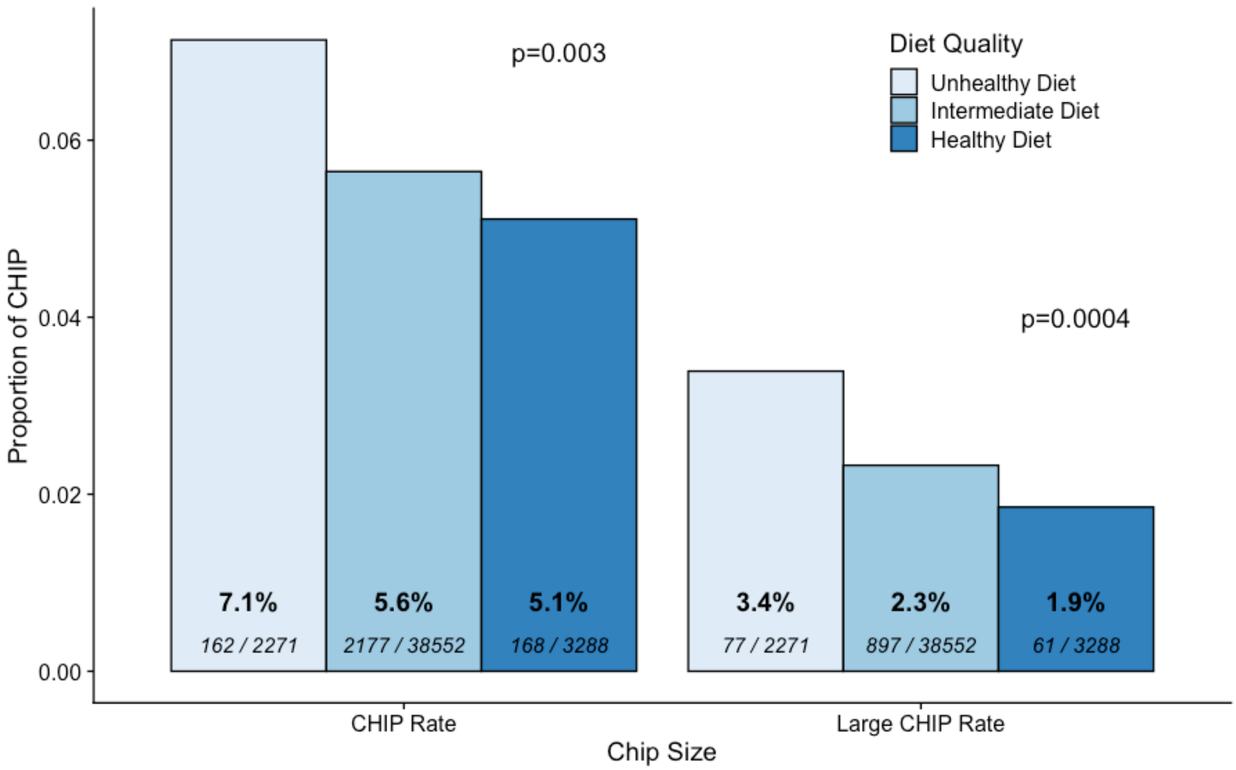
eFigure 4. Frequency of Individuals Harboring 1 or More CHIP Variation



# Mutations	1	2	3	4	5
Frequency	2300	185	18	5	2

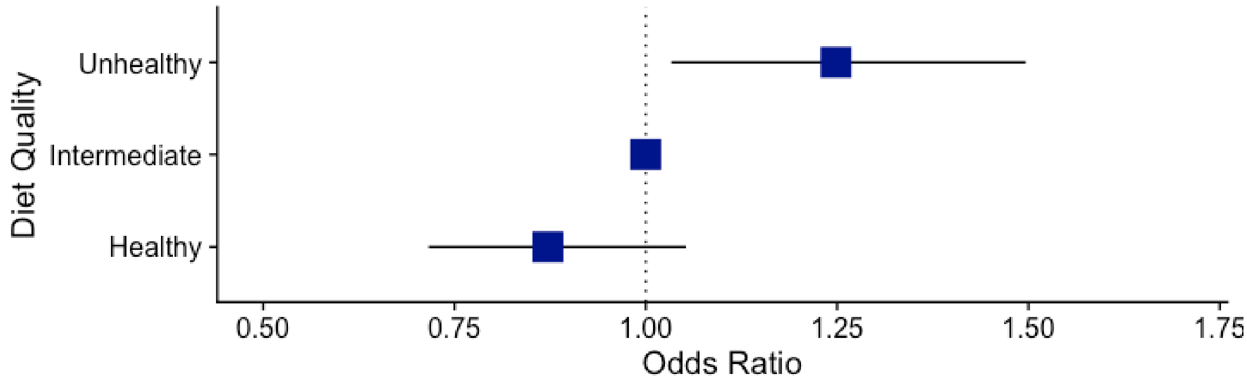
Among individuals with CHIP, the majority of individuals had only a single driver mutation, while a small number had multiple mutations. (CHIP, clonal hematopoiesis of indeterminate potential)

eFigure 5. Prevalence of CHIP (VAF >0.02) and Large CHIP Clones (VAF >0.10) With Improved Diet Quality



Unadjusted analyses show a significant test for trend. (CHIP, clonal hematopoiesis of indeterminate potential; VAF, variant allele fraction)

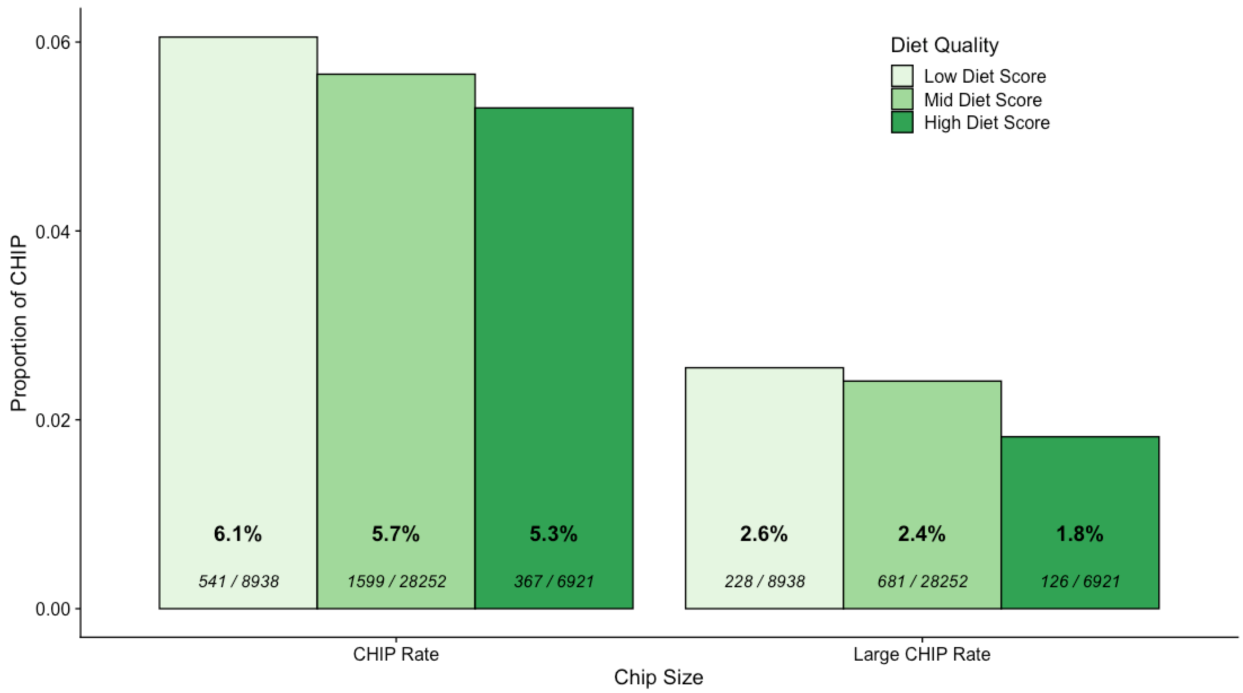
eFigure 6. Odds Ratios for CHIP Presence Across Diet Quality Categories



Diet Quality	CHIP	OR	95% CI	P-value
Unhealthy	162	1.25	1.03 - 1.50	0.019
Intermediate	2177	1	Ref.	Ref
Healthy	168	0.87	0.72 - 1.05	0.163

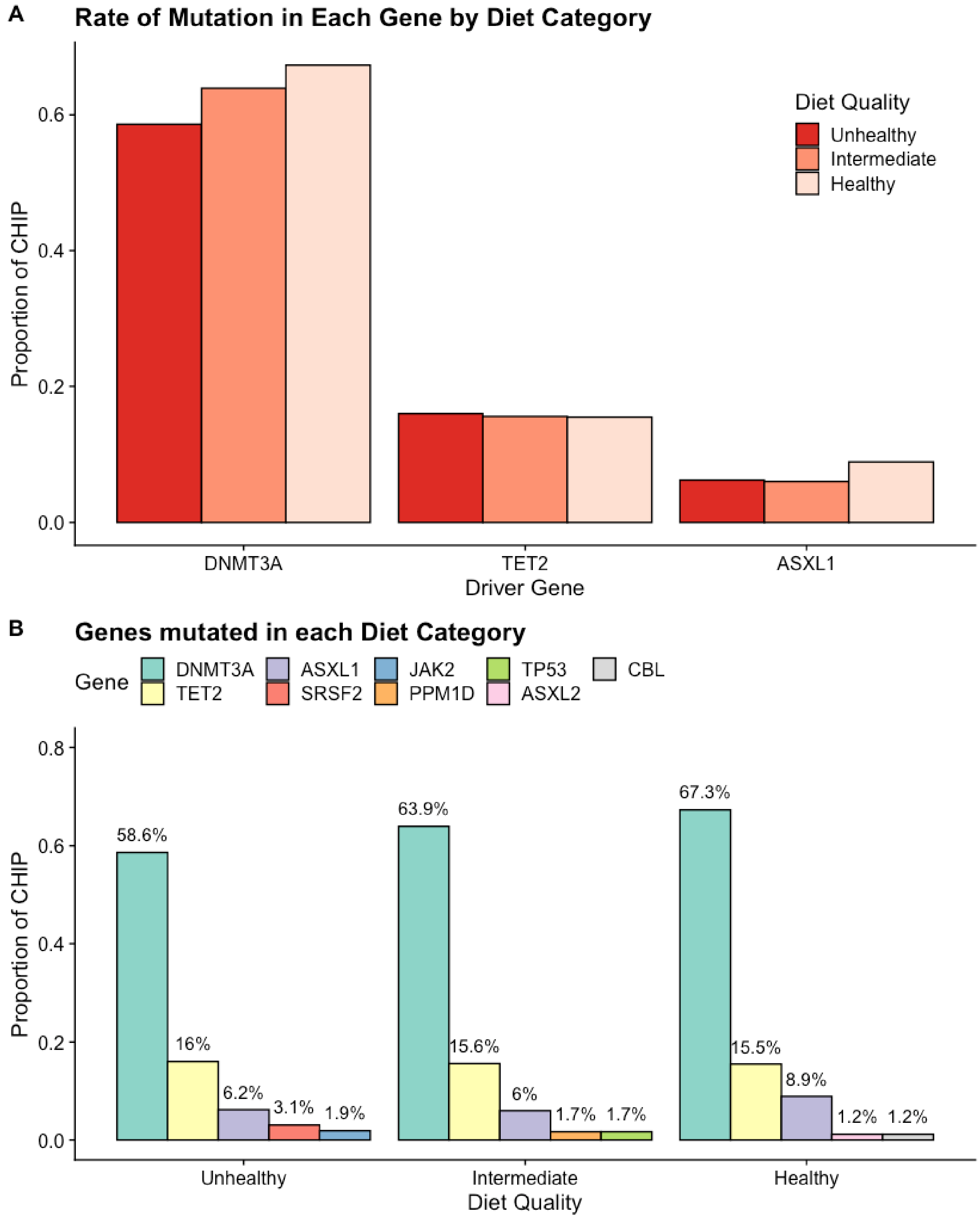
Forest plot of the odds ratio of CHIP presence by diet quality category when controlling for age, sex, smoking status, diabetes mellitus status, and the first ten principal components of genetic ancestry. When compared to a reference of an intermediate diet quality, an unhealthy diet was associated with a 25% increased odds of CHIP presence (P<0.02). (CHIP, clonal hematopoiesis of indeterminate potential)

eFigure 7. CHIP Prevalence by Diet Quality Treated as a Continuous Variable and Trichotomized



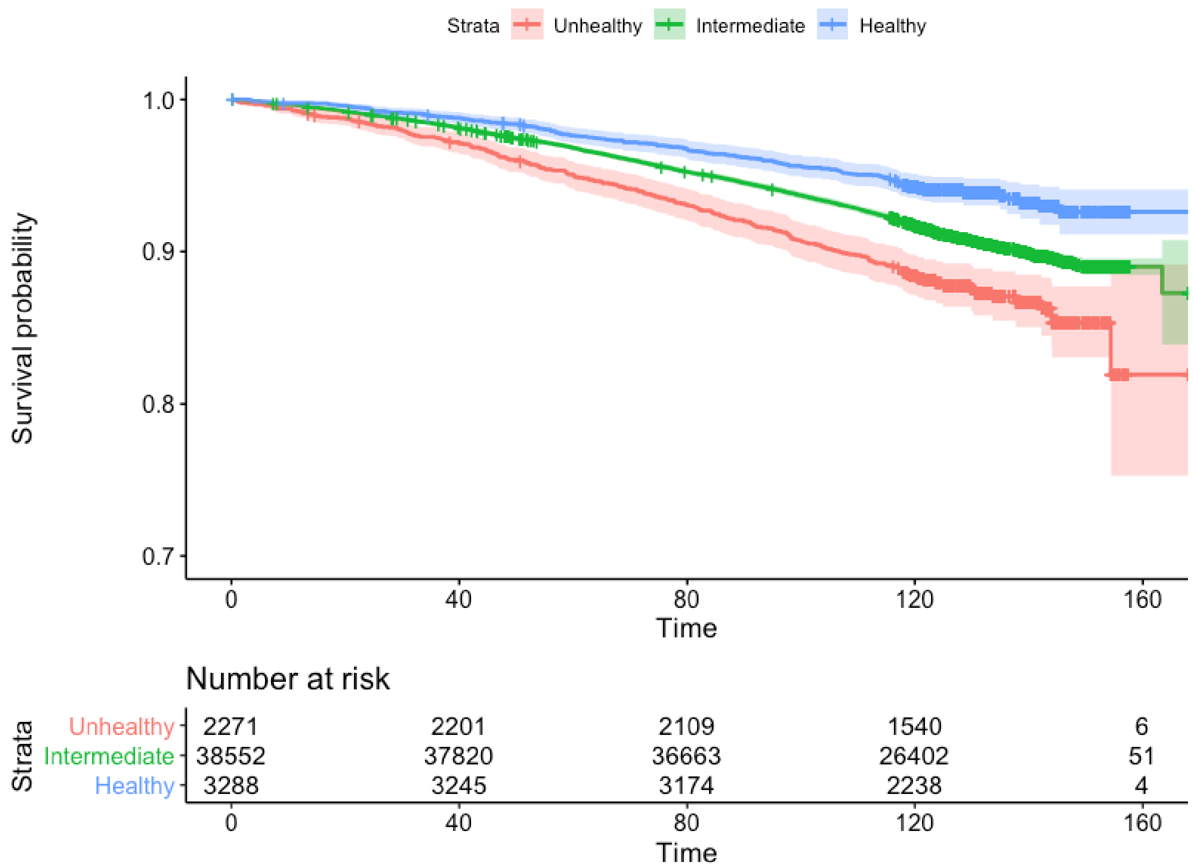
Diet quality was graded as a score of 1 to 3 in three different categories: (1) fruits and vegetables, (2) red and processed meats, (3) added salt. The top quintile was considered a “High Diet Score”. The bottom quintile was considered a “Low Diet Score” and the middle was considered a “Mid Diet Score”. CHIP prevalence decreased across diet quality classes. (CHIP, clonal hematopoiesis of indeterminate potential).

eFigure 8. CHIP Driver Variations in Diet Quality Categories



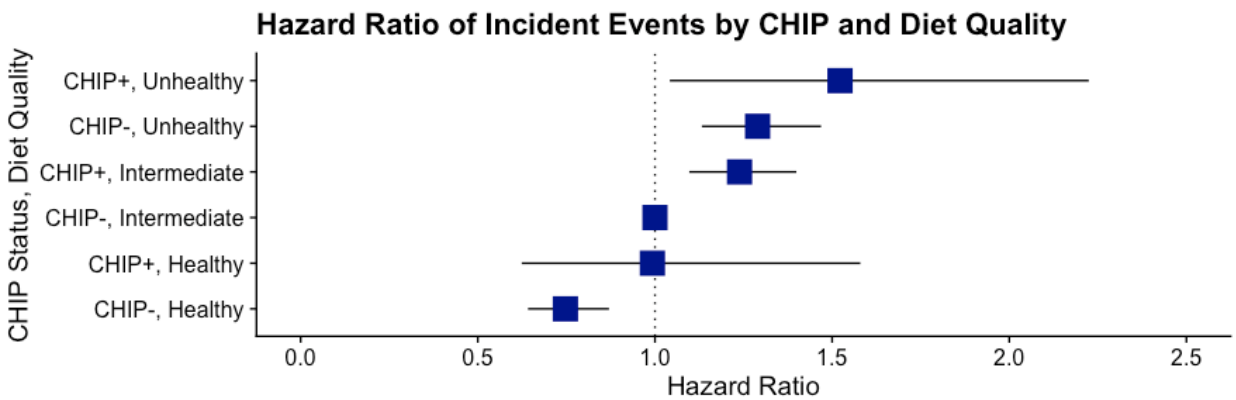
DNMT3A, TET2, and *ASXL1* were the top driver mutations in all diet classes. *DNMT3A* represented a numerically higher but non-significantly different proportion of CHIP cases as diet quality improved. Mutations in the remainder of driver genes were much less common. (CHIP, clonal hematopoiesis of indeterminate potential)

eFigure 9. Diet Quality Categories and Risk of Adverse Events With Simple Survival Analysis



Unadjusted survival analysis shows unhealthy diet is associated with greatest risk of adverse cardiovascular events or death, and a healthy diet is associated with the lowest risk of events.

eFigure 10. Forest Plot of Hazard Ratio of Incident Cardiovascular Events and Death Predicted by CHIP Status and Diet Quality Category



Adverse cardiovascular events or death was predicted using a multivariate Cox proportional hazard model adjusting for CHIP presence, diet quality, age, sex, smoking status, diabetes mellitus status, and the first ten principal components of genetic ancestry. (CHIP, clonal hematopoiesis of indeterminate potential)