Age and Ageing 2024; **53:** ii20–ii29
https://doi.org/10.1093/ageing/afae038

RESEARCH PAPER

Identifying the relation between food groups and biological ageing: a data-driven approach

Ynte Biemans[1], Daimy Bach[1], Pariya Behrouzi[2], Steve Horvath[3,4,5], Charlotte S. Kramer[1], Simin Liu[6,7], JoAnn E. Manson[8], Aladdin H. Shadyab[9], James Stewart[10], Eric A. Whitsel[10,11], Bo Yang[7], Lisette de Groot[1], Pol Grootswagers[1]

[1]Division of Human Nutrition and Health, Wageningen University and Research, Wageningen, The Netherlands
[2]Biometrics, Mathematical and Statistical Methods, Wageningen University and Research, Wageningen, The Netherlands
[3]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA
[4]Altos Labs, San Diego Institute of Science, San Diego, CA, USA
[5]Department of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, CA, USA
[6]Departments of Medicine and Surgery, Alpert School of Medicine, Brown University, Providence, RI, USA
[7]Department of Epidemiology and Center for Global Cardiometabolic Health, School of Public Health, Brown University, Providence, RI, USA
[8]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[9]Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, La Jolla, CA, USA
[10]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[11]Department of Medicine, School of Medicine, University of North Carolina, Chapel Hill, NC, USA

Address correspondence to: Ynte Biemans. Email: Yntebiemans@gmail.com

Abstract

Background: Heterogeneity in ageing rates drives the need for research into lifestyle secrets of successful agers. Biological age, predicted by epigenetic clocks, has been shown to be a more reliable measure of ageing than chronological age. Dietary habits are known to affect the ageing process. However, much remains to be learnt about specific dietary habits that may directly affect the biological process of ageing.
Objective: To identify food groups that are directly related to biological ageing, using Copula Graphical Models.
Methods: We performed a preregistered analysis of 3,990 postmenopausal women from the Women's Health Initiative, based in North America. Biological age acceleration was calculated by the epigenetic clock PhenoAge using whole-blood DNA methylation. Copula Graphical Modelling, a powerful data-driven exploratory tool, was used to examine relations between food groups and biological ageing whilst adjusting for an extensive amount of confounders. Two food group–age acceleration networks were established: one based on the MyPyramid food grouping system and another based on item-level food group data.
Results: Intake of eggs, organ meat, sausages, cheese, legumes, starchy vegetables, added sugar and lunch meat was associated with biological age acceleration, whereas intake of peaches/nectarines/plums, poultry, nuts, discretionary oil and solid fat was associated with decelerated ageing.
Conclusion: We identified several associations between specific food groups and biological ageing. These findings pave the way for subsequent studies to ascertain causality and magnitude of these relationships, thereby improving the understanding of biological mechanisms underlying the interplay between food groups and biological ageing.

Keywords: older people, biological ageing, epigenetic clocks, DNA methylation, food groups, Copula Graphical Models

Key Points

- Biological age is considered a more all-encompassing measure of ageing than chronological age, as it covers not only time, but also epigenetic, lifestyle and environmental factors.

- Data-driven approaches are suitable to indicate new interventional targets within the complex domain of diet and ageing.
- This study revealed multiple compelling interactions between food groups and biological age acceleration.
- Results provide a promising base for follow-up studies to determine causality and magnitude of relationships, thereby improving the understanding of biological mechanisms underlying the interplay between food groups and biological ageing.

Introduction

Although the ageing process seems relatively homogeneous at a population level, there are large differences in the pace of ageing and consequential disease susceptibility amongst individuals [1]. This heterogeneity makes biological age, which is the result of progressive structural and functional decline of tissues and organs because of damage accumulation over time, more appropriate to represent the ageing process than chronological age [2]. In order to find strategies to extend lifespan and health span in the progressively ageing population, an understanding of biological ageing and its accurate measures is required.

Biological age can be measured by various molecular or phenotypic biomarkers. Recent studies point towards the significance of epigenetic clocks, representing one's biological age, and their ability to accurately quantify human ageing [3]. Epigenetic clocks are powerful predictors of age or age-related phenotypes, based on DNA methylation (DNAm) values coupled with mathematical algorithms [4]. The majority of DNAm occurs on sites where a cytosine nucleotide is followed by guanine (CpG sites) [4]. Alterations in DNAm at specific sites across the genome are associated with age-related diseases and decreased functioning of cells, tissues and organs [5, 6]. Epigenetic clocks primarily target those 'predictable' CpG sites, where age-associated DNAm is relatively homogeneous between individuals [7]. Epigenetic clocks are trained via machine learning to predict an age-related outcome of interest, such as chronological age, morbidity or mortality risk [4]. In this process, the most predictive CpG sites are selected and then combined with a mathematical algorithm, thereby transforming the measured DNAm levels into an age prediction.

With the help of epigenetic clocks, the degree of age acceleration or deceleration can be estimated (i.e. the difference between biological and chronological age). Previous studies showed a slower age acceleration in long-lived individuals [8, 9], indicating a relationship between age acceleration and healthy ageing. Indeed, age acceleration is associated with decreased physical and cognitive functioning, and increased risk of cancer, cardiovascular diseases (CVD) and all-cause mortality [10, 11]. The extent of age acceleration or deceleration can differ tremendously amongst individuals, which raises the question of what the driving factors are behind this heterogeneity. According to recent pedigree analyses, true heritability of longevity is estimated at 10–16% [12, 13]. These relatively small numbers indicate that other factors, such as lifestyle, environment and air pollution, rather than genetics, play a dominant role in biological ageing. Notably, whilst the proportion of variation because of genetics seems relatively small at the population level, genetic disorders (e.g. progeroid syndromes) can have significant effects at the individual level, which makes it, independent of adherence to a healthy lifestyle, impossible to grow old [14].

Although the precise contribution of nutrition to the ageing process remains unclear, increasing research has shown the capacity of specific lifestyle interventions, such as the Mediterranean diet [15], caloric restriction [16] or other multimodal treatment programmes [17, 18], to slow down or even reverse ageing clocks. However, associations have been weak and inconsistent between studies. The use of small sample sizes, different epigenetic clocks and the inability to establish independent effects of lifestyle covariates because of collinearity has led to inconclusive results. Moreover, as nutrient databases often provide limited information on bioactive compounds, food matrices and potential other interactions amongst substances within foods, it is valuable to explore whether specific food groups relate to biological ageing. Therefore, the first step in the process of identifying lifestyle secrets of successful agers is to find associations between food groups and biological age in large databases. Instead of focussing on single nutrients in a hypothesis-driven way, as was done in previous studies, a data-driven approach can be more suitable to explore the interactive network of foods. Accordingly, in this study a large data set containing lifestyle and methylation data will be analysed using Copula Graphical Models (CGMs). This statistical method, with an emphasis on exploration and identification, as opposed to methods focused on magnitude and causality, allows for a comprehensive understanding of biological ageing. As CGMs enable independent analysis of covariates, the statistical tool is highly relevant for this type of complex and multivariate nutritional data [19]. The exploratory method can be applied to visualise conditional dependence relations between variables in a graphical network. To summarize, a data-driven approach in a suitable data set is needed to identify specific dietary habits in light of ageing in an objective and open-minded fashion. Therefore, the aim of this study is to discover food groups that are directly related to biological ageing, using CGMs.

Methods

A summary of the methods is provided below; further details of the study design, data collection and statistical analysis are available in Appendix 1 (Supplementary Methods).

Study design

The present study is a preregistered (Open Science Framework; osf.io/r24pj) secondary analysis of the Women's Health Initiative (WHI), a large cohort study of postmenopausal women based in North America. Baseline data from two WHI subsamples with available DNAm data (AS315 and BA23) were used in the present study for exploratory analysis [20–22]. The cross-sectional nutritional, lifestyle and socio-behavioural data were combined to construct exploratory models of variables associated with biological ageing.

The second-generation epigenetic clock PhenoAge is chosen predictor of biological age in this study, as it incorporates age-related and disease phenotypes in combination with chronological age, thereby being able to estimate health span outcomes and mortality (in contrast to the first-generation clocks HorvathAge and HannumAge) [3, 17]. DNAm was measured from whole-blood using the Illumina Infinium HumanMethylation450 BeadChip [21, 22], and used to calculate PhenoAge acceleration (DNAm PhenoAgeAccel) via an online software (https://dnamage.genetics.ucla.edu/new). This tool estimates DNAm PhenoAgeAccel adjusted for chronological age, by calculating the relative difference between chronological age and biological age estimated by PhenoAge. Dietary intake was assessed at baseline using the WHI Food Frequency Questionnaire (FFQ). For our analysis, two data sets containing FFQ data were used: one data set containing item-level data and the other one containing MyPyramid Equivalents Database 2.0 (MPED) food groups measures.

Statistical analysis

Descriptive statistics were stratified into quartiles of age acceleration and presented as means ± SD for continuous variables with normal distribution, as medians (25th–75th percentile) for skewed continuous variable, or as numbers (percentages) for categorical variables. Quartiles are used for reporting subject characteristics, to show the potential differences in characteristics for subjects with a decelerated or accelerated ageing. Descriptive analyses were performed using R version 4.1.3.

To examine the relations between the selected nutritional, lifestyle, sociodemographic variables and epigenetic age acceleration, CGMs were used. CGMs are probability structures that infer conditional dependence relationships between variables in a graph, where nodes (i.e. variables) are connected by edges [23]. By adjusting every association for all other variables in the data set, CGM deals with an extensive number of confounders in an objective way. This adjustment is especially important for examining the potential effects of lifestyle/demographic covariates (especially age, smoking status, BMI, waist-to-hip ratio (WHR), blood pressure and physical activity) on DNAm, as it deals with the potential collinearity of the covariates. The R package 'nutriNetwork' (version 0.1.2) was used to reconstruct the food group–age acceleration networks, and *selectnet* was used to select the most optimal (sparse) network based on extended Bayesian information criteria. Bootstrapping was performed to determine the uncertainty of each found association, and a sensitivity analysis was performed amongst BMI, physical activity level and data set-specific strata.

Results

Subject characteristics

The analytic sample consisted of 3,990 postmenopausal women from diverse ethnic backgrounds. Mean chronological age was 63.3 ± 7.1 years. Characteristics of the study population were stratified into quartiles of DNAm PhenoAgeAccel, as presented in Table 1. Whilst characteristics were comparable across quartiles, subjects with higher DNAm PhenoAgeAccel had an increased BMI, WHR, elevated blood glucose and insulin levels ($P < 0.001$; ANOVA), and a lower level of education ($P < 0.001$; chi-squared test). Average total energy intake was below the estimated calorie needs range from 1,600 to 2,400 kcal/day for adult women [24]. Notably, most subjects were overweight (median BMI 28.9 kg/m^2 (IQR: 25.3–33.1 kg/m^2)), had a relatively high systolic blood pressure (SBP) (130.3 ± 18.0 mmHg) and high prevalence of arthritis (49%). Most subjects had a low level of physical activity, with a median moderate to strenuous activity of 10 min per week.

Findings

The focus of this study was on partial correlations in the main CGM networks with a certainty of ≥95%. Table 2 gives an overview of the associations between DNAm PhenoAgeAccel and food groups based on the MyPyramid food group measure, lifestyle and sociodemographic variables and blood biomarkers**.** Positive direct associations (detrimental; red triangles) indicate that greater values of the variable (e.g. higher intake of the food group) were related to accelerated ageing, whereas negative direct associations (beneficial; green triangles) were related to decelerated ageing (i.e. PhenoAge deceleration). The partial correlation coefficients of the estimated CGM (rho as selected via computational procedure (Supplementary Methods) = 0.001) are presented in Supplementary Figure S2. Decelerated ageing was associated with poultry (96% certainty), nuts (96% certainty), discretionary fat (solid and oil, with both 99% certainty), diastolic blood pressure (DBP) (99% certainty), education (97% certainty) and osteoporosis (97% certainty) (Supplementary Table S3). Accelerated ageing was associated with eggs (96% certainty), organ meat (97% certainty), sausages (95% certainty), cheese (96% certainty), legumes (97% certainty), starchy vegetables (100% certainty), added sugar (97% certainty), SBP(97% certainty), BMI (100% certainty), WHR (100% certainty), being American-Indian or Alaskan Native (98% certainty), being Asian or Pacific Islander (96% certainty), having a race other than American-Indian/Alaskan

Table 1. Baseline characteristics and lifestyle variables, stratified by quartiles of PhenoAgeAccel. Values are mean (SD), median (IQR) or *n* (%)

	Total	Decelerated ageing		Accelerated ageing	
		Q1 (−33.0 to −4.4 years)	Q2 (−4.4 to −0.3 years)	Q3 (−0.3 to 3.9 years)	Q4 (3.9–34.4 years)
N	**3,990**	**998**	**997**	**997**	**997**
Age (years)	63 (7)	63 (7)	63 (7)	63 (7)	63 (7)
PhenoAgeAccel (years)	−0.2 (6.4)	−8.1 (3.4)	−2.3 (1.12)	1.6 (1.2)	8.0 (3.7)
Race/ethnicity (%)					
American Indian or Alaskan Native	49 (1.2)	10 (1.0)	6 (0.6)	19 (1.9)	14 (1.4)
Asian or Pacific Islander	128 (3.2)	24 (2.4)	45 (4.5)	38 (3.8)	21 (2.1)
Black or African American	1,119 (28.0)	288 (28.9)	258 (25.9)	279 (28.0)	294 (29.5)
Hispanic/Latino	666 (16.7)	164 (16.4)	165 (16.5)	155 (15.5)	182 (18.2)
White (not Hispanic)	1,995 (50.0)	506 (50.7)	516 (51.8)	494 (49.5)	479 (48.0)
Other	33 (0.8)	6 (0.6)	7 (0.7)	12 (1.2)	8 (0.8)
Education (%)					
Did not go to school	5 (0.1)	2 (0.2)	1 (0.1)	1 (0.1)	1 (0.1)
Grade school (1–4 years)	30 (0.8)	4 (0.4)	5 (0.5)	10 (1.0)	11 (1.1)
Grade school (5–8 years)	124 (3.1)	36 (3.6)	29 (2.9)	24 (2.4)	35 (3.5)
Some high school (9–11 years)	260 (6.6)	65 (6.6)	63 (6.4)	61 (6.2)	71 (7.2)
High school diploma or GED[a]	735 (18.6)	166 (16.8)	189 (19.1)	184 (18.6)	196 (19.8)
Vocational or training school	505 (26.6)	133 (13.5)	119 (12.0)	136 (13.8)	117 (11.8)
Some college or associate degree	1,052 (26.6)	245 (24.8)	266 (26.9)	266 (26.9)	275 (27.8)
College graduate or baccalaureate	347 (8.8)	91 (9.2)	89 (9.0)	91 (9.2)	76 (7.7)
Some post-graduate or professional	339 (8.6)	91 (9.2)	85 (8.6)	86 (8.7)	77 (7.8)
Master's degree	481 (12.2)	136 (13.8)	123 (12.4)	113 (11.4)	109 (11.0)
Doctoral degree	77 (1.9)	18 (1.8)	21 (2.1)	16 (1.6)	22 (2.2)
Smoking pack-years	0.0 (0.0–12.5)	0.0 (0.0–7.5)	0.0 (0.0–12.5)	0.0 (0.0–12.5)	0.0 (0.0–12.5)
Alcohol intake (servings/week)	0.0 (0.0–1.1)	0.0 (0.0–1.1)	0.0 (0.0–1.4)	0.0 (0.0–1.0)	0.0 (0.0–1.1)
Physical activity[b]	10.0 (0.0–100.0)	10.0 (0.0–105.0)	10.0 (0.0–100.0)	10.0 (0.0–100.0)	10.0 (0.0–90.0)
BMI (kg/m²)	29.6 (6.1)	28.8 (5.5)	29.4 (6.1)	29.9 (6.2)	30.6 (6.4)
Energy intake (kcal/day)	1,572 (1,180–2,028)	1,579 (1,192–1,984)	1,556 (1,170–2,083)	1,580 (1,164–1,989)	1,575 (1,177–2.071)
Total vegetable intake[c]	1.4 (0.7)	1.4 (0.7)	1.4 (0.6)	1.4 (0.6)	1.3 (0.7)
Total fruit intake[c]	1.4 (0.7)	1.5 (1.0)	1.4 (1.0)	1.4 (0.9)	1.4 (0.9)
Total dairy intakec	1.6 (1.2)	1.5 (1.1)	1.6 (1.3)	1.6 (1.3)	1.6 (1.3)
Total meat intake[d]	3.9 (2.4)	3.8 (2.2)	3.9 (2.6)	3.9 (2.4)	4.0 (2.3)
Total grain intake[d]	5.2 (2.8)	5.2 (2.8)	5.2 (2.8)	5.2 (2.9)	5.1 (2.8)
Systolic blood pressure (mmHg)	130.3 (18.0)	130.0 (17.9)	130.0 (17.8)	130.0 (18.5)	131.7 (17.6)
Diastolic blood pressure **(mmHg)**	76.2 (9.2)	76.2 (9.3)	76.2 (9.2)	76.2 (9.3)	76.2 (9.0)
WHR	0.83 (0.1)	0.82 (0.1)	0.82 (0.1)	0.83 (0.1)	0.83 (0.1)
Blood biomarkers (mg/dL)[e]					
n	***2,200***	***562***	***565***	***562***	***511***
Glucose	103.1 (31.7)	99.3 (28.0)	101.7 (27.0)	104.3 (34.1)	107.3 (36.7)
Insulin	12.2 (8.8)	10.6 (5.7)	11.8 (7.1)	12.3 (9.2)	14.4 (12.0)
Total cholesterol	223.9 (39.3)	228.2 (42.1)	225.1 (40.1)	221.8 (36.0)	220.0 (38.0)
HDL cholesterol	57.8 (15.0)	59.4 (14.8)	58.7 (16.0)	57.0 (14.4)	56.0 (14.5)
LDL cholesterol	136.0 (35.7)	139.7 (36.5)	135.2 (37.1)	134.4 (33.7)	134.1 (35.4)
Triglycerides	151.2 (85.2)	144.6 (77.5)	156.5 (104.0)	153.4 (82.9)	150.1 (71.1)

[a]General Educational Development programme (alternative to traditional high school programme). [b]In minutes per week of moderate to strenuous activity. [c]In number of cup equivalents. [d]In ounce equivalents. [e]Blood biomarkers were measured in only 55% of the subjects (*n* = 2,200).

Native, Asian/Pacific Islander, Black/African-American, Hispanic/Latino or White (96% certainty), smoking pack-years (99% certainty), CVD (97% certainty) and arthritis (97% certainty).

Supplementary Figure S3 shows the estimated main CGM network (rho as selected via computational procedure (Supplementary Methods) = 0.01) including food items, DNAm PhenoAgeAccel and lifestyle and sociodemographic variables. Corresponding partial correlation matrix is presented in Supplementary Figure S4. Based on bootstrapping, negative direct associations between DNAm PhenoAgeAccel and peaches, pizza and butter were found, with 99, 97 and 96% certainty, respectively (Supplementary Table S3). Positive direct associations were found between DNAm PhenoAgeAccel and lunch meat, fat, BMI and WHR, with certainty level of 97, 100, 100 and 99%, respectively.

After stratifying for BMI, physical activity, disease status and data set found associations with DNAm PhenoAgeAccel kept recurring in most of the sub-analyses (Tables 2 and 3), indicating high sensitivity of the food group–age acceleration

Table 2. Overview of the food groups based on MPED 2.0, lifestyle and demographic variables, and blood biomarkers associated with PhenoAgeAccel, depicted for the main CGM and subgroups of the sensitivity analyses. The green/downward-pointing and red/upward-pointing triangles indicate a negative (beneficial) and positive (detrimental) association with PhenoAgeAccel, respectively. The probability of being present in all bootstrap samples is shown per association, where 1 means that the association was consistent in all bootstrap samples (i.e. a certainty of 100%).

		Main model		Sensitivity analyses							
		Direction of association	Certainty	Low BMI	High BMI	Low PA[a]	High PA[a]	Diseased[b]	Non-diseased	AS315[c]	BA23[c]
Food groups based on MPED 2.0	Eggs	▲	**0.96**	0.97	0.99	0.95	0.97	0.97	0.99	0.99	0.96
	Organ meat	▲	**0.97**	1	0.95	1	0.98	0.98	0.99	1	0.98
	Sausages	▲	**0.95**	–	–	0.96	0.99	*0.94*	0.96	–	–
	Poultry	▼	**0.96**	0.97	*0.94*	0.95	*0.93*	0.96	1	0.98	0.96
	Nuts	▼	**0.96**	0.99	0.96	0.98	0.99	0.97	–	0.98	*0.92*
	Cheese	▲	**0.96**	–	0.97	–	0.96	1	–	*0.94*	0.95
	Legumes	▲	**0.97**	0.95	0.95	0.95	0.99	–	0.96	0.96	0.96
	Starchy vegetables	▲	**0.98**	0.95	0.99	0.96	*0.93*	0.96	0.98	0.97	0.96
	Discretionary oil	▼	**0.99**	–	0.99	–	–	0.96	*0.93*	0.99	0.97
	Discretionary solid fat	▼	**0.99**	0.98	0.96	0.95	1	0.98	0.96	0.98	*0.91*
	Added sugar	▲	**0.99**	–	1	–	0.97	–	0.98	0.96	*0.94*
Lifestyle and demographics	BMI	▲	**1**	*0.94*	1	0.99	1	1	1	1	0.99
	WHR	▲	**1**	1	1	1	1	0.98	1	1	1
	Systolic blood pressure	▲	**0.99**	0.99	0.98	0.99	–	0.95	1	1	–
	DBP	▼	**0.99**	0.97	*0.94*	0.98	0.98	*0.92*	0.97	0.99	0.99
	Smoking pack-years	▲	**1**	0.95	1	0.95	0.98	1	0.95	0.97	0.99
	Education	▼	**0.97**	1	–	1	0.98	–	0.98	0.98	–
	American-Indian race	▲	**0.97**	–	0.98	–	*0.93*	0.97	–	*0.93*	–
	Asian race	▲	**0.98**	–	0.98	–	*0.94*	0.96	–	0.99	–
	Other race	▲	**0.99**	–	0.98	–	–	0.98	–	*0.94*	–
	Osteoporosis	▼	**0.97**	0.95	0.97	0.97	0.98	0.97	0	0.98	0.98
	CVD	▲	**0.96**	–	*0.94*	–	–	*0.94*	0	0.95	*0.91*
	Arthritis	▲	**0.95**	*0.94*	0.97	0.97	0.98	–	0	0.99	–
Blood biomarkers[d]	Glucose	▲	**0.99**								
	HDL cholesterol	▼	**0.95**								
	Insulin	▲	**1**								
	Total cholesterol	▼	**0.95**								
	Triglycerides	▼	**1**								

[a]PA refers to physical activity. [b]The diseased subgroup includes subjects (ever) diagnosed with diseases that might influence DNAm levels, including cancer, arthritis, osteoporosis and CVD. [c]AS315 and BA23 refer to cohort subsamples. [d]The CGM network with blood biomarkers included only 55% of the subjects (n = 2,200), therefore sample size was too small for bootstrapping. Values in bold indicate the certainties of the main model, values in italics indicate certainties <0.95, empty boxes (–) indicate the absence of a partial correlation.

networks. Especially the relationships between DNAm PhenoAgeAccel and eggs, organ meat, WHR, smoking pack-years and osteoporosis showed a certainty of ≥0.95 in all subgroups. Conversely, certainty of associations between DNAm PhenoAgeAccel and sausages, butter, pizza, American-Indian race, Asian race, other race and CVD did not meet the certainty threshold of 0.95 in the majority of the sensitivity analyses.

Discussion

This study aimed to discover food groups that are directly associated with biological ageing in a multi-ethnic cohort of postmenopausal women. In contrast to previous hypothesis-driven studies, this study used CGM as a powerful exploratory statistical tool. CGM is able to reconstruct complex associations amongst variables in multivariate data, whilst adjusting for all other variables in the data set. We found that peaches, poultry, nuts, butter, discretionary oil and discretionary solid fat were associated with decelerated ageing. Eggs, organ meat, sausages, cheese, legumes, starchy vegetables, added sugar, lunch meat and fat added after cooking were associated with accelerated epigenetic ageing.

The relationships between increased PhenoAgeAccel and intake of eggs and organ meat were consistent in all sensitivity analyses. Both eggs and organ meat are high in cholesterol, protein and fat. Cholesterol and (saturated) fat are associated with negative health outcomes and play a critical

Table 3. Overview of the food items, lifestyle and demographic variables, and blood biomarkers associated with PhenoAgeAccel, depicted for the main CGM and subgroups of the sensitivity analyses. The green/downward-pointing and red/upward-pointing triangles indicate a negative (beneficial) and positive (detrimental) association with PhenoAgeAccel, respectively. The probability of being present in all bootstrap samples is shown per association, where 1 means that the association was consistent in all bootstrap samples (i.e. a certainty of 100%).

		Main model		Sensitivity analyses							
		Direction of association	Certainty	Low BMI	High BMI	Low PA[a]	High PA[a]	Diseased[b]	Non-diseased	AS315[c]	BA23[c]
Food items	Peaches, nectarines, plums	▼	**0.99**	*0.81*	0.99	0.99	0.84	0.99	0.85	0.96	0.83
	Lunch meat	▲	**0.97**	–	1	0.98	0.84	0.98	0.86	0.98	0.74
	Pizza	▼	**0.97**	*0.78*	*0.94*	*0.68*	*0.89*	*0.82*	*0.81*	*0.89*	*0.83*
	Butter	▼	**0.96**	0.98	*0.82*	0.98	*0.83*	*0.82*	*0.91*	*0.89*	*0.89*
	Fat[d]	▲	**1**	1	*0.88*	1	*0.94*	*0.87*	0.98	0.99	0.96
Life-style	BMI	▲	**1**	*0.59*	1	*0.91*	1	1	0.96	1	*0.75*
	WHR	▲	**0.99**	1	*0.86*	0.98	0.95	*0.86*	0.97	0.99	*0.7*

[a]PA refers to physical activity. [b]The diseased subgroup includes subjects (ever) diagnosed with diseases that might influence DNAm levels, including cancer, arthritis, osteoporosis and CVD. [c]AS315 and BA23 refer to cohort subsamples. [d]The CGM network with blood biomarkers included only 55% of the subjects (n = 2,200), therefore sample size was too small for bootstrapping. [d]Fats added after cooking (incl. butter, margarine, sour cream, oils added to vegetables, beans, rice and potatoes). Values in bold indicate the certainties of the main model, values in italics indicate a certaintycertainties <0.95, empty boxes (–) indicate the absence of a partial correlation.

role in the development of non-alcoholic fatty liver disease and CVD [25–27]. Protein is shown to be detrimental for metabolic health and longevity by stimulation of the mechanistic target of rapamycin [28]. Lunch meat showed an association with increased PhenoAgeAccel, which was consistent in half of all sensitivity analyses. These lunch meats are all processed, relatively fat and salty and commonly made of pork or beef. Consumption of processed meat is associated with telomere attrition, increased risk of cancer, type 2 diabetes, CVD and mortality [29, 30]. Further research is needed to investigate which processing factors (e.g. smoking, curing, dehydration of meat, use of additives and fillers) contribute most significantly to the detrimental health effects. Potatoes and starchy vegetables (e.g. corn, sweet potatoes, yams and cassava) showed a direct positive association with biological age acceleration. Although separately grouped in our study, potatoes and starchy vegetables can be classified in the same category, according to the USDA [31]. Compared with non-starchy vegetables, starchy vegetables contain less fibre, leading to less satiety and higher caloric intake [31]. In addition, starchy vegetables have a relatively high glycaemic load, which is related to several adverse health outcomes, such as higher insulin responses [32], weight gain [32], CVD [33], type 2 diabetes [34], certain cancer types [35] and mortality [33]. Besides, starchy vegetables such as corn often have a reduced antioxidant activity, because of industrial processing by canning [36]. Worth mentioning, in our study no distinguishment was made for preparation method of potatoes and/or starchy vegetables, making it difficult to draw firm conclusions about their relation with biological ageing.

An interesting finding is the direct association between decreased PhenoAgeAccel and the food group 'peaches', 'nectarines' and 'plums'. These three stone fruits are sources of vitamin A, C, E and phenolic compounds [37], which all act as antioxidants, thereby protecting cells from oxidative damage [38]. In addition, as visualised in Supplementary Figure S3, the variable 'PEACH' act as a hub for intake of other fruits. Since peach is not the most commonly eaten fruit, it can be argued that eating peaches represents a phenotype with high fruit intake, or even with a healthy lifestyle in general. The beneficial effects of poultry and nuts on healthy ageing were in line with previous studies [39–41]. The association between decelerated ageing and solid fats, such as butter and dairy-derived fats, that comprise primarily saturated fatty acids (SFAs), was rather unexpected. However, recent evidence has questioned the dietary recommendations to minimise SFA intake, as reducing SFA intake seems to have no beneficial effects on CVD and mortality [42, 43]. Rather, a multi-ethnic study in the USA showed that blood levels of pentadecanoic acid, the SFA that is strongly associated with self-reported butter intake, was associated with reduced CVD and CHD risk [44]. Notably, fat added after cooking (e.g. butter, margarine, sour cream, oils added to vegetables, beans, rice and potatoes) was associated with increased PhenoAgeAccel. This aligns with the findings that excessive intake of high-fat foods may contribute to higher body weight and metabolic dysfunction [45]. However, discerning the specific components of fats that impact biological ageing is challenging, since our study did not differentiate between various fat types. Different fatty acid profiles should be taken into account in future

research into fats and biological ageing. Lastly, the opposite direction of the SBP and DBP associations with PhenoAge on one hand, and blood glucose and lipids on the other hand, deserves some attention. The former might be explained by the increase in pulse pressure (i.e. the difference between SBP and DBP) with age, because of arteriosclerosis and vascular stiffening [46]. The latter is expected given how PhenoAge is estimated, as glucose is one of the nine clinical biomarkers of phenotypic age [3].

Strengths and limitations

To our knowledge, this is the first study that used network analyses to identify relationships between dietary habits and biological age acceleration. A major strength of this study is the use of an epigenetic clock to study the complex process of ageing. The first-generation clocks, HorvathAge and HannumAge, were strong predictors of chronological age, but showed only weak associations with age-related clinical phenotypes, such as physical function and cognitive performance [17]. Therefore, in this study we used the second-generation epigenetic clock PhenoAge, which incorporates age-related and disease phenotypes in combination with chronological age, enabling the estimate of health span outcomes and time-to-death [3]. The strong involvement of the epigenome in age-related diseases makes the epigenetic clock useful for clinical purposes: profiling of DNAm patterns may be used as an emerging tool to facilitate diagnosis, prognosis and prediction of pathological outcomes [47].

The objective explorative approach makes our study unique and of added value to existing literature. We used the mathematical method CGM for efficient network estimation. This method enabled us to use an extensive set of multivariate data. The large data set of almost 4,000 postmenopausal women has only strengthened these results. The associations between increased PhenoAgeAccel and BMI, WHR, smoking pack-years, systolic blood pressure and also sugar, were as expected and in accordance with previously reported findings [3, 40, 48–52], thereby being a form of validation of our CGM model. Our results revealed conditional dependencies between food groups and biological ageing. For each direct association, corrections were made to account for the influence of all other variables in the model. Since nutritional variables are often correlated with each other, this proposed exploratory tool is of great relevance in nutrition research, as it allows us to reveal the exact food groups that have an effect on ageing. Although we believe CGMs represent an excellent methodology to optimise extraction of information from cross-sectional data, this method is accompanied by some inevitable limitations. One of the main limitations of CGMs is the inability to infer causality or direction of relationships, as the edges lack orientation. However, the drifts in epigenetic age in these postmenopausal subjects likely reflect years of adherence to a particular lifestyle. The WHI-FFQ was developed to assess habitual intake over the past 3 months, which has been shown to correlate well with habitual intake over a longer period of time in the past [53]. Therefore, the likelihood of reverse causality, where changes in epigenetic age would alter dietary intake, appears to be limited. Still, causal GMs or interventions are needed to infer causality and direction of relationships. The second limitation of the CGM method is the presence of residual confounding. CGMs adjust every relationship between two variables for all other variables present in the model. This minimises subjective selection of confounders, which is a great advantage over alternative statistical methods. However, only the confounders that are measured are taken into account. Since there might be other variables influencing the relationships, such as processing of food, eating rate or level of income, inevitable residual confounding remains present. Therefore, investigating the replicability of our findings in an independent cohort is highly warranted. Even though we assessed robustness through bootstrapping and sensitivity analyses, including the analysis per WHI ancillary cohort, there remains a chance that our findings could be cohort-specific. Third, a limitation is present in the absence of not recalculating PhenoAgeAA by the residual method for different splits in our sensitivity analyses. Despite this, we maintain that this methodological limitation likely had minimal impact on our study's conclusions, given the large sample size and the division of the data set into only two groups, factors that, alongside the central limit theorem, suggest that variability in PhenoAgeAA calculations across the cohort would be minimised. Fourth, a limitation of the CGM method is the inability to acquire magnitude of relationships, as there will not be any beta coefficient of the observed associations. For these reasons, in contrast to other statistical methods such as regression analyses, it is difficult to understand the clinical value that can be attributed to the observed relationships. Importantly, CGMs offer a suitable tool to identify associations between dietary habits and epigenetic ageing, which may comprise a first step in developing targeted health-promoting interventions. Validation studies in which the WHI-FFQ was compared with other dietary assessment methods, such as food records or 24-h recalls, showed similar nutrient estimates. Moreover, the WHI-FFQ is considered to be a similar or even better measurement of nutrient intake compared with other FFQs in similar populations [53], and is highly feasible in large populations as in our study. Biomarkers of nutritional status might be a more objective and robust measure of dietary intake than questionnaires relying on self-report [54]. The study of Lu *et al*. [55] showed more robust associations between biological age and plasma biomarkers measuring dietary intake, than with self-reported dietary assessment methods. In our sample, some plasma nutritional biomarkers were assessed (e.g. cholesterol, triglycerides), but only in 55% of the subjects. In addition, as nutrient intake via supplements and medication use (e.g. lipid-lowering drugs) was not taken into account in this study, the plasma

biomarkers may not capture the complete dietary intake profile. Also the presence of diseases (e.g. osteoporosis and rheumatoid arthritis) was based on self-report. Therefore, further research in this area should focus on more objective markers of nutritional status and diseases, whether or not in combination with dietary assessment methods.

Conclusion

In conclusion, we identified several compelling associations between food groups and biological ageing, using CGMs as a powerful data-driven exploratory tool. Intake of eggs, organ meat, sausages, cheese, legumes, starchy vegetables, added sugar, lunch meat and fat added after cooking was associated with accelerated biological ageing, whereas intake of peaches/nectarines/plums, poultry, nuts, discretionary oil and solid fat was associated with decelerated ageing. Most results remained similar after stratifying for BMI, physical activity, data set or disease status, and were congruent with previous literature, indicating high sensitivity of the food group–age acceleration networks. Our findings provide a promising base for regression and intervention studies to determine causality and magnitude of relationships, thereby improving the understanding of biological mechanisms underlying the interplay between food groups and biological ageing.

Acknowledgements: The Women's Health Initiative (WHI) programme is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services through contracts 75N92021D00001, 75N92021D00002, 75N92021D00003, 75N92021D00004, 75N92021D00005. WHI investigators are listed @ https://www.whi.org/doc/WHI-Investigator-Long-List.pdf. WHI data are publicly available @ https://www.whi.org/md/working-with-whi-data. AS315 was funded by National Institute of Environmental Health Sciences grant R01-ES020836. BA23 was funded by the National Heart, Lung, and Blood Institute Broad Agency Announcement contract HHSN268201300006C.

Supplementary Data: Supplementary data mentioned in the text are available to subscribers in *Age and Ageing* online.

Declaration of Conflicts of Interest: None.

Declaration of Sources of Funding: P.G. was supported by a grant 162135 from the Regio Deal Foodvalley; S.H. was supported by grant 1U01AG060908-01 from National Institute on Ageing–National Institutes of Health; C.S.K. was supported by grant 457001001 from ZonMW; J.E.M was supported by HL34594, HL145386 and 75N92021D00002 from the National Heart, Lung, and Blood Institute; A.H.S. was supported by grant RF1AG074345 from the National Institute on Aging; J.D.S. was supported by grant R01ES020836 from the National Institute of Environmental Health Sciences and contract 75N92019R0031 from the National Heart, Lung, and Blood Institute; E.A.W. was supported by grant R01ES020836 from the National Institute of Environmental Health Sciences and contract 75N92019R0031 from the National Heart, Lung, and Blood Institute.

References

1. Hamczyk MR, Nevado RM, Barettino A, Fuster V, Andrés V. Biological versus chronological aging: JACC focus seminar. J Am Coll Cardiol 2020; 75: 919–30.
2. Dodig S, Čepelak I, Pavić I. Hallmarks of senescence and aging. Biochem Med (Zagreb) 2019; 29: 030501.
3. Levine ME, Lu AT, Quach A *et al.* An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY) 2018; 10: 573–91.
4. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet 2018; 19: 371–84.
5. Jung M, Pfeifer GP. Aging and DNA methylation. BMC Biol 2015; 13: 7.
6. Nishiyama A, Nakanishi M. Navigating the DNA methylation landscape of cancer. Trends Genet 2021; 37: 1012–27.
7. Galow AM, Peleg S. How to slow down the ticking clock: age-associated epigenetic alterations and related interventions to extend life span. Cell 2022; 11:468.
8. Horvath S, Pirazzini C, Bacalini MG *et al.* Decreased epigenetic age of PBMCs from Italian semi-supercentenarians and their offspring. Aging (Albany NY) 2015; 7: 1159–70.
9. Armstrong NJ, Mather KA, Thalamuthu A *et al.* Aging, exceptional longevity and comparisons of the Hannum and Horvath epigenetic clocks. Epigenomics 2017; 9: 689–700.
10. Perna L, Zhang Y, Mons U, Holleczek B, Saum KU, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. Clin Epigenetics 2016; 8: 64.
11. Jain P, Binder AM, Chen B *et al.* Analysis of epigenetic age acceleration and healthy longevity among older US women. JAMA Netw Open 2022; 5: e2223285.
12. Kaplanis J, Gordon A, Shor T *et al.* Quantitative analysis of population-scale family trees with millions of relatives. Science 2018; 360: 171–5.
13. Ruby JG, Wright KM, Rand KA *et al.* Estimates of the heritability of human longevity are substantially inflated due to assortative mating. Genetics 2018; 210: 1109–24.
14. Lessel D, Kubisch C. Hereditary syndromes with signs of premature aging. Dtsch Arztebl Int 2019; 116: 489–96.
15. Gensous N, Garagnani P, Santoro A *et al.* One-year Mediterranean diet promotes epigenetic rejuvenation with country- and sex-specific effects: a pilot study from the NU-AGE project. Geroscience 2020; 42: 687–701.
16. Belsky DW, Huffman KM, Pieper CF, Shalev I, Kraus WE. Change in the rate of biological aging in response to caloric restriction: CALERIE biobank analysis. J Gerontol A Biol Sci Med Sci 2017; 73: 4–10.
17. Fitzgerald KN, Hodges R, Hanes D *et al.* Potential reversal of epigenetic age using a diet and lifestyle intervention: a

pilot randomized clinical trial. Aging (Albany NY) 2021; 13: 9419–32.
18. Fiorito G, Caini S, Palli D *et al.* DNA methylation-based biomarkers of aging were slowed down in a two-year diet and physical activity intervention trial: the DAMA study. Aging Cell 2021; 20: e13439.
19. Behrouzi P, Arends D, Wit EC. Netgwas: an R package for network-based genome-wide association studies. *R journal* 2023; 14: 18–37.
20. WHI. Design of the Women's Health Initiative clinical trial and observational study. Control Clin Trials 1998; 19: 61–109.
21. Whitsel E. AS315-Epigenetic Mechanisms of PM-Mediated CVD Risk, 2018. https://sp.whi.org/researchers/data/WHIStudies/StudySites/AS315/Pages/home.aspx.
22. Assimes T *et al.* BA23–Integrative Genomics and Risk of CHD and Related Phenotypes in the Women's Health Initiative, 2018. https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001335.v2.p3.
23. Edwards D. Introduction to Graphical Modelling. Springer Science & Business Media, New York, 2012.
24. U.S. Department of Agriculture and U.S. Department of Health and Human Services. *Dietary Guidelines for Americans, 2020-2025*. 9th Edition. 2020. Available at DietaryGuidelines.gov.
25. Ioannou GN. The role of cholesterol in the pathogenesis of NASH. Trends Endocrinol Metab 2016; 27: 84–95.
26. Jung E, Kong SY, Ro YS, Ryu HH, Shin SD. Serum cholesterol levels and risk of cardiovascular death: a systematic review and a dose-response meta-analysis of prospective cohort studies. Int J Environ Res Public Health 2022; 19: 8272.
27. Meex RCR, Blaak EE. Mitochondrial dysfunction is a key pathway that links saturated fat intake to the development and progression of NAFLD. Mol Nutr Food Res 2021; 65: e1900942.
28. Kitada M, Ogura Y, Monno I, Koya D. The impact of dietary protein intake on longevity and metabolic health. EBioMedicine 2019; 43: 632–40.
29. Fretts AM, Howard BV, Siscovick DS *et al.* Processed meat, but not unprocessed red meat, is inversely associated with leukocyte telomere length in the strong heart family study. J Nutr 2016; 146: 2013–8.
30. Händel MN, Cardoso I, Rasmussen KM *et al.* Processed meat intake and chronic disease morbidity and mortality: an overview of systematic reviews and meta-analyses. PLoS One 2019; 14: e0223883.
31. Li X, Zhang T, Li H *et al.* Associations between intake of starchy and non-starchy vegetables and risk of hepatic steatosis and fibrosis. Hepatol Int 2022; 16: 846–57.
32. Bertoia ML, Mukamal KJ, Cahill LE *et al.* Changes in intake of fruits and vegetables and weight change in United States men and women followed for up to 24 years: analysis from three prospective cohort studies. PLoS Med 2015; 12: e1001878.
33. Jenkins DJA, Dehghan M, Mente A *et al.* Glycemic index, glycemic load, and cardiovascular disease and mortality. N Engl J Med 2021; 384: 1312–22.
34. Livesey G, Taylor R, Livesey HF *et al.* Dietary glycemic index and load and the risk of type 2 diabetes: a systematic review and updated meta-analyses of prospective cohort studies. Nutrients 2019; 11:1280.
35. Turati F, Galeone C, Augustin LSA, la Vecchia C. Glycemic index, glycemic load and cancer risk: an updated meta-analysis. Nutrients 2019; 11:2342.
36. Antonia Murcia M, Jiménez AM, Martínez-Tomé M. Vegetables antioxidant losses during industrial processing and refrigerated storage. Food Res Int 2009; 42: 1046–52.
37. Durst RW, Weaver GW. Nutritional content of fresh and canned peaches. J Sci Food Agric 2013; 93: 593–603.
38. Xu K, Peng R, Zou Y, Jiang X, Sun Q, Song C. Vitamin C intake and multiple health outcomes: an umbrella review of systematic reviews and meta-analyses. Int J Food Sci Nutr 2022; 73: 588–99.
39. Connolly G, Clark CM, Campbell RE, Byers AW, Reed JB, Campbell WW. Poultry consumption and human health: how much is really known? A systematically searched scoping review and research perspective. Adv Nutr 2022; 13: 2115–24.
40. Quach A, Levine ME, Tanaka T *et al.* Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. Aging (Albany NY) 2017; 9: 419–46.
41. Zuelch ML, Radtke MD, Holt RR *et al.* Perspective: challenges and future directions in clinical research with nuts and berries. Adv Nutr 2023; 14: 1005–28.
42. Health, U.D.o. and H. Services. Scientific Report of the 2015 Dietary Guidelines Advisory Committee. Added Sugars and Low-Calorie Sweeteners, New York, 2015; https://health.gov/dietaryguidelines/2015-scientific-report/11-chapter-6/d6-3. asp (19 November 2017, date last accessed).
43. Astrup A, Magkos F, Bier DM *et al.* Saturated fats and health: a reassessment and proposal for food-based recommendations: JACC state-of-the-art review. J Am Coll Cardiol 2020; 76: 844–57.
44. de Oliveira Otto MC, Nettleton JA, Lemaitre RN *et al.* Biomarkers of dairy fatty acids and risk of cardiovascular disease in the multi-ethnic study of atherosclerosis. J Am Heart Assoc 2013; 2: e000092.
45. Julibert A, Bibiloni MDM, Mateos D, Angullo E, Tur JA. Dietary fat intake and metabolic syndrome in older adults. Nutrients 2019; 11:1901.
46. Tang KS, Medeiros ED, Shah AD. Wide pulse pressure: a clinical review. J Clin Hypertens (Greenwich) 2020; 22: 1960–7.
47. Papanicolau-Sengos A, Aldape K. DNA methylation profiling: an emerging paradigm for cancer diagnosis. Annu Rev Pathol Mech Dis 2022; 17: 295–321.
48. Cassidy A, de Vivo I, Liu Y *et al.* Associations between diet, lifestyle factors, and telomere length in women. Am J Clin Nutr 2010; 91: 1273–80.
49. Lundgren S, Kuitunen S, Pietiläinen KH *et al.* BMI is positively associated with accelerated epigenetic aging in twin pairs discordant for body mass index. J Intern Med 2022; 292: 627–40.
50. Klopack ET, Carroll JE, Cole SW, Seeman TE, Crimmins EM. Lifetime exposure to smoking, epigenetic aging, and morbidity and mortality in older adults. Clin Epigenetics 2022; 14: 72.
51. Astuti Y, Wardhana A, Watkins J, Wulaningsih W, PILAR Research Network. Cigarette smoking and telomere length: a systematic review of 84 studies and meta-analysis. Environ Res 2017; 158: 480–9.
52. Xiao L, Zan G, Liu C *et al.* Associations between blood pressure and accelerated DNA methylation aging. J Am Heart Assoc 2022; 11: e022257.

53. Patterson RE, Kristal AR, Tinker LF, Carter RA, Bolton MP, Agurs-Collins T. Measurement characteristics of the Women's Health Initiative Food Frequency Questionnaire. Ann Epidemiol 1999; 9: 178–87.
54. Neuhouser ML, Prentice RL, Tinker LF, Lampe JW. Enhancing capacity for food and nutrient intake assessment in population sciences research. Annu Rev Public Health 2023; 44: 37–54.
55. Lu AT, Quach A, Wilson JG *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY) 2019; 11: 303–27.

Received 28 November 2023; editorial decision 8 February 2024