
Deep phenotyping of health–disease continuum in the Human Phenotype Project

In the format provided by the
authors and unedited

Supplementary Tables:

Measurements obtained prior to the initiation meeting	N (up to Dec 2024)
<u>Prior blood tests results from EHR</u> Results of blood tests prior to study initiation are obtained from the participant's EHR, including Complete Blood Count (CBC), Hemoglobin A1C (HbA1c), thyroid stimulating hormone (TSH), C-Reactive Protein (CRP), creatinine, glucose, total cholesterol, High-density lipoprotein (HDL), low-density lipoproteins (LDL), triglycerides, albumin, protein, aspartate transaminase (AST), alanine transaminase (ALT), sodium, potassium, chloride, calcium, phosphorus, vitamin B12 and vitamin D	1,020,678 blood test results (15,582 participants)
<u>Baseline questionnaires</u> Participants are asked to complete baseline questionnaires including information on demographics, income, occupation, educational level, stress, quality of life, well-being, nutritional habits, smoking habits, sleeping habits, physical activity, medical history, reproductive health, medication use and family medical history. In addition a set of questions are asked on previous diagnosis and medical conditions (self reporting)	36,450 (23,340 participants)
Measurements obtained in the initiation meeting at the research site	
Clinical measurements	
<u>Anthropometry</u> Height and weight are measured without shoes by Shekel stadiometer model H150-01-5. Waist, hip and neck circumference is measured with a measuring tape Examples of derived features: BMI (Body Mass Index), Waist to hip ratio	21,130 (13,072 participants)
<u>Blood pressure and heart rate</u> 2 measurements from the left arm and one from the right arm obtained after 5 minutes of chair-seated rest: OMRON HEM-907 blood pressure monitoring device Examples of derived features: Maximum blood pressure, average blood pressure, standard deviation of blood pressure, pulse pressure (systolic-diastolic pressure)	21,044 (12,995 participants)
<u>Orthostatic hypotension</u> One measurement obtained after 5 minutes laying down, followed by two measurements, one after 1 minute and another after 3 minutes standing up with OMRON blood pressure monitoring device. Self reporting of dizziness is recorded. Examples of derived features: Variation in heart rate, systolic blood pressure, and diastolic blood pressure between lying and standing positions	21,130 (12,327 participants)
<u>Electrocardiogram (ECG)</u> 1 ten seconds measurement with NORAV, 12 leads ECG Examples of derived features: Heart rate, P-wave duration, PR interval, QRS complex duration: QT interval, R-wave peak voltage in lead V1, S-wave peak voltage in lead V6	20,266 (12,828 participants)
<u>Liver Ultrasound</u> Liver ultrasound images and measurements of 2D shear wave elastography (2D-SWE) are performed from 3 different regions of the liver right lobe with Supersonic, Aixplorer MACH30.. Examples of derived features: Liver sound speed, Liver viscosity, Liver elasticity, Liver attenuation	17,278 (11,756 participants)
<u>Ankle-brachial pressure index (ABI)</u> 1 bilateral measurement with Viasonix - FALCON Examples of derived features: Right and left ankle and brachial systolic pressures	19,567 (12,603 participants)
<u>Carotid ultrasound for Intima Media Thickness (IMT)</u> Carotid artery ultrasound images and 2 measurements with automatic border detection and measurement, one for each side, with Supersonic- Aixplorer MACH30	17,278 (11,756 participants)
<u>Legs Pulse Wave velocity (PWV)</u> 1 bilateral measurement with Viasonix - FALCON of thigh-ankle pulse wave velocity. Examples of derived features: Pulse transient time	19,567 (12,603 participants)

<u>DXA (Dual energy X-ray Absorptiometry) scan - Bone density & total body composition scan</u> 1 scan with GE- Lunar Prodigy Advance. Examples of derived features: Measurements of fat mass, scanned vat mass and android tissue fat percent, bone mineral density Z and T scores.	16,683 (11,551 participants)
<u>Voice recording</u> 1 voice recording for a duration of 30 seconds conducted in the research center in a designated room The participant's voice is audio-recorded for 30 seconds while counting to thirty 30.	5,308 (4,601 participants)
<u>Hand grip test</u> 2 measurements, one for each hand, of maximal force hand grip strength by Jamar Plus Hand Dynamometer	21,130 (12,327 participants)
<u>Color fundus imaging</u> 2 scans for each eye, without pupil dilation, obtained by iCare DRSplus confocal fundus imaging system (iCare) ⁵⁵ from central and nasal views. Examples of derived features: Fundus vascular measurements include vessel density, the area of blood vessels as a fraction of the total area imaged; average width; fractal dimension, a measurement of vessel complexity; and three different tortuosity measures separated to veins, arteries and vessels ⁵⁶	10,954 (8,732 participants)
<u>Gait and balance</u> Testing protocol includes seven separate assessments (gait 3 km/hr, gait self-selected speed, stationary walk, sit to stand, Romberg balance test with closed and with open eyes) where each test is conducted per a predetermined measured time during the tested activities, the participant's movements are recorded by Newton system sensors. Examples of derived features: step length, sway width, speed, amplitude	1,300 (1,300 participants)
Biological samples	
<u>Buccal swabs</u> 1 buccal swab is used for the analyses of Human genetics and oral microbiome (Table S3)	10,571 (10,530 participants)
<u>Stool</u> 1 stool sample is self collected by participants with DNA Genotek OM-200 kit for the analysis of gut microbiome (Extended Data Table 3)	18,602 (11,702 participants)
<u>Vaginal swab</u> 1 vaginal sample is self collected by participants with a swab (Copan Diagnostics) for the analysis of vaginal microbiome (Extended Data Table 3).	1,652 (1,652 participants)
<u>Whole blood</u> Whole blood is collected in EDTA tubes. <u>Complete blood count (CBC)</u> is processed by a DxH 560 hematology analyzer. Additional samples are aliquotes and stored at -80°C freezer. Examples of derived features: Whole genome sequencing, epigenetics	9,816 15,209
<u>Serum</u> Serum samples are separated from 5 ml blood and 6 aliquots are stored at -80°C freezer for various multi omics analysis (Extended Data Table 3). Examples of derived features: Metabolomics, proteomics, antibodies, hormones.	19,450 (12,270 participants)
<u>PBMCs</u> Blood is collected in CPT tubes (BD) from which PBMCs are purified, aliquoted into 3 tubes and stored at -80°C freezer and liquid nitrogen. Examples of derived features: Bulk RNA-Seq, scRNA-Seq.	15,160 (9,570 participants)
Continuous Measurements	
<u>Continuous glucose measurements (CGM)</u> Connection to FreeStyle Libre Pro Flash continuous glucose monitoring (FSL-CGM) system for two weeks.	10,735 participants connected

Examples of derived features: Measures of glucose control and glucose variability including estimate HbA1C (eA1C), Mean Amplitude of Glycemic Excursions (MAGE), High Blood Glucose Index (HBGI) and Low Blood Glucose Index (LBGI) ⁵⁷ .	(For a total of 109,788 days)
Continuous real-time, food intake, medication intake, sleep times, and physical activity logging Logging is performed by a proprietary smartphone App (available at https://play.google.com/store/apps/details?id=il.org.project10k.app) Diet logging involves collecting data on food and drink consumption including information such as types of food, serving sizes, nutritional information and the times of ingestion. Examples of derived features: Total energy intake, micronutrient and macronutrient content (average carbohydrate, fat and protein consumption per day from the total energy), fiber content, processed food classification	12,923 participants logged in the app for a total of 574,000 days
Physical activity monitoring Performed by health and activity tracking application installed on the mobile phone, including steps count.	4,680 participants connected
Continuous sleep monitoring Performed by a home sleep apnea test (HSAT) WatchPAT™300 by Itamar Medical ⁴⁸ for 3 nights. Examples of derived features: Sleep duration, number of wakes during sleep; peripheral respiratory disturbance index (pRDI), respiratory-related arousal events during sleep, peripheral Apnea-Hypopnea Index (pAHI), peripheral Oxygen Desaturation Index (pODI), total rapid eye motion (REM) time, sleep stages (light/deep), number of snoring events and amplitude, sleeping positions, pulse rate, oxygen saturation values	14,233 connections to WatchPAT™ from 10,300 participants (41,899 days)

Table S1. Measurements obtained in the HPP study

Country	N
Poland	9379
Romania	3776
Morocco	3033
Russia	2976
Iraq	2882
Israel	2850
Ukraine	2840
Germany	1907
Yemen	1561
Hungary	1370
Iran	1309
Turkey	1068
Lithuania	1015
Tunisia	872
Belarus	830
Bulgaria	821
Libya	775

Czech Republic	581
Egypt	570
Syria	556
Austria	517
United States	433
United States of America	433
Moldova	381
Greece	361
Slovakia	314
United Kingdom	308
Argentina	308
Algeria	262
Latvia	258
Uzbekistan	224
India	210
Netherlands	196
Italy	180
Georgia	173
France	143
Azerbaijan	130
Afghanistan	123
South Africa	112
Lebanon	103
Spain	83
Serbia	60
Macedonia	58
Canada	57
Belgium	57
Brazil	47
Switzerland	43
Croatia	38
Chile	33

Denmark	29
Bosnia and Herzegovina	27
Kazakhstan	25
Uruguay	24
Ethiopia	16
Mexico	16
Australia	15
China	15
Ireland	15
Armenia	14
Estonia	14
Finland	13
Sweden	11

Table S2 Ancestry of the participants in the HPP cohort. The table displays the number of participants' grandparents born in each country. Only countries with a minimum of 10 reported grandparents born in them are included.

	N	Mean AUC-ROC	SD	Source
Ashkenaz	3760	0.985	0.006	Genetics
North African	574	0.999	0	Genetics
Middle Eastern	480	0.995	0.002	Genetics
Sephardi	176	0.97	0.015	Genetics
Yemen	156	0.999	0.001	Genetics
Ashkenaz	3497	0.617	0.005	Metabolomics
North African	524	0.565	0.024	Metabolomics
Middle Eastern	439	0.586	0.027	Metabolomics
Sephardi	166	0.49	0.03	Metabolomics
Yemen	149	0.613	0.051	Metabolomics
Ashkenaz	2694	0.658	0.017	Lifestyle habits
North African	377	0.565	0.03	Lifestyle habits
Middle Eastern	326	0.615	0.016	Lifestyle habits
Sephardi	118	0.576	0.089	Lifestyle habits
Yemen	109	0.539	0.067	Lifestyle habits

Ashkenaz	3604	0.51	0.025	Microbiome
North African	546	0.494	0.024	Microbiome
Middle Eastern	454	0.54	0.033	Microbiome
Sephardi	169	0.512	0.028	Microbiome
Yemen	144	0.467	0.041	Microbiome
Ashkenaz	2694	0.669	0.015	Lifestyle habits, metabolomics and microbiome
North African	377	0.561	0.038	Lifestyle habits, metabolomics and microbiome
Middle Eastern	326	0.638	0.031	Lifestyle habits, metabolomics and microbiome
Sephardi	118	0.56	0.042	Lifestyle habits, metabolomics and microbiome
Yemen	109	0.614	0.036	Lifestyle habits, metabolomics and microbiome

Table S3. The ability of a machine learning algorithm (LightGBM ¹², see Methods) to predict ethnicity from different data modalities in the main ethnic groups included in the study.

	Ashkenaz (N)	Mean	SD	Non-Ashkenaz (N)	Mean	SD	P-Value (Mann-Whitney U test)	FDR Adjusted P-Value
Total	2513			2510				
Diet composition								
Adherence to Kosher Diet (Yes)	1684	0.301	0.644	1619	0.657	0.838	< 0.001	< 0.001
Wine Consumption (g/day)	1971	17.04	36.074	2029	12.103	28.965	< 0.001	< 0.001
Ice Cream Consumption (g/day)	1971	16.914	34.714	2029	12.49	27.944	< 0.001	< 0.001
Potato Intake (g/day)	1971	20.588	27.159	2029	16.587	24.29	< 0.001	< 0.001
Butter as Preferred Oil for Frying (Yes)	1832	0.116	0.321	1774	0.068	0.251	< 0.001	< 0.001
Alcohol Consumption (g/day)	1527	2.345	1.161	1399	2.071	1.122	< 0.001	< 0.001
Butter Usage for Frying (Yes)	1832	0.204	0.403	1774	0.131	0.338	< 0.001	< 0.001
Consumption of Burekas (baked pastry) (g/day)	1971	8.843	23.382	2029	12.186	28.023	< 0.001	< 0.001
Potato Intake (g/day)	1971	20.588	27.159	2029	16.587	24.29	< 0.001	< 0.001
Canola Oil Usage for Frying (Yes)	1832	0.584	0.493	1774	0.658	0.475	< 0.001	< 0.001
White Wine (glasses per Week)	1317	1.128	1.927	1214	0.892	1.729	< 0.001	< 0.001
Consumption of Yellow Cheese (g/day)	1971	18.406	32.773	2029	14.11	24.948	< 0.001	< 0.001
Rice Consumption (g/day)	1971	31.282	37.517	2029	37.083	42.41	< 0.001	< 0.001
Olive Oil Usage for Frying (Yes)	1832	0.704	0.457	1774	0.642	0.48	< 0.001	0.001
Egg Consumption (g/day)	1971	16.507	24.641	2029	17.824	24.206	< 0.001	0.001
English Cake Consumption (g/day)	1971	7.228	21.044	2029	8.943	22.464	< 0.001	0.001
Coffee Consumption (g/day)	1971	18.845	29.517	2029	21.021	33.604	0.003	0.015

Cookie Consumption (g/day)	1971	11.839	24.409	2029	15.092	30.854	< 0.001	0.003
Granola Consumption (g/day)	1971	7.932	29.65	2029	6.323	24.415	0.001	0.006
Daily Water (glasses consumed per day)	1650	6.452	3.843	1574	6.884	4.022	0.001	0.008
Chocolate Consumption (g/day)	1971	11.301	24.935	2029	8.409	20.512	0.002	0.01
Pasta Consumption (g/day)	1971	13.934	28.163	2029	12.104	29.839	0.003	0.015
Banana Consumption (g/day)	1971	13.934	28.163	2029	12.104	29.839	0.003	0.015
Beer Consumption (g/day)	1971	8.353	20.617	2029	7.009	21.949	0.004	0.018
Bread roll Consumption (g/day)	1971	14.219	28.571	2029	16.579	31.271	0.005	0.023
White Wine (Glasses per month)	805	0.739	1.089	841	0.649	1.328	0.006	0.027
Weekly Liquor Consumption	1341	0.632	1.91	1218	0.458	1.223	0.007	0.03
Emmer Wheat Bread Consumption (g/day)	2508	102.982	8.719	2507	102.418	8.73	0.01	0.041
Cottage Cheese Consumption (g/day)	1971	13.656	29.518	2029	10.972	26.257	0.012	0.047
Blood test results								
Absolute Monocytes Count (cells/ μ L)	2370	0.525	0.755	2373	0.464	0.316	< 0.001	< 0.001
Absolute Neutrophils Count (cells/ μ L)	2347	3.742	1.171	2333	3.463	1.192	< 0.001	< 0.001
White Blood Cell Count (cells/ μ L)	2324	6.587	1.433	2318	6.21	1.462	< 0.001	< 0.001
Serum Creatinine Level (mg/dL)	2097	0.843	0.157	2058	0.81	0.165	< 0.001	< 0.001
Lymphocytes Percentage (%)	2374	32.296	6.782	2370	33.637	7.157	< 0.001	< 0.001
Total Serum Protein (g/dL)	1464	7.037	0.351	1392	7.121	0.359	< 0.001	< 0.001
Monocytes Percentage (%)	2354	7.676	1.79	2353	7.372	1.91	< 0.001	< 0.001
Mean Corpuscular Hemoglobin (MCH, pg)	2358	29.811	1.521	2343	29.406	2.008	< 0.001	< 0.001
Mean Platelet Volume (MPV, fL)	2315	10.106	1.263	2291	10.314	1.288	< 0.001	< 0.001
Mean Corpuscular Hemoglobin Concentration (MCHC, g/dL)	2371	33.362	0.959	2367	33.216	1.034	< 0.001	< 0.001
Alkaline Phosphatase Level (IU/L)	1802	72.753	19.466	1695	70.167	20.408	< 0.001	< 0.001
Absolute Eosinophils Count (cells/ μ L)	2343	0.205	0.2	2332	0.187	0.139	< 0.001	< 0.001
Mean Corpuscular Volume (MCV, fL)	2375	89.378	3.954	2357	88.832	4.454	< 0.001	< 0.001
Platelet Count ($\times 10^9$ /L)	2276	248.752	54.284	2242	242.568	54.771	< 0.001	< 0.001
Neutrophils Percentage (%)	2374	55.899	7.277	2366	54.88	7.949	< 0.001	< 0.001
Absolute Lymphocytes Count (cells/ μ L)	2374	2.109	0.672	2371	2.053	0.607	< 0.001	0.003
Thyroid-Stimulating Hormone (TSH, μ U/mL)	1684	2.288	1.396	1628	2.125	1.183	0.001	0.005
Serum Vitamin D Level (ng/mL)	1046	26.576	9.157	912	25.241	8.756	0.002	0.009
Estimated Glomerular Filtration Rate (eGFR, mL/min/1.73m ²)	885	92.111	13.123	772	94.247	12.201	0.002	0.012
Red Blood Cell Count (RBC, $\times 10^6$ / μ L)	2282	4.706	0.411	2258	4.758	0.47	0.003	0.015

Blood Urea Level (mg/dL)	1792	30.492	8.085	1711	29.802	7.824	0.008	0.032
Lifestyle habits								
Summer Outdoor Activity Time (hours)	1689	2.102	2.356	1607	2.726	2.906	< 0.001	< 0.001
Driving Hours per Day	1669	1.15	0.844	1602	1.416	1.472	< 0.001	< 0.001
Winter Outdoor Activity Time (hours)	1652	1.914	2.345	1598	2.331	2.733	< 0.001	< 0.001
Attendance at Pub or Social Club (once per week)	1832	0.088	0.283	1774	0.132	0.339	< 0.001	< 0.001
Workdays per Week	1626	3.588	1.842	1530	3.854	1.705	< 0.001	< 0.001
Transportation Mode: Car/Motorcycle	1832	0.748	0.434	1774	0.801	0.399	< 0.001	0.001
Daytime Naps (number per day)	1694	1.45	0.63	1642	1.379	0.607	< 0.001	0.002
Weekly Mobile Phone Usage Duration (hours)	1658	3.946	1.344	1620	4.106	1.328	0.001	0.004
Vigorous Physical Activity Days per Week	1832	0.039	0.193	1774	0.061	0.24	0.002	0.01
Daily Cigarette Consumption	435	1.115	1.755	428	0.787	1.505	0.002	0.01
Strenuous Physical Activity	1971	23.484	48.5	2029	18.964	37.486	0.002	0.011
Duration of High-Intensity Exercise	1971	8.353	20.617	2029	7.009	21.949	0.004	0.018
Transportation Mode: Public Transit or Train	936	3.232	1.139	893	3.082	1.14	0.005	0.021
Household Smoking Status*	1832	0.337	0.473	1774	0.294	0.456	0.005	0.024
Hours Watching Television per Day	1341	0.632	1.91	1218	0.458	1.223	0.007	0.03
Transportation Mode: Bicycle	1705	3.01	0.813	1647	2.934	0.826	0.009	0.035
Transportation Mode: Walking for Pleasure	1971	16.095	42.345	2029	20.017	49.246	0.011	0.043
Vital signs								
Systolic Blood Pressure (mmHg)	2508	121.104	17.41	2510	118.448	16.651	< 0.001	< 0.001
Systolic Blood Pressure While Lying Down (mmHg)	2298	120.093	14.974	2315	118.521	15.069	< 0.001	0.001
Systolic Blood Pressure After Standing for Three Minutes (mmHg)	2294	118.645	15.481	2305	117.014	15.74	< 0.001	0.003
Sitting Diastolic Blood Pressure (mmHg)	2509	78.025	10.664	2510	77.055	10.607	0.001	0.006
Systolic Blood Pressure After Standing for One Minute (mmHg)	2297	117.828	15.9	2309	116.453	16.295	0.001	0.007
Diastolic Blood Pressure After Standing for Three Minutes (mmHg)	885	92.111	13.123	772	94.247	12.201	0.002	0.012
Diastolic Blood Pressure After Standing for One Minute (mmHg)	2297	82.472	9.793	2309	81.656	9.709	0.007	0.029

Table S4. Comparison of clinical parameters between the Ashkenazi ethnic group (n=2,526) and a matched cohort from other ethnic groups (n=2,526), with matching based on sex, age, and BMI. Only parameters with statistically significant differences are presented. *The Household Smoking Status score was calculated based on survey responses: 0 indicates no household members smoke, 1 indicates one household member smokes, and 2

indicates more than one household member smokes. Statistical significance was determined by FDR adjusted P-Value ≤ 0.05 .

Cardiovascular age	Bottom 10th Percentile		Top 10th Percentile			
	Mean	SD	Mean	SD	Mann-Whitney P-Value	FDR Adjusted P-Value
Lying Blood Pressure Systolic (mmHg)	111.737	12.473	129.184	15.114	<0.001	<0.001
Sitting Blood Pressure Systolic (mmHg)	112.222	13.940	131.191	16.785	<0.001	<0.001
Standing 3-Min Blood Pressure Systolic (mmHg)	111.468	13.062	127.196	16.886	<0.001	<0.001
Sitting Blood Pressure Diastolic (mmHg)	74.417	8.826	83.408	10.596	<0.001	<0.001
Standing 1-Min Blood Pressure Systolic (mmHg)	110.933	13.450	125.627	17.287	<0.001	<0.001
Body Mass Index (kg/m ²)	24.533	3.533	27.893	4.412	<0.001	<0.001
Standing 3-Min Blood Pressure Diastolic (mmHg)	79.873	8.872	86.743	10.455	<0.001	<0.001
Waist Circumference (cm)	85.110	10.983	93.475	12.540	<0.001	<0.001
Waist Circumference (cm)	85.110	10.983	93.475	12.540	<0.001	<0.001
Standing 1-Min Blood Pressure Diastolic (mmHg)	79.307	8.610	85.832	10.784	<0.001	<0.001
Weight (kg)	70.254	12.809	80.363	15.649	<0.001	<0.001
Hip Circumference (cm)	99.939	7.822	105.434	8.654	<0.001	<0.001
Trunk Fat Mass (g)	11226.637	4412.275	15165.749	5395.410	<0.001	<0.001
Android Fat Mass (g)	1821.425	879.797	2587.193	1099.911	<0.001	<0.001
Total Fat Mass (g)	21740.684	7246.778	27720.141	8532.309	<0.001	<0.001
Visceral Fat Mass (g)	703.695	576.170	1138.701	745.191	<0.001	<0.001
Total Body Mass (kg)	70434.388	12795.242	80194.120	15244.280	<0.001	<0.001
Android Tissue Percent Fat (%)	0.341	0.102	0.403	0.094	<0.001	<0.001
Triglycerides (mg/dL)	96.415	47.993	128.064	74.007	<0.001	<0.001
Oxygen Desaturation Index (ODI, %)	2.977	4.261	5.818	8.157	<0.001	<0.001
Oxygen Desaturation Index (ODI)	2.977	4.261	5.818	8.157	<0.001	<0.001
REM Apnea-Hypopnea Index (AHI)	12.383	10.538	19.599	15.738	<0.001	<0.001
Apnea-Hypopnea Index (AHI)	8.119	7.742	13.028	11.258	<0.001	<0.001
Gynoid Fat Mass (g)	3789.662	1320.560	4541.750	1480.450	<0.001	<0.001

REM Oxygen Desaturation Index (ODI)	4.765	6.142	9.027	10.555	<0.001	<0.001
REM Respiratory Disturbance Index (RDI)	19.122	10.392	25.600	14.618	<0.001	<0.001
NREM Apnea-Hypopnea Index (AHI)	6.741	7.393	11.078	11.047	<0.001	<0.001
NREM Oxygen Desaturation Index (ODI)	2.390	4.063	4.812	8.004	<0.001	<0.001
Respiratory Disturbance Index (RDI)	13.827	8.488	18.461	10.844	<0.001	<0.001
Visceral Adipose Tissue Fat Percent (%)	0.030	0.020	0.040	0.021	<0.001	<0.001
Total Region Percent Fat (%)	0.308	0.079	0.344	0.076	<0.001	<0.001
Waist-to-Hip Ratio (WHR)	0.852	0.091	0.886	0.090	<0.001	<0.001
NREM Respiratory Disturbance Index (RDI)	12.127	8.727	16.230	10.863	<0.001	<0.001
Android/Gynoid Fat Ratio	1.006	0.260	1.118	0.247	<0.001	<0.001
Total Tissue Percent Fat (%)	0.319	0.081	0.356	0.077	<0.001	<0.001
Blood Glucose (mg/dL)	91.938	8.990	95.749	11.796	<0.001	<0.001
White Blood Cell Count (WBC, 10 ³ /μL)	6.223	1.436	6.734	1.548	<0.001	<0.001
High-Density Lipoprotein Cholesterol (HDL, mg/dL)	56.645	12.845	52.510	12.401	<0.001	<0.001
Minimum Oxygen Saturation Value (%)	88.510	3.734	86.808	5.148	<0.001	<0.001
Legs Lean Mass (g)	15243.585	3511.455	16693.917	3935.279	<0.001	<0.001
Appendicular Skeletal Muscle Mass (ASM, kg)	20476.683	5099.963	22435.866	5591.015	<0.001	<0.001
Absolute Lymphocyte Count (10 ³ /μL)	2.016	0.557	2.188	0.691	<0.001	<0.001
Alanine Aminotransferase (ALT/GPT, U/L)	20.283	9.361	23.767	12.728	<0.001	<0.001
Total Lean Mass (kg)	46174.131	9757.600	49858.093	10636.196	<0.001	<0.001
Non-HDL Cholesterol (mg/dL)	136.243	33.574	146.396	35.059	<0.001	<0.001
Arms Lean Mass (g)	5233.097	1732.591	5741.949	1849.891	<0.001	<0.001
Absolute Neutrophil Count (10 ³ /μL)	3.512	1.120	3.787	1.386	<0.001	<0.001
Uric Acid (mg/dL)	5.044	1.291	5.443	1.355	<0.001	<0.001
Mean Corpuscular Volume (MCV, fL)	89.708	4.136	88.858	4.257	<0.001	<0.001
Mean Corpuscular Hemoglobin (MCH, pg)	29.724	1.662	29.423	1.711	<0.001	<0.001
Spine L1-L4 Bone Mineral Density (BMD, g/cm ²)	1.133	0.160	1.173	0.161	<0.001	<0.001

Alkaline Phosphatase (U/L)	69.526	19.508	74.036	20.456	<0.001	<0.001
Femur Total Mean Z-Score	-0.189	0.968	0.041	0.943	<0.001	<0.001
Gynoid Tissue Percent Fat (%)	0.346	0.097	0.370	0.094	<0.001	<0.001
Neck Circumference (cm)	35.619	3.571	37.516	4.205	<0.001	<0.001
Absolute Eosinophil Count (10 ³ /μL)	0.186	0.131	0.218	0.255	<0.001	<0.001
REM Apnea-Hypopnea Index Central (AHI Central)	0.680	1.239	1.054	1.923	<0.001	<0.001
Apnea-Hypopnea Index Central (AHI Central)	0.599	1.128	0.957	1.865	<0.001	<0.001
Glycated Hemoglobin (HbA1C, %)	5.352	0.375	5.474	0.451	<0.001	<0.001
NREM Apnea-Hypopnea Index Central (AHI Central)	0.570	1.262	0.880	1.809	<0.001	<0.001
Lactate Dehydrogenase (LDH, U/L)	302.887	61.916	317.715	62.559	<0.001	0.001
Absolute Basophil Count (10 ³ /μL)	0.039	0.061	0.043	0.058	<0.001	0.001
Mean Oxygen Saturation Value (%)	95.237	1.107	95.035	1.194	<0.001	0.001
Number of Wakes	6.633	3.061	7.158	3.269	<0.001	0.001
Femur Total Mean T-Score	-0.327	1.085	-0.123	1.099	0.001	0.001
Total Protein (g/dL)	7.038	0.341	7.108	0.354	0.001	0.001
Absolute Monocyte Count (10 ³ /μL)	0.491	0.391	0.539	1.073	0.001	0.002
Red Blood Cell Count (RBC, 10 ⁶ /μL)	4.692	0.436	4.750	0.409	0.001	0.002
Total Cholesterol (mg/dL)	193.442	36.623	198.618	37.416	0.002	0.004
Monocytes (%)	7.653	1.815	7.514	2.996	0.004	0.006
Platelet Count (10 ³ /μL)	242.155	52.322	250.545	56.617	0.004	0.007
C-Reactive Protein (CRP, mg/L)	-0.315	1.979	-0.321	0.680	0.016	0.026
Femur Right Neck BMD (g/cm ²)	0.923	0.133	0.939	0.144	0.021	0.035
Femur Left Neck BMD (g/cm ²)	0.928	0.135	0.941	0.141	0.024	0.039
Vitamin D (ng/mL)	26.197	8.865	24.831	8.257	0.025	0.040

Table S5. Comparison of individuals predicted to have a high vs low cardiovascular age across several clinical measures, showing worse clinical values across several parameters for individuals predicted to have a higher cardiovascular age.

Gene expression age	Bottom 10th Percentile		Top 10th Percentile		Mann-Whitney P-Value	FDR Adjusted P-Value
	Mean	SD	Mean	SD		
Body Mass Index (kg/m ²)	25.14	3.78	27.23	4.54	<0.001	<0.001

Waist Circumference (cm)	87.03	11.68	92.19	13.10	<0.001	<0.001
Trunk Fat Mass (g)	11790.18	4649.80	14347.12	5634.92	<0.001	<0.001
Weight (kg)	72.02	13.25	77.94	16.10	<0.001	<0.001
Total Fat Mass (g)	22369.61	7240.78	26515.23	9120.34	<0.001	<0.001
Android Fat Mass (g)	1939.19	943.94	2427.42	1137.02	<0.001	<0.001
Hip Circumference (cm)	100.93	9.05	103.98	8.82	<0.001	<0.001
Triglycerides (mg/dL)	103.38	53.32	120.27	56.52	<0.001	<0.001
White Blood Cell Count (WBC, 10 ³ /μL)	6.23	1.41	6.80	1.62	<0.001	<0.001
Absolute Neutrophil Count (10 ³ /μL)	3.42	1.03	3.80	1.13	<0.001	<0.001
Android Tissue Percent Fat (%)	0.35	0.11	0.39	0.09	<0.001	<0.001
Gynoid Fat Mass (g)	3804.22	1310.00	4383.84	1522.08	<0.001	<0.001
Sitting Blood Pressure Systolic (mmHg)	118.64	15.76	123.54	17.14	<0.001	<0.001
Visceral Fat Mass (g)	801.54	635.28	1024.15	691.62	<0.001	<0.001
Total Region Percent Fat (%)	0.31	0.08	0.34	0.08	<0.001	<0.001
Total Tissue Percent Fat (%)	0.32	0.08	0.35	0.08	<0.001	<0.001
Total Body Mass (kg)	72220.05	13053.58	77516.76	15955.46	<0.001	0.001
High-Density Lipoprotein Cholesterol (HDL, mg/dL)	55.98	13.04	52.68	13.72	<0.001	0.001
Uric Acid (mg/dL)	5.12	1.30	5.55	1.33	<0.001	0.001
Lying Blood Pressure Systolic (mmHg)	117.99	14.74	122.16	15.35	<0.001	0.001
Standing 1-Min Blood Pressure Systolic (mmHg)	116.05	15.96	120.08	16.23	<0.001	0.001
Glycated Hemoglobin (HbA1C, %)	5.36	0.35	5.53	0.45	<0.001	0.001
Oxygen Desaturation Index (ODI)	3.71	5.11	5.46	7.45	<0.001	0.001
Oxygen Desaturation Index (ODI, %)	3.71	5.11	5.46	7.45	<0.001	0.001
Apnea-Hypopnea Index (AHI)	9.61	8.50	12.53	10.61	<0.001	0.002
Waist-to-Hip Ratio (WHR)	0.86	0.09	0.89	0.10	0.001	0.002
REM Oxygen Desaturation Index (ODI)	6.07	7.63	8.33	9.99	0.001	0.002
Sitting Blood Pressure Diastolic (mmHg)	77.39	10.21	79.65	10.21	0.001	0.003
NREM Oxygen Desaturation Index (ODI)	2.88	4.50	4.35	6.95	0.001	0.003
Mean Oxygen Saturation Value (%)	95.38	1.01	95.03	1.20	0.001	0.003
NREM Apnea-Hypopnea Index	7.91	7.83	10.59	10.22	0.001	0.003

(AHI)						
REM Apnea-Hypopnea Index (AHI)	14.48	11.40	18.01	13.35	0.001	0.003
Alanine Aminotransferase (ALT/GPT, U/L)	20.88	9.85	23.82	13.24	0.001	0.005
Standing 3-Min Blood Pressure Systolic (mmHg)	117.05	15.11	120.37	16.19	0.002	0.008
Femur Total Mean Z-Score	-0.26	0.91	-0.01	1.03	0.003	0.009
REM Respiratory Disturbance Index (RDI)	20.96	11.40	23.81	12.31	0.003	0.010
Gynoid Tissue Percent Fat (%)	0.34	0.10	0.37	0.09	0.003	0.011
Platelet Count (10 ³ /μL)	240.61	52.93	252.94	56.09	0.004	0.012
Respiratory Disturbance Index (RDI)	15.25	9.14	17.57	10.35	0.004	0.013
Blood Glucose (mg/dL)	93.49	9.40	96.09	12.02	0.005	0.015
Standing 1-Min Blood Pressure Diastolic (mmHg)	81.28	9.47	83.38	9.47	0.005	0.015
Android/Gynoid Fat Ratio	1.04	0.28	1.10	0.26	0.006	0.016
Non-HDL Cholesterol (mg/dL)	138.15	32.77	144.96	31.48	0.006	0.016
Absolute Monocyte Count (10 ³ /μL)	0.48	0.16	0.50	0.15	0.006	0.016
C-Reactive Protein (CRP, mg/L)	-0.44	0.29	-0.18	0.82	0.011	0.028
Visceral Adipose Tissue Fat Percent (%)	0.03	0.02	0.04	0.02	0.011	0.029
Total Protein (g/dL)	7.05	0.38	7.12	0.35	0.012	0.030
Neutrophils (%)	54.35	7.68	55.71	8.02	0.013	0.031
NREM Respiratory Disturbance Index (RDI)	13.25	8.95	15.41	10.49	0.014	0.033
Vitamin D (ng/mL)	27.03	9.30	24.58	8.22	0.017	0.039

Table S6. Comparison of individuals predicted to have a high vs low age by RNA expression across several clinical measures, showing worse clinical values across several parameters for individuals predicted to have a higher gene expression age.

	Low bone mass disorders	Fracture	Back pain	Anemia	Peptic ulcer	Anxiety	PCOS	NAFLD	HTN	Hyperlipoproteinaemia	Diabetes	OSA
Females												
Bone density	1.255	1.071										
Hematopoietic system				0.802				1.148				
Gut microbiome					1.189							

Cardiovascular system							1.076		1.218	1.058	1.100	
Blood lipids				0.920						1.170	1.081	
Insulin resistance								1.190		1.110	1.416	
Sleep characteristics						1.139		1.135	1.122	1.059	1.150	
Body composition	1.163							1.125	1.082	1.083	1.080	
Liver health								1.376	1.142	1.079		
Males												
Sleep characteristics						1.149			1.123	1.068		1.171
Bone density			1.053					1.071	1.057			1.090
Body composition								1.098	1.079	1.068		
Cardiovascular system								1.085	1.140	1.044	1.067	1.115
Blood lipids								1.112	1.096	1.131	1.090	
Insulin resistance								1.107	1.100	1.076	1.452	

Table S7. The association between biological age scores and self-reported medical diagnoses, separated into males and females, adjusted across sexes. Values of odds ratio for being diagnosed with a medical condition which were computed using a logistic regression model and remained statistically significant after adjustment for multiple hypotheses (see Methods) are shown. HTN – Hypertension; NAFLD – Non-Alcoholic Fatty Liver Disease; OSA – Obstructive Sleep Apnea; PCOS – Polycystic Ovary Syndrome.

Category	Human Phenotype Project (HPP)	UK Biobank ⁶²	All of Us Research Program ⁵	Million Veteran Program (MVP) ⁶³	China Kadoorie Biobank (CKB) ⁶⁴	Lifelines Cohort Study ⁶⁵
Phenotypes Studied & Data Modalities	Deep phenotyping: medical history, lifestyle, continuous glucose monitoring (CGM) with real-time meal logging, sleep monitoring, imaging, molecular profiling (genomics, transcriptomics, microbiome, metabolomics, immune system).	Broad phenotyping: EHR, anthropometrics, cognitive function, imaging (including MRI), whole genome sequencing (WGS), proteomics, metabolomics, lifestyle surveys (sleep & diet).	Broad phenotyping: genetics, EHR data, lifestyle surveys, wearable devices; limited imaging and omics data.	Genetic data linked to EHR, focus on mental health, chronic diseases, and military-specific health issues.	Chronic disease, genetics, environmental factors, and lifestyle in urban/rural Chinese populations.	Genetics, chronic diseases, environmental data, cognitive and mental tests, lung function, multi-generational follow-up.
Frequency of Visits	Baseline assessment, annual questionnaires, bi-annual in-person visits; planned 25-year follow-up.	Single baseline assessment; ~20,000 participants had a repeat visit; annual individual-specific surveys (e.g., chronic pain, digestive health, sleep, diet). Integration to longitudinal EHR	Annual updates; integrates EHR and wearable data for continuous monitoring.	Ongoing integration of genetic and EHR data.	Baseline assessments with follow-up for over 10 years.	Long-term, multi-generational follow-up; second in-person assessment conducted 1-12 years after baseline.

		data.				
Timing of Tests	All tests performed simultaneously at each clinic visit. Following the visit, participants are connected to a CGM device for 2 weeks and in parallel are asked to log food intake, exercise and drug consumption in real time. Participants are also connected to sleep monitoring for 3 nights.	Tests conducted at separate time points.	Tests linked to routine healthcare visits and EHR updates; separate wearable and survey data collection.	Based on clinical interactions and data availability.	Baseline tests conducted simultaneously.	Tests conducted simultaneously at recruitment.
Ethnic Diversity	Ethnically diverse cohort, including Ashkenazi, Middle Eastern, North African, and other ethnic groups.	Predominantly White European; increasing efforts to recruit diverse groups.	Designed for diversity: persons from underrepresented populations are prioritized for physical measurements and biospecimen collections.	Diverse military cohort with strong representation of underrepresented groups.	Chinese population with both urban and rural representation.	Predominantly Dutch, with some diversity in urban areas.
International Cohorts	Conducted in Israel, expanding to Japan and UAE; additional global expansion planned.	Predominantly UK-based, with limited international participation.	Limited to the US.	Limited to the US.	Limited to China.	Limited to the Netherlands.
Cohort Size	27,916 enrolled; 13,072 completed their first visit (Dec 2024).	~500,000 participants.	More than 500,000 enrolled by 2024. target of 1 million.	~1,000,000 participants.	~500,000 participants.	~167,000 participants.
Unique Features	Deep phenotyping, including CGM with real-time food logging (2 weeks), sleep monitoring (3 nights), microbiome (gut, vaginal, oral) and immunological profiling, AI-driven predictive modeling.	Large-scale imaging (MRI), whole-genome sequencing, extensive lifestyle surveys.	Emphasis on diversity; integration of EHRs and wearables.	Largest genomic-EHR linked database, military-specific health focus.	Rural vs. urban health comparisons in China.	Environmental data, multi-generational health focus.
Data Accessibility	Immediate access for academic and commercial researchers.	Open access with application and review.	Controlled access requiring training and approval.	Controlled access for approved researchers.	Restricted access for approved researchers.	Data available for approved academic researchers.
Follow-Up Period	Planned for 25 years, with longitudinal tracking of health outcomes.	Long-term follow-up, primarily through NHS health records.	10+ years, with ongoing EHR integration and wearable data updates.	Long-term follow-up with no fixed end date.	10+ years.	Multi-generational, long-term follow-up.

Table S8. Main characteristics of the HPP compared to other selected large-scale cohort studies. EHR -Electronic Health Records, CGM -Continuous Glucose Monitoring, MRI- Magnetic Resonance Imaging).

	Ashkenaz	Middle Eastern	North African	Other	Sephardi	Yemen	Mixed
N	6270	780	943	302	264	249	2852
Age - (years, mean± SD)	51.285 ± 8.677	52.969 ± 8.16	51.386 ± 8.16	46.666 ± 12.26	54.417 ± 8.19	52.285 ± 8.04	48.253 ± 8.8
Gender - Male %	45%	48%	48%	44%	44%	45%	48%
BMI - (kg/m2, mean± SD)	26.037 ± 4.18	25.982 ± 3.96	26.543 ± 4.26	25.724 ± 4.19	27.078 ± 4.44	26.069 ± 4.09	25.616 ± 4.12

Table S9. Baseline characteristics by ethnic groups

Diet	Description and diet-aligned foods
Mediterranean ⁶⁶	<p>Dietary pattern that emphasizes the consumption of fruits, vegetables, whole grains, legumes, nuts, olive oil, fish, and lean dairy while limiting red meat, processed foods, and refined sugars. Diet-aligned foods</p> <p>Include: Coffee, Potatoes, Grilled vegetables, Tahini, Apple, Mandarin, Cappuccino, Green Tea, Cooked vegetables, Pear, Tomatoes, Hummus Salad, Rye Bread, Dark Chocolate, Onion, Cucumber, Cooked green beans, Wine, Tomatoes Salad, Sweet potato, Honey, Natural Yogurt, Quinoa, Rice, Lentils, Wholemeal Bread, Omelette, Avocado, Cashew, Pumpkin seeds, Pecan, Pea soup, Tea, Grapefruit, Banana, Beet, Sourdough Bread, Vegetable Salad, Eggplant Salad, Orange, Peas, Pomegranate juice, Mushrooms, Mango, Dates, Salmon, Almonds, Hazelnuts, Brazil nuts, Walnuts, Olive oil, Egg, Soymilk, Green cabbage, Pickled cucumber, Soy sauce, Tuna Fish, ginger, Olives, Fennel, Sauteed vegetables, Steamed vegetables, Vegetable Soup, Prune, Chestnuts, Persimmon, Orange juice, Lettuce Salad, Apricot, Barley soup, Date honey, Peanut Butter, Mulberry, Raisins, Yogurt, Polenta, Pecans, Shakshouka, Carrot Salad, Wholemeal Roll, Pepper, Grapes, Cherry tomatoes, Boiled corn, Red Cabbage Salad, Zucchini, Lettuce, Edamame, emmer wheat bread, Goat Cheese, Wholemeal Crackers, Vegetable patties, Peanuts, Nectarine, Mixed nuts, Turkey, Prickly pears, Corn, Baked apple, Carrot Juice, Guacamole, Kiwi, Pistachio, Red pepper, Peach, Watermelon, Strawberry, Tabbouleh Salad, Lentil Soup, Pomegranate, Seabass Fish, Fennel Salad, Sheep milk Labaneh, Pomelo, Sheep Milk Yoghurt, Turkey and vegetables, Mustard, Parsley, Lemon juice, Tomato paste, Cooked cabbage, Sprouts, Oatmeal, Cinnamon, Beans black-eyed peas, Carrots, Popcorn, Dried celery, Celery, Dried banana, Cauliflower, Stuffed Vine leaves, Canned corn, Garlic, Spinach, Nuts, Yellow pepper, Cocoa Powder, Coconut, Bean soup, Melon, Chia seeds, Sesame seeds, Bulgur, Sunflower seeds, Black beans, Green Salad, Garlic sauce, Matbucha, Fruit salad, Cherries, Pine nuts, Kohlrabi, Acorn squash, Lemon, Chili pepper, Hummus, Blueberries, Beet Salad, Corn Soup, Artichoke, Goat Milk Yogurt, Buckwheat crackers, Majadra, Tofu, Coconut milk, Almond Beverage, Beans, Oat Drink, Sauerkraut, Tomato sauce, Halloumi Cheese, dried tomato spread, Cooked tomatoes, Portobello mushroom, Schug, Cooked zucchini, Pickled vegetables, Greek salad, Labneh Cheese, Lentil Patties, Roasted eggplant, Green beans, Cooked Pumpkin, Berries, Sweet Potato Soup, Olive spread, Seaweeds, Radish, Broth, Green onions, Almond flour, Carob, Zucchini Patties, Water, Curry Paste, Coconut water, Green pepper, Fig, Peas and carrots, Dried pineapple, Pumpkin soup, Flax seeds, Sprouted lentils, kale, Goji berry, Loquat, Tomato soup, Alfalfa sprouts, Coriander, Turkish Salad, Mixed Vegetables, Applesauce, Dried cranberries, Almond spread, Dried fig, Date spread, Zaatar, Canned tomatoes, Basil, Wholemeal Pita, Brown Rice, Leek, Beet Juice, Olive Bread, Celery Juice, Wholemeal Light Bread, Yellow beans, Cooked chili pepper, Soy Yogurt, Dried Mango, Stuffed Chard, Chard, Mark Hariri, Matcha Powder, Arugula, Yeast, Apple Vinegar, Shiitake mushrooms, black pepper, Mushroom soup, Cooked broccoli, Cauliflower Soup, Sweetsop, Chinese orange, Guava, Chinese cabbage, Harissa salad, Vinegar, Tom Yum soup, Asparagus, Chives, Spirulina, Borscht soup, Dried tomatoes, Gil Leben, Broccoli soup, Hummus flour, Moringa Powder, Caviar, and goldenberry</p>
Paleo ⁶⁷	<p>Diet based on eating foods that mimic those available to humans during the Paleolithic era and emphasizes whole, unprocessed foods such as lean meats, fish, fruits, vegetables, nuts, and seeds while excluding grains, legumes, dairy, refined sugars, and processed foods. Diet-aligned foods</p> <p>Include: Potatoes, Grilled vegetables, Tahini, Apple, Mandarin, Cooked vegetables, Fish balls, Pear, Tomatoes, Hummus Salad, Fried eggplant, Dark Chocolate, Onion, Chicken breast, Cucumber, Cooked green beans, Beef, Tomatoes Salad, Sweet potato, Honey, Avocado, Cashew, Pumpkin seeds, Pecan, Pea soup, Tea, Grapefruit, Banana, Beet, Vegetable Salad, Eggplant Salad, Orange,</p>

	<p>Pomegranate juice, Mushrooms, Mango, Dates, Salmon, Almonds, Hazelnuts, Brazil nuts, Walnuts, Olive oil, Egg, Coleslaw, Green cabbage, Pickled cucumber, Soy sauce, Tuna Fish, Ginger, Olives, Fennel, Chicken legs, Herring, Chicken Steak, Sauteed vegetables, Steamed vegetables, Vegetable Soup, Prune, Chestnuts, Persimmon, Lettuce Salad, Apricot, Mulberry, Raisins, Pecans, Shakshouka, Carrot Salad, Pepper, Grapes, Cherry tomatoes, Boiled corn, Red Cabbage Salad, Zucchini, Lettuce, Peanuts, Mixed nuts, Turkey, Prickly pears, Corn, Baked apple, Carrot Juice, Guacamole, Tuna Salad, Kiwi, Pistachio, Red pepper, Peach, Watermelon, Strawberry, Seabass Fish, Fennel Salad, Turkey and vegetables, Mustard, Parsley, Lemon juice, Tomato paste, Cooked cabbage, Sprouts, Cinnamon, Carrots, Dried celery, Celery, Dried banana, Cauliflower, Garlic, Spinach, Nuts, Yellow pepper, Cocoa Powder, Coconut, Melon, Chia seeds, Sesame seeds, Sunflower seeds, Green Salad, Matbucha, Fruit salad, Cherries, Pine nuts, Kohlrabi, Acorn squash, Lemon, Chili pepper, Blueberries, Beet Salad, Artichoke, Buckwheat crackers, Coconut milk, Tomato sauce, Dried tomato spread, Cooked tomatoes, Portobello mushroom, Schug, Cooked zucchini, Pickled vegetables, Roasted eggplant, Green beans, Cooked Pumpkin, Berries, Sweet Potato Soup, Olive spread, Seaweeds, Radish, Broth, Green onions, Almond flour, Carob, Water, Curry Paste, Coconut water, Green pepper, Fig, Peas and carrots, Dried pineapple, Pumpkin soup, Flax seeds, kale, Goji berry, Loquat, Tomato soup, Alfalfa sprouts, Coriander, Turkish Salad, Mixed Vegetables, Applesauce, Dried cranberries, Almond spread, Dried fig, Date spread, Zaatar, Canned tomatoes, Basil, Leek, Beet Juice, Celery Juice, Cooked chili pepper, Dried Mango, Chard, Mark Hariri, Matcha Powder, Arugula, Yeast, Apple Vinegar, Shiitake mushrooms, black pepper, Mushroom soup, Cooked broccoli, Cauliflower Soup, Sweetsop, Chinese orange, Guava, Chinese cabbage, Harissa salad, Vinegar, Asparagus, Chives, Spirulina, Borscht soup, Dried tomatoes, Broccoli soup, Moringa Powder, Caviar, and goldenberry</p>
Vegetarian	<p>Dietary pattern which excludes meat, poultry, and fish but includes dairy products and eggs. It emphasizes plant-based foods such as fruits, vegetables, legumes, nuts, seeds, grains, and plant-based proteins (e.g., tofu, tempeh, and legumes). Diet-aligned foods include: Coffee, Potatoes, Grilled vegetables, Tahini, Apple, Mandarin, Cooked vegetables, Pear, Tomatoes, Hummus Salad, Fried eggplant, Dark Chocolate, Onion, Cucumber, Sherbet, Wafers, Cooked green beans, Tomatoes Salad, Sweet potato, Quinoa, Rice, Lentils, Wholemeal Bread, Avocado, Cashew, Pumpkin seeds, Pecan, Pea soup, Onion Rings, Sweet Potato Fries, Tea, Grapefruit, Banana, Beet, Vegetable Salad, Eggplant Salad, Orange, Oatmeal Cereal, Peas, Pomegranate juice, Mushrooms, Mango, Dates, Almonds, Hazelnuts, Brazil nuts, Walnuts, Olive oil, Green cabbage, Pickled cucumber, Soy sauce, Ginger, Olives, Fennel, Sauteed vegetables, Steamed vegetables, Vegetable Soup, Prune, Chestnuts, Persimmon, Lettuce Salad, Apricot, Barley soup, Date honey, Mulberry, Raisins, Rice crackers, Pecans, Israeli couscous, Jam, Carrot Salad, Wholemeal Roll, Pepper, Grapes, Cherry tomatoes, Boiled corn, Red Cabbage Salad, Zucchini, Lettuce, Edamame, emmer wheat bread, Vegetable patties, Peanuts, Mixed nuts, Vegetable Pie, Prickly pears, Corn, Baked apple, Carrot Juice, Guacamole, Kiwi, Pistachio, Red pepper, Peach, Watermelon, Strawberry, Fennel Salad, Parsley, Lemon juice, Tomato paste, Cooked cabbage, Sprouts, Oatmeal, Cinnamon, Beans black-eyed peas, Carrots, Popcorn, Dried celery, Celery, Dried banana, Cauliflower, Stuffed Vine leaves, Canned corn, Garlic, Spinach, Nuts, Yellow pepper, Cocoa Powder, Coconut, Bean soup, Melon, Chia seeds, Sesame seeds, Bulgur, Sunflower seeds, Black beans, Green Salad, Garlic sauce, Matbucha, Fruit salad, Cherries, Pine nuts, Kohlrabi, Acorn squash, Lemon, Chili pepper, Hummus, Blueberries, Carrot Cake, Beet Salad, Corn Soup, Brown Sugar, Artichoke, Apple Cake, Buckwheat crackers, Granola bar, Mushroom Pie, Majadra, Tofu, Coconut milk, Almond Beverage, Beans, Oat Drink, Tomato sauce, Fried Sweet potato, Dried tomato spread, Cooked tomatoes, Portobello mushroom, Schug, Cooked zucchini, Pickled vegetables, Lentil Patties, Roasted eggplant, Green beans, Cooked Pumpkin, Berries, Sweet Potato Soup, Olive spread, Seaweeds, Radish, Rice Noodles, Broth, Green onions, Almond flour, Carob, Zucchini Patties, Water, Curry Paste, Coconut water, Green pepper, Nut cake, Fig, Peas and carrots, Dried pineapple, Pumpkin soup, Sweet Potato Pie, Flax seeds, Granola Cookies, Sprouted lentils, kale, Goji berry, Loquat, Tomato soup, Alfalfa sprouts, Coriander, Turkish Salad, Mixed Vegetables, Applesauce, Dried cranberries, Almond spread, Dried fig, Date spread, Zaatar, Canned tomatoes, Basil, Wholemeal Pita, Brown Rice, Leek, Fruit juice, Beet Juice, Olive Bread, Sesame Snack, Celery Juice, Cooked chili pepper, Dried Mango, Stuffed Chard, Chard, Mark Hariri, Matcha Powder, Arugula, Yeast, Apple Vinegar, Shiitake mushrooms, black pepper, Avocado Sandwich, Mushroom soup, Cooked broccoli, Cauliflower Soup, Sweetsop, Chinese orange, Guava, Chinese cabbage, Harissa salad, Vinegar, Asparagus, Cauliflower Pie, Cane sugar, Chives, Spirulina, Borscht soup, Dried tomatoes, Pea protein powder, Broccoli soup, Hummus flour, Moringa Powder, Hemp Oil, Agave Syrup, Seitan, Coffee Substitute, goldenberry, Bread, Biscuit, Falafel, Cookies Oreo, Halva, Pasta with tomato sauce, Low fat Milk, Tomato soup and rice, Ciabatta, Peanut Butter, Cereals, Salty Cookies, Bamba, Bread Crumbs, Fried cauliflower, Pesto, Baklava, Tortit Bar, Vegetable patties, Mixed nuts, Cracker, Wholemeal Roll, Protein bar, Bagel, Vegan Cheese, Gin and tonic, Papaya, Brandy, Corn crackers, Fried zucchini, Cocktail, Grapefruit juice, Rice paper, Soy Pudding, Potato pancakes, Muffin, Cooked Okra, Fenugreek, Peas and carrots, Amba, Chili sauce, Dried pineapple, Doritos, Malt beverage, Sourdough Bread, Pistachio, Yellow Cheese, Ricotta Cheese, Mozzarella Cheese, Kashkaval, Cream Cheese, Greek Yogurt, Flavored Yogurt, Cottage cheese, Pancake, Granola, Egg, Natural Yogurt, Light Yellow Cheese, Low fat Milk, Yogurt, Cheese Pie, Lasagna, Pasta With Cheese, Light Cream Cheese, Cheddar, Feta Cheese, Butter Cookies, Cheesecake, Whipped Cream Pudding, Goat Cheese, Camembert or Brie, Labneh Cheese, Egg Roll, Egg Salad, Pasta, Biscuit, Croissant, Tzfatit</p>

	Cheese, Manchego Cheese, Parmesan cheese, Blue Cheese, Brie Cheese, Pecorino, Ice cream, Butter, Cream Cake, Milk, Shakshouka, Mousse Cake, Chocolate spread, Pesto, Chocolate, Chocolate Cookies, Chocolate Milk, Malabi, English cake, Sponge Cake, Honey Cake, Baklava, Poppy seed cake, Croissant, Tortit Bar, Cookies, Falafel, Mayonnaise, Rice drink, Tomato soup and rice, Omelette, Scrambled eggs, Meringue Kisses, Cappuccino, Meringue Kisses, Fruit flavored Yogurt, Fruit syrup, Pastels, Mille-feuille, Strawberry Yogurt, Granola bar, Muesli, Vegetarian Mincemeat, Tiramisu, Cheese Pastry, Rugelach, Cake, Pudding, Custard, Cream Puffs, Sweetened Condensed Milk, Sweetened Cocoa Powder, Dulce de Leche, Caramel, Whipped Cream, White Chocolate, Marshmallow, Sour Cream, Cream Cheese Spread, Cottage and topping, and Macaroni and Cheese
Vegan	A plant-based dietary approach that excludes all animal-derived products, including meat, fish, dairy, eggs, and animal byproducts like honey or gelatin. Diet-aligned foods are as specified in Vegetarian diet, <i>excluding</i> Yellow Cheese, Ricotta Cheese, Mozzarella Cheese, Kashkaval, Cream Cheese, Greek Yogurt, Flavored Yogurt, Cottage cheese, Pancake, Granola, Egg, Natural Yogurt, Light Yellow Cheese, Low fat Milk, Yogurt, Cheese Pie, Lasagna, Pasta With Cheese, Light Cream Cheese, Cheddar, Feta Cheese, Butter Cookies, Cheesecake, Whipped Cream Pudding, Goat Cheese, Camembert or Brie, Labneh Cheese, Egg Roll, Egg Salad, Pasta, Biscuit, Croissant, Tzfatit Cheese, Manchego Cheese, Parmesan cheese, Blue Cheese, Brie Cheese, Pecorino, Ice cream, Butter, Cream Cake, Milk, Shakshouka, Mousse Cake, Chocolate spread, Pesto, Chocolate, Chocolate Cookies, Chocolate Milk, Malabi, English cake, Spongecake, Honey Cake, Baklava, Poppy seed cake, Croissant, Tortit Bar, Cookies, Falafel, Mayonnaise, Rice drink, Tomato soup and rice, Omelette, Scrambled eggs, Meringue Kisses, Cappuccino, Meringue Kisses, Fruit flavored Yogurt, Fruit syrup, Pastels, Mille-feuille, Strawberry Yogurt, Granola bar, Muesli, Vegetarian Mincemeat, Tiramisu, Cheese Pastry, Rugelach, Cake, Pudding, Custard, Cream Puffs, Sweetened Condensed Milk, Sweetened Cocoa Powder, Dulce de Leche, Caramel, Whipped Cream, White Chocolate, Marshmallow, Sour Cream, Cream Cheese Spread, Cottage and topping, and Macaroni and Cheese

Table S10. Overview of various diets and their compatible foods

Package Name	Version
Python	3.9.13
Babel	2.13.1
biopython	1.81
geopandas	1.0.1
idna	3.40
iglu-py	1.1.0
lifelines	0.30.0
lightgbm	4.5.0
llvmlite	0.43.0
matplotlib-inline	0.1.6
matplotlib	3.8.2
mistune	3.0.2
numpy	1.26.2
pandas-flavor	0.6.0
pandas	2.2.2
pingouin	0.5.5
python-dateutil	2.8.2

python-json-logger	2.0.7
python-utils	3.9.1
pytz	2023.3.post1
rpds-py	0.12.0
rpy2	3.5.16
scikit-learn	1.6.1
scipy	1.11.4
seaborn	0.11.2
shap	0.46.0
shapely	2.0.6
six	1.16.0
slicer	0.0.8
sniffio	1.3.0
soupsieve	2.50
stack-data	0.6.3
statannot	0.2.3
statannotations	0.6.0
statsmodels	0.14.3
sympy	1.12
tables	3.9.2
tabulate	0.9.0
torch	2.1.1
torchaudio	2.1.1
torchvision	0.16.1
tornado	6.3.3

Table S11. The software packages and their corresponding versions used in the analysis and code implementation for this manuscript. All computations were performed using Python 3.9.13.