Supplementary Information, Data S1

Supplementary Notes

Note S1 Physical indicators

The following 18 blood serum indicators, 24 blood cell-related indicators and 7 body indexes were collected based on routine physical examination.

The 18 blood serum indicators include alanine transaminase (ALT, U/L), aspartate transaminase (AST, U/L), total bilirubin (TBIL, μmol/L), direct bilirubin (DBIL, μmol/L), indirect bilirubin (IBIL, μmol/L), gamma glutamyl transpeptidase (GGT, U/L), alkaline phosphatase (ALP, U/L), total protein (TP, g/L), albumin (ALB, g/L), globulin (GLB, g/L), ratio of albumin to globulin (A/G), blood urea nitrogen (Urea, mmol/L), creatinine (CREA, μmol/L), uric acid (UA, μmol/L), total cholesterol (CHO, mmol/L), triglyceride (TG, mmol/L), high density lipoprotein cholesterol (HDL-C, mmol/L) and low density lipoprotein cholesterol (LDL-C, mmol/L).

The 24 blood cell indicators include WBC (white blood cell, 10^9/L), RBC (red blood cell, 10^12/L), HGB (hemoglobin, g/L), HCT (hematocrit, %), MCV (mean corpuscular volume, fL), MCH (mean corpuscular hemoglobin, pg), MCHC (mean corpuscular hemoglobin concentration, g/L), PLT (platelet, 10^9/L), LYM% (lymphocyte rate, %), MONO% (monocyte rate, %), NEUT% (neutrophil ratio, %), EO% (eosinophil ratio, %), BA% (basophilic granulocyte ratio, %), LYM# (lymphocyte number, 10^9/L), MONO# (monocyte count, 10^9/L), NEUT# (neutrophil number, 10^9/L), EO# (eosinophil number, 10^9/L), BA# (basophilic granulocyte number, 10^9/L), RDW-CV (red blood cell distribution width coefficient of variation, %), RDW-SD (red blood cell distribution width standard deviation, fL), PDW (platelet distribution width, fL), MPV (mean platelet volume, fL), P-LCR (platelet-large cell ratio, %) and PCT (plateletcrit, %).

The 7 body indexes include body height (centimeter), body weight (kg),

systolic blood pressure (mmHg), diastolic blood pressure (mmHg), systolic diastolic difference (mmHg), pulse (beats/minute) and BMI (body mass index, kg/m^2).

Note S2 Definition of facial morphological features

The 17 landmarks are right eye outer corner, right eye inner corner, left eye inner corner, left eye outer corner, right lip corner, left lip corner, nose tip, nasion, right alare, left alare, lip center, upper lip, lower lip, subnasale, pogonion, right earlobe tip and left earlobe tip. The following morphological features were defined based on these landmarks.

Distance between the nose and mouth: distance between landmark q and s;

Mouth width: the distance between landmark v and w;

Upper lip thickness: the distance between landmark s and r.

Lower lip thickness: the distance between landmark r and t.

Nose width: the distance between landmark m and n.

Interocular (endocanthions) distance: the distance between right and left eye inner corner (landmark i and j).

External canthi width: the distance between right and left eye outer corner (landmark k and l).

Left eye corner slope: first the line from the inner to the outer corner of the left eye was used to describe the direction of the left eye. Then the intersection angle of this line with the horizontal plane was used to represent the left eye corner slope of each sample.

Right eye corner slope: similar to left eye corner slope, except a minus sign to the slope was added to make the slopes comparable between left and right eyes.

Asymmetry of two eye corner slopes: the absolute difference between the left and right eye corner slopes.

Note S3 Age prediction analysis

The efficiency of age prediction was evaluated by the mean absolute deviation and can be expressed as:

$$\text{MAD} = \frac{\sum_i |Y_i - Y_i'|}{N}$$

where N is the number of samples, Y is the chronological age and Y' is the predicted age.

Note S4 Clustering analysis of various features

Clustering analysis was based on 10 quantified facial features, 18 blood serum indicators, 24 blood cell indicators, 7 body indexes and the first 3 components from PLS. All of the samples were first sorted according to their ages from young to old. Then the median of each sliding window with a step size of 5 samples was used to reduce the influence of outliers. Finally, all features were z-score normalized within each feature before clustered by the BIC-SK algorithm.