Appendix S1

Antropometric measurements

All clinical health examinations took place and all measurements were taken at the age of 46 years. The examination included body weight in light clothing which was measured with a digital scale (calibrated regularly). Height was measured twice using a standard and calibrated stadiometer and the mean of the two measurements was used. Body mass index (BMI) was calculated as the ratio of weight to height squared (kg/m^2). Waist and hip circumferences were measured twice (mean of the two measurements was used) and the waist–hip ratio was assessed as the ratio between circumferences of the waist (at the level midway between lowest rib margin and the iliac crest) and the hip (at the widest trochanters). Body fat mass, fat percentage, muscle mass and visceral fat area were measured by InBody 720 bioelectrical impedance analyser (Biospace Co.. Ltd. Seoul. Korea). All measurements were done after overnight (12 h) fasting period.

Cardiovascular measurements

Systolic and diastolic blood pressure was measured three times with 1 min interval after 15 min of rest on the right arm of the seated participants using an automated oscillometric blood pressure device and appropriately sized cuff (Omron Digital Automatic Blood Pressure Monitor Model M10-IT. Japan). Finally, the mean of two lowest systolic values and their diastolic values was used in the analyses. The previously diagnosed hypertension was defined according to self-reported diagnoses and medications, hospital outpatient and inpatient registers and medication registers from the Social Insurance Institution of Finland.

Oral glucose tolerance test and diabetes

Venous blood samples were used to determine fasting plasma glucose and HbA1c –values of every eligible participant. Blood samples were taken at the cohort laboratory between 7:00 and 11:00 a.m. after an over-night (12h) fast. Cases whose fasting plasma glucose was <8.0 mmol/L and had no previous diagnosis of diabetes underwent an OGTT with a 75 g glucose load, after which a 2-hour plasma sample was collected. Both serum insulin and plasma glucose were measured at baseline and 30. 60 and 120 min after 75 g glucose intake. Glucose tolerance status was classified according to WHO criteria: normal glucose tolerance (NGT) was defined as having fasting plasma glucose (FPG) level <6.1 mmol/l and 2-h glucose level <7.8 mmol/l. impaired fasting glucose (IFG) as having FPG level 6.1–6.9 mmol/l and 2-h glucose level <7.8 mmol/l. impaired glucose tolerance (IGT) as having FPG level <7.0 mmol/l and 2-h glucose level 7.8–11.0 mmol/l and screen detected diabetes mellitus (scDM) as having FPG level ≥7.0 mmol/l and/or 2-h glucose level ≥11.1 mmol/l. Fasting glucose and insulin values were used to calculate fasting indices: HOMA-IR index (Homeostasis Model Assessment—insulin resistance) (FPG × FSI / 22.5). HOMA2-β index (Homeostasis Model Assessment—beta-cell function) ((20 × FSI)/(FPG – 3.5) × 100).

Previously diagnosed diabetes was defined according to self-reported diagnoses and medications. hospital outpatient and inpatient registers and medication registers from the Social Insurance Institution of Finland. The concentration of HbA1c and the concentration of total haemoglobin were measured by immunochemical assay method. Glucose were analysed using an enzymatic hexokinase/glucose-6-phosphate dehydrogenase method. (both method: Advia 1800; Siemens Healthcare Diagnostics Inc.. Tarrytown. Ny. USA). The samples were analysed in NordLab Oulu. a testing laboratory (T113) accredited by Finnish Accreditation Service (FINAS) (EN ISO 15189).

Other biochemical measurements

Blood samples were taken after an overnight fasting period, centrifuged immediately and stored firstly at −20°C and later at −80°C. All blood samples were analysed in the laboratory of the University Hospital of Oulu according to a standardized protocol. Serum total cholesterol, high-density lipoprotein (HDL) and low-density lipoprotein (LDL). and triglycerides were determined using an enzymatic assay method (Advia 1800; Siemens Healthcare Diagnostics Inc.. Tarrytown. NY. USA). Serum samples for testosterone (T) were analysed using Agilent triple quadrupole 6410 LC/MS equipped with electrospray ionisation source operating with positive-ion mode (Agilent Technologies. Wilmington. DE. USA). Multiple reaction monitoring was used to quantify testosterone by d3-testosterone with the following transitions: m/z 289.2 to 97 and 289.2 to 109 for testosterone and 292.2 to 97 and 292.2 to 109 for d3-testosterone. The intra-assay CVs of the method were 5.3. 1.6 and 1.2% for testosterone at 0.6. 6.6 and 27.7 nmol/l. respectively. The interassay CVs were 5.3. 4.2 and 1.0% for the respective concentrations. Serum sex hormone-binding globulin (SHBG) was analysed by chemiluminometric immunoassay (Immulite 2000. Siemens Healthcare Diagnostica Inc.. Llanberis. UK). Free androgen index (FAI) was determined (T/SGBG). High sensitivity C-reactive protein (hs-CRP) was analysed by an immune nefelometric assay (BN ProSpec. Siemens Healthcare Diagnostics Inc.. Newark. DE. USA).

CVD risk assessment scoring

Three CVD risk assessment tools were used. Framingham Risk Score [2], CVD SCORE [3] and [4, 5]FINRISK. The Framingham Risk Score estimates a 10-year risk of developing coronary heart disease, cerebrovascular events, peripheral artery disease or heart failure. It bases its risk-percentage result on the following factors: gender, age, smoking, total cholesterol, HDL-cholesterol, systolic blood pressure, requiring treatment for raised blood pressure, and diabetes. As a result the Framingham risk assessment tool reports points ranging from ≤−2 to ≥21 that refer to a risk percentage ranging from <1 to >30%[2] . The CVD SCORE estimates a 10-year risk of fatal CVD on the basis of gender, age, smoking, total cholesterol and systolic blood pressure. It results risk percentages ranging from <1 to ≥15%[3]. FINRISK estimates the 10-year risk of cardiovascular diseases and includes the diabetes in model[5].

Fatty liver index

Fatty liver disease is strongly associated to obesity and it can be predicted using a FLI algorithm, which is based on BMI, central obesity measured by waist circumference, triglyceride and gamma-glutamyl-transferase (GGT) levels [6]. FLI is calculated as follows: FLI = (e0.953 * loge (triglycerides) + 0.139 * BMI + 0.718 * loge (ggt) + 0.053 * waist circumference − 15.745) / (1 + e0.953 * loge (triglycerides) + 0.139 * BMI + 0.718 * loge (ggt) + 0.053 * waist circumference − 15.745) * 100. FLI varies between 0 and 100, with cut offs at 30 and 60; score <30 rules out fatty liver and a score ≥60 is considered a strong predictor for fatty liver.

References

1. Shanewise. J. S. *et al*. ASE/SCA guidelines for performing a comprehensive intraoperative multiplane transesophageal echocardiography examination: recommendations of the American Society of Echocardiography Council for Intraoperative Echocardiography and the Society of Cardiovascular Anesthesiologists Task Force for Certification in Perioperative Transesophageal Echocardiography. *J. Am. Soc. Echocardiogr.* **12**. 884-900 (1999).

2. Parikh. N. I. *et al*. A risk score for predicting near-term incidence of hypertension: the Framingham Heart Study. *Ann. Intern. Med.* **148**. 102-110 (2008).

3. Conroy. R. *et al*. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur. Heart J.* **24**. 987-1003 (2003).

4. Bhopal. R. *et al*. Predicted and observed cardiovascular disease in South Asians: application of FINRISK. Framingham and SCORE models to Newcastle Heart Project data. *Journal of public health* **27**. 93-100 (2005).

5. Vartiainen. E.. Laatikainen. T.. Peltonen. M. & Puska. P. Predicting coronary heart disease and stroke: The FINRISK Calculator. *Global heart* **11**. 213-216 (2016).

6. Bedogni. G. *et al*. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC gastroenterology* **6**. 33 (2006).