**Supplementary Materials**

**Circulating Biomarker Score for Visceral Fat and Risks of Incident Colorectal and Postmenopausal Breast Cancer: The Multiethnic Cohort Adiposity Phenotype Study**

Loïc Le Marchand, Lynne R. Wilkens, Ann M. Castelfranco, Kristine R. Monroe, Bruce S. Kristal, Iona Cheng, Gertraud Maskarinec, Meredith A. Hullar, Johanna W. Lampe, John A. Shepherd, Adrian Franke, Thomas Ernst, Unhee Lim

**Supplementary Methods:**

*Anthropometric and Body Composition Assessment for the MEC Adiposity Phenotype Study (APS)*

Trained technicians obtained measurements of height, weight and circumferences of the waist and hip. Body composition was determined by whole-body DXA using a fan-beam densitometer [Hologic Discovery A densitometer (Bedford, MA) at UH and USC]. Fat mass and lean mass were estimated for the whole body, trunk, arms and legs. An abdominal MRI scan was acquired on 3-T scanners [TIM Trio [Erlangen, Germany] at University of Hawaii (UH); General Electric HDx [Milwaukee, WI] at University of Southern California (USC)] to quantify visceral and subcutaneous fat areas (cm2) at 4 abdominal intervertebral sections (L1–L2, L2–L3, L3–L4, L4–L5). Visceral and subcutaneous fat measures were obtained from an axial gradient-echo sequence with water suppression with breath holds (1). The processing of participant MRI scan data was performed centrally at UH with a standardized protocol. The visceral fat area measures were highly correlated across intervertebral segments. Therefore, the average VAT area across segments was used as the primary VAT measure. Sixty of the 1,861 participants were excluded for invalid MRI imaging (due to implants, motion artefacts or presence of a visceral mass). MRI measures were calibrated for scanner differences between the two study sites based on 15 healthy volunteers (BMI 21.8–39.6 kg/m2), who were scanned at both sites within one week during the course of the study. The repeat measures were highly correlated (*r*> 0.99). The USC values were calibrated to the UH values to correct for a small systematic difference (1).

*Obesity-related blood biomarkers*

Circulating biomarkers selected for their reported associations to obesity-caused metabolic, hormonal and inflammation dysfunctions or their being lipid-soluble micronutrients were measured in plasma or serum centrally at the UH Cancer Center’s Analytical Biochemistry Shared Resource (ABSR) laboratory. The list of biomarkers assessed on the APS participants for potential inclusion in the VAT score is presented in **Supplementary Table 1**, along with assay and reproducibility information. A Cobas MiraPlus chemistry autoanalyzer (Roche, Indianapolis, IN) was used to measure concentrations of serum total and HDL-cholesterol, glucose, high-sensitivity C-reactive protein (CRP), triglycerides, and alanine aminotransferase (ALT). Enzyme-linked immunosorbent assay (ELISA) was used to measure insulin and sex hormone binding globulin (SHBG). LDL cholesterol was derived from the Friedewald equation for TG < 400 mg/dL (2) and based on formulas in Dansethakul et al (3) for TG ≥ 400 mg/dL. Serum concentrations of lipid soluble micronutrients were determined by high pressure liquid chromatography (HPLC) with photo-diode array detection (4,5). Analyte concentrations were determined by peak areas obtained from monitoring at 450 nm for carotenoids, at 325 nm for retinol and at 295 nm for tocopherols and CoQ10, by using external calibration curves obtained from authentic standards, and by adjusting for recovery of the appropriate internal standards. Unconjugated estradiol, unconjugated estrone, and unconjugated testosterone were analyzed according to a sensitive LCMS assay (6). Free steroids were calculated by the Vermeulen equation that uses the levels of the unconjugated steroids, measured SHBG levels and assumed constant albumin levels (7,8). Commercial ELISA assays (**Supplementary Table 1**) were used following the manufacturer protocol for SHBG, lipopolysaccharide binding protein (LBP), leptin, adiponectin, insulin-like growth factor-I (IGF-I), insulin-like growth factor binding protein 2 (IGFBP2), and insulin-like growth factor binding protein 3 (IGFBP3). The APS samples were arranged randomly so that each batch was reasonably balanced by sex and race. Approximately 10% blind quality control duplicate pairs were added to each batch to assess reproducibility (**Supplementary Table 1**). For 500 APS participants, randomly selected within sex, race/ethnicity and BMI strata, samples from the MEC Biospecimen repository (established an average of 10.9 years earlier than APS) were assessed for the same analyte panel in order to evaluate the prospective association with VAT of the biomarkers score.

For cost efficiency, we conducted the laboratory assays in three stages by retaining only a subgroup of the biomarkers for the subsequent stage based on their predictions of liver fat or visceral fat. The clustering of biomarkers within panels was also a consideration for inclusion. In particular, an entire panel of biomarkers was retained, such as lipid soluble micronutrients, when there was strong evidence of important biomarkers within the cluster. This stepwise approach began with the assay of 133 markers in the first subset of 30 0 sex/ethnicity-balanced APS samples, followed by the assay of 78 markers in the next 700 samples and of 39 markers in the remaining samples, which provided data on 39 biomarkers in all participants, 39 additional biomarkers in 1000 participants and 36 additional biomarkers in 300 participants. We used random forest (R, randomForest) for this prioritization in order to take advantage of this method’s robust handling of different scales and outliers (9). We observed fairly consistent rankings of the biomarkers, e.g., for increase in mean squared error given random imputation of the specific biomarker, across the assay stages (See **Supplementary Table 1 for the associations with visceral fat**).

The samples for cases and controls in the nested studies of breast and colorectal cancer were analyzed by the ABSR for the final APS analyte panel using the same methods. Batches included case-control pairs and QC duplicates, and laboratory staff were blinded to the status of the samples. Case-control pairs were shuffled to provide similar sex and race representation across batches.

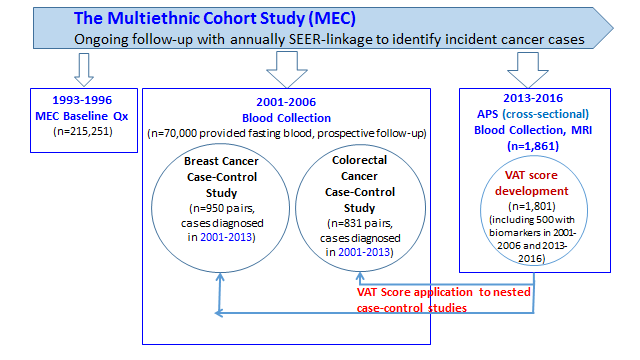
**References**

1. Lim U, Monroe KR, Buchthal S, et al. Propensity for Intra-abdominal and Hepatic Adiposity Varies Among Ethnic Groups. Gastroenterology. 2019 Mar;156(4):966-975.
2. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem. 1972;18(6):499-502.
3. Dansethakul P, Thapanathamchai L, Saichanma S, Worachartcheewan A, Pidetcha1 P. Determining A New Formula For Calculating Low-Density Lipoprotein Cholesterol: Data Mining Approach.

EXCLI Journal 2015; 14:478-483; March 26, 2015.

1. Franke AA, Custer LJ, Cooney RV. Synthetic carotenoids as internal standards for plasma micronutrient analysis by high-performance liquid chromatography. J Chromatogr B 1993;614(1):43-57.
2. Franke AA, Custer LJ, Morrison CM, Li X, Lai JF. Simultaneous analysis of circulating 25-hydroxy-vitamin D3, 25-hydroxy-vitamin D2, retinol, tocopherols, carotenoids, and oxidized and reduced coenzyme Q10 by HPLC with photo diode-array detection using C18 and C30 columns alone or in combination. J Chromatogr A 2013;1301:1-9.
3. Li X, Franke AA. Improved profiling of estrogen metabolites by orbitrap LC/MS. Steroids 2015;99:84-90.
4. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. J Clin Endocrinol Metab 1999;84:3666-3672.
5. Rinaldi S, Geay A, Dechaud H, Biessy C, Zeleniuch-Jacquotte A, Akhmedkhanov A, et al. Validity of free testosterone and free estradiol determinations in serum samples from postmenopausal women by theoretical calculations. Cancer Epidemiol Biomarkers Prev 2002;11:1065-1071.
6. Breiman L. Random Forests. Machine Learning 2001;45:5-32.

**Supplementary Figure 1**: Flowchart of the study design



**Supplementary Table 1.** List of measured and derived biomarkers

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Biomarkers** | **Specimen** | **Assay** | **Assay Specifics** | Random Forest Importance in Assay Stages | | | **Total QC** | |
| Stage 1 (n=266) | Stage 1+2 (n=896) | Included in Final Stage (n=1861) |
| **% ΔMSE**b | **% ΔMSE**b |  | **%CV**a | **%ICC**a |
| **Alanine transaminase** (ALT), U/L | serum | Cobas | Randox Laboratories, AL1205 | 1.71E-03 | 1.32E-0­3 | X | 4.4% | 82% |
| **C-reactive protein** (CRP), mg/L | serum | Cobas | Core Lab Supplies, C7568 | 1.00E-03 | 1.69E-03 | X | 13.0% | 88% |
| **Cholesterol total**, mg/dL | serum | Cobas | Pointe Scientific, H7510 | -1.14E-04 | 5.21E-04 | X | 1.1% | 64% |
| **Cholesterol high-density lipoprotein** (HDLC), mg/dL | serum | Cobas | Pointe Scientific, H7545 | 1.87E-03 | 2.47E-0­3 | X | 1.7% | 78% |
| **Cholesterol low-density lipoprotein** (LDLC), mg/dL | serum | Derived | from cholesterol, HDLC and TG | 4.60E-04 | 3.05E-0­5 | X | 2.2% | 62% |
| **Glucose**, mg/dL | serum | Cobas | Randox Laboratories, GL1611 | 1.78E-03 | 1.64E-0­3 | X | 0.9% | 86% |
| Hemoglobin A1c (HbA1c), % | RBC | Cobas |  | 1.24E-04 | N/A |  | 1.8% | 99% |
| **Triglycerides** (TG), mg/dL | serum | Cobas | Pointe Scientific, T7532 | 8.25E-03 | 6.10E-03 | X | 1.8% | 92% |
| Cytokines 10-plex, pg/mL | serum | Luminex |  |  |  |  |  |  |
| Granulocyte macrophage colony stimulating factor (GM-CSF) |  |  |  | 3.15E-04 | -9.68E-0­5 |  | 33.0% | 87% |
| Interferon gamma (IFNg) |  |  |  | 8.01E-04 | 6.08E-0­4 |  | 30.3% | 73% |
| Interleukin-1 beta (IL1b) |  |  |  | 6.46E-05 | 3.83E-04 |  | 9.4% | 16% |
| Interleukin-2 (IL2) |  |  |  | -8.56E-05 | 1.52E-04 |  | 20.9% | 81% |
| Interleukin-4 (IL4) |  |  |  | 6.42E-05 | 1.46E-05 |  | 38.1% | 78% |
| Interleukin-5 (IL5) |  |  |  | 2.38E-05 | 3.05E-0­4 |  | 22.4% | 87% |
| Interleukin-6 (IL6) |  |  |  | -7.81E-05 | 2.44E-0­4 |  | 12.0% | 87% |
| Interleukin-8 (IL8) |  |  |  | 2.18E-04 | 1.62E-0­4 |  | 6.1% | 69% |
| Interleukin-10 (IL10) |  |  |  | -4.22E-04 | 5.55E-0­4 |  | 10.8% | 91% |
| Tumor necrosis factor alpha (TNFa) |  |  |  | 2.05E-04 | 6.98E-0­4 |  | 34.2% | 85% |
| **Adiponectin total**, ng/mL | serum | ELISA | R&D, DRP300 | 1.70E-02 | 1.57E-0­2 | X | 0.7% | 98% |
| Complement component 3 (C3), mcg/mL | serum | ELISA |  | -5.61E-04 | N/A |  | 2.1% | 84% |
| Cortisol, ng/mL | serum | ELISA |  | -7.26E-05 | N/A |  | 2.0% | 94% |
| Fibroblast growth factor 21 (FGF21), ng/mL | plasma | ELISA |  | 2.88E-05 | N/A |  | 4.7% | 99% |
| Ghrelin, pg/mL | serum | ELISA |  | 2.64E-04 | N/A |  | 5.9% | 99% |
| **Insulin**, microU/mL | serum | ELISA | EMD Millipore, EZH1-14K | 6.82E-02 | 7.17E-0­2 | X | 1.8% | 95% |
| **Insulin-like growth factor-1** (IGF1), ng/mL | serum | ELISA | R&D, DG100 | -3.21E-04 | -3.43E-05 | X | 1.4% | 89% |
| **Insulin-like growth factor binding protein-1** (IGFBP1), ng/mL | serum | ELISA | R&D, DGB200 | 6.76E-03 | 1.83E-03 |  | 6.4% | 86% |
| **Insulin-like growth factor binding protein-2** (IGFBP2), ng/mL | serum | ELISA | R&D, DGB200 | 4.37E-03 | 6.80E-03 | X | 1.2% | 97% |
| **Insulin-like growth factor binding protein-3** (IGFBP3), ng/mL | serum | ELISA | R&D, DGB300 | 3.79E-04 | 4.18E-04 |  | 0.7% | 99% |
| **Leptin** , ng/mL | serum | ELISA | R&D, DLP00 | -4.37E-04 | 1.29E-02 | X | 2.5% | 94% |
| Leptin receptor soluble (sLEPR), ng/mL |  |  |  | 8.49E-05 | N/A |  | 1.7% | 94% |
| **Lipopolysaccharide binding protein** (LBP), ng/mL | plasma | ELISA | Cell Sciences, CKH113 | 9.62E-04 | 1.13E-0­3 |  | 0.7% | 80% |
| Omentin 1, ng/mL | plasma | ELISA |  | 9.52E-05 | N/A |  | 7.7% | 61% |
| Plasminogen activator inhibitor-1 (PAI1), ng/mL | plasma | ELISA |  | 4.36E-04 | N/A |  | 3.0% | 99% |
| **Sex hormone binding globulin** (SHBG), nmol/L | serum | ELISA | R&D, DSHBG0B | 7.72E-03 | 1.07E-0­2 |  | 1.9% | 93% |
| Uric acid, micromol/L | serum | ELISA |  | 4.43E-03 | N/A |  | 0.4% | 99% |
| **Lipid-soluble micronutrients**, ng/mL | serum | HPLC [7] |  |  |  |  |  |  |
| **Carotene cis-beta** | serum | HPLC |  | 9.43E-04 | 1.43E-03 | X | 5.2% | 57% |
| **Carotene trans-alpha** | serum | HPLC |  | 2.77E-03 | 2.25E-03 | X | 2.0% | 97% |
| **Carotene trans-beta** | serum | HPLC |  | 4.01E-02 | 9.90E-0­3 | X | 1.3% | 98% |
| **Cryptoxanthin cis-beta** | serum | HPLC |  | 1.24E-03 | 1.07E-03 | X | 2.0% | 92% |
| **Cryptoxanthin trans-alpha** | serum | HPLC |  | 3.45E-03 | 2.30E-03 | X | 2.0% | 90% |
| **Cryptoxanthin trans-beta** | serum | HPLC |  | 1.68E-03 | 7.92E-04 | X | 1.1% | 99% |
| **Lutein cis1** | serum | HPLC |  | 5.66E-04 | 4.90E-04 | X | 1.6% | 94% |
| **Lutein cis2** | serum | HPLC |  | -3.51E-04 | 6.85E-04 | X | 4.1% | 75% |
| **Lutein cis-anhydro** | serum | HPLC |  | 9.57E-05 | 2.43E-04 | X | 1.5% | 94% |
| **Lutein trans** | serum | HPLC |  | 2.90E-04 | 1.46E-03 | X | 1.4% | 95% |
| **Lutein trans-anhydro** | serum | HPLC |  | 1.73E-03 | 8.70E-03 | X | 1.3% | 96% |
| **Lycopene dihydro** | serum | HPLC |  | -6.68E-05 | 1.05E-03 | X | 2.2% | 79% |
| **Lycopene total** | serum | HPLC |  | 3.19E-04 | 4.27E-04 | X | 1.1% | 89% |
| **Retinol** | serum | HPLC |  | -3.37E-04 | 8.02E-0­4 | X | 1.2% | 89% |
| **Tocopherol alpha** | serum | HPLC |  | 1.03E-03 | 3.32E-0­4 | X | 0.6% | 89% |
| **Tocopherol beta+gamma** | serum | HPLC |  | 2.14E-04 | 4.20E-04 | X | 0.7% | 95% |
| **Tocopherol delta** | serum | HPLC |  | -4.40E-05 | 2.17E-04 | X | 1.0% | 83% |
| **Tocopherol total** | serum | HPLC |  | 3.78E-04 | 3.42E-04 | X | 1.1% | 12% |
| **Ubiquinone (oxidized Coenzyme Q10)** | serum | HPLC |  | 8.75E-04 | 3.65E-0­4 |  | 3.1% | 76% |
| **Ubiquinol (reduced Coenzyme Q10)** | serum | HPLC |  | -6.01E-04 | 5.87E-0­5 |  | 8.8% | 81% |
| **Vitamin D 25-hydroxyvitamin D3** | serum | HPLC |  | -7.47E-04 | 3.11E-0­4 |  | 2.9% | 83% |
| **Zeaxanthin** | serum | HPLC |  | 8.52E-04 | 1.59E-03 |  | 2.1% | 93% |
| Alkylresorcinol, ng/mL | plasma | LC/MS |  |  |  |  |  |  |
| Alkylresorcinol C17 | plasma | LC/MS |  | 6.94E-05 | N/A |  | 49.5% | 51% |
| Alkylresorcinol C19 | plasma | LC/MS |  | -6.61E-04 | N/A |  | 24.4% | 68% |
| Alkylresorcinol C21 | plasma | LC/MS |  | -2.71E-04 | N/A |  | 13.3% | 83% |
| Alkylresorcinol C23 | plasma | LC/MS |  | 1.52E-04 | N/A |  | 36.6% | 60% |
| Alkylresorcinol C25 | plasma | LC/MS |  | -1.30E-04 | N/A |  | 57.1% | 63% |
| Bile acids, ng/mL | serum | LC/MS |  |  |  |  |  |  |
| Cholic acid (CA) | serum | LC/MS |  | 8.31E-05 | N/A |  | 3.7% | 87% |
| Chenodeoxycholic acid (CDCA) | serum | LC/MS |  | -4.49E-04 | N/A |  | 5.5% | 72% |
| Deoxycholic acid (DCA) | serum | LC/MS |  | -1.60E-05 | N/A |  | 3.1% | 78% |
| Hyodeoxycholic acid (HDCA) | serum | LC/MS |  | -3.96E-04 | N/A |  | 10.1% | 63% |
| Lithocholic acid (LCA) | serum | LC/MS |  | 1.17E-04 | N/A |  | 35.3% | 68% |
| Ursodeoxycholic acid (UDCA) | serum | LC/MS |  | 1.33E-04 | N/A |  | 9.6% | 83% |
| Fatty acids, ng/mg | RBC | LC/MS |  |  |  |  |  |  |
| Arachidonic acid | RBC | LC/MS |  | -8.18E-06 | N/A |  | 1.8% | 63% |
| Docosahexaenoic acid (DHA) | RBC | LC/MS |  | -3.25E-04 | N/A |  | 2.4% | 76% |
| Docosapentaenoic acid (DPA) | RBC | LC/MS |  | 4.19E-04 | N/A |  | 3.2% | 73% |
| Eicosadienoic acid (EDA) | RBC | LC/MS |  | -4.64E-04 | N/A |  | 5.6% | 78% |
| Eicosapentaenoic acid (EPA) | RBC | LC/MS |  | -1.38E-04 | N/A |  | 3.1% | 87% |
| Linoleic acid | RBC | LC/MS |  | 3.96E-04 | N/A |  | 5.4% | 78% |
| Linolenic acid alpha | RBC | LC/MS |  | 2.03E-05 | N/A |  | 1.8% | 56% |
| Linolenic acid gamma | RBC | LC/MS |  | -1.95E-04 | N/A |  | 6.2% | 86% |
| Myristic acid | RBC | LC/MS |  | -2.79E-04 | N/A |  | 3.6% | 72% |
| Oleic acid | RBC | LC/MS |  | 2.88E-04 | N/A |  | 1.8% | 58% |
| Palmitic acid | RBC | LC/MS |  | 7.48E-04 | N/A |  | 1.7% | 57% |
| Palmitoleic acid | RBC | LC/MS |  | 3.98E-04 | N/A |  | 3.5% | 88% |
| Stearic acid | RBC | LC/MS |  | 2.38E-04 | N/A |  | 1.7% | 61% |
| Isoprostanes, pg/mL |  |  |  |  |  |  |  |  |
| 8-isoprostane | plasma | LC/MS |  | -2.58E-04 | N/A |  | 10.8% | 63% |
| Prostaglandin F2-alpha | plasma | LC/MS |  | -1.37E-04 | N/A |  | 6.0% | 80% |
| Sex steroid hormones, pg/mL | serum | LC/MS [8,9] |  |  |  |  |  |  |
| Estradiol total | serum | LC/MS |  | 4.85E-03 | 3.91E-03 | X | 18.2% | 92% |
| **Estradiol unconjugated** | serum | LC/MS |  | 1.72E-03 | 3.21E-03 | X | 19.9% | 94% |
| **Estradiol free** | serum | Derived | from unconjugated estradiol & SHBG [10] | 6.11E-03 | 1.06E-03 | X | 24.4% | 90% |
| Estrone total | serum | LC/MS |  | 5.93E-04 | 2.98E-04 | X | 10.2% | 73% |
| **Estrone unconjugated** | serum | LC/MS |  | -1.55E-04 | 3.72E-0­5 | X | 9.3% | 87% |
| **Estrone free** | serum | Derived | from unconjugated estrone & SHBG | 3.41E-04 | 6.65E-04 | X | 15.7% | 85% |
| Testosterone total | serum | LC/MS [9] |  | 5.64E-03 | 9.05E-0­3 | X | 3.8% | 88% |
| **Testosterone unconjugated** | serum | LC/MS |  | 2.75E-03 | 1.45E-02 | X | 2.1% | 97% |
| **Testosterone free** | serum | Derived | from unconjugated testosterone & SHBG | 1.91E-03 | 2.13E-02 | X | 6.1% | 99% |
| Sugars, mg/dL |  |  |  |  |  |  |  |  |
| Fructose | serum | LC/MS |  | 4.57E-05 | N/A |  | 23.0% | 83% |
| Glucose | serum | LC/MS |  | 1.52E-04 | N/A |  | 1.1% | 99% |
| Trimethylamine N-oxide metabolites, micromol/L |  |  |  |  |  |  |  |  |
| Betaine | plasma | LC/MS |  | -2.83E-04 | 1.94E-04 |  | 1.7% | 89% |
| Choline | plasma | LC/MS |  | 1.11E-03 | 5.70E-04 |  | 2.6% | 87% |
| Trimethylamine N-oxide (TMAO) | plasma | LC/MS |  | 2.30E-04 | 1.18E-03 |  | 4.5% | 91%F |

CV (coefficient of variation), ELISA (enzyme-linked immunosorbent assay), HPLC (high-pressure liquid chromatography), ICC (intra-class correlation coefficient), LC/MS (liquid chromatography/mass spectrometry), MSE (mean squared error), RBC (red blood cell).

a Approximately 10% blinded QC samples were inserted among study samples at random locations and analyzed to obtain CV (from all QC replicates) and ICC (from QC replicates included in multiple batches).

b Biomarkers were filtered in 3 successvie stages for cost efficiency. In the first stage, 300 APS samples were analyzed for 133 candidate biomarkers (and 6 additional derived variables) of body fat distribution in the literature. Based on their performance to predict visceral fat and liver fat in Random Forest models, a subgroup of the biomarkers (78 out of 133) were selected and analyzed in 700 APS samples in the second stage. Percent increase in mean squared error (% ΔMSE) without the given biomarker (e.g., larger, positive values indicate higher prediction ability) was used as the selection criterion. Random Forest modeling was similarly repeated after the second stage, and 39 markers (and 5 additional derived markers), listed in bold, were selected and analyzed in the remaining 860 APS samples.

**Supplementary Table 2.** Results of log VAT prediction model for all biomarkers, MEC Adiposity Phenotype Study

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Males (n=886)** | |  | **Females (n=915)** | |
| **Variable\*** | **Beta in log units** | **Standardized Beta** |  | **Beta in log units** | **Standardized Beta** |
| Intercept | -22.8026 | 5.1881 |  | -14.4247 | 4.7870 |
| BMI (kg/m2) | 14.3151 | 2.1892 |  | 9.5425 | 1.7413 |
| Adiponectin (ng/mL) | -0.0370 | -0.0267 |  | -0.0697 | -0.0472 |
| Alanin aminotransferase (U/L) | 0.0343 | 0.0162 |  | 0.0183 | 0.0086 |
| HDL Cholesterol (mg/dL) | 0.0895 | 0.0383 |  | 0.0589 | 0.0266 |
| LDL Cholesterol (calculated, mg/dL) | 0.2393 | 0.0967 |  | 0.2495 | 0.0955 |
| Total Cholesterol (mg/dL) | -0.4137 | -0.1070 |  | -0.5131 | -0.1321 |
| Glucose (mg/dL) | 0.1121 | 0.0268 |  | 0.0813 | 0.0195 |
| High-sensitivity C-reactive protein (mg/L) | 0.0126 | 0.0077 |  | 0.0060 | 0.0038 |
| IGF1 (ng/mL) | -0.0412 | -0.0139 |  | -0.0177 | -0.0058 |
| IGF binding protein 2 (ng/mL) | 0.0164 | 0.0078 |  | 0.0000 | 0.0000 |
| Insulin (microU/mL) | 0.0824 | 0.0490 |  | 0.1040 | 0.0580 |
| Leptin (ng/mL) | 0.0643 | 0.0568 |  | 0.0017 | 0.0015 |
| Triglycerides (mg/dL) | 0.1037 | 0.0490 |  | 0.2249 | 0.1069 |
| Total carotene (ng/mL) | -0.0775 | -0.0555 |  | -0.0352 | -0.0247 |
| Reduced CoQ10 (ubiquinol) (ng/mL) | 0.0171 | 0.0210 |  | 0.0013 | 0.0015 |
| Cryptoxanthin (ng/mL) | 0.0009 | 0.0006 |  | 0.0023 | 0.0015 |
| Lutein (ng/mL) | -0.0341 | -0.0154 |  | 0.0035 | 0.0017 |
| Anhydrolutein (ng/mL) | 0.0529 | 0.0220 |  | -0.0894 | -0.0404 |
| Lycopene total (ng/mL) | 0.0075 | 0.0035 |  | 0.0505 | 0.0243 |
| Retinol (ng/mL) | 0.0203 | 0.0071 |  | 0.0683 | 0.0249 |
| Alpha-tocopherol (ng/mL) | 0.0612 | 0.0217 |  | 0.0496 | 0.0171 |
| Beta+gamma tocopherol (ng/mL) | 0.0102 | 0.0042 |  | -0.0042 | -0.0017 |
| Zeaxanthin (ng/mL) | -0.0303 | -0.0132 |  | -0.0028 | -0.0014 |
| Sex hormone binding globulin (nmol/L) | -0.0971 | -0.0460 |  | -0.0391 | -0.0217 |
| Height (m) | 4.4786 | 0.1879 |  | 4.0481 | 0.1772 |
| BMI squared (kg2/m4) | -1.9051 | -1.9345 |  | -1.2174 | -1.4764 |
| Height squared (m2) | -4.0109 | -0.1819 |  | -4.5163 | -0.1816 |
|  |  |  |  |  |  |
| R2 | 0.65 |  |  | 0.68 |  |

* All variables were log transformed
* Average of parameters of 100 bootstrap LASSO models with alpha=0.9 and 10-fold cross-validation.

**Supplementary Table 3.** Goodness of fit statistics for the of final log VAT prediction model\* by sex and ethnicity, MEC Adiposity Phenotype Study

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Males (n=886)** | | | | **Females (n=915)** | | | |
| **Race/Ethnicity** | **N** | **R2** | **% VAT ≥ 150 cm2** | **AUROC\*\*** | **N** | **R2** | **% VAT ≥ 150 cm2** | **AUROC** |
|  |  |  |  |  |  |  |  |  |
| African American | 127 | 0.46 | 65.4% | 0.86 | 175 | 0.57 | 25.4% | 0.84 |
| Native Hawaiian | 134 | 0.67 | 69.4% | 0.91 | 155 | 0.74 | 40.6% | 0.85 |
| Japanese American | 228 | 0.72 | 67.1% | 0.93 | 202 | 0.74 | 39.1% | 0.91 |
| Latino | 189 | 0.49 | 85.2% | 0.87 | 187 | 0.66 | 47.6% | 0.85 |
| European American | 208 | 0.73 | 62.5% | 0.92 | 196 | 0.69 | 29.6% | 0.87 |

\*Final VAT model as described in Table 2.

\*\* Area under the Receiver Operating Curve for VAT ≥ 150 cm2.

**Supplementary Table 4.** Breast Cancer Odds Ratios and 95% Confidence Intervals for BMI and Waist Circumference, 950 case-control pairs nested in the MEC

|  |  |  |
| --- | --- | --- |
|  | Model 1 | Model 2 |
|  |  |  |
| **BMI categories at baseline** |  |  |
| BMI < 25 kg/m2 | 1.00 | 1.00 |
| BMI 25-29.9 kg/m2 | 1.64  (1.32-2.03) | 1.63  (1.27-2.09) |
| BMI 30+ kg/m2 | 1.36  (1.05-1.75) | 1.46  (1.07-1.98) |
|  |  |  |
| **BMI continuous at baseline** |  |  |
| P-value for log BMI kg/m2 | 0.004 | 0.01 |
|  |  |  |
| **Waist1 circumference tertiles at 10-year follow-up** |  |  |
| Waist < 33 inches | 1.00 | 1.00 |
| Waist 33-37 inches | 1.29  (1.00-1.68) | 1.15  (0.85-1.54) |
| Waist 38+ inches | 1.53  (1.16-2.03) | 1.58  (1.13-2.20) |
|  |  |  |
| **Waist1 continuous at 10-year follow-up** |  |  |
| P-value for log waist | 0.02 | 0.03 |
| P-value for log waist, adjusted for log BMI | 0.34 | 0.18 |

Model 1 adjusted for matching factors and age at blood draw

Model 2 further adjusted for menopausal hormone therapy, pack-years of smoking, moderate-vigorous activity, family history of breast cancer, type and age of menopause, age at first live birth, number of children, ethanol (g/day), log energy (kcal/day)

1Waist circumference was available on the 714 cases and 703 controls who provided waist circumference in the 10-year follow-up survey.

**Supplementary Table 5.** Odds Ratios and 95% Confidence Interval for Breast Cancer and Colorectal Cancer Associated with the VAT Predictive BMI-Adjusted1 Score, among those with Waist Circumference

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Breast Cancer | | | Colorectal Cancer | | |
| VAT Score1 | Nb of cases / Nb of controls | Multivariate Adjusted2 | Further Adjusted for Waist Circumference | Nb of cases / Nb of controls | Multivariate Adjusted3 | Further Adjusted for Waist Circumference |
|  | 686/668 |  |  | 525/581 |  |  |
| Tertile 1 (low) | 216/252 | 1.00 | 1.00 | 182/208 | 1.00 | 1.00 |
| Tertile 2 | 218/233 | 1.08  (0.82-1.42) | 1.03  (0.78-1.36) | 165/196 | 0.88  (0.65-1.20) | 0.86  (0.63-1.17) |
| Tertile 3 (high) | 330/263 | 1.59  (1.21-2.11) | 1.51  (1.14-2.01) | 178/177 | 1.06  (0.77-1.46) | 1.02  (0.74-1.41) |
| P**-v**alue for log VAT Score Trend (continuous) |  | 0.002 | 0.01 |  | 0.96 | 0.77 |

1The VAT score is adjusted for BMI by the method of residuals.

2Multivariate model is adjusted for matching factors and age at blood draw, menopausal hormone therapy, pack-years of smoking, moderate-vigorous activity, family history of breast cancer, type and age of menopause, age at first live birth, number of children, ethanol (g/day), and log energy (kcal/day).

3Multivariate model is adjusted for matching factors and age at blood draw, menopausal hormone therapy (women only), pack-years of smoking, vigorous activity, multivitamin use, history of polyps, NSAID use, family history of colorectal cancer, and log intakes of alcohol, dietary fiber, Dietary Folate Equivalents from food or supplements, calcium from food or supplements and energy.

**Supplementary Table 6.** Odds Ratios and 95% Confidence Interval for Breast Cancer Associated with the VAT Predictive BMI-Adjusted Score, by Menopausal Hormone Therapy Use, among case-control pairs nested in MEC

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Ever Used Menopausal Hormone Therapy  (477 cases, 480 controls) | | | | Never Used Menopausal Hormone Therapy  (462 cases, 458 controls) | | |
| VAT Score | Model 1 |  | Model 2 |  | Model 1 |  | Model 2 |
| Tertile 1 (low) | 1.00 |  | 1.00 |  | 1.00 |  | 1.00 |
| Tertile 2 | 1.10  (0.80-1.52) |  | 1.06  (0.75-1.49) |  | 1.12  (0.80-1.56) |  | 1.14  (0.80-1.64) |
| Tertile 3 (high) | 1.42  (1.02-1.97) |  | 1.45  (1.03-2.06) |  | 1.59  (1.14-2.22) |  | 1.55  (1.08-2.21) |
| P-value for log VAT Score Trend (continuous) | 0.06 |  | 0.05 |  | 0.002 |  | 0.004 |

The VAT score is adjusted for BMI by the method of residuals.

Model 1 is adjusted for matching factors and age at blood draw.

Model 2 is further adjusted for age at blood draw, pack-years of smoking, moderate-vigorous activity, family history of breast cancer, type and age of menopause, age at first live birth, number of children, ethanol (g/day), and log energy (kcal/day).

The p-value for the Wald test (1 degree of freedom) for interaction between the VAT score and menopausal hormonal therapy was 0.38 for Model 1 and 0.35 for Model 2.

**Supplementary Table 7.**  Odds Ratios and 95% Confidence Interval for Breast Cancer Associated with the VAT Predictive BMI-Adjusted Score, by Estrogen Receptor Status (753 ER+ cases, 130 ER- cases and 950 controls)\*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ER + | | | | ER- | | | |
| VAT Score | Model 1 |  | Model 2 |  | Model 1 |  | Model 2 |  |
| Tertile 1 (low) | 1.00 |  | 1.00 |  | 1.00 |  | 1.00 |  |
| Tertile 2 | 1.16  (0.91-1.47) |  | 1.18  (0.92-1.51) |  | 0.89  (0.56-1.43) |  | 0.86  (0.53-1.40) |  |
| Tertile 3 (high) | 1.46  (1.14-1.86) |  | 1.54  (1.19-2.00) |  | 1.46  (1.15-1.86) |  | 1.34  (0.84-2.15) |  |
| P-value for log VAT Score Trend (continuous) | 0.008 |  | 0.004 |  | 0.05 |  | 0.06 |  |

\* Based on a polytomous logistic regression.

The VAT score is adjusted for BMI by the method of residuals.

Model 1 is adjusted for matching factors and age at blood draw.

Model 2 is further adjusted for age at blood draw, menopausal hormone therapy, pack-years of smoking, moderate-vigorous activity, family history of breast cancer, type and age of menopause, age at first live birth, number of children, ethanol (g/day), and log energy (kcal/day).

The p-value for the Wald test (1 degree of freedom) for heterogeneity of the odds ratios across the ER status case groups in a polytomous logistic model was 0.59 for Model 1 and 0.75 for Model 2.

**Supplementary Table 8.** Odds Ratios and 95% Confidence Interval for Breast Cancer Associated with the VAT Predictive BMI-Adjusted Score, by Race/Ethnicity, among case-control pairs nested in MEC

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | African American  (150 cases/150 controls) | | Native Hawaiian  (106 cases / 106 controls) | | Japanese American (307 cases / 308 controls) | | Latino  (195 cases / 193 controls) | | European American (181 cases / 181 controls) | |
| Predicted VAT from Final Model | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 |
| Tertile 1 (low) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Tertile 2 | 0.94  (0.55-1.59) | 0.74  (0.40-1.37) | 0.78  (0.38-1.62) | 0.72  (0.31-1.70) | 1.17  (0.77-1.76) | 1.14  (0.74-1.75) | 1.25  (0.74-2.13) | 1.33  (0.72-1.41) | 1.34  (0.83-2.18) | 1.52  (0.89-2.60) |
| Tertile 3 (high) | 0.89  (0.50-1.60) | 0.87  (0.44-1.71) | 1.15  (0.60-2.20) | 0.99  (0.47-2.10) | 2.13  (1.42-3.22) | 2.13  (1.37-3.32) | 1.42  (0.84-2.38) | 1.54  (0.84-2.80) | 1.23  (0.72-2.10) | 1.12  (0.60-2.09) |
| P-value for log VAT Score Trend (continuous) | 0.64 | 0.59 | 0.41 | 0.60 | 0.0001 | 0.0003 | 0.48 | 0.31 | 0.31 | 0.49 |

The VAT score is adjusted for BMI by the method of residuals.

Model 1 is adjusted for matching factors and age at blood draw.

Model 2 is further adjusted for age at blood draw, menopausal hormone therapy, pack-years of smoking, moderate-vigorous activity, family history of breast cancer, type and age of menopause, age at first live birth, number of children, ethanol (g/day), and log energy (kcal/day).

The p-value for the Wald test (4 degrees of freedom) for interaction between the VAT score and race/ethnicity was 0.08 for Model 1 and 0.08 for Model 2.

**Supplementary Table 9.** Colorectal Cancer Odds Ratios and 95% Confidence Intervals for BMI and Waist Circumference, 831 case-control pairs

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Men and Women | | Men | | Women | |
|  | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 |
|  |  |  |  |  |  |  |
| **BMI categories at baseline** |  |  |  |  |  |  |
| BMI < 25 kg/m2 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| BMI 25-29.9 kg/m2 | 1.26  (1.00-1.59) | 1.30  (1.02-1.67) | 1.24  (0.90-1.71) | 1.30  (0.92-1.84) | 1.26  (0.90-1.77) | 1.34  (0.93-1.94) |
| BMI 30+ kg/m2 | 1.65  (1.24-2.21) | 1.60  (1.15-2.21) | 1.41  (0.92-2.17) | 1.51  (0.94-2.43) | 1.89  (1.27-2.81) | 1.70  (1.08-2.70) |
|  |  |  |  |  |  |  |
| **BMI continuous at baseline** |  |  |  |  |  |  |
| P-value for log BMI kg/m2 | 0.002 | 0.004 | 0.20 | 0.07 | 0.00 | 0.02 |
|  |  |  |  |  |  |  |
| **Waist1 circumference tertiles at 10-year follow-up** |  |  |  |  |  |  |
| Waist < 36 inches for men and < 33 inches for women | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Waist 36-39 inches for men and 33-37 for women | 1.24  (0.93-1.67) | 1.27  (0.93-1.74) | 1.16  (0.77-1.76) | 1.18  (0.75-1.83) | 1.32  (0.87-2.01) | 1.43  (0.91-2.24) |
| Waist ≥ 40 inches for men and ≥ 38 for women | 1.33  (0.97-1.82) | 1.37  (0.96-1.93) | 1.13  (0.72-1.76) | 1.18  (0.72-1.92) | 1.58  (1.00-2.48) | 1.70  (1.01-2.85) |
|  |  |  |  |  |  |  |
| **Waist1 continuous at 10-year follow-up** |  |  |  |  |  |  |
| P-value for log waist | 0.07 | 0.03 | 0.66 | 0.39 | 0.05 | 0.02 |
| P-value for log waist, adjusted for log BMI | 0.81 | 0.73 | 0.98 | 0.97 | 0.79 | 0.42 |

Model 1: Adjusted for matching factors and age at blood draw.

Model 2 is further adjusted for hormone replacement therapy (for women only), pack-years of smoking, vigorous activity, multivitamin use, history of polyps, NSAID use, family history of colorectal cancer, and log intakes of alcohol, dietary fiber, Dietary Folate Equivalents from food or supplements, calcium from food or supplements and energy.

1Among the 559 cases and 608 controls who provided waist circumference in the 10-year follow-up survey.

**Supplementary Table 10.** Odds Ratios and 95% Confidence Interval for Colorectal Cancer Associated with the VAT Predictive BMI-Adjusted Score by Sex

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Males (422 cases / 420 controls) | | | | Females (393 cases / 394 controls) | | | |
| Predicted VAT from Final Model |  | Model 1 | Model 2 |  |  | Model 2 | Model 3 |  |
| Tertile 1 (low) |  | 1.00 | 1.00 |  |  | 1.00 | 1.00 |  |
| Tertile 2 |  | 1.13  (0.86-1.49) | 0.97  (0.68-1.39) |  |  | 1.04  (0.71-1.52) | 1.03  (0.70-1.52) |  |
| Tertile 3 (high) |  | 1.07  (0.75-1.54) | 1.00  (0.69-1.46) |  |  | 1.13  (0.76-1.68) | 1.10  (0.74-1.64) |  |
|  |  |  |  |  |  |  |  |  |
| P-value for log VAT Score Trend (continuous) |  | 0.71 | 0.80 |  |  | 0.30 | 0.36 |  |
| P-value for test of interaction between sexes for VAT Score |  |  |  |  |  | 0.41 | 0.49 |  |

The VAT score is adjusted for BMI by the method of residuals.

Model 1 is further adjusted for menopausal hormone therapy (for women only), pack-years of smoking, vigorous activity, multivitamin use, history of polyps, NSAID use, family history of colorectal cancer, and log intakes of alcohol, dietary fiber, Dietary Folate Equivalents from food or supplements, calcium from food or supplements and energy

Model 2 is additionally adjusted for log body mass index.

The p-value for the Wald test (1 degree of freedom) for interaction between the VAT score and sex was 0.41 for Model 1 and 0.49 for Model 2.

**Supplementary Table 11.** Odds Ratios and 95% Confidence Interval for Colorectal Cancer Associated with the VAT Predictive BMI-Adjusted Score, by Race/Ethnicity\*, for the cases diagnosed ≥7 years after blood draw and all controls nested in MEC

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | African American  (70 cases / 174 controls) | | Japanese American  (91 cases / 266 controls) | | Latino  (71 cases / 205 controls) | | European American  (47 cases / 115 controls) | |
| Predicted VAT from Final Model | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 |
| Tertile 1 (low) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Tertile 2 | 0.59  (0.29-1.22) | 0.77  (0.33-1.79) | 0.94  (0.49-1.80) | 0.94  (0.45-1.93) | 1.52  (0.72-3.19) | 1.51  (0.61-3.72) | 0.94  (0.40-2.21) | 1.32  (0.46-3.79) |
| Tertile 3 (high) | 0.80  (0.39-1.64) | 0.99  (0.42-2.31) | 1.61  (0.88-2.97) | 1.85  (0.95-3.61) | 1.36  (0.66-2.81) | 1.29  (0.53-3.14) | 1.43  (0.58-3.55) | 1.25  (0.36-4.39) |
| P-value for log VAT Score Trend (continuous) | 0.99 | 0.42 | 0.05 | 0.05 | 0.32 | 0.25 | 0.43 | 0.58 |

\*Native Hawaiians were excluded due to a small sample size (n=17 cases).

The VAT score is adjusted for BMI by the method of residuals.

Model 1 is adjusted for matching factors and age at blood draw.

Model 2 is further adjusted for menopausal hormone therapy (for women only), pack-years of smoking, vigorous activity, multivitamin use, history of polyps, NSAID use, family history of colorectal cancer, and log intakes of alcohol, dietary fiber, Dietary Folate Equivalents from food or supplements, calcium from food or supplements and energy.

The p-value for the Wald test (3 degrees of freedom) for interaction between the VAT score and race/ethnicity was 0.08 for Model 1 and 0.08 for Model 2.