



Published in final edited form as:

Br J Nutr. 2016 November 28; 116(10): 1787–1798. doi:10.1017/S0007114516003755.

Development and validation of empirical indices to assess the insulinemic potential of diet and lifestyle

Fred K. Tabung^{*,1,2}, Weike Wang^{*,1,2}, Teresa T. Fung^{1,3}, Frank B. Hu^{1,2,4,5}, Stephanie A. Smith-Warner^{1,2}, Jorge E. Chavarro^{1,2,4,5}, Charles S. Fuchs^{4,5,6}, Walter C. Willett^{1,2,4,5}, Edward L. Giovannucci^{1,2,4,5}

¹Department of Nutrition, Harvard T.H. Chan School of Public Health

²Department of Epidemiology, Harvard T.H. Chan School of Public Health

³Department of Nutrition, Simmons College

⁴Channing Division of Network Medicine, Brigham and Womens' Hospital

⁵Department of Medicine, Harvard Medical School

⁶Department of Medical Oncology, Dana-Farber Cancer Institute

Abstract

The glycemic and insulin indices assess postprandial glycemic and insulin response to foods respectively, which may not reflect the long-term effects of diet on insulin response. We developed and evaluated the validity of four empirical indices to assess the insulinemic potential of usual diets and lifestyles, using dietary, lifestyle and biomarker data from the Nurses' Health Study (NHS, n=5,812 for hyperinsulinemia, n=3,929 for insulin resistance). The four indices were: the empirical dietary index for hyperinsulinemia (**EDIH**) and empirical lifestyle index for hyperinsulinemia (**ELIH**); empirical dietary index for insulin resistance (**EDIR**) and empirical lifestyle index for insulin resistance (**ELIR**). We entered 39 food frequency questionnaire-derived food groups in stepwise linear regression models and defined indices as the patterns most predictive of fasting plasma C-peptide, for the hyperinsulinemia pathway (EDIH and ELIH); and of the triglyceride/high density lipoprotein-cholesterol (TG/HDL) ratio, for the insulin resistance pathway (EDIR and ELIR). We evaluated the validity of indices in two independent samples from NHS-II and Health Professionals Follow-up Study (HPFS) using multivariable-adjusted linear regression analyses to calculate relative concentrations of biomarkers. EDIH is comprised of 18 food groups; 13 were positively associated with C-peptide, five inversely. EDIR is comprised of 18 food groups; ten were positively associated with TG/HDL and eight inversely. Lifestyle indices had fewer dietary components, and included BMI and physical activity as components. In the

ADDRESS CORRESPONDENCE TO: Fred K. Tabung, Department of Nutrition, Harvard T. H. Chan School of Public Health, Building #2, Room 351A, Boston, MA 02115; ftabung@hsph.harvard.edu.

*Co-first author

CONFLICT OF INTEREST

All authors declare no conflict of interest

AUTHORSHIP

FKT, WW, and ELG designed research; FKT and WW conducted research and performed statistical analysis; TTF, FBH, SAS, JEC, CSF and WCW analyzed and interpreted findings and provided critical input; FKT and WW wrote the paper; ELG provided study oversight, and had primary responsibility for final content; all authors read and approved final content.

validation samples, all indices significantly predicted biomarker concentrations, e.g., the relative concentrations (95%CI) of the corresponding biomarkers comparing extreme index quintiles in HPFS were: EDIH, 1.29(1.22, 1.37); ELIH, 1.78(1.68, 1.88); EDIR, 1.44(1.34, 1.55); ELIR, 2.03(1.89, 2.19); all P-trend<0.0001. The robust associations of these novel hypothesis-driven indices with insulin response biomarker concentrations suggests their usefulness in assessing the ability of whole diets and lifestyles to stimulate and/or sustain insulin secretion.

Keywords

hypothesis-driven; dietary patterns; lifestyle; hyperinsulinemia; insulin resistance; insulin secretion; C-peptide; triglycerides; high-density lipoprotein cholesterol

INTRODUCTION

Hyperinsulinemia and insulin resistance are considered important underlying mechanisms linking poor dietary and lifestyle behaviors to the development of multiple chronic diseases and conditions. For example, studies suggest that hyperinsulinemia is associated with higher risk of colorectal adenomas⁽¹⁾ and colorectal cancer independent of adiposity^(2; 3), and insulin resistance has been consistently linked to obesity, inflammation, heart disease and type 2 diabetes^(4; 5; 6). Although specific dietary factors have been shown to influence insulin resistance and secretion^(7; 8); dietary patterns or indices that include multiple dietary factors and account for the complex interactions among nutrients and foods may be more predictive of diet-disease associations^(9; 10). Other lifestyle factors that have been linked to hyperinsulinemia and insulin resistance are body weight and physical activity^(11; 12; 13; 14). Physical activity plays an important role in the prevention of insulin insensitivity⁽¹⁴⁾, while increased body weight has a direct association with insulin resistance⁽¹¹⁾. Therefore, combining diet, exercise, and body weight in a lifestyle index would likely be more predictive of hyperinsulinemia and insulin resistance than each of these factors considered separately.

Currently, the most common dietary index used to assess the ability of diets to stimulate insulin secretion is the glycemic index (GI). The GI classifies carbohydrate-containing foods by their ability to raise the postprandial blood glucose concentration relative to glucose or white bread⁽¹⁵⁾ and therefore indirectly assesses immediate insulin responses to food intake. However, it neglects dietary factors such as proteins and fats that are also important in insulin secretion. Moreover the GI does not quantify the long-term effects of diet on glycemia. As an improvement on the GI, our group previously developed a food insulin index to directly quantify the postprandial insulin response⁽¹⁶⁾. However this index was not predictive of C-peptide concentrations⁽¹⁶⁾. The lack of predictive ability may be because the insulin index, similar to the glycemic index, assesses postprandial insulin response to the intake of specific foods and therefore is limited to quantifying short-term insulin response rather than the long-term effects of whole diets on insulinemia. Hence we developed dietary and lifestyle patterns that assess the insulinemic potential of usual diets and lifestyles to reflect long term insulin exposure and overall insulin resistance, the more relevant exposure for chronic disease prevention.

Previously, our group derived a dietary pattern associated with hyperinsulinemia and found this pattern to be significantly associated with colorectal cancer risk⁽¹⁷⁾. However, the sample size used to derive this pattern was small (n=833) and the pattern was applied in the same cohort. Our objectives in the current study were three-fold: first, we updated the previously developed dietary pattern using the currently available larger sample of women and additionally developed separate dietary and lifestyle patterns predictive of hyperinsulinemia, as well as insulin resistance. Second, in validation studies, we evaluated how well these patterns predicted concentrations of insulin response biomarkers in independent samples of men and women; and third, we examined the joint influence of diet, body weight and physical activity on clinically relevant hyperinsulinemia and insulin resistance.

METHODS

Study populations

The Nurses' Health Study (NHS), Nurses' Health Study-II (NHS-II) and Health Professionals Follow-up Study (HPFS) are ongoing prospective cohorts established in 1976, 1989 and 1986 respectively. The NHS (n=121,701) enrolled female registered nurses aged 30–55 years, while the NHS-II (n=116,430) enrolled younger female registered nurses 25 to 42 years⁽¹⁸⁾. The HPFS (n=51,529) enrolled male health professionals aged 40–75 years. Blood samples were collected from subpopulations of the NHS (n=32,826) in 1989–1990, NHS-II (n=29,611) between 1996 and 1999 and HPFS (n=18,225) from 1993 to 1994⁽¹⁹⁾. Blood collection was conducted using similar protocols for all cohorts. The procedures, including collection, handling and storage, have been previously summarized⁽²⁰⁾. In the current study, we used data from previous matched case-control studies nested within each of the three cohorts that measured fasting concentrations of plasma C-peptide, triglycerides (TG) and high density lipoprotein cholesterol (HDL). In the NHS, 5,812 women with C-peptide data and 3,929 women with data on TG and HDL were included in the development of the dietary and lifestyle indices. For the validation studies, there were 4,002 men with C-peptide data and 3,559 men with TG and HDL data in the HPFS cohort; and 1,717 women with C-peptide data and 1,008 women with TG, HDL data in the NHS-II cohort. The Institutional Review Boards at Brigham and Women's Hospital and at Harvard T.H. Chan School of Public Health approved this study.

Biomarker assessment

For the current analysis, we utilized fasting plasma C-peptide concentrations to assess hyperinsulinemia. Compared to insulin, C-peptide has proven to be a better measure of beta-cell secretory activity as it is not extracted by the liver, has a slower metabolic clearance rate, and does not cross-react with antibodies to insulin⁽²¹⁾. To assess insulin resistance, we utilized the ratio of fasting triglyceride to fasting high-density lipoprotein-cholesterol (TG/HDL) which has been shown to be significantly correlated with insulin resistance⁽²²⁾. TG/HDL is also a simple and clinically useful way to identify apparently healthy individuals who are insulin resistant^(23; 24; 25).

Procedures for the measurement of fasting plasma insulinemic markers (C-peptide, TG and HDL) in the NHS, NHS-II, and HPFS have been described^(26; 27). C-peptide was measured by ELISA (Diagnostic Systems Laboratories/Beckman Coulter, Webster, TX). HDL cholesterol and TG were measured by standard methods with reagents from Roche Diagnostics (Indianapolis, IN) and Genzyme (Cambridge, MA)^(26; 27). The intra-assay coefficient of variation from blinded quality control samples were <12% for C-peptide and <1.8% for TG and HDL across batches.

In the nested case-control studies in which these biomarkers were measured, samples from cases and their matched controls were analyzed in the same batch. Quality control samples were randomly interspersed among the case-control samples, and laboratory personnel were blinded to quality control and case-control status for all assays. Biomarkers were measured in multiple batches over several years. There may be differences in mean biomarker levels by batch due to different reagents, technicians, or laboratories, but also due to differences in the participants in each batch. We therefore used a 3-step method previously described by Rosner et al.⁽²⁸⁾, to recalibrate biomarker concentrations across several batches to the value of an "average batch" accounting for true variability across batches due to different distributions of predictors of the biomarker across batches: i) we constructed a linear regression model with biomarker levels as the dependent variable and batch indicators as well as variables that may vary by biomarker levels and by batch (regular aspirin/NSAIDs use, age at blood draw, physical activity, smoking status, diabetes, other chronic diseases/conditions, case-control status, and menopausal status, postmenopausal hormone use in women) as the independent variables, ii) next we calculated the average batch beta coefficient (β) by summing the batch indicator β s and dividing by the total number of batches, iii) lastly we calculated the difference between each batch β and average β and recalibrated biomarker concentrations by subtracting this difference from the original biomarker concentration. The recalibrated biomarkers were then used in analyses. The correlations between the recalibrated and uncalibrated TG/HDL was 0.96, and 0.85 for C-peptide in the NHS, therefore we used the uncalibrated TG/HDL and calibrated C-peptide in the primary analyses and conducted sensitivity analyses with the recalibrated TG/HDL and uncalibrated C-peptide.

Assessment of dietary and non-dietary data

Dietary data are updated every four years in the NHS (since 1980), NHS-II (since 1991) and in the HPFS (since 1986) with a validated semi quantitative food frequency questionnaire (FFQ) that assessed diet intake in the previous one year^(29; 30; 31). We used dietary data from the questionnaires closest to the blood draw. That is, the 1990 FFQ for the NHS, 1999 FFQ for NHS-II, and the 1994 FFQ for HPFS. Participants with excessive missing items (> 70) on the FFQs or implausibly low or high energy intake (<600 or >3500 kcal/d for women and <800 or >4,200 kcal/d for men) were excluded⁽³²⁾.

All three cohorts collected nondietary data (e.g., medical history and health practices) and updated the data through biennial self-administered questionnaires. We calculated participants' body mass index (BMI -kg/m²) using height (meters) reported at baseline for each cohort, and weight (kg) reported on the questionnaire closest to blood draw.

Participants reported smoking status (never, former, current), and we calculated physical activity, expressed in metabolic equivalent (MET)-hours per week by summing the average MET-hours/week for the following activities: tennis/squash/racquetball, rowing, calisthenics, walking, jogging, running, bicycling, and swimming. The reproducibility and validity of the physical activity questionnaire have been evaluated.^(33; 34) Regular use of aspirin or other non-steroidal anti-inflammatory drugs (NSAID) was defined as use of 2 standard tablets (325-mg) of aspirin or 2 tablets of NSAIDs per week. We derived a chronic disease comorbidity score by summing the presence=1/absence=0 of the following chronic diseases/conditions: hypercholesterolemia, cancer, high blood pressure, heart disease, and rheumatoid/other arthritis).

Development of the indices of lifestyle and dietary insulinemic potential

We developed four indices to assess the insulinemic potential of whole diets and lifestyles: the empirical dietary index for hyperinsulinemia (**EDIH**) and empirical lifestyle index for hyperinsulinemia (**ELIH**) that also includes BMI and physical activity as components; the empirical dietary index for insulin resistance (**EDIR**) and empirical lifestyle index for insulin resistance (**ELIR**) that also includes BMI and physical activity as components.

Of the three cohorts, the NHS had the largest sample of participants with biomarker data, therefore we used dietary, lifestyle and biomarker data (C-peptide, TG and HDL) in the NHS to develop the indices, and based the scores on food groups rather than nutrients, to approximate how people perceive dietary intake. We first calculated daily intakes per 1000 kcal of 39 previously defined food groups⁽³²⁾ from the 1990 FFQ. The grouping scheme was based on the similarity of the nutrient profiles or culinary usage among the foods.⁽³²⁾ We then used four separate stepwise multivariable-adjusted linear regression analyses to identify the most important component food groups and lifestyle factors contributing to hyperinsulinemia (with C-peptide concentrations as the dependent variable) and to insulin resistance (with TG/HDL as the dependent variable), with the 39 food groups as independent variables, and a significance level of $P=0.1$ for entry into, and retention in the model. BMI and physical activity were added to the list of the 39 food group predictors in models to develop the lifestyle indices. Intakes of the food groups identified in the stepwise linear regression analyses were weighted by the regression coefficients derived from the final stepwise linear regression model, and then summed to constitute the indices. All four index scores assess the insulinemic potential of diet on a continuum from maximally low insulinemic potential to maximally high insulinemic potential, with higher (more positive) scores indicating more highly insulinemic diets or lifestyles (hyperinsulinemia or insulin resistance) while lower (more negative) scores indicate low insulinemic or insulin sensitive diets or lifestyles.

Sensitivity analyses

In sensitivity analyses, we created three potential alternative versions of both the EDIH and EDIR by: i) using uncalibrated C-peptide and calibrated TG/HDL; ii) using unweighted components, thus assuming that all components contribute equally to the total score; iii) constructing the indices among only control subjects of the nested case-control studies

(although all the nested case-control studies that generated the data for the current study used prediagnostic blood samples from chronic disease-free participants).

Additionally, we compared the predictive ability of the previously developed C-peptide dietary pattern. This pattern was high in red meat, high energy beverages, fish and creamy soup intake, and low in coffee, high-fat dairy and whole grains intake. Lastly, we compared the predictive ability of EDIH and EDIR with that of the previously developed insulin index. The insulin index has been described; its values compare the postprandial plasma insulin response of a specific food relative to a reference food⁽¹⁶⁾.

Statistical analysis

Where it is not explicitly stated, the analyses described for EDIH and EDIR were also applied to their respective lifestyle versions. We described participants' characteristics using means (standard deviations) for continuous variables or geometric means (coefficient of variation) for log transformed variables, and frequencies (%) for categorical variables. Concentrations of all biomarkers were back transformed to their original units (e^x where x is the transformed biomarker value) because biomarkers were log transformed using natural logarithms prior to analyses.

In NHS, we calculated correlation coefficients between the EDIH or EDIR, their alternative versions and the insulinemic markers. We also assessed the distribution of the absolute average concentrations of C-peptide across quintiles of EDIH, and TG/HDL across quintiles of EDIR, stratified by joint categories of BMI and physical activity (PA) as follows: lean and active (BMI <25kg/m² and PA ≥ median PA), lean and sedentary (BMI <25kg/m² and PA <median PA), overweight/obese and active (BMI ≥ 25kg/m² and PA ≥ median PA) and overweight/obese and sedentary (BMI ≥ 25kg/m² and PA <median PA). The multivariable models were adjusted for the following covariates: age at blood draw (years, continuous), physical activity (MET-hours/week, continuous), smoking status (never, former, current), regular aspirin/NSAIDs use (yes/no), case-control status, history of diabetes (yes/no), chronic disease comorbidity score and additionally for menopausal status and postmenopausal hormone use. BMI was not controlled for in the multivariable models because it has been shown to mediate^(35; 36) and/or modify⁽¹⁷⁾ the association between diet and insulin markers, thus controlling for BMI could result in attenuation of true associations or loss of statistical power to detect true associations.

In the validation studies in which we evaluated how well the indices predicted concentrations of the insulin response biomarkers in the HPFS and NHS-II samples, we calculated scores for the EDIH and EDIR and their potential alternative versions, and estimated correlations among the index scores and biomarkers (C-peptide for hypersinsulinemia) and (TG/HDL for insulin resistance). Also, we assessed the distribution of the absolute average concentrations of C-peptide across quintiles of EDIH, and TG/HDL across quintiles of EDIR, stratified by joint categories of BMI, PA described above. To determine if there were clinically relevant differences in the insulinemic potential of diet between these categories, we used clinically relevant cut points; 1.8ng/mL for C-peptide^(37; 38) and 3 for TG/HDL^(25; 39) (values considered to be the upper limit of normal) to dichotomize the biomarkers. Participants with values ≥ 1.8ng/mL were classified as having

high C-peptide concentrations while those with TG/HDL >3 had high TG/HDL ratio. We then calculated proportions of participants with clinically high levels of biomarkers across dietary index quintiles in each category of BMI,PA.

The associations between EDIH or EDIR and their respective outcome biomarkers was assessed in multivariable-adjusted linear regression models using relative concentrations of the biomarkers predicted in higher EDIH or EDIR quintiles, with the lowest quintile as reference (e.g., concentration in quintile 5 / concentration in quintile 1). We used the continuous index adjusted for multiple covariates to assess the trend of biomarker concentrations across quintiles of the categorized index. All multivariable models were adjusted for the previously described potential confounding variables.

In sensitivity analyses, we applied each of the three alternative versions of the EDIH or EDIR (scores developed using uncalibrated C-peptide and calibrated TG/HDL, scores developed using unweighted components, scores developed in control subjects only) in multivariable-adjusted linear regression models to predict relative concentrations of the biomarkers. In addition, we compared the predictive ability of the previously developed C-peptide dietary pattern and the insulin index with that of the EDIH and EDIR. Though participants were free from diabetes at blood collection, we excluded participants identified to have diabetes during the nested case-control studies, and compared findings with those from all participants.

All analyses were conducted using SAS version 9.3 for UNIX. All tests were 2-sided and 95% confidence intervals not including 1 were considered to indicate statistically significant results.

RESULTS

Of the 39 food groups examined, 18 were identified as significant contributors to the EDIH, with 13 of them positively associated and five inversely associated with C-peptide concentrations (Table 1). EDIH had 14 components; seven components including BMI were positively associated with C-peptide while the remaining seven components including physical activity were inversely associated with C-peptide concentrations. Common to both the dietary and lifestyle hyperinsulinemia indices were red meat, margarine, creamy soups, butter (positive associations); high fat dairy, wine, coffee and whole fruit (inverse associations). The EDIR had 18 components; ten were positively associated while eight were inversely associated with TG/HDL. EDIR had 17 components: 11 including BMI, were positively associated with TG/HDL, while the remaining six including physical activity were inversely associated with TG/HDL. Common to both the dietary and lifestyle insulin resistance indices were margarine, red meat, refined grains, processed meat, tomatoes, other vegetables, low energy beverages (positive associations); coffee, wine, high fat dairy, liquor and green leafy vegetables (inverse associations) (Table 1). The potential alternative versions were similar and mainly differed from EDIH and EDIR in the number of components (Supplemental Table 1).

In the NHS, the proportion of overweight women in the highest quintile of both the EDIH and EDIR was ≈ 2 times higher than the proportion in the lowest quintile. Similarly, the proportion of lean and active participants was highest in quintile 1 and lowest in quintile 5. The proportion of participants with ≥ 3 chronic diseases/conditions in the highest quintile was >2 times higher than in the lowest quintile (Table 2). Both dietary indices showed moderate correlations with biomarkers. For example, the Spearman correlation coefficient was 0.21 for EDIH and C-peptide and 0.32 for EDIR and TG/HDL. The correlations were stronger for the two lifestyle indices, with correlations coefficients of 0.47 between ELIH and C-peptide and 0.46 between ELIR TG/HDL (Table 3). Also, the EDIH and EDIR were highly correlated with their potential alternative versions but correlations with the insulin index were low, e.g., while the EDIH had a correlation coefficient of 0.90 with the version developed in control subjects, its correlations with the insulin index was -0.07 . Corresponding correlations for the EDIR were 0.89 and 0.14 respectively (Supplemental Table 2).

In multivariable-adjusted models in the NHS, the EDIH and EDIR were significantly associated with C-peptide and TG/HDL. The C-peptide concentration of women in the highest quintile of the EDIH was 40% (95%CI; 34%, 46%; P -trend <0.0001) higher than that of women in the lowest quintile. Similarly, women in the highest quintile of the EDIR had a 67% (95%CI; 55%, 80%; P -trend <0.0001) higher concentration of TG/HDL than women in the lowest quintile. The corresponding contrasts for the ELIH and ELIR were 97% (95%CI; 89%, 106%) and 127% (95%CI; 111%, 145%), respectively (Table 4). Multivariable-adjusted analyses excluding women with diabetes were not materially different (Supplemental Table 3). In stratified analyses, there were large differences in C-peptide concentrations in EDIH quintiles across combinations of BMI,PA. Women in the overweight/obese and sedentary category had the highest concentrations of C-peptide while those in the lean, active category had the lowest concentrations. Also, there were significant trends of increasing TG/HDL concentrations within joint strata of BMI and physical activity (Figure 1).

In the validation studies using HPFS and NHS-II data, we observed similar trends in participant characteristics as in the NHS. Concentrations of C-peptide and TG/HDL increased monotonically across quintiles of their respective dietary and lifestyle indices. For example, between extreme index quintiles in the HPFS, there was a 25 and 82 percent increase in C-peptide for EDIH and ELIH respectively, and a 60 and 132 percent increase in TG/HDL for EDIR and ELIR respectively (Supplemental Table 4 for EDIH and EDIR and Supplemental Table 5 for ELIH and ELIR). Also, we found similar correlation patterns for the indices and biomarkers in the HPFS and NHS-II samples as in the NHS. That is, moderate correlations between dietary indices and biomarkers, stronger correlations between lifestyle indices and biomarkers (Table 3), and very strong correlations between dietary indices and potential alternative versions but low to moderate correlations with the insulin index and previously developed C-peptide dietary pattern (Supplemental Table 2). The insulin index was inversely correlated with C-peptide and with EDIH. In the HPFS the correlation between the EDIH and EDIR was 0.63.

All four indices were significantly associated with their respective biomarkers in HPFS and NHS-II, with stronger associations observed for the two lifestyle indices than their diet-only counterparts (Table 4). For example, in HPFS, the relative concentration of C-peptide was 29% (95% CI; 22%, 27%; P -trend<0.0001) higher in the highest quintile of EDIH compared to the lowest quintile; while the concentration of TG/HDL was 44% (95% CI; 34%, 55%; P -trend<0.0001) higher in quintile 5 of EDIR compared to quintile 1. Corresponding associations for the lifestyle indices were: 78% (95% CI; 68%, 88%; P -trend<0.0001 for ELIH and 103% (95% CI; 89%, 119%; P -trend<0.0001 for ELIR (Table 4). Excluding participants with diabetes did not materially change these findings (Supplemental Table 3). In HPFS, there were differences in concentrations of C-peptide and TG/HDL across index quintiles and in categories of BMI,PA, with overweight/obese and sedentary men having the highest biomarker levels compared with overweight/obese and active men or to lean, active or sedentary men (Figure 2). The proportion of participants with clinically high C-peptide concentrations across each EDIH quintile was 1.5 to 2 times higher among overweight/obese and sedentary men than among lean and active men, while the proportion with high TG/HDL levels was 2 to 3 times higher with EDIR quintile among overweight/obese and sedentary men than among lean and active men. Among men classified as lean and active, a higher proportion of those consuming diets with high insulinemic potential had clinically high biomarker levels than those consuming insulin sensitive diets (Figure 3).

Results from the sensitivity analyses in both men and women showed that associations between the dietary patterns developed only in control subjects and those with uncalibrated C-peptide and uncalibrated TG/HDL with biomarkers were reasonably similar to the associations obtained with the EDIH or EDIR. However, associations for the unweighted versions and the previously developed C-peptide pattern were smaller in magnitude. In contrast, the insulin index was not predictive of C-peptide concentrations in both men and women. Relative concentrations were: 0.94 (95% CI; 0.89, 1.00; P -trend=0.03) for men and 0.99 (95% CI; 0.91, 1.09; P -trend<0.90) for women, comparing extreme index quintiles, though there was a trend towards an inverse association in men. The insulin index however had a direct (but smaller compared to EDIR) association with TG/HDL in men 1.20 (95% CI; 1.11, 1.29; P -trend<0.0001) but not in women 1.12 (95% CI; 0.99, 1.26; P -trend=0.06). The previously developed C-peptide dietary pattern also had direct associations (though smaller in magnitude) with C-peptide concentrations in both men and women (Supplemental Table 6).

DISCUSSION

We developed two dietary and two lifestyle indices in a large cohort of women and evaluated their validity in two large independent cohorts of men and women. In all cohorts, the indices were predictive of both the absolute and relative concentrations of the insulin response biomarkers though the lifestyle indices were more predictive than the dietary indices. When we applied cut points that have been shown to discriminate between clinically high and low biomarker concentrations in adults, we found a consistently higher proportion of participants with high biomarker concentrations across index quintiles within subgroups defined by joint categories of BMI,PA, and across BMI,PA categories within index quintiles. These dietary indices assess the long term insulinemic potential of whole diets, which is in contrast

to the assessment of the acute postprandial glycemc or insulinemic potential of specific foods as has been done previously. In addition, the use of the TG/HDL ratio to derive the insulin resistance dietary pattern is novel. While our group previously used C-peptide concentrations to derive a hyperinsulinemia dietary pattern⁽¹⁷⁾, in the current study, we updated and strengthened this pattern by validating it in two independent cohorts of men and women. Several sensitivity analyses supported the robustness of the EDIH and EDIR.

The dietary patterns though empirical, align well with current knowledge. In concordance with the inverse associations found for whole fruits, green leafy vegetables, and coffee with hyperinsulinemia, other studies have shown that higher coffee intake as well as a plant-based diet that is high in fiber, fruit and whole grains is associated with lower concentrations of C-peptide^{(7; 40) (8)}. The dietary pattern predictive of insulin resistance is simultaneously influenced by factors that affect both triglycerides and high-density lipoprotein cholesterol. We found margarine, refined grains, processed meats, creamy soups and fruit juice to be positively associated with insulin resistance while nuts, alcohol, and green leafy vegetables to be inversely associated. Similarly, in previous studies, diets consisting of refined carbohydrates, and sweeteners, large amounts of saturated fats and trans fats (as in many cream-based sauces) have been associated with higher triglycerides concentrations, while higher intake of omega-3-fats such as in nuts and the moderate use of alcohol have been linked to higher levels of HDL cholesterol^(7; 41).

We found clinically relevant differences in biomarker concentrations both across dietary index quintiles and across BMI,PA categories. For example 73% of overweight/obese and sedentary men consuming the most pro-insulinemic diets had high C-peptide concentrations (1.8ng/mL) compared to only 37% of lean and active men consuming the least pro-insulinemic diets. Also, 72% of overweight/obese and sedentary men consuming the most insulin resistant diets had high TG/HDL levels (>3) compared to only 19% of lean and active men consuming the most insulin sensitive diets. These differences further strengthen the idea that these dietary indices can be useful in identifying populations at risk of hyperinsulinemia or insulin resistance. Our approach to create lifestyle indices (ELIH and ELIR) is complementary to the stratification of the diet-only indices (EDIH and EDIR) by BMI,PA combinations. Lifestyle indices assess the joint influence of diet, body weight and physical activity on hyperinsulinemia and insulin resistance, which is important for public health interventions. The indices assess the insulinemic potential of diet/lifestyle on a continuum from maximally low insulinemic to maximally high insulinemic potential with no optimal cut point for classifying individuals as absolutely high or low. Stratifying the diet-only indices by BMI,PA combinations accordingly to established clinically relevant biomarker cut points provides further insight on subgroups to target with specific dietary and or lifestyle interventions to reduce hyperinsulinemia and/or insulin resistance.

The differences between participants with clinically high and low biomarker levels within quintiles of the dietary indices were observed despite the low to moderate correlations between the indices and biomarkers. In previous studies, hypothesis-driven dietary patterns have shown low to moderate correlations with the biomarkers used to derive the patterns, yet these dietary patterns have shown robust associations with disease risk in independent populations^(42; 43). For example, Fung et al. reported a correlation coefficient of 0.23

between the dietary pattern predictive of C-peptide and C-peptide concentrations in NHS, though the pattern showed a significant positive association with colon cancer risk⁽¹⁷⁾. Also, a dietary inflammatory index showed low correlations with inflammatory markers yet strong associations with chronic diseases including cancer^(44; 45; 46). This suggests that correlations with biomarkers may not be a direct assessment of the performance of the dietary pattern in disease prediction or clinical significance. For example, among lean and active men, comparing the highest quintile of EDIR to the lowest, the prevalence of clinically high TG/HDL levels can potentially be reduced by >50% through diet interventions even though the EDIR had a low correlation ($r=0.15$) with TG/HDL. A low/moderate correlation may also be due to the dietary patterns not capturing other lifestyle behaviors that are associated with the biomarker. Interestingly, when lifestyle factors such as BMI and physical activity were included, the correlations between the lifestyle indices and biomarkers were >2 higher than that between the diet-only indices and biomarkers.

Our group previously created the dietary insulin index to quantify the short term (postprandial) insulin-secreting ability of specific foods⁽¹⁶⁾. This index was associated with higher triglycerides and lower HDL levels, with an indicative inverse association with C-peptide concentrations⁽¹⁶⁾. In the current study we compared the predictive ability of the four indices with the insulin index in sensitivity analyses. The insulin index was directly associated with TG/HDL, which is expected in the context of prevalent insulin resistance, but the correlation was much lower than that of our empirical indices with TG/HDL. Moreover, the index also showed an inverse trend of association with C-peptide concentrations, which at first seemed counterintuitive but may be understood in the context of our cross-sectional study design using fasting plasma samples. For example, in participants who may usually be consuming a high EDIH/high GI diet; such a diet will elicit higher insulin secretion to reduce the acute postprandial glycemia. The lowered glucose level will down-regulate further insulin secretion⁽⁴⁷⁾, and blood drawn a couple of hours into the fasting period will therefore show an inverse association (temporarily) between the insulin index (postprandial insulinemia) and insulin secretion (C-peptide concentration) which may not persist longitudinally.

Our study is not without limitations. We only had one measurement of the insulin markers, which may underestimate validity assessed by correlation coefficients.⁽⁴⁸⁾ Given that food intake was self-reported, some measurement error is inevitable, though the validation data showed reasonably good correlations between FFQ and diet records suggesting that dietary intake is generally well measured in our cohorts^(29; 30; 31). The composition of food groups may not be uniform across studies, which would limit the ability to apply the indices across studies in a standardized manner, though investigators may be able to create unified food groups in pooled analyses of primary data or in multi-center studies and thus enhance the usefulness of these hypothesis-driven dietary patterns in large scale epidemiologic research. Study participants in all three cohorts are mostly Caucasian health professionals, but the distributions of most participant characteristics in the three cohorts are generally similar to that of the larger US multi-racial/ethnic population. It is important however to further apply the indices in multi-racial/ethnic populations. Other lifestyle factors include smoking and exogenous hormone use but we focused mainly on BMI and PA in the lifestyle indices because these have been shown to be strongly associated with circulating insulin

markers^(11; 12; 13; 14). We adjusted for a large number of potential confounding variables including a history of diabetes and other chronic diseases/conditions, but these variables were self-reported, thus allowing the possibility of residual confounding. However, results from the age-adjusted and multivariable-adjusted models were very similar in all cohorts, suggesting that any confounding would have been very minimal.

CONCLUSION

These novel hypothesis-driven empirically derived dietary and lifestyle indices assess dietary and lifestyle quality based on insulinemic potential. Their robust associations with the insulin response biomarkers in independent samples suggest their usefulness in assessing the ability of whole diets and lifestyles to stimulate and/or sustain insulin secretion. The indices can be useful in identifying populations at high risk for hyperinsulinemia or insulin resistance. Additionally, the indices may be calculated in a standardized and reproducible manner across different populations thus circumventing a major limitation of dietary patterns derived in the same study in which they are applied. Moreover, studies without insulin markers data may calculate the index scores to investigate associations between dietary and lifestyle insulinemic potential and disease outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

FINANCIAL SUPPORT

Drs. Jorge E. Chavarro and Frank B. Hu were supported by National Institutes of Health (NIH) grants P30DK046200 and U54 CA155426. The HPFS, NHS and NHS-II cohorts are supported by the following NIH grants: UMI CA 167552, UMI CA 186107 and UMI CA 176726 respectively.

References

1. Yoon Y, Keum N, Zhang X, Cho E, Giovannucci EL. 2015; Hyperinsulinemia, insulin resistance and colorectal adenomas: A meta-analysis. *Metabolism - Clinical and Experimental*. 64: 1324–1333. [PubMed: 26169471]
2. Chen L, Li L, Wang Y, Li P, Luo L, Yang B, Wang H, Chen M. 2013; Circulating C-peptide level is a predictive factor for colorectal neoplasia: evidence from the meta-analysis of prospective studies. *Cancer Causes & Control*. 24: 1837–1847. [PubMed: 23846284]
3. Giovannucci E. 1995; Insulin and colon cancer. *Cancer Causes and Control*. 6: 164–179. [PubMed: 7749056]
4. Fernández-Veledo S, Nieto-Vazquez I, Vila-Bedmar R, Garcia-Guerra L, Alonso-Chamorro M, Lorenzo M. 2009; Molecular mechanisms involved in obesity-associated insulin resistance: therapeutical approach. *Archives of Physiology and Biochemistry*. 115: 227–239. [PubMed: 19673658]
5. Kahn B, Flier JS. 2000; Obesity and insulin resistance. *Journal of Clinical Investigation*. 106: 473–481. [PubMed: 10953022]
6. Reaven GM. 1995; Pathophysiology of insulin resistance in human disease. *Physiological Reviews*. 75: 473–486. [PubMed: 7624391]

7. Jensen M, Koh-Banerjee P, Franz M, Sampson L, Grønbaek M, Rimm EB. 2006; Whole grains, bran, and germ in relation to homocysteine and markers of glycemic control, lipids, and inflammation. *The American Journal of Clinical Nutrition*. 83: 275–283. [PubMed: 16469984]
8. Wu T, Willett WC, Hankinson SE, Giovannucci E. 2005; Caffeinated Coffee, Decaffeinated Coffee, and Caffeine in Relation to Plasma C-Peptide Levels, a Marker of Insulin Secretion, in U.S. Women. *Diabetes Care*. 28: 1390–1396. [PubMed: 15920057]
9. Hu FB. 2002; Dietary pattern analysis: a new direction in nutritional epidemiology. *Current Opinion in Lipidology*. 13: 3–9. [PubMed: 11790957]
10. Jacques PF, Tucker KL. 2001; Are dietary patterns useful for understanding the role of diet in chronic disease? *The American Journal of Clinical Nutrition*. 73: 1–2. [PubMed: 11124739]
11. Garca-Estévez D, Araújo-Vilar D, Saavedra-González A, Fiestras-Janeiro G, Cabezas-Cerrato J. 2004; Analysis of the relationship between body mass index, insulin resistance, and beta-cell function: A cross-sectional study using the minimal model. *Metabolism - Clinical and Experimental*. 53: 1462–1466. [PubMed: 15536602]
12. Tatsumi Y, Morimoto A, Miyamatsu N, Noda M, Ohno Y, Deura K. 2015; Effect of Body Mass Index on Insulin Secretion or Sensitivity and Diabetes. *American Journal of Preventive Medicine*. 48: 128–135. [PubMed: 25442224]
13. Dubé JJ, Allison KF, Rousson V, Goodpaster BH, Amati F. 2012; Exercise dose and insulin sensitivity: relevance for diabetes prevention. *Medicine and Science in Sports and Exercise*. 44: 793–799. [PubMed: 22051572]
14. Borghouts L, Keizer HA. 2000; Exercise and insulin sensitivity: a review. *International Journal of Sports Medicine*. 21: 1–12. [PubMed: 10683091]
15. Jenkins D, Wolever TM, Taylor RH, Barker H, Fielden H, Baldwin JM, Bowling AC, Newman HC, Jenkins AL, Goff DV. 1981; Glycemic index of foods: a physiological basis for carbohydrate exchange. *The American Journal of Clinical Nutrition*. 34: 362–366. [PubMed: 6259925]
16. Nimptsch K, Brand-Miller JC, Franz M, Sampson L, Willett WC, Giovannucci E. 2011; Dietary insulin index and insulin load in relation to biomarkers of glycemic control, plasma lipids, and inflammation markers. *The American Journal of Clinical Nutrition*. 94: 182–190. [PubMed: 21543531]
17. Fung TT, Hu FB, Schulze M, Pollak M, Wu T, Fuchs CS, Giovannucci E. 2012; A dietary pattern that is associated with C-peptide and risk of colorectal cancer in women. *Cancer Causes & Control*. 23: 959–965. [PubMed: 22535146]
18. Colditz G, Hankinson SE. 2005; The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer*. 5: 388–396. [PubMed: 15864280]
19. Pai J, Pischon T, Ma J, Manson JE, Hankinson SE, Joshipura K, Curhan GC, Rifai N, Cannuscio CC, Stampfer MJ, Rimm EB. 2004; Inflammatory markers and the risk of coronary heart disease in men and women. *New England Journal of Medicine*. 351: 2599–2610. [PubMed: 15602020]
20. Hankinson SE, Willett WC, Manson JE, et al. 1995; Alcohol, height, and adiposity in relation to estrogen and prolactin levels in postmenopausal women. *J Natl Cancer Inst*. 87: 1297–1302. [PubMed: 7658481]
21. Bonser A, Garcia-Webb P. 1984; C-Peptide Measurement: Methods and Clinical Utility. *Critical Reviews in Clinical Laboratory Sciences*. 19: 297–352. [PubMed: 6373142]
22. Olson K, Hendricks B, Murdock DK. 2012; The triglyceride to HDL ratio and its relationship to insulin resistance in pre- and postpubertal children: observation from the Wausau SCHOOL Project. *Cholesterol*. 2012: 794252. [PubMed: 22811895]
23. Murguía-Romero M, Jiménez-Flores JR, Sigrist-Flores SC, Espinoza-Camacho MA, Jiménez-Morales M, Piña E, Méndez-Cruz AR, Villalobos-Molina R, Reaven GM. 2013; Plasma triglyceride/HDL-cholesterol ratio, insulin resistance, and cardiometabolic risk in young adults. *Journal of Lipid Research*. 54: 2795–2799. [PubMed: 23863983]
24. Salazar M, Carbajal HA, Espeche WG, Leiva Sisniegues CE, March CE, Balbín E, Dulbecco CA, Aizpurúa M, Marillet AG, Reaven GM. 2013; Comparison of the abilities of the plasma triglyceride/high-density lipoprotein cholesterol ratio and the metabolic syndrome to identify insulin resistance. *Diabetes and Vascular Disease Research*. 10: 346–352. [PubMed: 23624761]

25. McLaughlin T, Abbasi F, Cheal K, Chu J, Lamendola C, Reaven G. 2003; Use of metabolic markers to identify overweight individuals who are insulin resistant. *Annals of Internal Medicine*. 139: 802–809. [PubMed: 14623617]
26. Shai I, Rimm EB, Hankinson SE, Curhan G, Manson JE, Rifai N, Stampfer MJ, Ma J. 2004; Multivariate assessment of lipid parameters as predictors of coronary heart disease among postmenopausal women: potential implications for clinical guidelines. *Circulation*. 110: 2824–2830. [PubMed: 15492318]
27. Willett W, Stampfer M, Chu NF, Spiegelman D, Holmes M, Rimm E. 2001; Assessment of questionnaire validity for measuring total fat intake using plasma lipid levels as criteria. *American Journal of Epidemiology*. 154: 1107–1112. [PubMed: 11744515]
28. Rosner B, Cook N, Portman R, Daniels S, Falkner B. 2008; Determination of blood pressure percentiles in normal-weight children: some methodological issues. *American Journal of Epidemiology*. 167: 653–666. [PubMed: 18230679]
29. Willett W, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, Speizer FE. 1985; Reproducibility and validity of a semiquantitative food frequency questionnaire. *American Journal of Epidemiology*. 122: 51–65. [PubMed: 4014201]
30. Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB, Willett WC. 1993; Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *Journal of the American Dietetic Association*. 93: 790–796. [PubMed: 8320406]
31. Rimm E, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. 1992; Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *American Journal of Epidemiology*. 135: 1114–1126. [PubMed: 1632423]
32. Hu FB, Rimm E, Smith-Warner SA, Feskanich D, Stampfer MJ, Ascherio A, Sampson L, Willett WC. 1999; Reproducibility and validity of dietary patterns assessed with a food frequency questionnaire. *The American Journal of Clinical Nutrition*. 69: 243–249. [PubMed: 9989687]
33. Chasan-Taber S, Rimm EB, Stampfer MJ, Spiegelman D, Colditz GA, Giovannucci E, Ascherio A, Willett WC. 1996; Reproducibility and Validity of a Self-Administered Physical Activity Questionnaire for Male Health Professionals. *Epidemiology*. 7: 81–86. [PubMed: 8664406]
34. Wolf AM, Hunter DJ, Colditz GA, Manson JE, Stampfer MJ, Corsano KA, Rosner B, Kriska A, Willett WC. 1994; Reproducibility and Validity of a Self-Administered Physical Activity Questionnaire. *International Journal of Epidemiology*. 23: 991–999. [PubMed: 7860180]
35. Mason C, Foster-Schubert KE, Imayama I, Kong A, Xiao L, Bain C, Campbell KL, Wang CY, Duggan CR, Ulrich CM, Alfano CM, Blackburn GL, McTiernan A. 2011; Dietary weight-loss and exercise effects on insulin resistance in postmenopausal women. *American journal of preventive medicine*. 41: 366–375. [PubMed: 21961463]
36. Torjesen PABK, Anderssen SA, Hjermann I, Holme I, Urdal P. 1997; Lifestyle changes may reverse development of the insulin resistance syndrome. *The Oslo Diet and Exercise Study: a randomized trial*. *Diabetes Care*. 20: 26–31. [PubMed: 9028689]
37. Berger B, Stenström G, Sundkvist G. 2000; Random C-peptide in the classification of diabetes. *Scandinavian Journal of Clinical and Laboratory Investigation*. 60: 687–693. [PubMed: 11218151]
38. Jones AG, Hattersley AT. 2013; The clinical utility of C-peptide measurement in the care of patients with diabetes. *Diabetic Medicine*. 30: 803–817. [PubMed: 23413806]
39. Palamaner SSG, Kumar AA, Kahan S, Irukulla PK, Cheskin LJ. 2013; Triglyceride/HDL ratio as a screening tool for predicting success at reducing anti-diabetic medications following weight loss. *PLoS ONE*. 8: e69285. [PubMed: 23869240]
40. Bobe G, Murphy G, Rogers CJ, Hance KW, Albert PS, Laiyemo AO, Sansbury LB, Lanza E, Schatzkin A, Cross AJ. 2010; Serum adiponectin, leptin, C-peptide, homocysteine, and colorectal adenoma recurrence in the Polyp Prevention Trial. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 19: 1441–1452.
41. Dhingra R, Sullivan L, Jacques PF, Wang TJ, Fox CS, Meigs JB, D'Agostino RB, Gaziano JM, Vasani RS. 2007; Soft drink consumption and risk of developing cardiometabolic risk factors

- and the metabolic syndrome in middle-aged adults in the community. *Circulation*. 116: 480–488. [PubMed: 17646581]
42. Fung TT, Schulze MB, Hu FB, Hankinson SE, Holmes MD. 2012; A dietary pattern derived to correlate with estrogens and risk of postmenopausal breast cancer. *Breast cancer research and treatment*. 132: 1157–1162. [PubMed: 22218885]
43. Harris HR, Bergkvist L, Wolk A. 2015; An estrogen-associated dietary pattern and breast cancer risk in the Swedish Mammography Cohort. *International Journal of Cancer*. 137: 2149–2154. [PubMed: 25924604]
44. Shivappa N, Bosetti C, Zucchetto A, Serraino D, La Vecchia C, Hébert JR. 2015; Dietary inflammatory index and risk of pancreatic cancer in an Italian case-control study. *British Journal of Nutrition*. 113: 292–298. [PubMed: 25515552]
45. Tabung FK, Steck SE, Ma Y, Liese AD, Zhang J, Caan B, Hou L, Johnson KC, Mossavar-Rahmani Y, Shivappa N, Wactawski-Wende J, Ockene JK, Hebert JR. 2015; The association between dietary inflammatory index and risk of colorectal cancer among postmenopausal women: results from the Women's Health Initiative. *Cancer Causes & Control*. 26: 399–408. [PubMed: 25549833]
46. Tabung FK, Steck SE, Zhang J, Ma Y, Liese AD, Agalliu I, Hou L, Hurley TG, Hingle M, Jiao L, Martin LW, Millen EA, Park HL, Rosal CM, Shikany JM, Shivappa N, Ockene JK, Hebert JR. 2015; Construct validation of the dietary inflammatory index among postmenopausal women. *Annals of Epidemiology*. 25: 398–405. [PubMed: 25900255]
47. Last AR, Wilson SA. 2006; Low-carbohydrate diets. *Am Fam Physician*. 73: 1942–1948. [PubMed: 16770923]
48. Perrier F, Giorgis-Allemand Li, Slama R, Philippat C. 2016; Within-subject pooling of biological samples to reduce exposure misclassification in biomarker-based studies. *Epidemiology*. 27: 378–388. [PubMed: 27035688]
49. Quan H, Zhang J. 2003; Estimate of standard deviation for a log-transformed variable using arithmetic means and standard deviations. *Statistics in Medicine*. 22: 2723–2736. [PubMed: 12939782]

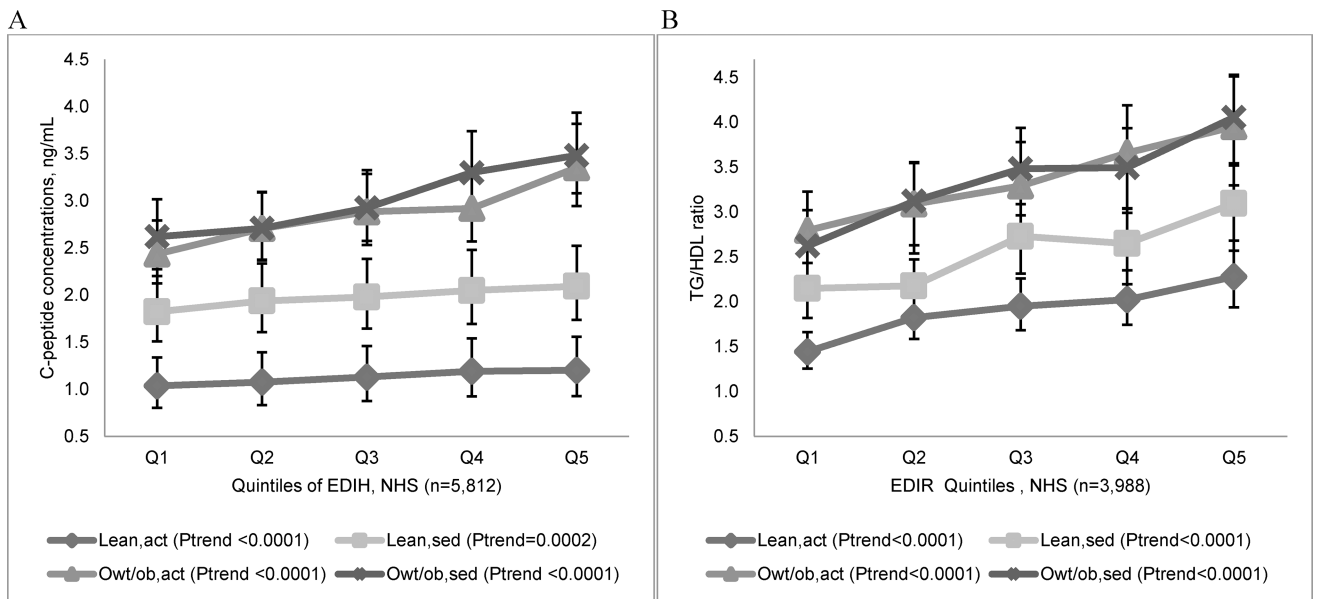


Figure 1. Multivariable-adjusted biomarker concentrations across quintiles of the (A) empirical dietary index for hyperinsulinemia (EDIH) and (B) empirical dietary index for insulin resistance (EDIR), stratified by joint categories of body mass index (BMI) and physical activity (PA) in the Nurses’ Health Study (NHS), 1990. Values are back transformed (e^x , where x is the transformed biomarker value) predicted mean fasting plasma biomarker concentrations, obtained from linear regression models, adjusted for regular aspirin/NSAIDs use, age, smoking status, physical activity, menopausal status, postmenopausal hormone use, diabetes, other chronic diseases/conditions, case-control status. The *P*-value for trend was the *P*-value of the dietary index as a continuous index variable adjusted for all covariates. Categories of BMI and PA combinations were created as follows: lean and active (lean,act; BMI <25kg/m² and PA ≥ median PA), lean and sedentary (lean, sed; BMI <25kg/m² and PA <median PA), overweight/obese and active (owt/ob, sed; BMI ≥ 25kg/m² and PA ≥ median PA) and overweight/obese and sedentary (owt/ob, act; BMI ≥ 25kg/m² and PA <median PA). Median PA=10.2 MET-hrs/week for women with C-peptide data, and 9.10 MET-hrs/week for those with TG/HDL data.

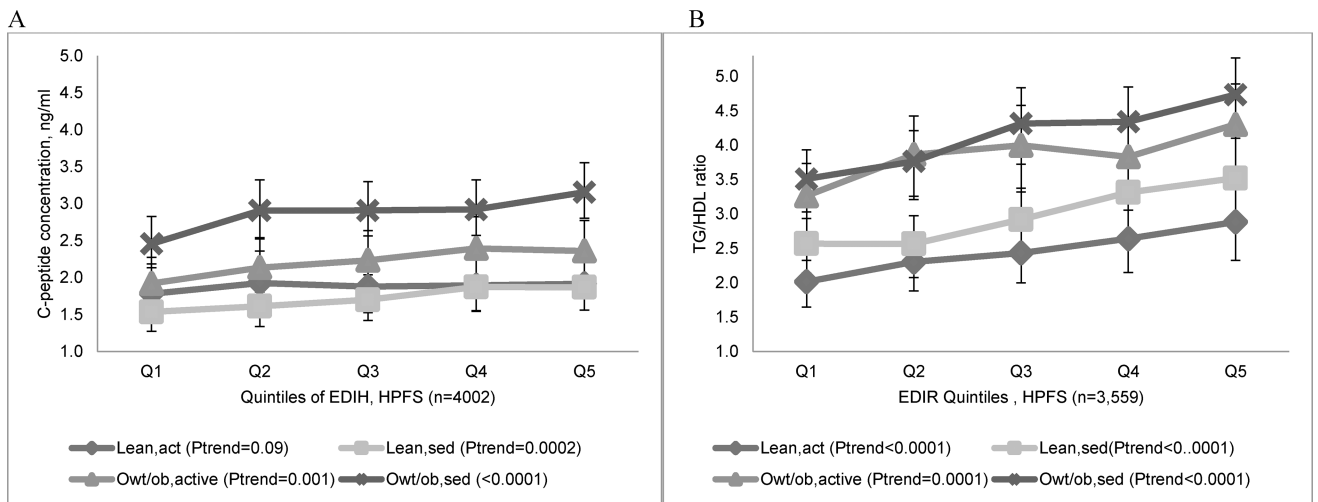


Figure 2. Multivariable-adjusted biomarker concentrations across quintiles of the (A) empirical dietary index for hyperinsulinemia (EDIH) and (B) empirical dietary index for insulin resistance (EDIR), stratified by joint categories of body mass index (BMI) and physical activity (PA) in the Health Professional Follow-up Study (HPFS), 1994. Values are back transformed (e^x , where x is the transformed biomarker value) predicted mean fasting plasma biomarker concentrations, obtained from linear regression models, adjusted for regular aspirin/NSAIDs use, age, smoking status, physical activity, diabetes, other chronic diseases/conditions, case-control status. The P -value for trend was the P -value of the dietary index as a continuous index variable adjusted for all covariates. Categories of BMI and PA combinations were created as follows: lean and active (lean,act; BMI $<25\text{kg/m}^2$ and PA \geq median PA), lean and sedentary (lean,sed; BMI $<25\text{kg/m}^2$ and PA $<$ median PA), overweight/obese and active (owt/ob,act; BMI $\geq 25\text{kg/m}^2$ and PA \geq median PA) and overweight/obese and sedentary (owt/ob,sed; BMI $\geq 25\text{kg/m}^2$ and PA $<$ median PA). Median PA=28.1 MET-hrs/week for men with C-peptide data, and 24.8 MET-hrs/week for men with TG/HDL data.

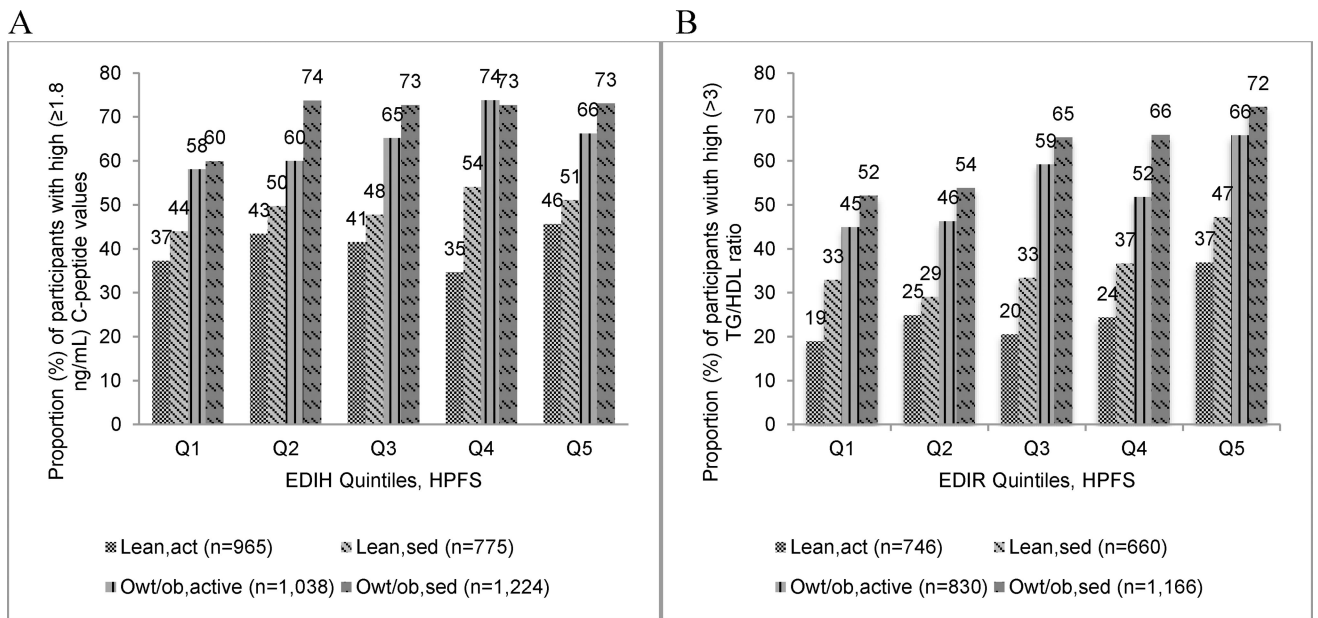


Figure 3. Distribution of participants (%) with clinically high levels of biomarkers in quintiles (Q) of dietary indices and in joint categories of body mass index/physical activity (PA) combinations, Health Professionals Follow-up Study (HPFS), 1994. Categories of BMI and PA combinations were created as follows: lean and active (lean,act; BMI <25kg/m² and PA ≥median PA), lean and sedentary (lean,sed; BMI <25kg/m² and PA <median PA), overweight/obese and active (owt/ob,act; BMI ≥25kg/m² and PA ≥median PA) and overweight/obese and sedentary (owt/ob,sed; BMI ≥25kg/m² and PA <median PA). Median PA=28.1 MET-hrs/week for men with C-peptide data, and 24.8 MET-hrs/week for men with TG/HDL data.

Table 1

Components of the indices to assess the insulinemic potential of diet and lifestyle; Nurses' Health Study; 1990

| Empirical dietary index for hyperinsulinemia (EDIH) | Empirical lifestyle index for hyperinsulinemia (ELIH) | Empirical dietary index for insulin resistance (EDIR) | Empirical lifestyle index for insulin resistance (ELIR) | | | | | | |
|---|---|---|---|------------------------------|-----------|--------------------------------|---------|------------------------------|--|
| Food group* | Weight† | R-square‡ | Food group* | Weight† | R-square‡ | Food group* | Weight† | R-square‡ | |
| <i>Positive associations</i> | | | | <i>Positive associations</i> | | | | <i>Positive associations</i> | |
| Red meat | 0.250 | 0.008 | Body mass index (kg/m ²) | 0.051 | 0.187 | Low energy beverages | 0.116 | 0.014 | |
| Low energy beverages | 0.053 | 0.004 | Margarine | 0.041 | 0.001 | Margarine | 0.121 | 0.013 | |
| Cream soups | 0.787 | 0.003 | Liquor | 0.072 | 0.001 | Red meat | 0.328 | 0.009 | |
| Processed meat | 0.199 | 0.002 | Cream soups | 0.536 | 0.001 | Refined grains | 0.102 | 0.006 | |
| Margarine | 0.054 | 0.002 | Butter | 0.058 | 0.001 | processed meats | 0.327 | 0.004 | |
| Poultry | 0.183 | 0.002 | Red meat | 0.089 | 0.001 | Tomatoes | 0.145 | 0.002 | |
| Butter | 0.094 | 0.001 | Fruit juice | 0.042 | 0.001 | Other vegetables | 0.126 | 0.001 | |
| <i>Inverse associations</i> | | | | <i>Inverse associations</i> | | | | <i>Inverse associations</i> | |
| French fries | 0.581 | 0.001 | Coffee | -0.020 | 0.002 | Fruit juice | 0.052 | 0.001 | |
| Other fish | 0.172 | 0.001 | Whole fruit | -0.029 | 0.002 | Creamy soups | 0.519 | 0.001 | |
| High energy beverages | 0.104 | 0.001 | Wine | -0.071 | 0.002 | Other fish | 0.155 | 0.001 | |
| Tomatoes | 0.095 | 0.001 | Physical activity (MET-hrs/wk) | -0.001 | 0.001 | Other fish | 0.155 | 0.001 | |
| Low fat dairy | 0.025 | 0.001 | High fat diary | -0.054 | 0.001 | Coffee | -0.070 | 0.018 | |
| Eggs | 0.124 | 0.001 | Snacks | -0.024 | 0.001 | Wine | -0.261 | 0.011 | |
| <i>Inverse associations</i> | | | | <i>Inverse associations</i> | | | | <i>Inverse associations</i> | |
| Wine | -0.165 | 0.009 | Salad dressing | -0.059 | 0.001 | Liquor | -0.204 | 0.006 | |
| Coffee | -0.035 | 0.005 | High fat dairy | -0.024 | 0.001 | Beer | -0.210 | 0.002 | |
| Whole fruits | -0.029 | 0.003 | Physical activity (MET-hrs/wk) | -0.066 | 0.001 | Green leafy vegetables | -0.076 | 0.001 | |
| High fat dairy | -0.046 | 0.001 | Dark yellow vegetables | -0.103 | 0.001 | High fat dairy | -0.064 | 0.001 | |
| Green leafy vegetables | -0.055 | 0.001 | Nuts | -0.078 | 0.001 | Physical activity (MET-hrs/wk) | -0.001 | 0.001 | |
| | | | | | | Green leafy vegetables | -0.064 | 0.001 | |

*The food groups (servings/d) retained were defined as follows: **red meats** (4–6 oz beef, 4–6 oz pork, 4–6 oz lamb, 1 patty hamburger); **processed meat** (1 piece or 1 slice processed meats, 2 slices bacon, 1 hot dog); **low energy beverages** (1 glass, 1 bottle or 1 can of low-energy cola, other low-energy carbonated beverages); **cream soups** (1 cup chowder or cream soup); 1 pat **margarine**; **poultry** (4–6 oz chicken or turkey with or without skin); **high energy beverages** (1 glass, 1 bottle or 1 can of cola with sugar, other carbonated beverages with sugar, fruit punch drinks); 1 pat **butter**; 4 oz **french fries**; **other fish** (3–5 oz canned tuna, shrimp, lobster, scallops, fish and other seafood other than dark meat fish); **low-fat dairy products** (8oz glass skim or low-fat milk, ½ cup sherbet or ice milk, 1 cup yogurt); **tomatoes** (1 fresh tomato, 1 small glass of tomato juice, ½ cup of tomato sauce); 1 **egg**; **cruciferous vegetables** (1/2 cup of broccoli, coleslaw and uncooked cabbage; cooked cabbage; cauliflower; brussels sprouts; kale, mustard, and chard greens; sauerkraut); **wine** (4 oz glass of red wine, white wine); 1 cup **coffee**; **high-fat dairy products** (8 oz glass whole milk, cream, 1 tbs sour cream, ½ cup ice cream, 1 oz cream cheese, 1 oz or 1 slice other cheese); **green leafy vegetables** (1/2 cup spinach, serving of iceberg or head lettuce, serving of romaine or leaf lettuce); **whole fruit** (1oz or small pack raisins, 1/2 grapes, 1 avocado, 1 banana, ¼ melon cantaloupe, 1 slice watermelon, 1 fresh apple or pear or ½ cup canned, 1 orange, ½ grapefruit, ½ cup strawberries, ½ cup blueberries, 1 fresh or ½ cup canned apricots or plums); **dark yellow vegetables** (1/2 cup carrots, ½ cup yellow (winter) squash, ½ cup yams, ½ cup sweet potatoes); **snacks** (1 small bag or 1 oz potato chips, corn chips or popcorn, 1 crackers); 1 pat **butter**; **fruit juice** (1 small glass of apple juice or cider, orange juice, grapefruit juice, other fruit juice); **liquor** (1 drink or 1 shot whiskey, gin etc); **salad dressing** (1 tbs oil and vinegar salad dressing); **refined grains** (1 slice white bread, 1 English muffin, 1 bagel or roll, 1 muffin or biscuit, 1 cup white rice, 1 cup pasta, 1 serving of

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

pancakes or waffles); **beer** (1 glass, 1 bottle or 1 can); 1 cup **coffee**; 1 cup **tea** (not herbal tea); **Potatoes** (1 baked or boiled, 1 cup mashed); ; **other vegetables** (4 inch stick celery, 1 fresh, cooked or canned mushroom, ½ green pepper, 1 ear or ½ cup frozen or canned corn, ½ cup mixed vegetables, eggplant, ½ cup zucchini, ½ cup alfalfa sprouts, ¼ cucumber); **whole grains** (1 cup cooked oatmeal, 1 cup other cooked breakfast cereal, 1 slice dark bread, 1 cup brown rice, 1 cup other grains, 1 tbs bran added to food, 1 tbs wheat germ).

⁷Weights are regression coefficients derived from the final step of the stepwise linear regression models. Each weight represents the contribution of the corresponding index component to the total weighted index score;

⁸The partial R-square represents the proportion of variance in biomarkers explained by the index component;

Participant characteristics in quintiles (Q) of the insulin dietary patterns; Nurses' Health Study; 1990

Table 2

| | Empirical dietary index for hyperinsulinemia-EDIH (n=5,812) | | | | Empirical dietary index for insulin resistance-EDIR (n=3,929) | | | |
|--|--|--------------------------------|-------------------------------|------------------------------|--|-------------------------------|------------------------------|------------------------------|
| | Q1 (n=1162) (-0.57 to <0.09) | Q3 (n=1162) (0.16 to <0.21) | Q5 (n=1162) (0.28 to 0.75) | Q5 (n=786) (0.39 to 1.25) | Q1 (n=785) (-1.15 to <0.06) | Q3 (n=786) (0.19 to <0.28) | Q5 (n=786) (0.39 to 1.25) | Q5 (n=786) (0.39 to 1.25) |
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Fasting C-peptide, ng/mL * | 1.8 | 0.9 | 2.2 | 0.9 | 2.6 | 0.9 | NA | NA |
| Fasting TG/HDL * | NA | NA | NA | NA | NA | NA | 1.0 | 3.19 |
| Insulin index | 42.5 | 5.8 | 43.0 | 4.8 | 41.8 | 4.3 | 43.4 | 4.3 |
| Age, years | 59.3 | 6.5 | 58.2 | 6.9 | 56.2 | 7.4 | 59.6 | 6.3 |
| Physical activity, MET-hr/wk | 18.6 | 18.7 | 17.5 | 30.5 | 12.9 | 21.4 | 17.4 | 18.7 |
| Alcohol, servings/d [†] | 0.7 | 1.0 | 0.4 | 0.7 | 0.3 | 0.7 | 1.1 | 1.3 |
| Body mass index (BMI), kg/m ² | 24.0 | 3.6 | 25.7 | 4.3 | 27.7 | 5.6 | 24.3 | 3.9 |
| | n | % | n | % | n | % | n | % |
| Overweight/obese (> 25kg/m ²) | 373 | 32.1 | 570 | 49.1 | 728 | 62.7 | 279 | 35.5 |
| Current smokers | 125 | 10.8 | 116 | 10.0 | 146 | 12.6 | 247 | 31.5 |
| Regular aspirin/NSAIDs users | 393 | 33.8 | 436 | 37.5 | 484 | 41.7 | 277 | 35.3 |
| BMI-physical activity combinations [‡] | | | | | | | | |
| Lean and active | 479 | 41.2 | 334 | 28.7 | 203 | 17.5 | 281 | 35.8 |
| Lean and sedentary | 310 | 26.7 | 258 | 22.2 | 231 | 19.9 | 225 | 28.7 |
| Overweight/obese and active | 211 | 18.2 | 272 | 23.4 | 259 | 22.3 | 165 | 21.0 |
| Overweight/obese and sedentary | 162 | 13.9 | 298 | 25.7 | 469 | 40.4 | 114 | 14.5 |
| Chronic disease/comorbidities comorbidity score [§] | | | | | | | | |
| no chronic disease/condition | 517 | 44.5 | 521 | 44.8 | 517 | 44.5 | 356 | 45.4 |
| 1 chronic disease/condition | 415 | 35.7 | 376 | 32.4 | 374 | 32.2 | 261 | 33.3 |
| 2 chronic diseases/conditions | 184 | 15.8 | 183 | 15.8 | 199 | 17.1 | 116 | 14.8 |
| 3 chronic diseases/conditions | 46 | 4.0 | 82 | 7.1 | 72 | 6.2 | 52 | 6.6 |
| Diabetes (yes) | 6 | 0.5 | 17 | 1.5 | 23 | 2.0 | 46 | 5.9 |
| | | | | | | | 108 | 13.7 |
| | | | | | | | 259 | 33.0 |

| | Empirical dietary index for hyperinsulinemia-EDIH (n=5,812) | | Empirical dietary index for insulin resistance-EDIR (n=3,929) | | | | | | | | | |
|-----------------------------|--|--------------------------------|--|--------------------------------|-------------------------------|------------------------------|-----|------|-----|------|-----|------|
| | Q1 (n=1162) (-0.57 to <0.09) | Q3 (n=1162) (0.16 to <0.21) | Q5 (n=1162) (0.28 to 0.75) | Q1 (n=785) (-1.15 to <0.06) | Q3 (n=786) (0.19 to <0.28) | Q5 (n=786) (0.39 to 1.25) | | | | | | |
| | Mean | SD | Mean | SD | Mean | SD | | | | | | |
| Postmenopausal women | 1024 | 88.1 | 974 | 83.8 | 889 | 76.5 | 726 | 92.5 | 714 | 90.8 | 697 | 88.7 |
| Postmenopausal hormone user | 679 | 58.4 | 633 | 54.5 | 535 | 46.0 | 490 | 62.4 | 455 | 57.9 | 424 | 53.9 |

* Geometric means (coefficient of variation, CV) are presented for the biomarkers (fasting plasma samples) because all biomarkers were log transformed prior to analyses; The Quan-Zhang formula, $CV=(e^{SD_1})/2$ (49) was used to calculate CV's;

† Alcohol was the sum of wine (4 oz glass), beer (1 bottle, can or glass) and liquor (1 drink or shot) intake;

‡ Categories of BMI and physical activity (PA) combinations were created as follows: lean and active (BMI <25kg/m² and PA median PA), lean and sedentary (BMI <25kg/m² and PA <median PA), overweight/obese and active (BMI ≥25kg/m² and PA median PA) and overweight/obese and sedentary (BMI ≥25kg/m² and PA <median PA). Median PA=10.2 MET-hrs/week for women with C-peptide data, and 9.10 MET-hrs/week for those with TG/HDL data;

§ Chronic diseases/conditions included in the score were hypercholesterolemia, cancer, diabetes, high blood pressure, heart disease, and rheumatoid/other arthritis; NA=not applicable;

Table 3

Spearman correlations coefficients among the insulinemic dietary and lifestyle patterns and fasting plasma biomarker concentrations in the three cohorts

| Empirical dietary indices for hyperinsulinemia | C-peptide | | |
|--|-----------|--------|-------|
| | NHS | NHS-II | HPFS |
| C-peptide | 1 | 1 | 1 |
| EDIH | 0.21 | 0.20 | 0.14 |
| ELIH | 0.47 | 0.43 | 0.36 |
| Unweighted EDIH | 0.16 | 0.16 | 0.09 |
| Unweighted ELIH | 0.28 | 0.24 | 0.19 |
| EDIH in controls | 0.20 | 0.19 | 0.14 |
| EDIH with unadjusted C-peptide | 0.20 | 0.21 | 0.13 |
| Previously developed C-peptide dietary pattern | 0.11 | 0.12 | 0.09 |
| Insulin index | -0.03 | -0.03* | -0.06 |
| Empirical dietary indices for insulin resistance | TG/HDL | | |
| | NHS | NHS-II | HPFS |
| TG/HDL | 1 | 1 | 1 |
| EDIR | 0.32 | 0.16 | 0.21 |
| ELIR | 0.46 | 0.35 | 0.39 |
| Unweighted EDIR | 0.28 | 0.10 | 0.19 |
| Unweighted ELIR | 0.27 | 0.24 | 0.16 |
| EDIR in controls | 0.28 | 0.16 | 0.18 |
| EDIR with adjusted TG, HDL | 0.31 | 0.18 | 0.21 |
| Insulin index | 0.06 | 0.07 | 0.05 |

NHS, Nurses' Health Study, 1990; NHS-II, Nurses' Health Study-II, 1999; HPFS, Health Professional Follow-up Study, 1994; EDIH, empirical dietary index for hyperinsulinemia; ELIH, empirical lifestyle index for hyperinsulinemia; EDIR, empirical dietary index for insulin resistance; ELIR, empirical lifestyle index for insulin resistance; TG, triglyceride, HDL, high density lipoprotein cholesterol;

* $P > 0.05$

Adjusted* relative concentrations† of biomarkers in quintiles of insulinemic dietary and lifestyle patterns in the three cohorts

Table 4

| | Quintile 1 (reference) | Quintile 2 | Quintile 3 | Quintile 4 | Quintile 5 | P-trend‡ |
|---|---------------------------|-------------------|-------------------|-------------------|-------------------|----------|
| Empirical dietary index for hyperinsulinemia (EDIH) | | | | | | |
| C-peptide (NHS, n=5,812) | | | | | | |
| Age-adjusted | 1 | 1.09 (1.04, 1.15) | 1.20 (1.14, 1.25) | 1.28 (1.23, 1.35) | 1.44 (1.38, 1.51) | <0.0001 |
| Multivariable-adjusted | 1 | 1.09 (1.04, 1.14) | 1.18 (1.13, 1.24) | 1.27 (1.21, 1.32) | 1.40 (1.34, 1.46) | <0.0001 |
| C-peptide (HPFS, n=4,002) | | | | | | |
| Age-adjusted | 1 | 1.13 (1.07, 1.20) | 1.16 (1.09, 1.23) | 1.21 (1.14, 1.29) | 1.29 (1.22, 1.37) | <0.0001 |
| Multivariable-adjusted | 1 | 1.12 (1.06, 1.19) | 1.16 (1.09, 1.23) | 1.22 (1.15, 1.29) | 1.29 (1.22, 1.37) | <0.0001 |
| C-peptide (NHS-II, n=1,717) | | | | | | |
| Age-adjusted | 1 | 1.05 (0.96, 1.15) | 1.15 (1.05, 1.26) | 1.19 (1.09, 1.30) | 1.37 (1.25, 1.50) | <0.0001 |
| Multivariable-adjusted | 1 | 1.05 (0.96, 1.15) | 1.13 (1.04, 1.24) | 1.16 (1.06, 1.27) | 1.32 (1.21, 1.45) | <0.0001 |
| Empirical lifestyle index for hyperinsulinemia (ELIH) | | | | | | |
| C-peptide (NHS, n=5,812) | | | | | | |
| Age-adjusted | 1 | 1.10 (1.06, 1.15) | 1.26 (1.21, 1.32) | 1.54 (1.48, 1.60) | 2.04 (1.96, 2.13) | <0.0001 |
| Multivariable-adjusted | 1 | 1.10 (1.05, 1.14) | 1.25 (1.19, 1.30) | 1.51 (1.44, 1.57) | 1.97 (1.89, 2.06) | <0.0001 |
| C-peptide (HPFS, n=4,002) | | | | | | |
| Age-adjusted | 1 | 1.20 (1.13, 1.27) | 1.31 (1.24, 1.38) | 1.46 (1.38, 1.54) | 1.83 (1.73, 1.94) | <0.0001 |
| Multivariable-adjusted | 1 | 1.19 (1.12, 1.25) | 1.29 (1.22, 1.36) | 1.43 (1.35, 1.51) | 1.78 (1.68, 1.88) | <0.0001 |
| C-peptide (NHS-II, n=1,717) | | | | | | |
| Age-adjusted | 1 | 1.16 (1.06, 1.26) | 1.31 (1.21, 1.43) | 1.41 (1.29, 1.54) | 1.96 (1.80, 2.14) | <0.0001 |
| Multivariable-adjusted | 1 | 1.16 (1.06, 1.26) | 1.31 (1.21, 1.43) | 1.41 (1.29, 1.54) | 1.90 (1.74, 2.08) | <0.0001 |
| Empirical dietary index for insulin resistance (EDIR) | | | | | | |
| TG/HDL (NHS, n=3,929) | | | | | | |
| Age-adjusted | 1 | 1.19 (1.11, 1.28) | 1.41 (1.32, 1.52) | 1.56 (1.45, 1.67) | 1.98 (1.84, 2.12) | <0.0001 |
| Multivariable-adjusted | 1 | 1.18 (1.10, 1.26) | 1.34 (1.25, 1.44) | 1.43 (1.33, 1.54) | 1.67 (1.55, 1.80) | <0.0001 |
| TG/HDL (HPFS, n=3,559) | | | | | | |
| Age-adjusted | 1 | 1.09 (1.01, 1.18) | 1.24 (1.15, 1.34) | 1.35 (1.25, 1.45) | 1.59 (1.48, 1.72) | <0.0001 |
| Multivariable-adjusted | 1 | 1.11 (1.03, 1.19) | 1.21 (1.13, 1.30) | 1.29 (1.20, 1.39) | 1.44 (1.34, 1.55) | <0.0001 |
| TG/HDL (NHS-II, n=1,008) | | | | | | |

| | Quintile 1 (reference) | Quintile 2 | Quintile 3 | Quintile 4 | Quintile 5 | P-trend [‡] |
|---|---------------------------|-------------------|-------------------|-------------------|-------------------|----------------------|
| Age-adjusted | 1 | 1.02 (0.90, 1.14) | 1.12 (0.99, 1.26) | 1.32 (1.17, 1.49) | 1.32 (1.17, 1.49) | <0.0001 |
| Multivariable-adjusted | 1 | 0.98 (0.87, 1.10) | 1.08 (0.96, 1.22) | 1.23 (1.09, 1.38) | 1.19 (1.05, 1.34) | 0.001 |
| Empirical lifestyle index for insulin resistance (ELIR) | | | | | | |
| TG/HDL (NHS, n=3,929) | | | | | | |
| Age-adjusted | 1 | 1.24 (1.15, 1.32) | 1.57 (1.46, 1.68) | 1.98 (1.85, 2.12) | 2.60 (2.43, 2.78) | <0.0001 |
| Multivariable-adjusted | 1 | 1.24 (1.16, 1.33) | 1.54 (1.44, 1.65) | 1.84 (1.72, 1.98) | 2.27 (2.11, 2.45) | <0.0001 |
| TG/HDL (HPFS, n=3,559) | | | | | | |
| Age-adjusted | 1 | 1.28 (1.19, 1.37) | 1.58 (1.47, 1.70) | 1.92 (1.79, 2.06) | 2.34 (2.18, 2.52) | <0.0001 |
| Multivariable-adjusted | 1 | 1.23 (1.15, 1.32) | 1.49 (1.39, 1.60) | 1.76 (1.64, 1.89) | 2.03 (1.89, 2.19) | <0.0001 |
| TG/HDL (NHS-II, n=1,008) | | | | | | |
| Age-adjusted | 1 | 1.12 (1.00, 1.26) | 1.29 (1.15, 1.44) | 1.64 (1.46, 1.83) | 1.90 (1.69, 2.13) | <0.0001 |
| Multivariable-adjusted | 1 | 1.06 (0.95, 1.19) | 1.21 (1.08, 1.36) | 1.49 (1.32, 1.68) | 1.67 (1.48, 1.89) | <0.0001 |

NHS, Nurses' Health Study, 1990; NHS-II, Nurses' Health Study-II, 1999; HPFS, Health Professional Follow-up Study, 1994;

* Multivariable-adjusted models were adjusted for regular aspirin/NSAIDs use, age, physical activity, smoking status, diabetes, other chronic diseases/conditions, case-control status; NHS-II models were additionally adjusted for menopausal status and postmenopausal hormone use;

[‡] Values are relative concentrations of fasting plasma biomarkers (i.e., ratio of concentration in higher index quintiles to concentration in lowest quintile 1 as reference). All values were back transformed (e^x where x is the transformed biomarker value) since all biomarkers were transformed using natural log prior to analyses;

[‡] The P-value for trend was the P-value of the index as a continuous variable, adjusted for all covariates listed in footnote * ;