**Supplementary Table 1. Odds ratios and 95% confidence intervals for CRC risk by NAE**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **No. of controls (%)** | **No. of cases (%)** | **Model I** | **Model II** |
| **NAE (mEq/day)a** |
| ***Total (n=2757)*** |  |  |  |  |
|  T1 (<40.86) | 611 (33.3) | 199 (21.6) | 1.00 | 1.00 |
|  T2 (40.86-51.40) | 613 (33.4) | 385 (41.8) | 1.93 (1.57-2.37) | 1.83 (1.43-2.34) |
|  T3 (≥51.40) | 611 (33.3) | 338 (36.6) | 1.70 (1.38-2.09) | **1.68 (1.28-2.21)** |
|  *P* for trend |  |  | <0.001 | **<0.001** |
| Per 1-SD increase | 1835 (100) | 922 (100) | 1.28 (1.18-1.40) | **1.30 (1.16-1.46)** |
| ***Males (n=1866)a*** |  |  |  |  |
|  T1 (<45.35) | 413 (33.3) | 155 (24.8) | 1.00 | 1.00 |
|  T2 (45.35-53.88) | 415 (33.4 | 256 (41.0) | 1.64 (1.29-2.09) | 1.64 (1.23-2.20) |
|  T3 (≥53.88) | 414 (33.3) | 213 (34.2) | 1.37 (1.07-1.76) | 1.33 (0.98-1.80) |
|  *P* for trend |  |  | 0.009 | 0.054 |
| Per 1-SD increase | 1242 (100) | 624 (100) | 1.17 (1.06-1.30) | 1.15 (1.02-1.31) |
| ***Females (n=891*)a** |  |  |  |  |
|  T1 (<32.41) | 197 (33.2) | 34 (11.4) | 1.00 | 1.00 |
|  T2 32.41-43.43) | 198 (33.4) | 124 (41.6) | 3.63 (2.37-5.56) | 2.72 (1.67-4.43) |
|  T3 (≥ 43.43) | 198 (33.4) | 140 (47.0) | 4.10 (2.68-6.26) | **3.00 (1.84-4.88)** |
|  *P* for trend |  |  | <0.001 | **<0.001** |
| Per 1-SD increase | 593 (100) | 298 (100) | 1.66 (1.41-1.97) | **1.45 (1.19-1.76)** |

*NAE, net acid excretion.*

*a Participants with missing values for weight and height were excluded.*

***Model I****: unadjusted model;* ***Model II****: adjusted for age, education, first-degree family history of CRC, occupation, physical activity, income, marital status, smoking status, alcohol consumption, and BMI. In all subjects, model II was additionally adjusted for sex.*

**Supplementary Table 2. Model performance of dietary acid score and sociodemographic characteristics in the classification of CRC status**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Accuracy**  | **F1-score** | **Average precision score** | **AUC**  |
| Random forest  | 0.794 | 0.677 | 0.775 | 0.859 |
| XGBoost | 0.802 | 0.714 | 0.804 | 0.876 |
| Decision tree | 0.762 | 0.687 | 0.617 | 0.768 |
| AdaBoost classifier | 0.796 | 0.706 | 0.771 | 0.861 |
| Gaussian naive Bayes | 0.690 | 0.654 | 0.693 | 0.808 |