**Table S1: Overview of missing information in EPIC-Potsdam (N=25,392)**

|  |  |
| --- | --- |
| **Variable** | **Missing frequency****N (%)** |
| Incident diabetes case | 0 |
| Follow-up time (years) | 0 |
| Age (years) | 0 |
| Prevalent hypertension (yes/no) | 0 |
| Physical activity (h/week) | 0 |
| Former smoker (% < 20 units/d) | 6 (0.02) |
| Former smoker (% ≥ 20 units/d) | 6 (0.02) |
| Current smoker (% < 20 units/d) | 6 (0.02) |
| Current smoker (% ≥ 20 units/d) | 6 (0.02) |
| Wholegrain intake (bread, muesli) (50g portion/d) | 16 (0.1) |
| Coffee consumption (150g portion/d) | 16 (0.1) |
| Intake of red meat (150g portion/d) | 16 (0.1) |
| Waist circumference (cm) | 36 (0.1) |
| Waist-to-hip ratio\* | 43 (0.2) |
| Body weight (kg)\* | 68 (0.3) |
| Body height (cm) | 142 (0.6) |
| Body mass index (kg/m²)\* | 187 (0.7) |
| One parent with diabetes (yes/no) | 3366 (13.3) |
| Both parents with diabetes (yes/no) | 3366 (13.3) |
| At least one sibling with diabetes (yes/no) | 3366 (13.3) |
| GDRS score points\* | 3509 (13.8) |
| Hemoglobin A1C (%) | 22464 (88.5) |

\* Variables only used in the imputation models

**Table S2: Overview of missing information in the GNHIES98 cohort (N=3,717)**

|  |  |
| --- | --- |
| **Variable** | **Missing frequency****N (%)** |
| Weighting factor | 0 |
| Clustering | 0 |
| Incident diabetes case | 0 |
| Age (years) | 0 |
| Sex | 0 |
| Prevalent hypertension (yes/no) | 0 |
| Body height (cm) | 11 (0.3) |
| Waist circumference (cm) | 25 (0.7) |
| Smoking status (never, former, current) | 41 (1.1) |
| Wholegrain intake (bread, muesli) (g/d) | 47 (1.3) |
| Intake of red meat (g/d) | 48 (1.3) |
| Coffee consumption (g/d) | 52 (1.4) |
| Sport activity (no sport, <1 h/week, 1–2 h/week, 2–4 h/week, >4 h/week) | 54 (1.5) |
| Educational level (low, middle, high) | 55 (1.5) |
| Numbers of cigarettes currently smoked (cig./d) | 70 (1.9) |
| Numbers of cigarettes previously smoked (cig./d) | 91 (2.4) |
| Hemoglobin A1C (%) | 135 (3.6) |
| Mother with diabetes (yes/no) | 366 (9.8) |
| Father with diabetes (yes/no) | 479 (12.9) |

****

E/O ratio:

1.51 (1.22-1.87)

**Supplementary Figure 1: Calibration plot for the clinical GDRS in the GNHIES98 cohort – before recalibration.**

Predicted risks were determined using score points, mean score points and the baseline survival as derived from EPIC-Potsdam. Predicted risks were plotted against observed risks in four predefined risk groups (low risk: <2%, still low risk: 2 to <5%, increased risk: 5 to <10%, high or very high risk ≥10% 5-year risk). Expected-to-observed (E/O) ratio was calculated as the number of expected cases (as mean predicted risk multiplied with the overall sample size) divided by the number of observed cases; 95% confidence intervals were determined assuming a Poisson variance.