DOI: 10.26599/FSHW.2026.9251129

Contents lists available at SciOpen

Food Science and Human Wellness

journal homepage: https://www.sciopen.com/journal/2097-0765

China Food Publishing Co.

Prospective Associations between the Sulfur Microbial Diet and Type 2 Diabetes Risk across Two Diverse Cohorts

Jigen Na [a,b], Xiaojin Shi [a,b], Zijian Dai [a,b], Yangyang Tang [a,b], Da Zhang [a,b], Yucheng Yang [a,b], Yuefeng Tan [a,b], Zhihui Li [a,b], John S. Ji [a,b], Ai Zhao [a,b,*]

[a] *Vanke School of Public Health, Tsinghua University, Beijing, China.*

[b] *Institute for Healthy China, Tsinghua University, Beijing, China.*

ABSTRACT: Objective: To investigate the association between adherence to the sulfur microbial diet (SMD) and type 2 diabetes risk, and to explore dose–response relationships, effect modification, and underlying biological mechanisms.**Methods:** Data were obtained from two large-scale cohorts: 104,684 participants in the UK Biobank and 16,345 in the China Health and Nutrition Survey (CHNS). SMD scores were derived from dietary data based on food groups correlated with 43 sulfur-metabolizing taxa. Incident type 2 diabetes was ascertained through medical records (UK Biobank) and physician-diagnosed self-reports (CHNS). Cox models estimated hazard ratios (HRs) and 95% confidence intervals (CIs), with restricted cubic splines for dose–response relationships. Stratified and mediation analyses assessed potential modifiers and mediating biomarkers.**Results:** Higher SMD scores were consistently associated with a higher risk of developing type 2 diabetes (highest vs. lowest quartile: UK Biobank: HR = 1.31, 95% CI: 1.17–1.47; CHNS: HR = 3.26, 95% CI: 2.52–4.23). Spline analyses indicated a linear association in the UK Biobank and a J-shaped association in CHNS. Higher physical activity attenuated the association in CHNS (p for interaction = 0.002). Mediation analyses identified 15 circulating biomarkers spanning inflammatory, metabolic, hormonal, and hepatic pathways.**Conclusion:** High adherence to the SMD was associated with an increased risk of type 2 diabetes, partly mediated through diverse biological pathways. Physical activity attenuated this risk. These findings highlight the potential role of sulfur-metabolizing bacteria in glucose regulation and underscore the importance of dietary and lifestyle modification for diabetes prevention and care.

Keywords: sulfur microbial diet, type 2 diabetes, physical activity, gut microbiota, multi-cohort study

1. Introduction

Type 2 diabetes has become one of the fastest-growing global health challenges, with rapidly rising prevalence and substantial contributions to premature mortality and health care costs [1]. Beyond its direct clinical consequences, type 2 diabetes serves as an upstream determinant in the causal pathways of multiple chronic diseases [2]. Understanding its causes is therefore of major public health importance and essential for improving diabetes prevention and care.

***Corresponding author**
aizhao18@mail.tsinghua.edu.cn

Received 6 November 2025
Received in revised from 27 January 2026
Accepted 12 February 2026

Dietary patterns remain central to both the prevention and management of type 2 diabetes [3], and growing attention has been directed toward the gut microbiome as a mediator linking diet to host metabolism [4]. The sulfur microbial diet (SMD), a novel dietary pattern reflecting the potential to alter sulfur-metabolizing gut bacteria, has recently been associated with increased risks of several chronic diseases [5, 6]. Unlike conventional dietary patterns defined by food combinations, SMD is microbiome-informed, capturing dietary influences on sulfur-metabolizing taxa. This approach enhances cross-population applicability, as distinct cultural diets may converge on similar microbial functional profiles [7]. Our prior work suggested that type 2 diabetes might partly mediate the link between higher SMD adherence and elevated dementia risk in the UK Biobank [8]. However, whether SMD itself is prospectively associated with type 2 diabetes risk and the biological pathways involved remains unclear.

To address these questions, we leveraged two large, population-based cohorts to prospectively examine the association between SMD and type 2 diabetes and to validate the findings across populations with distinct genetic and socioeconomic contexts. By elucidating diet-microbiota-metabolism pathways relevant to glycemic regulation, our study aims to inform dietary strategies and lifestyle interventions for more effective diabetes prevention and care.

2. Methods

2.1 Study design and participants

We used data from the UK Biobank and the China Health and Nutrition Survey (CHNS) (Fig. 1). The UK Biobank was approved by the Northwest Multi-Centre Research Ethics Committee (21/NW/0157), and CHNS by the institutional review committees of the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health, Chinese Center for Disease Control and Prevention. Informed consent was obtained from all participants at recruitment by the original cohort studies, and the present analyses were conducted using de-identified data in accordance with the Declaration of Helsinki.

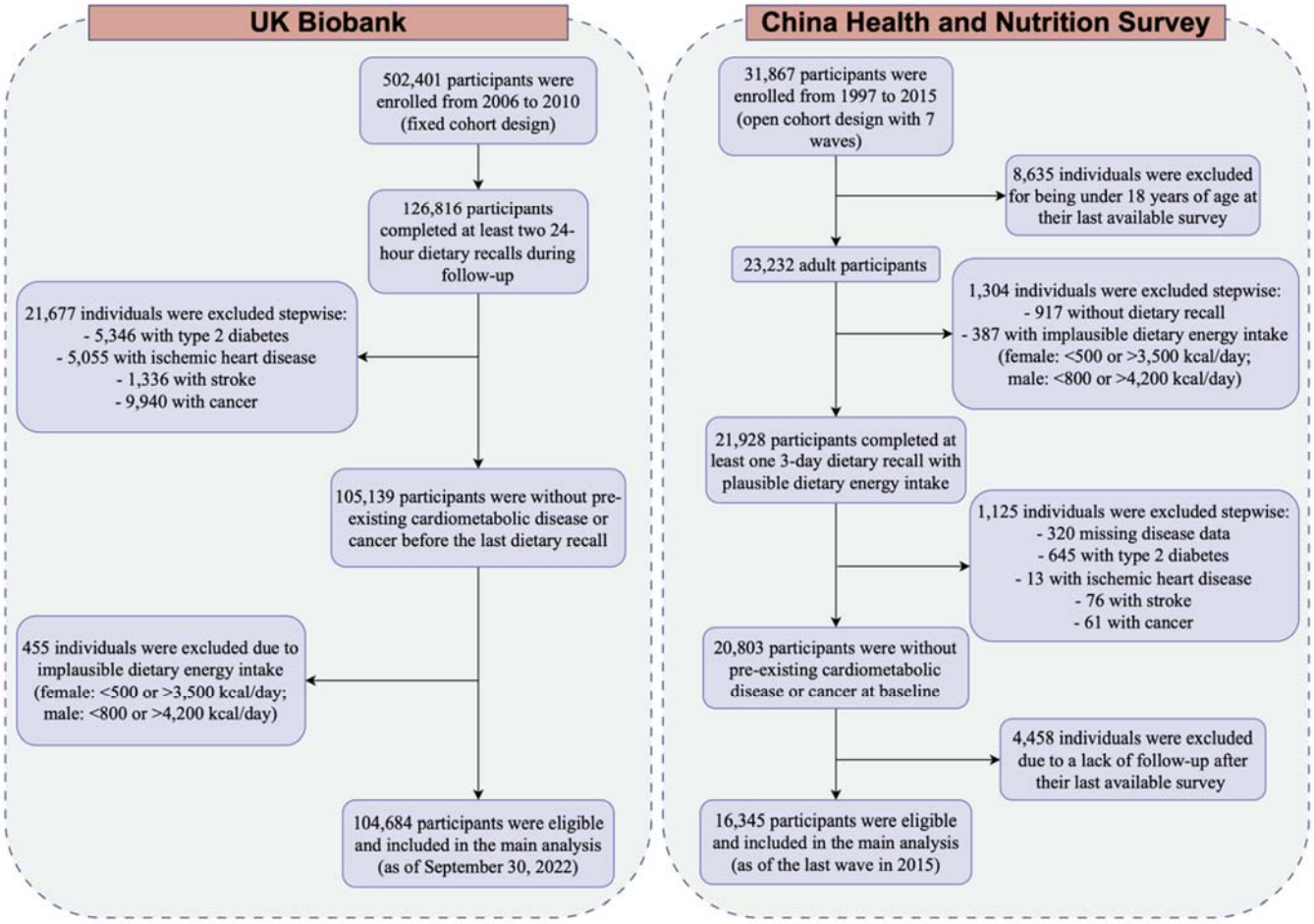

Figure 1. Flowchart of participant inclusion and exclusion in the UK Biobank (left) and China Health and Nutrition Survey (right).

The UK Biobank has been described previously [8, 9]. Briefly, it recruited ~500,000 adults aged 40–69 years across the UK, with follow-up through linkage to hospital and death records. We included participants with at least two 24-hour dietary recalls who were free of type 2 diabetes, ischemic heart disease, stroke, and cancer to reduce reverse causation. Those with implausible energy intake (<500 or >3,500 kcal/day for females; <800 or >4,200 kcal/day for males) were excluded [10]. The date of the last dietary recall was considered baseline, and participants were followed until September 30, 2022.

The CHNS is an ongoing cohort initiated in 1989 with follow-ups every 2–4 years [11]. We included adults (≥ 18 years, non-pregnant) who had completed at least one 3-day dietary recall, had plausible energy intake [10], and were free of diabetes, ischemic heart disease, stroke, or cancer at baseline. Those without disease data or follow-up were excluded. The first eligible survey entry was defined as baseline, and participants were followed until 2015 or their most recent survey.

2.2 Assessment of SMD

In the UK Biobank, dietary data were obtained from repeated 24-hour recalls via a validated self-administered digital questionnaire covering >200 food items. In the CHNS, dietary recalls were collected face-to-face by trained interviewers on three consecutive days. Details of the food item coding used in each

cohort are provided in Supplementary Table S1. In both cohorts, average daily food intake (servings/day) was derived using a standardized weight-to-serving conversion chart [12].

The SMD was defined by eight food groups identified and validated in relation to 43 gut sulfur-metabolizing bacteria [5]. The score was calculated as the weighted sum of their intakes, with β-coefficients from the validation study as weights, according to the following formula [5, 6]:

$$\text{SMD score} = \sum_{i=1}^{n} (\beta_i \times Serving\ size_i)$$

Specifically, the included food groups and their corresponding weights were processed meats (0.64), liquor (0.31), low-calorie drinks (0.38), beer (-0.54), fruit juice (-0.21), legumes (-0.64), other vegetables (-0.30), and sweets & desserts (-0.23) [5]. Higher scores indicate a greater dietary potential to promote sulfur-metabolizing bacteria.

To compare SMD composition between cohorts, we used violin plots to show score distributions, radar plots to illustrate standardized relative intakes (with the larger cohort mean set to 1), and stacked bar charts to display the percentage contribution of each food group, calculated as its absolute score divided by the sum of all absolute scores.

2.3 Assessment of Type 2 Diabetes

In the UK Biobank, incident type 2 diabetes was identified by the Outcome Adjudication Group using the International Classification of Diseases, 10th Revision code E11 (non-insulin-dependent diabetes mellitus). E14 (unspecified diabetes) was also classified as type 2 diabetes, as unspecified diabetes in older adults is mostly type 2 [13]. Sensitivity analysis was conducted with cases restricted to E11 only. Between 2012 and 2013, blood glucose and HbA1c were measured in a subset of participants between 2012 and 2013. Using self-reported time since last meal (8–16 hours) to approximate fasting status, these biomarker data enabled identification of WHO-defined type 2 diabetes (fasting glucose ≥7.0 mmol/L or HbA1c ≥48 mmol/mol [6.5%]) [14]. As biomarker measurements were only available in a subset of participants and not throughout the entire follow-up, biomarker-based definitions were used in sensitivity analyses only to minimize potential selection bias.

In the CHNS, cases were primarily identified based on self-reported physician diagnosis and/or diabetes medication use. In 2009, fasting glucose and HbA1c were measured in a subset of participants, allowing WHO-defined type 2 diabetes [14]. Considering the unavailability of these data in other waves, biomarker-based definitions were included in sensitivity analyses only.

2.4 Assessment of Covariates

Covariates were selected based on a directed acyclic graph (Supplementary Fig. S1) and previous literature [15, 16]. Covariates included sex, age, body mass index (BMI), socioeconomic position, education, smoking (never, former, or current), alcohol consumption, physical activity (low, moderate, or high, assessed by the International Physical Activity Questionnaire [17]), and total energy intake. Family history and a

polygenic risk score (PRS) for type 2 diabetes were additionally adjusted for in the UK Biobank analyses, whereas these variables were unavailable in CHNS. Ethnicity was not adjusted for due to missing values; sensitivity analyses were restricted to primary populations (Caucasians in the UK Biobank, Han Chinese in CHNS).

In the UK Biobank, socioeconomic position was assessed by the Townsend Deprivation Index (TDI), education was measured by the highest qualification, and alcohol consumption was categorized as never, former, or current. In the CHNS, socioeconomic position was represented by inflation-adjusted personal income (2015 values), education by highest degree completed, and alcohol consumption categorized as never, ≤2 times/month, or ≥1 time/week.

2.5 Assessment of Mediators

Multiple clinical blood biomarkers were explored as potential mediators in the association between SMD scores and the risk of type 2 diabetes.

In the UK Biobank, non-fasting venous blood samples were collected at recruitment and assayed under standardized procedures with quality control. Thirty-four biomarkers were analyzed, grouped as: inflammatory markers (white blood cell [WBC], lymphocyte, monocyte, neutrophil, eosinophil, and basophil counts; C-reactive protein; rheumatoid factor); liver enzymes and function (alanine aminotransferase [ALT], aspartate aminotransferase, alkaline phosphatase [ALP], gamma-glutamyl transferase [GGT], direct and total bilirubin); renal and metabolic (urea, creatinine, cystatin C, urate); minerals (calcium, phosphate); lipids and apolipoproteins (total cholesterol, triglyceride, high-density lipoprotein cholesterol [HDL-C], low-density lipoprotein cholesterol [LDL-C], apolipoprotein A [ApoA], apolipoprotein B [ApoB], lipoprotein(a)); serum proteins (albumin, total protein); and hormones and binding proteins (insulin-like growth factor 1 [IGF-1], sex hormone-binding globulin [SHBG], testosterone, estradiol, vitamin D).

In the CHNS, fasting blood samples collected in 2009 were analyzed by standardized biochemical assays. Nineteen biomarkers were analyzed, grouped as: inflammatory markers (WBC count, C-reactive protein, ferritin); liver enzyme (ALT); renal and metabolic (urea, creatinine, uric acid); mineral (magnesium); lipids and apolipoproteins (total cholesterol, triglyceride, HDL-C, LDL-C, ApoA, ApoB, lipoprotein(a)); and serum proteins and nutrition markers (albumin, total protein, transferrin, transferrin receptor).

2.6 Statistical Analysis

Participants were categorized into four groups (Q_1–Q_4) by cohort-specific quartiles. Missing covariates were imputed via random forest multiple imputation (see Supplementary Table S2 for missing data proportions). Continuous variables are shown as mean ± standard deviation and compared by analysis of variance, and categorical variables as counts (percentages) and compared by chi-square tests. No between-cohort comparisons were performed due to distinct study designs.

Cox models estimated hazard ratios (HRs) and 95% confidence intervals (CIs) for SMD score quartiles and incident type 2 diabetes, with Q_1 as reference. The proportional hazards (PH) assumption was tested using

Schoenfeld residuals, and time-varying Cox models were fitted when violated. Models were adjusted for predefined covariates. Between-cohort heterogeneity was assessed by Cochran's Q and I^2 statistics.

Restricted cubic splines (three knots) assessed dose–response relationships, and non-linearity was tested by likelihood ratio tests versus linear models. Stratified analyses examined effect modification by sex, age (<65 or ≥65 years), BMI, PRS (<0 or ≥0; UK Biobank only), physical activity, smoking, and alcohol consumption. Interactions were tested by likelihood ratio tests comparing models with and without cross-product terms. Mediation analyses (Cox-based) assessed potential biomarker mediation and were restricted to participants in Q_4 versus Q_1.

Sensitivity analyses tested robustness and minimized reverse causation, including: (A) re-estimating associations via stepwise covariate selection; (B) redefining type 2 diabetes by E11 in the UK Biobank; (C) combining type 2 diabetes ascertainment based on fasting glucose and HbA1c data; (D) performing competing risk analyses (death as competing event) using Fine-Gray models; (E) excluding participants with <1-year follow-up; (F) excluding participants with <2-year follow-up; (G) restricting analyses to primary ethnic groups; (H) applying time-varying Cox models with log (time) interactions when PH assumption was violated; (I) additionally adjusting for carbohydrate, protein, and fat intakes; and (J) additionally adjusting for baseline hypertension status in both cohorts and baseline hypertriglyceridemia status in the UK Biobank only (not available in the CHNS).

All statistical analyses were conducted in R version 4.2.3 using the mice, survival, rms, MASS, and CMAverse packages. A two-tailed *p*-value <0.05 was considered statistically significant.

3. Results

3.1 Baseline characteristics

Table 1 summarizes the baseline characteristics of 104,684 participants in the UK Biobank and 16,345 in the CHNS. In both cohorts, participants were less likely to be current smokers but more likely to report alcohol consumption. UK Biobank participants were older and had higher BMI than those in the CHNS. UK Biobank participants were less likely to report low physical activity, whereas CHNS participants had higher daily energy intake. These findings indicate clear differences in sociodemographic and lifestyle characteristics between the two populations.

Table 1. Characteristics of participants in the UK Biobank and the China Health and Nutrition Survey.

Characteristics [a]	**UK Biobank** (n = 104,684)	**CHNS** (n = 16,345)
Sex		
Female	60,490 (57.8%)	8,370 (51.2%)
Male	44,194 (42.2%)	7,975 (48.8%)
Age (year)	55.40 ± 7.83	42.81 ± 15.76
BMI (kg/m^2)	26.45 ± 4.41	22.72 ± 3.32
Socioeconomic position [UKB/CHNS]		
TDI/Income (×1,000 CNY, 2015)	-1.64 ± 2.84	11.54 ± 18.16
Qualification/Educational level [UKB/CHNS]		
CSEs/Primary school	3,807 (3.6%)	3,188 (19.5%)
O levels or GCSEs/Junior high school	20,382 (19.5%)	5,185 (31.7%)

A or AS levels/Senior high school	14,254 (13.6%)	2,298 (14.1%)
NVQ, HND, or HNC/Technical school	4,817 (4.6%)	1,121 (6.9%)
College or university	50,501 (48.2%)	1,070 (6.5%)
Others/Postgraduate	4,932 (4.7%)	29 (0.2%)
None of the above/No formal schooling	5,991 (5.7%)	3,454 (21.1%)
Smoking status		
Never	61,495 (58.7%)	11,211 (68.6%)
Former	35,928 (34.3%)	324 (2.0%)
Current	7,261 (6.9%)	4,810 (29.4%)
Drinking status/Frequency [UKB/CHNS]		
Never	2,866 (2.7%)	10,658 (65.2%)
Former/≤ 2 times/month	2,835 (2.7%)	1,924 (11.8%)
Current/≥ 1 time/week	98,983 (94.6%)	3,763 (23.0%)
Physical activity level		
Low	18,825 (18.0%)	5,316 (32.5%)
Moderate	45,158 (43.1%)	4,786 (29.3%)
High	40,701 (38.9%)	6,243 (38.2%)
Dietary energy intake (kcal/day)	2,057.06 ± 497.47	2,152.50 ± 635.76
Family history of type 2 diabetes		
No	85,591 (81.8%)	NA
Yes	19,093 (18.2%)	NA
Polygenic risk score for type 2 diabetes	-0.21 ± 0.95	NA

Abbreviation: CHNS, China Health and Nutrition Survey; BMI, body mass index; UKB, UK Biobank; TDI, Townsend Deprivation Index; NA, not available/applicable.

[a] Categorical variables were shown as numbers (percentages); numeric variables were shown as mean ± standard deviation. Variables annotated with [UKB/CHNS] indicate differences between the UK Biobank and CHNS due to questionnaire structure.

Across both cohorts, several baseline characteristics varied across SMD score quartiles (Supplementary Tables S3 and S4). Higher quartiles were associated with lower daily energy intake, a higher prevalence of low physical activity, and lower educational attainment. Socioeconomic gradients were observed, with higher SMD score levels corresponding to greater deprivation in the UK Biobank and lower income in CHNS. Age and BMI showed opposite patterns: in the UK Biobank, higher SMD quartiles were characterized by younger age and higher BMI, whereas in the CHNS, they were associated with older age and lower BMI. Other characteristics, including smoking, drinking, and sex distribution, differed only modestly across quartiles. Genetic and family history variables, available only in the UK Biobank, showed minimal variation across SMD categories.

3.2 Distribution of SMD

As shown in Fig. 2a, SMD scores ranged from -6.99 to 3.82 in the UK Biobank and from -16.56 to 8.16 in the CHNS. Participants were categorized into quartiles using cohort-specific cut-offs: -1.36, -0.85, and -0.38 for the UK Biobank, and -2.21, -1.57, and -0.98 for the CHNS.

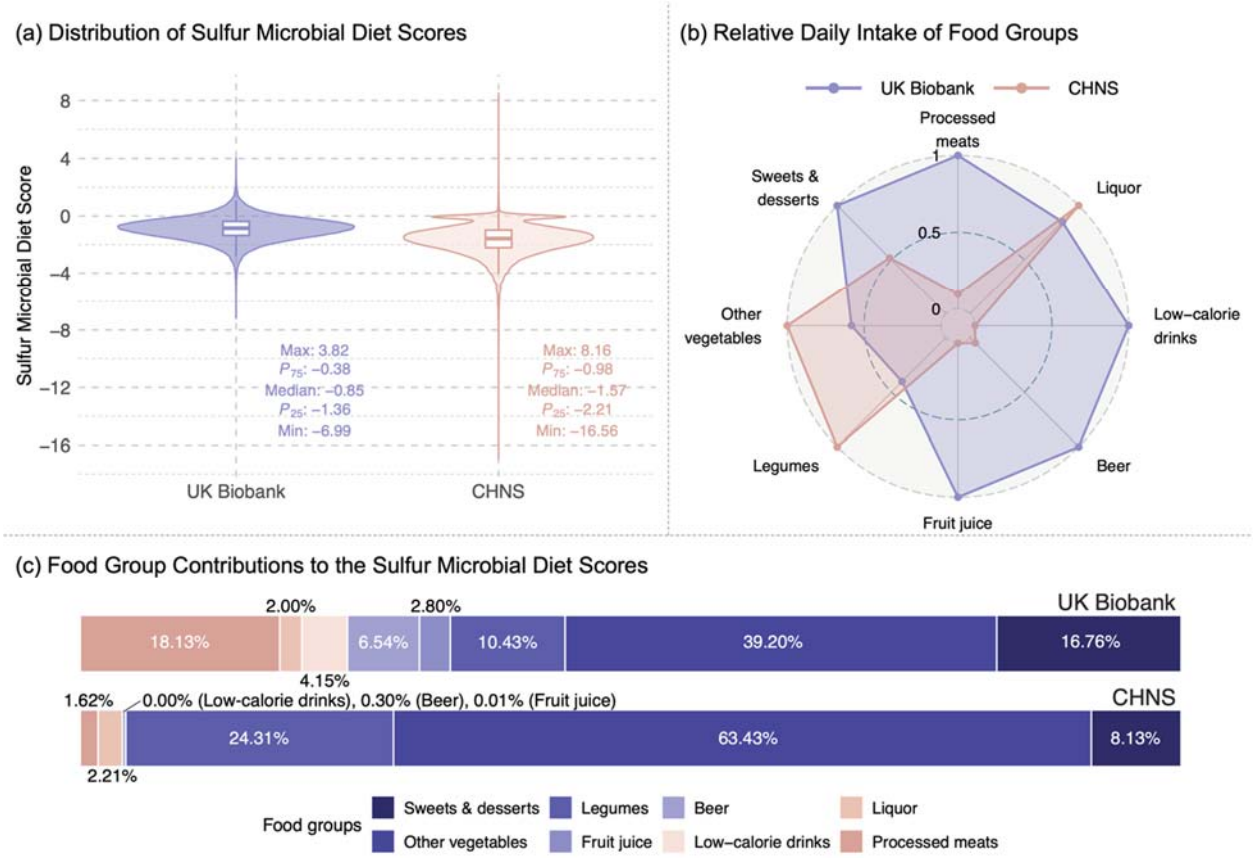

Figure 2. Comparison of (a) sulfur microbial diet (SMD) score distribution; (b) relative daily intakes standardized to the higher cohort-specific mean (set to 1); and (c) food group contributions expressed as the percentage of absolute SMD scores relative to the total, between the UK Biobank and China Health and Nutrition Survey (CHNS).

The distribution and food group composition of the SMD differed notably between cohorts (Fig. 2b,c). UK Biobank participants reported higher intakes of processed meats, beer, low-calorie drinks, fruit juice, and sweets & desserts, whereas CHNS participants consumed more legumes and other vegetables. Liquor intake was similar between cohorts, with slightly higher levels in the CHNS. Consistent with these patterns, the main contributors to SMD scores were other vegetables and processed meats in the UK Biobank, and other vegetables and legumes in the CHNS.

3.3 Associations between SMD Score Quartiles and Type 2 Diabetes

During a median follow-up of 13.23 years (interquartile range [IQR]: 12.61–14.01) in the UK Biobank and 9 years (IQR: 4–15) in the CHNS, 2,490 and 445 incident type 2 diabetes cases were identified, respectively.

The associations between SMD score quartiles and incident type 2 diabetes are shown in Table 2. Higher quartiles were associated with increased risk of type 2 diabetes in both cohorts after adjusting for covariates. In the UK Biobank, participants in the highest quartile had a 31% higher risk of type 2 diabetes compared with those in the lowest quartile (HR = 1.31, 95% CI: 1.17–1.47), with elevated risks also observed in the second (HR = 1.15, 95% CI: 1.02–1.29) and third quartiles (HR = 1.14, 95% CI: 1.01–1.28). No evidence of proportional hazards violation was observed (p for PH assumption = 0.932).

Table 2. Associations between sulfur microbial diet scores and risk of type 2 diabetes in the UK Biobank and the China Health and Nutrition Survey.

SMD scores quartiles	Type 2 diabetes cases	Incident rate per 1000 person years	Adjusted HR (95% CI) [a]
UK Biobank			
Q_1	535	1.55	1.00 (Reference)
Q_2	564	1.64	1.15 (1.02, 1.29)
Q_3	582	1.69	1.14 (1.01, 1.28)
Q_4	809	2.36	1.31 (1.17, 1.47)
CHNS			
Q_1	90	2.34	1.00 (Reference)
Q_2	75	1.53	0.70 (0.51, 0.95)
Q_3	91	1.93	0.86 (0.64, 1.16)
Q_4	189	6.57	3.26 (2.52, 4.23)

Abbreviation: SMD, sulfur microbial diet; HR, hazard ratio; CI, confidence interval; CHNS, China Health and Nutrition Survey.

[a] Cox models were adjusted for sex, age, BMI, socioeconomic position, qualification/educational level, smoking status, drinking status/frequency, physical activity level, and dietary energy intake in both the UK Biobank and the CHNS. The model for the UK Biobank was additionally adjusted for family history and polygenic risk score for type 2 diabetes.

In the CHNS, risk estimates were less stable but showed a similar pattern, with a more than threefold higher risk in the highest versus lowest quartile (HR = 3.26, 95% CI: 2.52–4.23). The PH assumption was violated ($P < 0.001$), indicating that the effect of SMD quartiles attenuated over time. The observed non-proportional hazards were mainly driven by Q_2 and Q_3 (Supplementary Fig. S2). When time-varying effects were modeled using log (time) interactions, the associations for Q_4 vs. Q_1 remained approximately constant (HR = 3.31, 95% CI: 1.54–7.08).

Restricted cubic spline analyses indicated an approximately linear dose–response relationship in the UK Biobank (P for non-linearity = 0.752), whereas in the CHNS, the pattern appeared J-shaped (P for non-linearity < 0.001) and was stronger at higher SMD scores (Fig. 3).

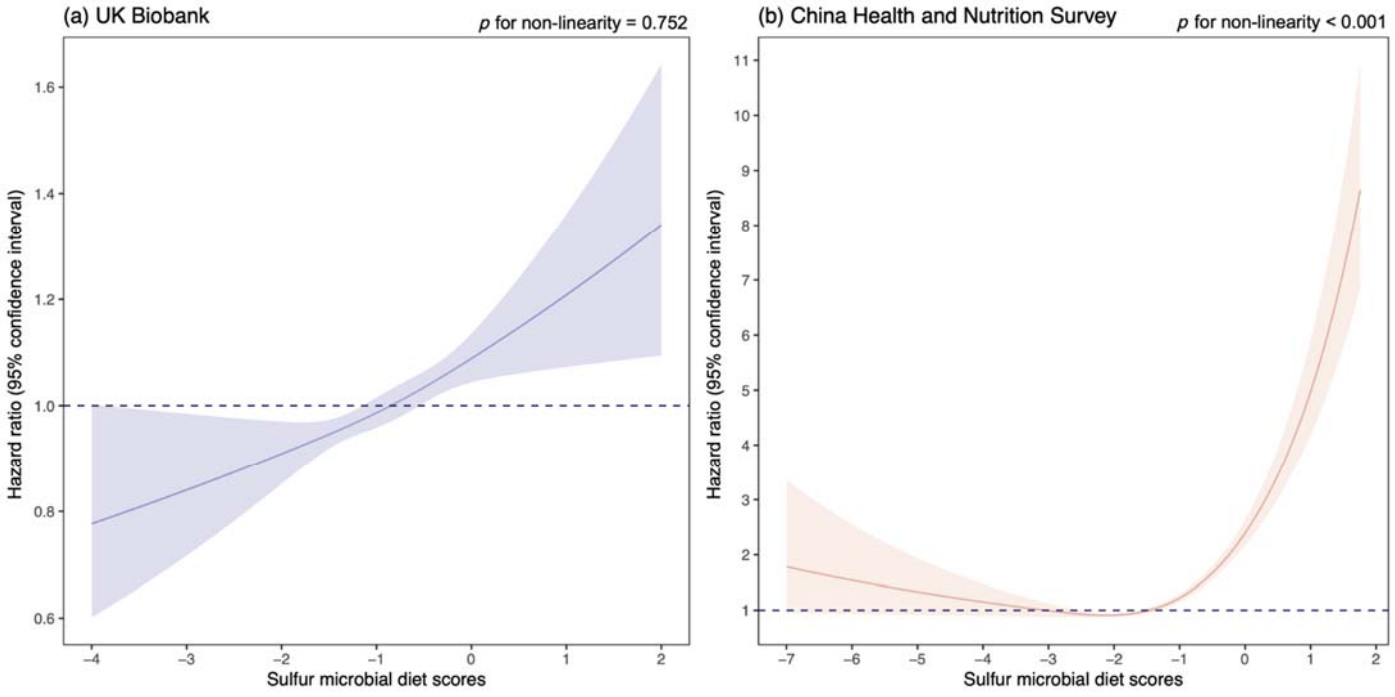

Figure 3. Dose–response associations (with the median as the reference value) between sulfur microbial diet scores and type 2 diabetes risk in the (a) UK Biobank and (b) China Health and Nutrition Survey.

Although the overall direction of associations was similar across cohorts, particularly for the highest versus lowest quartiles, the effect sizes and temporal patterns differed markedly ($I^2 = 93.5\%$, $P < 0.001$).

Because of this heterogeneity, results were presented separately by cohort rather than combined in a meta-analysis.

3.4 Stratified Analyses

Fig. 4 shows the results of stratified analyses. In the UK Biobank, no significant effect modification was observed, indicating that the association between SMD and type 2 diabetes risk was consistent across sex, age, BMI, genetic risk, and lifestyle factors (all p for interaction > 0.05).

In the CHNS, subgroup differences were identified. Physical activity modified the association (p for interaction = 0.002): positive associations were observed among participants with low or moderate activity, but not among those with high activity, suggesting that higher physical activity may attenuate the adverse association between high SMD adherence and type 2 diabetes risk. The interaction with smoking status (p = 0.033) appeared to be driven by former smokers, a small subgroup representing 2% of the cohort (n = 324). No consistent pattern of effect modification was found across other smoking categories.

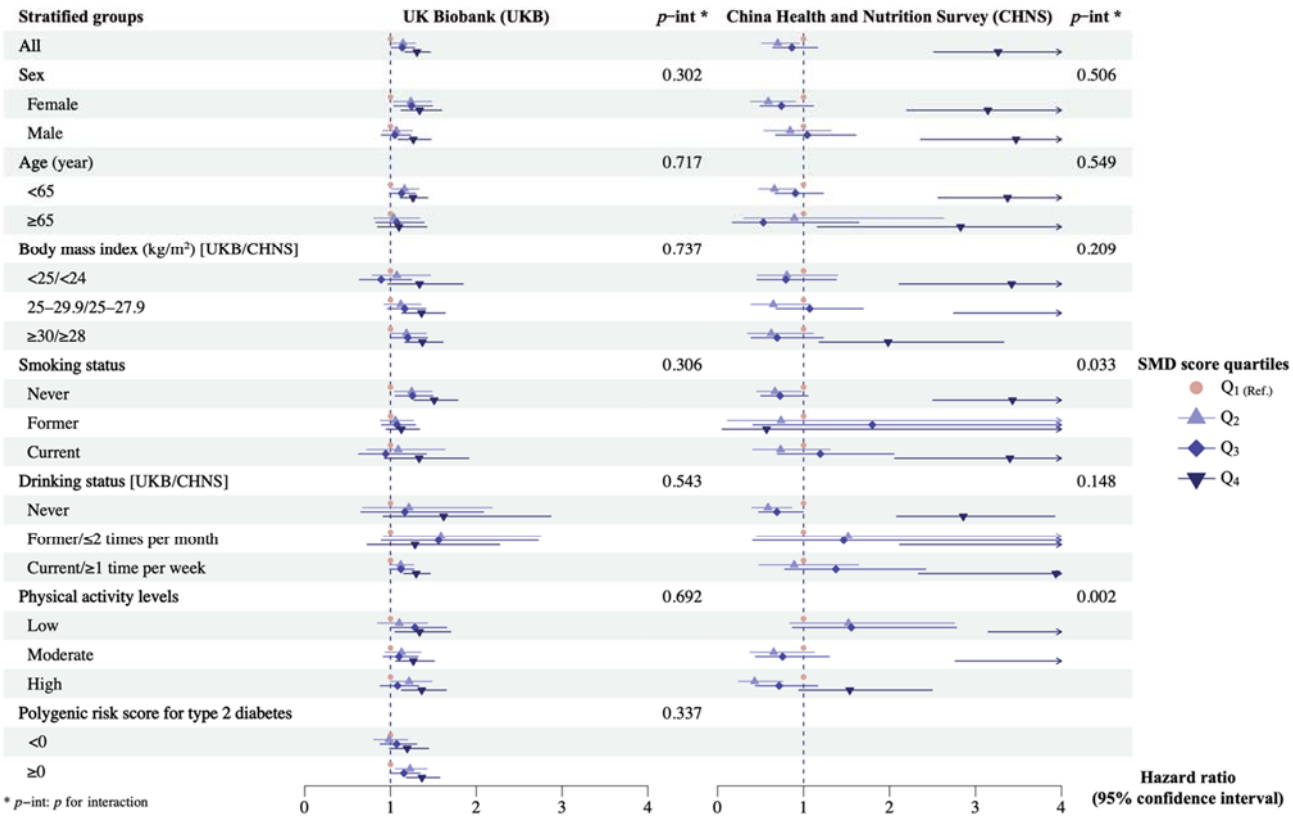

Figure 4. Stratified associations between sulfur microbial diet (SMD) scores and type 2 diabetes risk across subgroups in the UK Biobank and China Health and Nutrition Survey.

3.5 Mediation Analyses

Several blood biomarkers were identified as mediators of the association between high SMD adherence and type 2 diabetes risk (Fig. 5). Among UK Biobank participants, high SMD adherence was associated with increased type 2 diabetes risk, partly through elevated chronic inflammation, reflected by higher counts of WBC, lymphocyte, monocyte, neutrophil, and eosinophil. Decreases in SHBG, IGF-1, and testosterone were also related to a higher risk. Conversely, lower levels of ALP, calcium, and total protein, together with higher

creatinine, contributed to inverse mediation effects. In the CHNS, ferritin also acted as a positive mediator with a smaller magnitude. Among these biomarkers, SHBG explained the largest proportion of the association (proportion mediated [PM] = 5.99%, 95% CI: 2.79%–11.50%), whereas most others accounted for no more than ±3%.

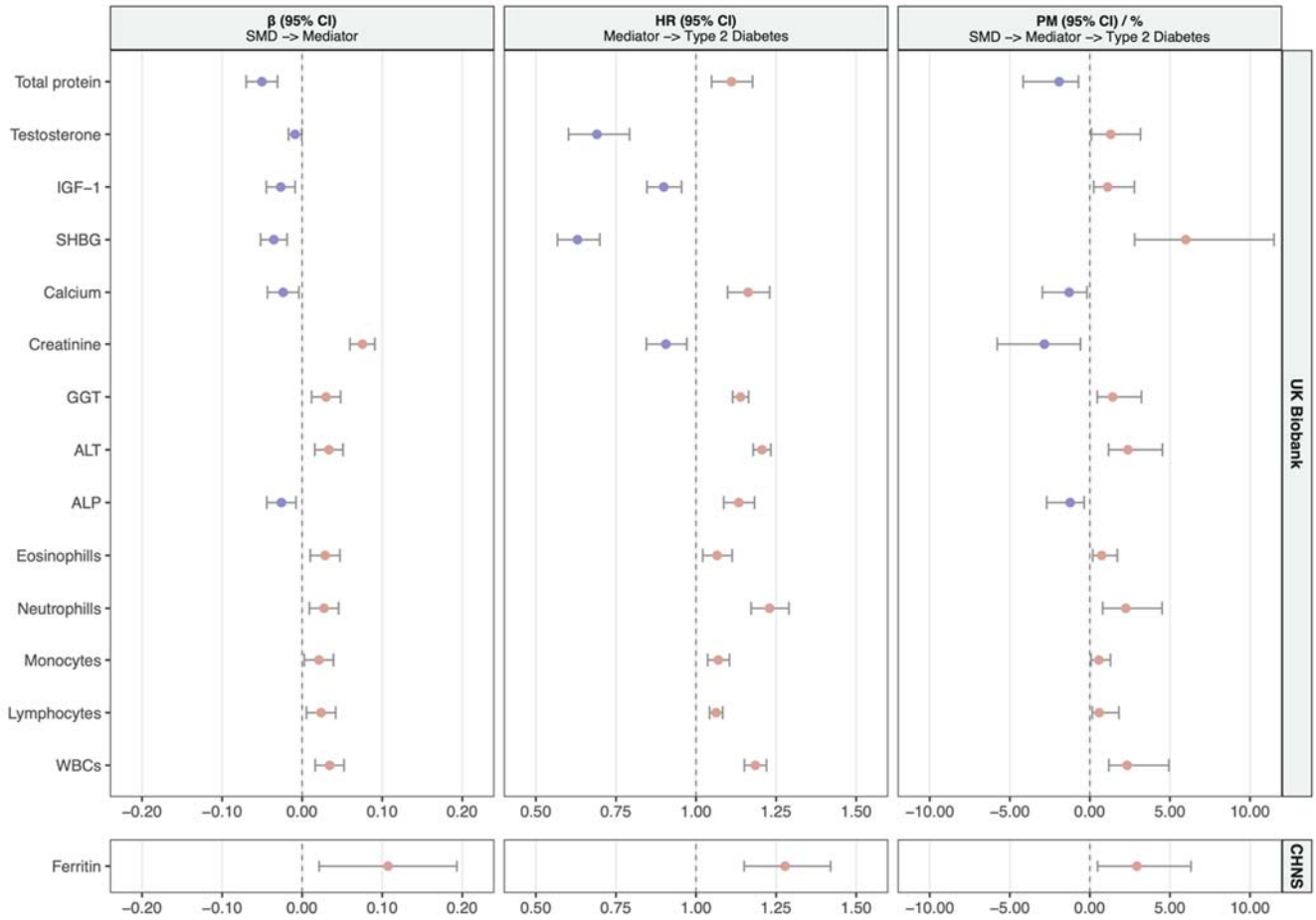

Figure 5. Significant mediators of the association between high adherence to the sulfur microbial diet (SMD) and type 2 diabetes risk in the UK Biobank and China Health and Nutrition Survey (CHNS).

Full results for all biomarkers are provided in Supplementary Table S5. No additional statistically significant mediators were identified beyond those described above.

3.6 Sensitivity Analyses

Results were consistent across multiple sensitivity analyses (Supplementary Table S6), including stepwise covariate selection, alternative definitions of type 2 diabetes, competing risk models, exclusion of early events, restriction to primary ethnic groups, and time-varying Cox models. The overall associations remained similar in direction and magnitude, with excess risk in the highest quartile showing particular stability across models, supporting the robustness of the main findings. In contrast, estimates for the other quartiles in the CHNS were less stable, particularly for Q_2 and Q_3 in time-varying Cox models.

4. Discussion

This study examined the prospective association between SMD and incident type 2 diabetes across two large-scale, population-based cohorts with distinct genetic and socioeconomic profiles. Higher SMD

adherence was associated with increased risk of type 2 diabetes, showing linear or J-shaped dose–response relationships. Among Chinese participants, vigorous physical activity appeared to attenuate this association, whereas no effect modification was observed in the UK Biobank. Mediation analyses identified several blood biomarkers, suggesting inflammatory, metabolic, hormonal, and hepatic pathways involved.

Previous studies have investigated associations between individual food groups and type 2 diabetes risk [18]. Most have focused on single dietary components or conventional dietary patterns such as the Mediterranean or Western diets, often overlooking their influence on the gut microbiota and metabolic outcomes. The glycemic response to diet varies across individuals, partly due to gut microbial differences [19]. Diets higher in processed meats and low in vegetables have been linked to greater abundances of sulfur-metabolizing bacteria [20, 21], which may influence glucose metabolism through hydrogen sulfide production [22]. Because direct microbiome profiling is often infeasible in large cohorts, the SMD serves as a validated proxy for these bacteria and reflects the diet-microbiota-metabolism axis distinct from conventional patterns such as the Western diet (Spearman $\rho = -0.009$) [5]. This approach underscores the potential role of sulfur-metabolizing bacteria in the diet-diabetes network and supports further mechanistic and interventional research.

In both cohorts, higher SMD score quartiles were associated with elevated type 2 diabetes risk, although the dose–response shapes and effect sizes differed. These differences likely reflect variations in demographic characteristics, dietary habits (notably processed meats, low-calorie drinks, beer, and fruit juice), and diabetes ascertainment methods. Given these inherent discrepancies, HRs were not pooled to avoid misleading overall estimates; instead, results were cross-validated across cohorts.

Estimates for intermediate quartiles in the CHNS were unstable in time-varying models, likely due to sparse events and violation of the PH assumption, which can yield imprecise estimates. Such irregularities may also reflect stage-specific effects, for example, early protective influence offset by long-term adverse outcomes or survivor bias over extended follow-up [23]. Biologically, these patterns align with the biphasic role of hydrogen sulfide in glucose metabolism, where low or transient exposure may be beneficial, whereas chronic or excessive levels appear harmful [22]. However, the thresholds at which hydrogen sulfide shifts from beneficial to harmful remain uncertain and require further study.

By contrast, the highest quartile results were consistent across sensitivity analyses, lending confidence to the main conclusion. Although the mechanistic role of hydrogen sulfide in glucose regulation remains complex, excessive hydrogen sulfide is generally detrimental [22], consistent with our findings. Increased abundances of sulfur-metabolizing taxa have been linked to type 2 diabetes and metabolic dysfunction [24]. One plausible mechanism involves bacterial hydrogen sulfide overproduction, which can disrupt enteroendocrine cells that secrete glucagon-like peptide-1 (GLP-1) [25], a key regulator of insulin and glucose homeostasis [26]. Elevated bacterial hydrogen sulfide may suppress GLP-1 synthesis and receptor expression, impairing metabolic function [25].

Beyond direct effects on glucose metabolism, sulfur-metabolizing bacteria and hydrogen sulfide may also contribute indirectly to type 2 diabetes. Excess hydrogen sulfide can compromise gut barrier integrity, trigger immune activation, and promote systemic inflammation [27], a known driver of diabetes [28]. Consistent with this, higher inflammatory cell counts and ferritin levels mediated the SMD-diabetes association. Hormonal pathways may also be involved: gut microbiota influence circulating IGF-1 and sex hormones [29, 30], and lower levels of IGF-1, SHBG, and testosterone are linked to insulin resistance and diabetes risk [31, 32]. SMD was inversely associated with these molecules, supporting hormonal involvement, although no sex-specific modification was observed. The liver may also play a role. Prior studies have linked SMD to nonalcoholic fatty liver disease [33], and high SMD adherence was associated with elevated GGT and ALT, both predictors of diabetes [34]. Conversely, inverse associations of SMD with ALP, creatinine, and total protein suggest potential counteracting influences, underscoring the biological complexity.

High physical activity, as assessed by the IPAQ, attenuated SMD-associated diabetes risk in the Chinese population, reinforcing the protective role of exercise. Physical activity mitigates adverse dietary effects [35], through reduced inflammation and improved liver function, insulin sensitivity, and hormonal balance [36, 37]. Lack of similar findings in the UK Biobank may reflect population heterogeneity, including differences in activity domains or genetic predisposition to insulin resistance [38-40]. The interaction between ethnicity, diet, and physical activity warrants further study.

Taken together, this study has several notable strengths. First, the use of two large, population-based cohorts with prospective designs provided high statistical power, and cross-validation across distinct populations enhanced both the robustness and generalizability of the results. Second, repeated dietary assessments across multiple days helped reduce misclassification bias in exposure measurement. Third, a wide range of sensitivity analyses consistently supported the associations, strengthening the reliability of the conclusions. Fourth, the integration of blood biomarker data offered additional depth, allowing exploration of potential biological pathways. Finally, the careful exclusion of participants with pre-existing cardiometabolic diseases or cancer at baseline minimized reverse causation and ensured a clear temporal relationship between SMD and incident type 2 diabetes.

Several limitations should also be acknowledged. First, fasting glucose and HbA1c measurements were not available for all participants across cohorts and follow-up periods, precluding a standardized biomarker-based assessment of type 2 diabetes; therefore, type 2 diabetes was ascertained using alternative definitions, which may have resulted in some degree of misclassification. In the UK Biobank, type 2 diabetes cases were identified through health record-based diagnoses, which may have missed non-hospitalized or clinically unrecorded cases. In the CHNS, type 2 diabetes cases were identified primarily through self-report physician diagnosis and/or medication use, which may have introduced recall bias and underdiagnosis. Second, inclusion of log (time) interactions in the CHNS led to time-dependent HRs that were difficult to interpret; therefore, we reported main findings using standard Cox models, acknowledging that cohort-wide HRs in the CHNS may be influenced by early follow-up dynamics. Third, the mediation analyses relied on

biomarker data from the UK Biobank baseline and the CHNS 2009, making it impossible to establish the precise temporal ordering between exposure, mediators, and outcomes. Fourth, dietary assessments were based on average intakes from multiple surveys, which may not fully capture long-term dietary changes and could be affected by recall or reporting bias. Fifth, as in all observational studies, residual confounding from unmeasured or imperfectly measured factors cannot be completely excluded. Finally, we were unable to directly validate the SMD score against gut microbiota profiles in the present cohorts, which limits causal inference regarding the role of sulfur-metabolizing bacteria.

In conclusion, from a public health perspective, our findings highlight a potential role of dietary patterns linked to sulfur-metabolizing bacteria in the development of type 2 diabetes, suggesting that modifiable dietary patterns may be relevant targets for prevention. Limiting processed meats and low-calorie drinks while increasing legumes and vegetables may benefit populations with more Westernized diets, whereas moderate intake of fresh fruit juice may be favorable for populations with lower baseline fruit consumption. Reducing liquor intake may be advantageous across populations. Future studies integrating microbiome and multi-omics data across diverse populations are warranted to clarify causality and inform translational strategies.

Acknowledgments

The authors are deeply grateful to the investigators, staff, and participants of the UK Biobank and China Health and Nutrition Survey for their invaluable contributions.

Data Availability

Access to data from the UK Biobank (https://biobank.ndph.ox.ac.uk/ukb/) was granted. All bona fide researchers can apply for access through the UK Biobank according to its standard procedures and criteria. Data from the CHNS are available upon application at https://chns.cpc.unc.edu/.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest Statement

The authors declare that they have no financial or non-financial relationships or activities that could appear to have influenced the submitted work.

References

[1] X. Lin, Y. Xu, X. Pan, et al., Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025, Sci Rep. 10 (2020) 14790. https://doi.org/10.1038/s41598-020-71908-9.

[2] X. Lu, Q. Xie, X. Pan, et al., Type 2 diabetes mellitus in adults: pathogenesis, prevention and therapy, Signal Transduct Target Ther. 9 (2024) 262. https://doi.org/10.1038/s41392-024-01951-9.

[3] M. Gao, S.A. Jebb, P. Aveyard, et al., Associations Between Dietary Patterns and Incident Type 2 Diabetes: Prospective Cohort Study of 120,343 UK Biobank Participants, Diabetes Care. 45 (2022) 1315-1325. https://doi.org/10.2337/dc21-2258.

[4] T. Zhu, M.O. Goodarzi, Metabolites Linking the Gut Microbiome with Risk for Type 2 Diabetes, Curr Nutr Rep. 9 (2020) 83-93. https://doi.org/10.1007/s13668-020-00307-3.

[5] L.H. Nguyen, W. Ma, D.D. Wang, et al., Association Between Sulfur-Metabolizing Bacterial Communities in Stool and Risk of Distal Colorectal Cancer in Men, Gastroenterology. 158 (2020) 1313-1325. https://doi.org/10.1053/j.gastro.2019.12.029.

[6] X. Liu, X. Wan, L. Zhang, et al., The sulfur microbial diet and increased risk of obesity: Findings from a population-based prospective cohort study, Clin Nutr. 42 (2023) 764-772. https://doi.org/10.1016/j.clnu.2023.03.011.

[7] B.D. Muegge, J. Kuczynski, D. Knights, et al., Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans, Science. 332 (2011) 970-974. https://doi.org/10.1126/science.1198719.

[8] J. Na, C. Yang, M. Na, et al., Association between sulfur microbial diet and the risk of dementia: a large-scale prospective cohort study, Nutr J. 24 (2025) 104. https://doi.org/10.1186/s12937-025-01172-y.

[9] C. Sudlow, J. Gallacher, N. Allen, et al., UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age, PLoS Med. 12 (2015) e1001779. https://doi.org/10.1371/journal.pmed.1001779.

[10] R.K. Kelly, T.Y.N. Tong, C.Z. Watling, et al., Associations between types and sources of dietary carbohydrates and cardiovascular disease risk: a prospective cohort study of UK Biobank participants, BMC Med. 21 (2023) 34. https://doi.org/10.1186/s12916-022-02712-7.

[11] B.M. Popkin, S. Du, F. Zhai, et al., Cohort Profile: The China Health and Nutrition Survey--monitoring and understanding socio-economic and health change in China, 1989-2011, Int J Epidemiol. 39 (2010) 1435-1440. https://doi.org/10.1093/ije/dyp322.

[12] A. Perez-Cornago, Z. Pollard, H. Young, et al., Description of the updated nutrition calculation of the Oxford WebQ questionnaire and comparison with the previous version among 207,144 participants in UK Biobank, Eur J Nutr. 60 (2021) 4019-4030. https://doi.org/10.1007/s00394-021-02558-4.

[13] H. Luo, Q. Zhang, K. Yu, et al., Long-term exposure to ambient air pollution is a risk factor for trajectory of cardiometabolic multimorbidity: A prospective study in the UK Biobank, EBioMedicine. 84 (2022) 104282. https://doi.org/10.1016/j.ebiom.2022.104282.

[14] W.H. Organization. Use of Glycated Haemoglobin (HbA1c) in the Diagnosis of Diabetes Mellitus: Abbreviated Report of a WHO Consultation. Geneva: World Health Organization; 2011.

[15] A.S. Thompson, C.J. Candussi, A. Tresserra-Rimbau, et al., A healthful plant-based diet is associated with lower type 2 diabetes risk via improved metabolic state and organ function: A prospective cohort study, Diabetes Metab. 50 (2024) 101499. https://doi.org/10.1016/j.diabet.2023.101499.

[16] Z. Ying, M. Fu, Z. Fang, et al., Mediterranean diet lowers risk of new-onset diabetes: a nationwide cohort study in China, Nutr J. 23 (2024) 131. https://doi.org/10.1186/s12937-024-01036-x.

[17] C.L. Craig, A.L. Marshall, M. Sjostrom, et al., International physical activity questionnaire: 12-country reliability and validity, Med Sci Sports Exerc. 35 (2003) 1381-1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB.

[18] L. Schwingshackl, G. Hoffmann, A.M. Lampousi, et al., Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies, Eur J Epidemiol. 32 (2017) 363-375. https://doi.org/10.1007/s10654-017-0246-y.

[19] H. Tily, E. Patridge, Y. Cai, et al., Gut Microbiome Activity Contributes to Prediction of Individual Variation in Glycemic Response in Adults, Diabetes Ther. 13 (2022) 89-111. https://doi.org/10.1007/s13300-021-01174-z.

[20] C. Mu, Y. Yang, Z. Luo, et al., The Colonic Microbiome and Epithelial Transcriptome Are Altered in Rats Fed a High-Protein Diet Compared with a Normal-Protein Diet, J Nutr. 146 (2016) 474-483. https://doi.org/10.3945/jn.115.223990.

[21] L. Teigen, A. Biruete, A. Khoruts, Impact of diet on hydrogen sulfide production: implications for gut health, Curr Opin Clin Nutr Metab Care. 26 (2023) 55-58. https://doi.org/10.1097/MCO.0000000000000881.

[22] A. Untereiner, L. Wu, Hydrogen Sulfide and Glucose Homeostasis: A Tale of Sweet and the Stink, Antioxid Redox Signal. 28 (2018) 1463-1482. https://doi.org/10.1089/ars.2017.7046.

[23] F.W. Dekker, R. de Mutsert, P.C. van Dijk, et al., Survival analysis: time-dependent effects and time-varying risk factors, Kidney Int. 74 (2008) 994-997. https://doi.org/10.1038/ki.2008.328.

[24] J. Qin, Y. Li, Z. Cai, et al., A metagenome-wide association study of gut microbiota in type 2 diabetes, Nature. 490 (2012) 55-60. https://doi.org/10.1038/nature11450.

[25] Q. Qi, H. Zhang, Z. Jin, et al., Hydrogen sulfide produced by the gut microbiota impairs host metabolism via reducing GLP-1 levels in male mice, Nat Metab. 6 (2024) 1601-1615. https://doi.org/10.1038/s42255-024-01068-x.

[26] T.D. Muller, B. Finan, S.R. Bloom, et al., Glucagon-like peptide 1 (GLP-1), Mol Metab. 30 (2019) 72-130. https://doi.org/10.1016/j.molmet.2019.09.010.

[27] R. Wu, Z. Yu, P. Guo, et al., Desulfovibrio vulgaris exacerbates sepsis by inducing inflammation and oxidative stress in multiple organs, Front Microbiol. 16 (2025) 1574998. https://doi.org/10.3389/fmicb.2025.1574998.

[28] M.Y. Donath, S.E. Shoelson, Type 2 diabetes as an inflammatory disease, Nat Rev Immunol. 11 (2011) 98-107. https://doi.org/10.1038/nri2925.

[29] X. Zheng, Y. Qian, L. Wang, Causal relationship between gut microbiota and insulin-like growth factor 1: a bidirectional two-sample Mendelian randomization study, Front Cell Infect Microbiol. 14 (2024) 1406132. https://doi.org/10.3389/fcimb.2024.1406132.

[30] J.A. Santos-Marcos, M. Mora-Ortiz, M. Tena-Sempere, et al., Interaction between gut microbiota and sex hormones and their relation to sexual dimorphism in metabolic diseases, Biol Sex Differ. 14 (2023) 4. https://doi.org/10.1186/s13293-023-00490-2.

[31] N. Friedrich, B. Thuesen, T. Jorgensen, et al., The association between IGF-I and insulin resistance: a general population study in Danish adults, Diabetes Care. 35 (2012) 768-773. https://doi.org/10.2337/dc11-1833.

[32] E.L. Ding, Y. Song, V.S. Malik, et al., Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis, JAMA. 295 (2006) 1288-1299. https://doi.org/10.1001/jama.295.11.1288.

[33] Z. Liu, H. Huang, J. Ruan, et al., The sulfur microbial diet and risk of nonalcoholic fatty liver disease: a prospective gene-diet study from the UK Biobank, Am J Clin Nutr. 119 (2024) 417-424. https://doi.org/10.1016/j.ajcnut.2023.11.012.

[34] A. Fraser, R. Harris, N. Sattar, et al., Alanine aminotransferase, gamma-glutamyltransferase, and incident diabetes: the British Women's Heart and Health Study and meta-analysis, Diabetes Care. 32 (2009) 741-750. https://doi.org/10.2337/dc08-1870.

[35] B. Hemmingsen, G. Gimenez-Perez, D. Mauricio, et al., Diet, physical activity or both for prevention or delay of type 2 diabetes mellitus and its associated complications in people at increased risk of developing type 2 diabetes mellitus, Cochrane Database Syst Rev. 12 (2017) CD003054. https://doi.org/10.1002/14651858.CD003054.pub4.

[36] S. Almuraikhy, A. Doudin, A. Domling, et al., Molecular regulators of exercise-mediated insulin sensitivity in non-obese individuals, J Cell Mol Med. 28 (2024) e18015. https://doi.org/10.1111/jcmm.18015.

[37] E. Trefts, A.S. Williams, D.H. Wasserman, Exercise and the Regulation of Hepatic Metabolism, Prog Mol Biol Transl Sci. 135 (2015) 203-225. https://doi.org/10.1016/bs.pmbts.2015.07.010.

[38] J. Tarp, M. Luo, M.A. Sanchez-Lastra, et al., Leisure-time physical activity and all-cause mortality and cardiovascular disease in adults with type 2 diabetes: Cross-country comparison of cohort studies, J Sport Health Sci. 13 (2024) 212-221. https://doi.org/10.1016/j.jshs.2023.10.004.

[39] T.A. Matthews, X. Liu, L. Chen, et al., Prospective associations of occupational and leisure-time physical activity with risk of diabetes: a cohort study from the United States, Ann Work Expo Health. 68 (2024) 581-592. https://doi.org/10.1093/annweh/wxae034.

[40] R.C. Ma, J.C. Chan, Type 2 diabetes in East Asians: similarities and differences with populations in Europe and the United States, Ann N Y Acad Sci. 1281 (2013) 64-91. https://doi.org/10.1111/nyas.12098.