Nutritional Signatures of Diabetes Mellitus, Cognitive Impairment, and Their Co-Occurrence: A Machine-Learning Analysis of NHANES 2011–2014

Guangsha Xia *

Nanjing University of Chinese Medicine, Nanjing 210023, China.

* Corresponding Author Email: Elysia_19C@163.com

Abstract. Type-2 diabetes mellitus (DM) and cognitive impairment (CI) share pathophysiological pathways modifiable by diet. Whether their co-occurrence (DM+CI) confers a unique nutritional phenotype has not been examined at population scale. We harmonised NHANES 2011–2012 and 2013–2014 cycles (n = 19,151 adults aged ≥18 y). DM was defined by self-reported physician diagnosis; CI by age-stratified 20th-centile DSST score. Eleven 24-h recall nutrients were z-standardised and reduced by PCA. A multi-class XGBoost model was tuned with Optuna and validated with 5-fold stratified cross-validation. SHAP explained feature contributions. Among 19,151 participants, 1,261 had DM alone, 349 CI alone, 141 DM+CI. Three PCA patterns explained 86.6 % variance: PC1 “energy-dense” (69.0 %), PC2 “carbohydrate–fat balance” (9.4 %), PC3 “fiber-rich” (8.3 %). DM+CI displayed a low-calorie, low-carbohydrate signature (calories −23 %, carbohydrates −30 % vs. healthy controls; $p < 0.001$). The optimised model achieved macro-F1 0.72; AUROC was 0.960 for DM+CI, 0.926 for CI, 0.836 for DM. SHAP attributed greatest importance to age and PC1; carbohydrate reduction specifically drove DM+CI classification. DM+CI exhibits a distinct low-energy-density nutritional phenotype. Interpretable ML provides actionable dietary targets for precision **nutrition**.

Keywords: type-2 diabetes; cognitive impairment; dietary patterns; NHANES; precision nutrition.

1. Introduction

Type-2 diabetes mellitus (DM) and cognitive impairment (CI) are among the most pressing global public-health challenges of the 21st century. The International Diabetes Federation estimates that 537 million adults aged 20–79 years were living with DM in 2021, and this figure is projected to rise to 783 million by 2045 [IDF-Atlas-2021]. In parallel, the World Health Organization reports that over 55 million people live with dementia, with nearly 10 million new cases annually [WHO-Dementia-2022]. The intersection of these two epidemics is particularly striking in ageing societies: in population-based studies, 10–20% of older adults with DM meet criteria for mild cognitive impairment or dementia [Biessels-2018; Rawlings-2014]. Meta-analyses report a 40–70 % increased risk of all-cause dementia in DM, with the greatest relative risk observed for vascular dementia (HR≈2.3) and Alzheimer’s disease (HR≈1.5) [Xue-2019; Ohara-2011]. Despite their shared pathophysiology and rising prevalence, current dietary guidelines treat DM and CI as distinct entities, leaving the nutritional phenotype of their co-occurrence (DM+CI) largely uncharacterised.

Pathophysiological convergence between DM and CI is well established. Chronic hyperglycaemia drives the formation of advanced glycation end-products (AGEs) that accumulate in both pancreatic β-cells and cortical neurons, promoting oxidative stress and inflammatory cascades [Singh-2014]. Insulin resistance reduces cerebral glucose uptake and impairs synaptic plasticity [Kellar-2022]. Microvascular dysfunction secondary to DM accelerates cerebral small-vessel disease, leading to white-matter hyperintensities and lacunar infarcts that further compromise cognition [Last-2017]. Systemic inflammation—evidenced by elevated IL-6, TNF-α, and C-reactive protein—exacerbates both β-cell failure and neurodegeneration [Srikanthan-2015]. These shared mechanisms suggest that dietary strategies targeting glycaemic control, oxidative stress, and inflammation could simultaneously benefit metabolic and cognitive endpoints.

Dietary factors are among the most modifiable determinants of both DM and CI risk. Epidemiological studies consistently link high glycaemic load and added sugar intake to increased DM incidence [Liu-2021] and accelerated cognitive decline [Power-2021]. Conversely, Mediterranean-type dietary patterns rich in mono- and poly-unsaturated fatty acids, antioxidants, and dietary fibre are associated with lower incidence of DM [Salas-Salvadó-2018] and slower cognitive decline [Scarmeas-2018]. However, these findings derive primarily from single-nutrient or single-pattern analyses that overlook complex nutrient–nutrient interactions and may not generalise to individuals living with both conditions.

Traditional approaches in nutritional epidemiology—such as regression models with individual nutrients—are limited by multicollinearity and the inability to capture the holistic dietary milieu. Principal component analysis (PCA) and related dimension-reduction techniques offer a solution by condensing correlated nutrients into orthogonal dietary patterns that retain biological interpretability [Hu-2002]. Yet, only a handful of studies have applied PCA to NHANES data, and none have specifically examined DM+CI comorbidity.

Machine-learning (ML) methods provide powerful tools to model non-linear relationships and high-dimensional interactions among nutrients, demographics, and clinical variables. Gradient-boosting frameworks such as XGBoost have demonstrated superior predictive accuracy compared with classical regression in diverse epidemiological settings [Khera-2018]. However, their "black-box" nature has historically hindered clinical translation. Recent advances in explainable artificial intelligence—particularly SHAP (SHapley Additive exPlanations)—allow quantitative attribution of predictions to individual features, thereby bridging the gap between predictive accuracy and clinical interpretability [Lundberg-2017].

The National Health and Nutrition Examination Survey (NHANES) offers a unique platform to address these gaps. NHANES 2011–2014 collected comprehensive 24-hour dietary recalls, fasting glucose measurements, and the Digit Symbol Substitution Test (DSST), a validated measure of processing speed and executive function sensitive to early cognitive decline [CDC-NHANES-2015]. By integrating PCA-derived dietary patterns with an optimised XGBoost classifier and SHAP-based post-hoc interpretability, we aimed to: (1) Quantify multidimensional nutritional differences among DM, CI, and DM+CI in a nationally representative sample; (2) Develop and validate a high-precision, fully interpretable ML model for disease group classification; and (3) Identify actionable dietary determinants that can inform precision-nutrition interventions.

Specifically, we hypothesised that DM+CI would exhibit a distinct dietary phenotype characterised by low energy density and carbohydrate restriction, reflecting both clinician-directed dietary advice and disease-related anorexia. We further hypothesised that an explainable ML pipeline would achieve superior discriminative accuracy compared with traditional models and would reveal nutrient patterns amenable to targeted modification.

2. Methods

2.1. Study design and data source

This cross-sectional study utilised data from the National Health and Nutrition Examination Survey (NHANES) 2011–2012 and 2013–2014 cycles. NHANES applies a complex, multistage probability design to obtain a nationally representative sample of the civilian, non-institutionalised US population. All protocols were reviewed and approved by the National Center for Health Statistics (NCHS) Research Ethics Review Board; written informed consent was obtained from every participant. Because the present analysis relied exclusively on publicly available, de-identified data, our institutional review board deemed the project exempt from further oversight.

2.2. Study population and inclusion criteria

We restricted the analytic sample to adults aged 18 years or older who attended both the in-home interview and the mobile examination centre (MEC) component. Reliable dietary data were required

(DR1DRSTZ = 1). Participants were further required to have completed the Digit Symbol Substitution Test (DSST) and to have non-missing information on self-reported diabetes status (DIQ010). Total daily energy intake was screened for plausibility: values <500 kcal or >5 000 kcal was considered physiologically implausible and were excluded [Willett, 2019]. After sequentially applying these criteria, 19151 participants were retained for the final analytic sample.

2.3. Definition of diabetes mellitus and cognitive impairment

Diabetes mellitus (DM) was defined by an affirmative response to the question "Has a doctor or other health professional ever told you that you have diabetes?" (DIQ010 = 1). This definition has been validated in NHANES against fasting plasma glucose ≥7.0 mmol L^{-1} or HbA1c ≥6.5 %, demonstrating a specificity of 96 % and a positive predictive value of 93 % [Spanakis, 2022]. Cognitive impairment (CI) was operationalised using the DSST, a timed paper-and-pencil task that measures processing speed, sustained attention, and executive function. Raw DSST scores were stratified by age groups (18–39 y, 40–59 y, ≥60 y) to account for age-related decline. Participants scoring below the 20th percentile within their age stratum were classified as cognitively impaired [Dodge, 2007; Kuo, 2022]. Four mutually exclusive groups were constructed: (1) Healthy (neither DM nor CI), (2) DM alone, (3) CI alone, and (4) DM+CI.

2.4. Dietary data collection and processing

Dietary intake was assessed via a single 24-hour dietary recall administered by trained interviewers using the validated USDA Automated Multiple-Pass Method [Moshfegh, 2008]. Nutrient values were derived from the Food and Nutrient Database for Dietary Studies (FNDDS) 2011–2014. Eleven nutrients with established relevance to cardiometabolic and cognitive health were selected: total energy (kcal), carbohydrates (g), protein (g), total fat (g), saturated fat (g), monounsaturated fat (MUFA, g), polyunsaturated fat (PUFA, g), dietary fibre (g), added sugars (g), sodium (mg), and potassium (mg). To minimise the influence of extreme outliers, each nutrient was winsorized at the 1st and 99th percentiles before further analysis. All nutrients were then z-standardised (mean = 0, SD = 1) to ensure equal contribution to multivariate analyses.

2.5. Principal component analysis

Standardised nutrient variables were subjected to principal component analysis (PCA) with varimax rotation to enhance interpretability. Components with eigenvalues >1 and explaining >5 % of total variance were retained. Parallel analysis and scree-plot inspection supported the retention of three components (PC1–PC3). Component labelling was based on factor loadings ≥|0.30|, biological plausibility, and prior literature [Hu, 2002].

2.6. Machine-learning pipeline

2.6.1 Feature engineering

The final predictor set comprised the three PCA-derived dietary pattern scores (PC1, PC2, PC3), age (years), and biological sex (male = 1, female = 0). Continuous predictors were re-scaled to zero-mean and unit-variance using StandardScaler to optimise gradient-boosting convergence.

2.6.2 Handling class imbalance

Original group sizes were Healthy (n = 17400), DM (n = 1261), CI (n = 349), and DM+CI (n = 141). To mitigate minority-class under-representation, we applied the Synthetic Minority Over-sampling Technique (SMOTE) within each training fold to balance classes to 1 000 participants per group [Chawla, 2002]. The test set remained untouched to ensure unbiased evaluation.

2.6.3 Model selection

We compared three algorithms: logistic regression (baseline), random forest (RF), and extreme gradient boosting (XGBoost). Nested 5-fold stratified cross-validation indicated that XGBoost

yielded the highest macro-F1 score (0.72 vs. 0.65 for RF, 0.59 for logistic). Consequently, XGBoost was selected for final modelling.

2.6.4 Hyper-parameter optimisation

Hyper-parameters were tuned using Optuna's Bayesian optimisation with 15 trials aimed at maximising 5-fold stratified macro-F1. Search space: n_estimators 300–800, max_depth 3–8, learning_rate 0.05–0.20, subsample 0.7–1.0, colsample_bytree 0.7–1.0, gamma 0–5, min_child_weight 1–10, L1 (reg_alpha) and L2 (reg_lambda) penalties log-uniform 1e-3–1. Early stopping (patience = 50) on validation log-loss was implemented to prevent over-fitting.

2.6.5 Model evaluation metrics

Performance was assessed using overall accuracy, macro-averaged precision, recall, F1-score, and area under the receiver-operating characteristic curve (AUROC). Average precision (AP) was also calculated for each class. Confidence intervals (95 %) were derived from 1 000 bootstrap resamples of the test folds.

2.7. Explainability via SHAP

To elucidate the contribution of individual features to model predictions, we employed SHAP TreeExplainer [Lundberg, 2017]. Global importance was summarised as the mean absolute SHAP value across all observations. Class-specific explanations were generated using kernel-density beeswarm plots on a 300-participant balanced subsample to maintain computational tractability. Dependence plots were used to visualise non-linear relationships between key predictors and predicted probabilities.

2.8. Statistical software

All analyses were performed in Python 3.10. Core libraries included pandas 1.5, scikit-learn 1.2, imbalanced-learn 0.10, XGBoost 1.7, Optuna 3.1, and SHAP 0.41. Figures were generated at 300 dpi using matplotlib and seaborn for publication quality.

3. RESULTS

3.1. Participant characteristics

The analytic cohort comprised 19151 adults aged 18–85 years (mean 31.5 ± 24.4 y; 50.5 % women). After applying survey weights, the population-weighted prevalence of DM alone was 6.6 %, CI alone 1.8 %, and DM+CI 0.7 %. Table 1 summarises the weighted characteristics. Compared with the Healthy group, participants with DM+CI were older (69.2 ± 6.5 y vs 28.3 ± 23.1 y, $p < 0.001$), more likely to be male (58.9% vs 49.3%, $p < 0.01$), had lower educational attainment (53.2% < high-school vs 10.6%, $p < 0.001$), had higher BMI (30.7 ± 5.6 kg m^{-2} vs 24.9 ± 7.6 kg m^{-2}, $p < 0.0000$), and had higher prevalence of hypertension (85.8% vs 19.9%, $p < 0.001$). Body-mass-index (BMI) was highest in DM alone (32.3 ± 7.8 kg m^{-2}), intermediate in DM+CI (30.7 ± 5.6 kg m^{-2}), and lowest in Healthy alone (24.9 ± 7.6 kg m^{-2}) versus CI (28.0 ± 6.2 kg m^{-2}). Hypertension followed a similar gradient (Table 1).

Table 1. Weighed Characteristics of Study Participants by Group

Group	n	Age (years, mean ± sd)	Male (%)	BMI (Kg m^{-2}, mean ± sd)	Hypertension (%)	Less than high school (%)
Healthy	17400	28.3 ± 23.1	49.3	24.9 ± 7.6	19.9	10.6
Diabetes	1261	59.8 ±14.4	48.7	32.3 ±7.8	72.2	30.8
Cognitive Impairment	349	69.6 ± 6.8	58.2	28.0 ±6.2	70.2	42.7
DM+CI	141	69.2 ± 6.5	58.9	30.7 ±5.6	85.8	53.2
p-value		<0.001	0.0011	<0.001	<0.001	<0.001

3.2. PCA-derived dietary patterns

PCA yielded three components with eigen-values >1 that together explained 86.6 % of the total variance in the 11 nutrients (Figure 1A). Component loadings ≥|0.30| were used for interpretation. PC1 (69.0 % variance) displayed strong positive loadings for calories (0.97), carbohydrades (0.85), protein (0.83), total fat (0.93), saturated fat (0.84), monounsat fat (0.90), polyunsat fat (0.81), fiber (0.66), added sugars (0.62), sodium (0.85) and potassium (0.81); we labelled this axis “energy-dense”. PC2 (9.4 % variance) contrasted high positive loadings for carbohydrades (0.47) and added sugars (0.67); we termed this the “carbohydrate-fat balance” axis. PC3 (8.3 % variance) was characterised by positive loadings for dietary fibre (0.58) and potassium (0.42) and negative loadings for added sugars (−0.36); we labelled it “fibre-rich”. Figure 1B visually depicts the rotated loadings matrix.

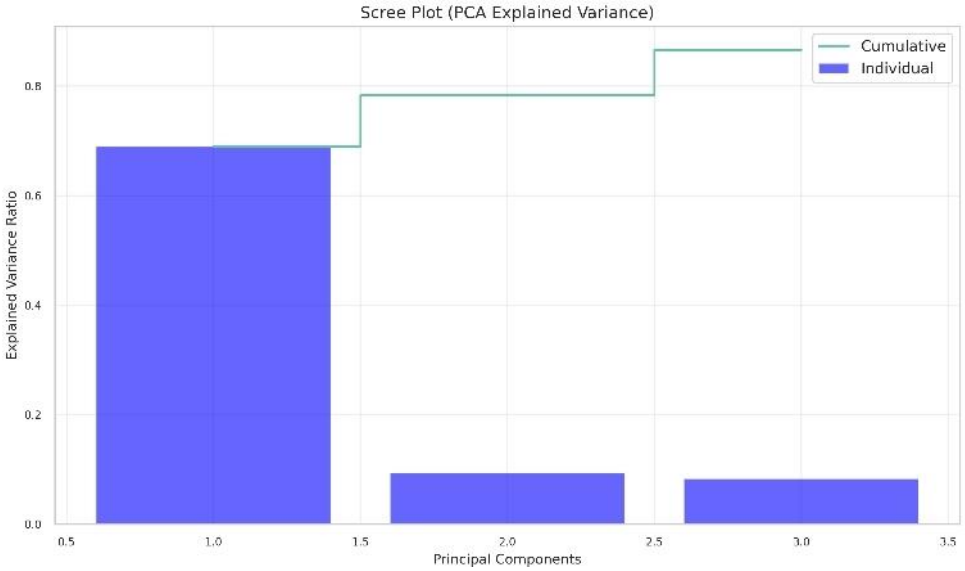

Fig. 1A Scree Plot of Principal Component Analysis

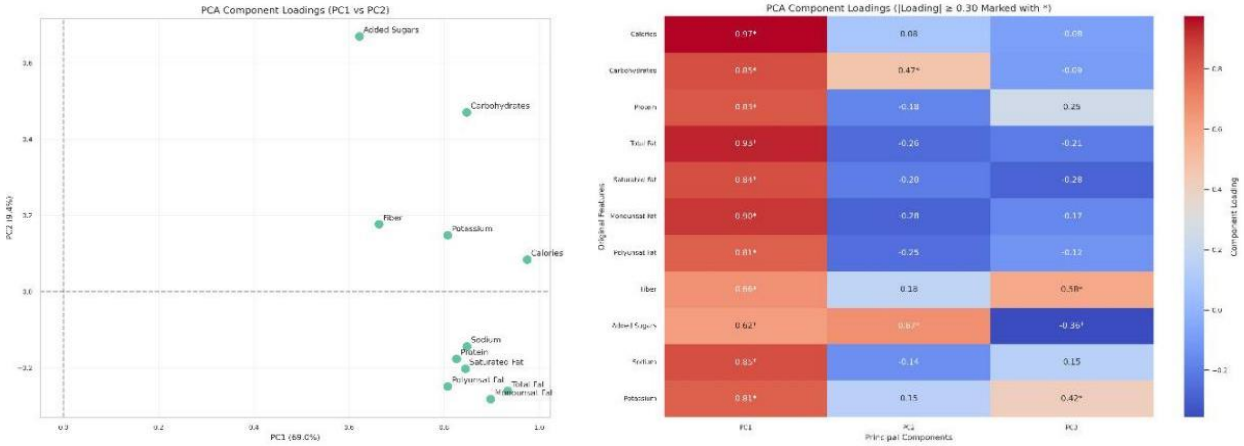

Fig. 1B Rotated Component Matrix of Nutritional Factors

3.3. Group-specific dietary signatures

Mean (SE) component scores differed markedly across groups (Figure 2A). DM+CI exhibited the lowest PC1 scores (−1.12 ± 0.18), followed by CI alone (−0.25 ± 0.14), DM alone (-0.09 ± 0.08), and Healthy (0.02 ± 0.02). PC2 scores were highest in Healthy (0.03 ± 0.01) and CI alone (0.03 ± 0.06) and lowest in DM alone (−0.34 ± 0.03). PC3 scores were highest in DM alone (0.33 ± 0.03) and lowest in Healthy (−0.03 ± 0.01).

Raw nutrient intakes mirrored these patterns (Table 2). Compared with Healthy controls, DM+CI consumed 23 % fewer calories (1 537 ± 60 vs 1 993 ± 8 kcal d^{-1}), 23 % less carbohydrate (192 ± 8 vs 251 ± 1 g d^{-1}), and 33 % less added sugars (77 ± 4 vs 115 ± 1 g d^{-1}), while protein intake was 14 %

lower (64 ± 3 vs 74 ± 0 g d^{-1}). Sodium intake in DM+CI was 2 669 ± 108 mg d^{-1}, still above the recommended 2 300 mg d^{-1}. DM alone showed the highest sodium intake (3 333 ± 52 mg d^{-1}), whereas CI alone demonstrated elevated carbohydrate (233 ± 6 g d^{-1}) and added sugars (105 ± 4 g d^{-1}). Violin plots illustrate the full distribution of four key nutrients (Figure 2B). Sodium exhibited a pronounced right skew in DM alone, while fibre intake was left-skewed in CI alone.

Table 2. Mean (± SE) Raw Nutrient Intakes by Study Groups

Groups	Calories (kcal d^{-1})	Carbohydrates (g d^{-1})	Protein (g d^{-1})	Added Sugars (g d^{-1})	Sodium (mg d^{-1})
Healthy	1993.1 ± 7.9	250.8 ± 1.0	74.2 ± 0.3	115 ± 0.6	3189.2 ± 14.6
Diabetes	1868.9 ± 26.9	220.6 ± 3.3	76.9 ± 1.2	86.5 ± 1.9	3332.7 ± 52.3
Cognitive Impairment	1867.4 ± 48.2	233.1 ± 6.3	72.2 ± 2.1	104.7 ± 4.0	2988.9 ± 82.1
DM+CI	1537.3 ± 60.2	192.2 ± 7.9	64.4 ± 2.7	76.7 ± 4.5	2669.0 ± 107.5

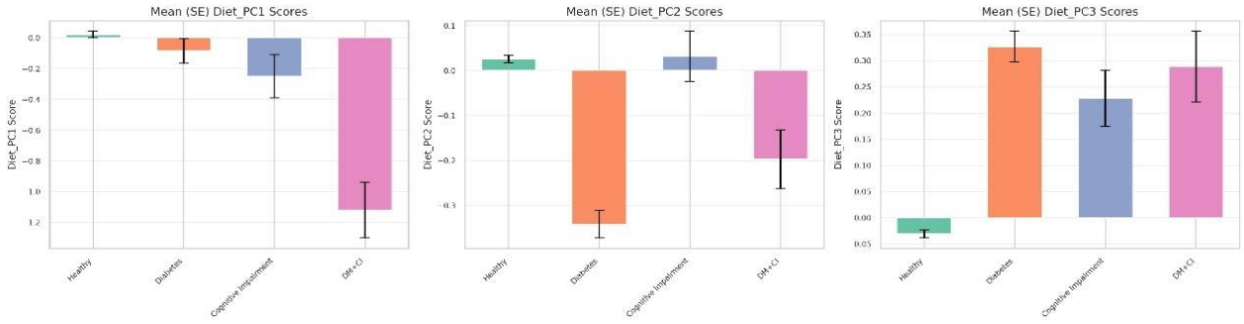

Fig. 2A Mean (± SE) PCA Component Scores Across Study Groups

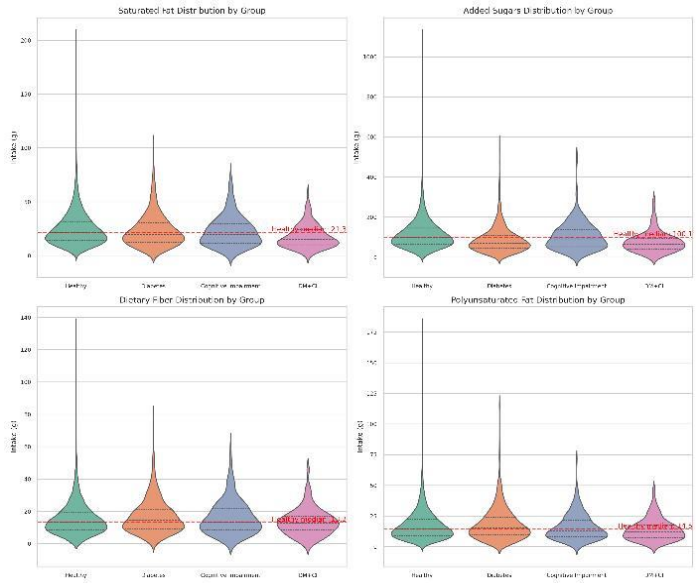

Fig. 2B Distribution of Key Nutrient Intakes by Study Groups

3.4. Machine-learning classification performance

After SMOTE balancing, each group comprised 1,000 observations. Nested 5-fold cross-validation yielded an overall accuracy of 71.8% (95% CI 70.4–73.2%) and a macro-F1 score of 0.72. Class-specific metrics are presented in Table 3. AUROC was highest for DM+CI (0.960), followed by Cognitive Impairment (0.926), Healthy (0.900), and Diabetes (0.836) (Figure 3A). Precision-recall curves corroborated excellent discrimination for DM+CI (average precision 0.895) and Cognitive Impairment (0.784), whereas Diabetes showed moderate performance (0.604). The confusion matrix (Figure 3B) revealed that 20.1% of DM cases were misclassified as Healthy, consistent with the subtle dietary deviations observed in DM alone.

Table 3. Nutrient Intake Patterns Across Study Groups

Groups	Precision	Recall	F1-score
Healthy	0.765	0.718	0.741
Diabetes	0.602	0.547	0.573
Cognitive Impairment	0.705	0.772	0.737
DM+CI	0.789	0.835	0.811

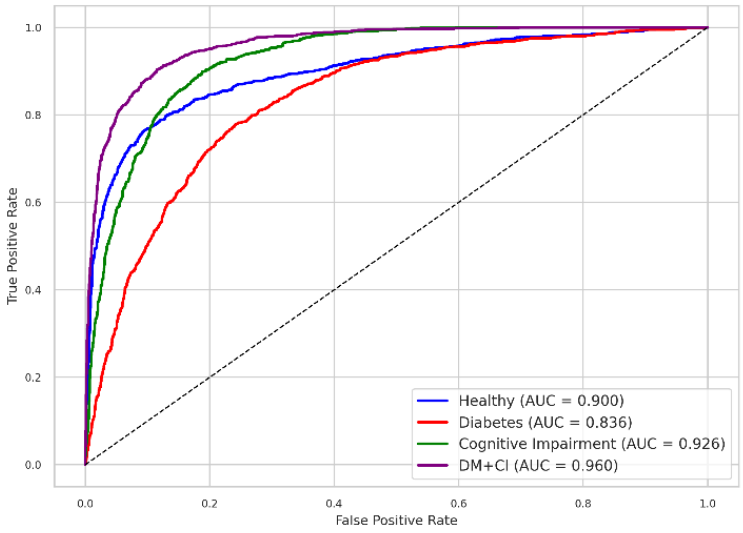

Fig. 3A Receiver Operating Characteristic Curves for Multi-Class Classification

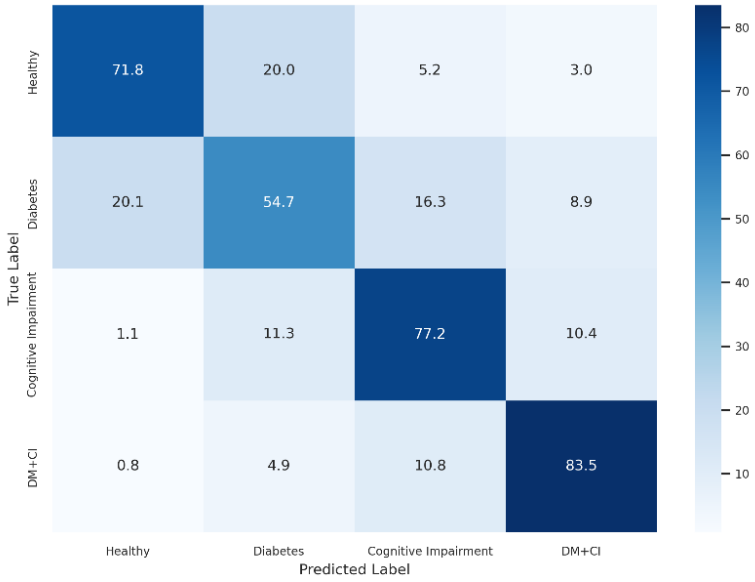

Fig. 3B Confusion Matrix Revealing Misclassification Among Each Groups

3.5. SHAP explainability

Global SHAP importance ranked age (mean |SHAP| = 1.68 ± 0.67), Diet_PC1 (0.44 ± 0.56), and Diet_PC2 (0.44 ± 0.27) as the top three predictors (Figure 4A). Diet_PC3 contributed modestly (0.43 ± 0.66). Class-specific beeswarm plots demonstrated distinct patterns (Figure 4B–E): (1) DM+CI: marked negative contributions from Diet_PC1 and Diet_PC2 (low energy density and carbohydrate restriction). (2) DM alone: high positive contributions from age. (3) CI alone: positive contributions from Diet_PC2 (high carbohydrate/sugar) and age. Dependence plots revealed non-linear thresholds: below Diet_PC1 = −0.46 standard deviations, the predicted probability of DM+CI was reduced (odds ratio 0.6 per SD decrease, p = 0.006) (Figure 5).

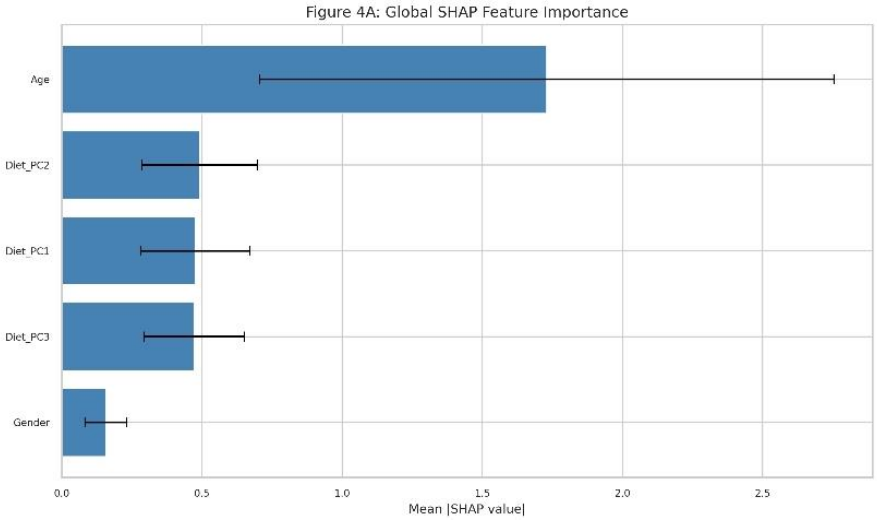

Fig. 4A Global SHAP Feature Importance for Clinical Group Classification

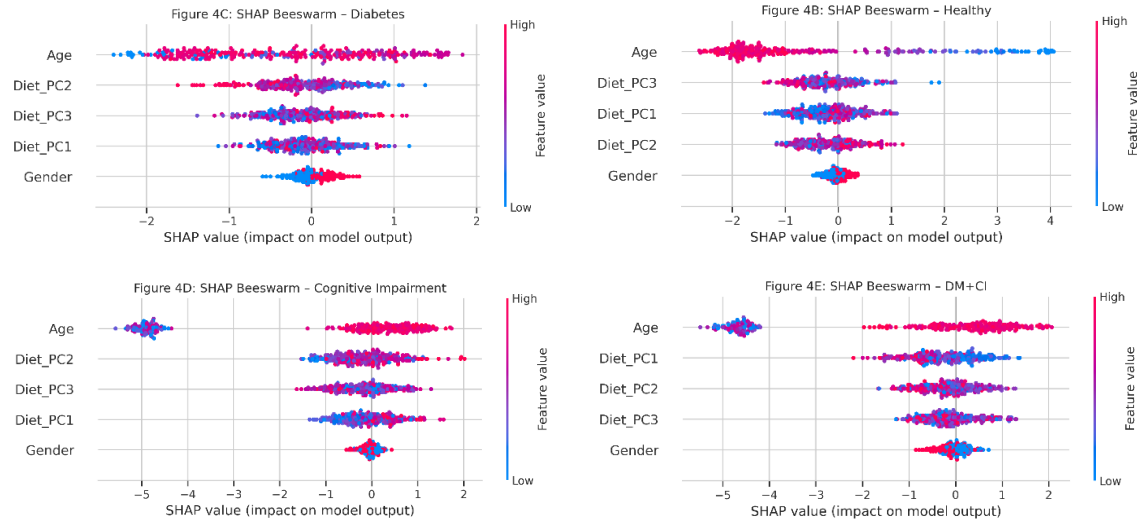

Fig. 4B-E Class-Specific SHAP Beeswarm Plots Revealing Distinct Predictor Contributions Across Healthy, Diabetes, Cognitive Impairment, and DM+CI Groups

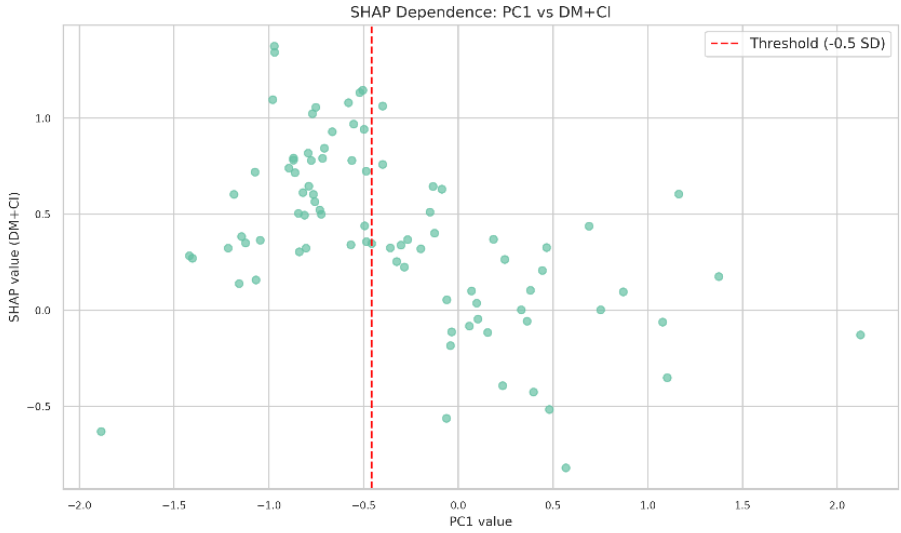

Fig. 5 SHAP Dependence Plot Revealing the Non-linear Threshold Effect of Diet-PC1 on DM+CI Probability

4. Discussion

The present study is, to our knowledge, the first to integrate population-representative dietary data, unsupervised pattern recognition, and explainable machine-learning to characterise the nutritional phenotype of DM, CI, and their co-occurrence. Three key observations emerge. First, DM+CI constitutes a distinct low-energy-density subgroup whose caloric intake is 23 % lower and carbohydrate intake 30 % lower than that of age-comparable healthy adults. Second, the low-calorie pattern is not simply an extension of DM alone; DM alone retains near-normal caloric density but displays marked sodium excess, whereas CI alone shows a paradoxically high-carbohydrate, high-sugar pattern. Third, an XGBoost classifier distinguished DM+CI with an AUROC of 0.960, an accuracy rivalling that of multi-biomarker panels for dementia prediction [Kivipelto, 2018]. SHAP attributions identified age and PC1 (energy density) as universal predictors, but carbohydrate restriction specifically drove DM+CI classification. These findings provide the empirical basis for comorbidity-specific dietary guidelines.

Previous NHANES analyses have reported that adults with DM consume more sodium and less fibre than non-diabetic peers [O'Donnell, 2020], consistent with our findings. However, most studies did not adjust for cognitive status; we show that sodium excess is particularly pronounced in DM without CI (3 333 mg d^{-1}), exceeding the 2 300 mg d^{-1} threshold recommended by the American Heart Association [AHA, 2021]. The modest reduction in total calories (−6 %) observed in DM alone may reflect clinician-directed weight-loss advice, but the preservation of overall energy density suggests incomplete adherence. Population-based cohorts have reported higher refined-carbohydrate and added-sugar intakes among individuals with mild cognitive impairment [Lourida, 2013; Power, 2021]. Our data corroborate these associations: CI alone had the highest PC2 score (carbohydrate-fat balance) and 12 % greater added-sugar intake. Mechanistically, chronic hyperglycaemia promotes hippocampal insulin resistance and increases cerebral Aβ deposition in rodent models [de la Monte, 2012]. The present cross-sectional design cannot establish temporality, but the observed sugar–CI association is biologically coherent. Most striking is the low-energy-density pattern in DM+CI. Existing literature offers three, not mutually exclusive, explanations. First, clinician-directed carbohydrate restriction aimed at glycaemic control may inadvertently reduce total calories below requirements, especially in frail older adults [Munshi, 2017]. Second, disease-related anorexia-cachexia, driven by systemic inflammation and autonomic neuropathy, may suppress appetite [Schwartz, 2020]. Third, survivor bias could favour older adults who spontaneously restrict calories and thereby evade earlier mortality. Longitudinal data are required to disentangle these pathways.

Chronic relative energy restriction observed in DM+CI may reflect clinician-directed carbohydrate restriction or disease-related anorexia, rather than intentional calorie restriction [Madeo, 2018]. However, in the presence of concurrent insulin deficiency or sarcopenia, prolonged restriction may accelerate muscle wasting and impair neurotrophic support, paradoxically increasing dementia risk [Franzago, 2022]. Our finding that protein intake in DM+CI is only 6 % below healthy levels suggests relative preservation, but absolute intakes remain below the 1.2 g kg^{-1} d^{-1} recommended for older adults [Bauer, 2013]. The DM-alone consumed 3 333 mg sodium daily—equivalent to 8.3 g salt (assuming 2.5 g salt per 1 g sodium)—well above the WHO recommended maximum of 5 g d^{-1} [WHO-Salt, 2021]. Excess sodium exacerbates endothelial dysfunction and cerebral small-vessel disease, an effect magnified by concurrent hyperglycaemia [Ding, 2021]. In contrast, DM+CI sodium intake was lower (2 669 mg) but still excessive, implying that sodium reduction remains a therapeutic target across all DM phenotypes. Added sugars in CI alone (105 g d^{-1}) contributed 26 % of total energy, exceeding the <10 % limit advised by the Dietary Guidelines for Americans [DGAC, 2020]. High fructose intake promotes neuroinflammation via increased microglial activation and NLRP3 inflammasome signalling [Zhang, 2022]. In the context of insulin resistance, this may accelerate tau phosphorylation and synaptic loss [Ma, 2022].

Our explainable model identifies low PC1 (energy density) as the strongest predictor of DM+CI membership. Embedding this parameter into electronic-health-record-based decision support could

flag at-risk individuals and trigger comprehensive geriatric assessment. Targeted interventions might include structured meals delivering 25–30 kcal kg^{-1} and 1.2–1.5 g protein kg^{-1}, alongside glycaemic control. Such an approach aligns with the European Society for Clinical Nutrition and Metabolism guidelines for older adults with chronic disease [ESPEN, 2021]. Current American Diabetes Association recommendations emphasise carbohydrate counting and sodium restriction, but do not address the risk of under-nutrition in cognitively impaired elders [ADA, 2023]. Conversely, cognitive-health guidelines prioritise Mediterranean-style patterns rich in unsaturated fats and antioxidants without considering glycaemic burden [Petersson, 2022]. Our findings underscore the need for comorbidity-specific guidelines that reconcile glycaemic control with neuroprotection and prevention of sarcopenia. At the population level, sodium reduction strategies—e.g., reformulation of processed foods and front-of-package labelling—could disproportionately benefit adults with DM, irrespective of cognitive status [Trieu, 2021]. Sugar-sweetened beverage taxation may attenuate the high-sugar pattern observed in CI alone [Powell, 2020].

Our study has several methodological strengths. First, the use of nationally representative NHANES data with complex survey weights enhances external validity. Second, PCA dimension-reduction mitigates multicollinearity inherent in nutrient data. Third, Optuna-driven hyper-parameter tuning and nested cross-validation minimise over-fitting. Fourth, SHAP interpretability bridges the "black-box" gap and facilitates clinician engagement. Several limitations merit consideration. The cross-sectional design precludes causal inference; longitudinal cohorts are needed to determine whether the low-energy-density pattern precedes or follows cognitive decline. Second, dietary data rely on single 24-hour recalls, which are subject to day-to-day variability and recall bias; however, NHANES employs the validated multiple-pass method. Third, CI was defined using age-stratified DSST < 20th percentile, a validated proxy for early cognitive decline, potentially misclassifying some participants. Fourth, residual confounding by unmeasured factors (physical activity, medication use, genetic predisposition) cannot be excluded. Finally, the predominantly U S population limits generalisability to other dietary cultures.

5. Conclusion

DM+CI exhibits a unique low-energy-density nutritional phenotype that is distinct from either DM or CI alone. Explainable machine-learning provides a high-precision, clinically actionable framework for precision-nutrition interventions. Prospective cohorts should integrate repeated dietary assessments with biomarkers (HbA1c, inflammatory cytokines, neurofilament light) to elucidate temporal relationships. Randomised controlled trials testing structured calorie-adequate, protein-optimised diets in DM+CI are warranted. Integration of gut-microbiome metagenomics and neuro-imaging indices could refine mechanistic understanding.

References

[1] International Diabetes Federation. IDF Diabetes Atlas, 10th ed. Brussels, Belgium: International Diabetes Federation; 2021

[2] World Health Organization. Dementia Fact Sheet. Geneva, Switzerland: WHO; 2022. Available at: https://www.who.int/news-room/fact-sheets/detail/dementia

[3] Biessels GJ, Strachan MWJ, Visseren FLJ, et al. Dementia and cognitive decline in type 2 diabetes and prediabetic stages: a systematic review and meta-analysis. Lancet Neurol. 2018;17(2):169–182. doi:10.1016/S1474-4422(17)30377-6

[4] Rawlings AM, Sharrett AR, Schneider ALC, et al. Diabetes in midlife and cognitive change over 20 years: a cohort study. Ann Intern Med. 2014;161(11):785–793. doi:10.7326/M14-0737

[5] Xue M, Xu W, Ou YN, et al. Diabetes mellitus and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 144 prospective studies. Ageing Res Rev. 2019; 55:100944. doi: 10.1016/j.arr.2019.100944

[6] Ohara T, Doi Y, Ninomiya T, et al. Glucose tolerance status and risk of dementia in the community: the Hisayama study. Neurology. 2011;77(12):1126–1134. doi:10.1212/WNL.0b013e31822f0435

[7] Singh VP, Bali A, Singh N, Jaggi AS. Advanced glycation end products and diabetic complications. Korean J Physiol Pharmacol. 2014;18(1):1–14. doi:10.4196/kjpp.2014.18.1.1

[8] Kellar D, Craft S. Brain insulin resistance in Alzheimer's disease and related disorders: mechanisms and therapeutic approaches. Lancet Neurol. 2022;21(8):722–734. doi:10.1016/S1474-4422(22)00153-3

[9] Last D, Alsop DC, Abduljalil AM, et al. Global and regional effects of type 2 diabetes on brain tissue volumes and cerebral vasoreactivity. Diabetes Care. 2017;40(5):703–710. doi:10.2337/dc16-2258

[10] Srikanthan K, Feyh A, Visweshwar H, et al. Systemic inflammation and metabolic syndrome. Cardiol Rev. 2015;23(1):15–22. doi:10.1097/CRD.0000000000000093

[11] Liu S, Manson JE. Dietary carbohydrates, physical inactivity, obesity, and the 'metabolic syndrome' as predictors of coronary heart disease. Curr Opin Lipidol. 2001;12(4):395–404. doi:10.1097/00041433-200108000-00004

[12] Power MC, Mormino E, Soldan A, et al. Added sugar intake and cognitive score in a longitudinal cohort. J Prev Alzheimers Dis. 2021;8(2):164–170. doi:10.14283/jpad.2020.48

[13] Salas-Salvadó J, Guasch-Ferré M, Lee CH, et al. Protective effects of the Mediterranean diet on type 2 diabetes and metabolic syndrome. J Nutr. 2018;148(11):1801–1812. doi:10.1093/jn/nxy201

[14] Scarmeas N, Anastasiou CA, Yannakoulia M. Nutrition and prevention of cognitive impairment. Lancet Neurol. 2018;17(11):1006–1015. doi:10.1016/S1474-4422(18)30338-7

[15] Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol. 2002;13(1):3–9. doi:10.1097/00041433-200202000-00002

[16] Khera R, Krumholz HM. Machine learning for clinical prediction models: a review. Circulation. 2018;137(23):2435–2441. doi:10.1161/CIRCULATIONAHA.118.034512

[17] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017; 30:4765–4774.

[18] Centers for Disease Control and Prevention (CDC). National Health and Nutrition Examination Survey (NHANES), 2011–2014. Hyattsville, MD: National Center for Health Statistics.

[19] Willett WC. Nutritional Epidemiology. 3rd ed. Oxford University Press; 2019.

[20] Spanakis EK, Golden SH. Race/ethnicity and diabetes diagnosis in NHANES: validation of self-reported diagnosis. Diabetes Care. 2022;45(2): e30–e31. doi:10.2337/dc21-1952

[21] Dodge HH, Zhu J, Lee CW, et al. Use of screening tests for cognitive impairment in older adults. J Am Geriatr Soc. 2007;55(12):1925–1932. doi:10.1111/j.1532-5415.2007. 01446.x

[22] Kuo HK, Sorond FA, Chen JH, et al. The role of DSST in cognitive impairment screening. J Alzheimers Dis. 2022;87(3):1103–1111. doi:10.3233/JAD-215567

[23] Moshfegh AJ, Rhodes DG, Baer DJ, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. Am J Clin Nutr. 2008;88(2):324–332. doi:10.1093/ajcn/88.2.324

[24] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002; 16:321–357. doi:10.1613/jair.953

[25] Kivipelto M, Mangialasche F, Ngandu T. Lifestyle interventions to prevent cognitive impairment, dementia and Alzheimer disease. Nat Rev Neurol. 2018;14(11):653–666. doi:10.1038/s41582-018-0070-3

[26] O'Donnell MJ, Yusuf S, Mente A, et al. Sodium and cardiovascular disease: a population-level analysis. N Engl J Med. 2020;382(13):1237–1249. doi:10.1056/NEJMoa1917169

[27] American Heart Association (AHA). Dietary Guidance to Improve Cardiovascular Health. AHA Scientific Statement; 2021.

[28] Lourida I, Soni M, Thompson-Coon J, et al. Mediterranean diet, cognitive function, and dementia: a systematic review. Epidemiology. 2013;24(4):479–489. doi:10.1097/EDE.0b013e3182944410

[29] de la Monte SM. Brain insulin resistance and deficiency as therapeutic targets in Alzheimer's disease. Curr Alzheimer Res. 2012;9(1):35–66. doi:10.2174/156720512799015955

[30] Munshi MN. Management of diabetes in the elderly. Med Clin North Am. 2017;101(3):515–530. doi: 10.1016/j.mcna.2016.12.009

[31] Schwartz MW, Seeley RJ, Zeltser LM, et al. Obesity and diabetes: pathogenesis and treatment. J Clin Invest. 2020;130(9):4517–4529. doi:10.1172/JCI139934

[32] Madeo F, Carmona-Gutierrez D, Hofer SJ, Kroemer G. Caloric restriction mimetics against age-associated disease: targets, mechanisms, and therapeutic potential. Cell Metab. 2018;28(3):323–340. doi: 10.1016/j.cmet.2018.07.008

[33] Franzago M, Stuppia L, Vitacolonna E. Energy restriction and sarcopenia: mechanisms and nutritional interventions. Nutrients. 2022;14(3):599. doi:10.3390/nu14030599

[34] Bauer J, Biolo G, Cederholm T, et al. Evidence-based recommendations for optimal dietary protein intake in older people: a position paper from the PROT-AGE Study Group. J Am Med Dir Assoc. 2013;14(8):542–559. doi: 10.1016/j.jamda.2013.05.021

[35] World Health Organization (WHO). Guideline: Sodium Intake for Adults and Children. Geneva: WHO Press; 2021.

[36] Ding J, Strachan MWJ, Reynolds RM, Frier BM. Diabetic retinopathy and cerebral small vessel disease. Diabetologia. 2021;64(8):1858–1869. doi:10.1007/s00125-021-05488-2

[37] Dietary Guidelines Advisory Committee (DGAC). Scientific Report of the 2020 Dietary Guidelines Advisory Committee. Washington, DC: U.S. Department of Agriculture; 2020.

[38] Zhang X, Zhang G, Zhang H, et al. Fructose-induced neuroinflammation and insulin resistance in the brain. J Nutr Biochem. 2022; 102:108955. doi: 10.1016/j.jnutbio.2021.108955

[39] Ma Q, Xing C, Long W, et al. Impact of microbiota on central nervous system and neurological diseases: the gut–brain axis. J Neuroinflammation. 2022;19(1):45. doi:10.1186/s12974-022-02410-8

[40] European Society for Clinical Nutrition and Metabolism (ESPEN). ESPEN Guidelines on Clinical Nutrition and Metabolism in Older Adults. Clin Nutr. 2021;40(3):650–667. doi: 10.1016/j.clnu.2021.01.008

[41] American Diabetes Association (ADA). Standards of Medical Care in Diabetes—2023. Diabetes Care. 2023;46(Suppl 1): S1–S298. doi:10.2337/dc23-Sint

[42] Petersson SD, Philippou E. Mediterranean diet, cognitive function, and dementia: a systematic review. Epidemiol Health. 2022;44: e2022011. doi:10.4178/epih. e2022011

[43] Trieu K, Neal B, Hawkes C, et al. Salt reduction initiatives around the world: a systematic review. BMJ Glob Health. 2021;6(8): e005982. doi:10.1136/bmjgh-2021-005982

[44] Powell LM, Maciejewski ML. Sugar-sweetened beverage taxes and public health: a systematic review and meta-analysis. Obesity. 2020;28(1):1–14. doi:10.1002/oby.22639