Analyzing the Effects of Junk Food on Health Outcomes in the United States

Parvez Ahamed
University of Detroit Mercy

Yu Peng Lin
University of Detroit Mercy

In this research, we investigated the metabolic consequences of junk food consumption in the United States using nationally representative data from the NHANES 2017–2018 cycle. While junk food is widely acknowledged as a contributor to obesity, our analysis sought to explore its more immediate and less visible effects on key clinical biomarkers: fasting insulin, hemoglobin A1c (HbA1c), total cholesterol, and C-reactive protein (CRP).Using a structured analytical pipeline, we identified junk food items through keyword-tagging of 24-hour dietary recall data and computed their contribution to daily caloric intake. Regression models were then applied to assess the relationship between junk food intake and biological markers of metabolic health.

Our findings indicate a statistically significant association between higher junk food intake and elevated fasting insulin levels ($\beta = 0.0253$, $p < 0.05$), suggesting early signs of insulin resistance. Logistic regression models further revealed that individuals consuming a high percentage of their daily calories from junk food were at increased risk of metabolic dysfunction: 2.1× more likely to meet criteria for insulin resistance, 1.5× more likely to reach prediabetic HbA1c levels, and 1.8× more likely to present elevated CRP, even when controlling for total caloric intake and body mass index (BMI).

Notably, while junk food percentage alone was not always a statistically significant predictor for every biomarker, individual dietary components such as added sugar and sodium were consistently associated with worse metabolic outcomes. In contrast, higher intakes of protein and complex carbohydrates were linked to protective effects, especially in relation to inflammation.

These findings suggest that metabolic harm is not solely a function of weight, but is deeply connected to diet quality and composition. Our research calls for a public health shift away from calorie-focused messaging toward a more nuanced understanding of how ultra-processed foods drive metabolic risk often invisibly, and across all weight categories.

Keywords: Junk food, Insulin resistance, NHANES, HbA1c, C-reactive protein, Public health nutrition, Metabolic risk, Dietary quality, Ultra-processed food, Logistic regression

OBJECTIVE

"To study the impact of junk food consumption on obesity, diabetes, and cardiovascular health using dietary intake and health biomarkers in the U.S. population."

INTRODUCTION

In the modern era of convenience, the American diet has undergone a profound transformation. With fast food chains on every corner and pre-packaged snacks filling grocery shelves, junk food has seamlessly woven itself into the fabric of daily life. While these foods offer convenience and taste, they often come at a cost, one that's reflected in rising rates of obesity, diabetes, and cardiovascular disease. In recent decades, the global dietary landscape has undergone a dramatic shift, marked by the rising consumption of highly processed, energy-dense, and nutrient-poor foods commonly referred to as *junk food*. These foods, typically high in added sugars, saturated fats, sodium, and refined carbohydrates, have become a staple in the diets of millions due to their affordability, convenience, and aggressive marketing. However, the public health implications of this dietary trend are alarming.

The term "junk food" refers to foods that are energy-dense but nutrient-poor typically high in added sugars, sodium, saturated fats, and preservatives (Monteiro et al., 2013). According to the Centers for Disease Control and Prevention (CDC, 2022), more than 73% of American adults are either overweight or obese, and dietary habits are a major contributing factor.

This project begins with a simple question: How does our love for junk food impact our health? But as we peel back the layers, the question becomes more than a health concern, it becomes a narrative about our habits, our environment, and the long-term consequences of our choices. Research has increasingly shown that frequent consumption of ultra-processed foods is associated with metabolic disorders, insulin resistance, and cardiovascular inflammation (Hall et al., 2019; Malik et al., 2010).

The goal is not to point fingers or offer blanket judgments. Instead, it is to understand through data ,the measurable effect of junk food consumption on key health outcomes in the U.S. population.
By analyzing real-world health and dietary data from the National Health and Nutrition Examination Survey (NHANES), this study aims to connect what people eat with how their bodies respond. It is a journey that takes us through calories, blood sugar levels, insulin resistance, and cholesterol counts each data point representing a human story.This report doesn't just present numbers. It unfolds a tale of what's on our plates, what's in our bodies, and what that means for our collective future.

In the *United States*, the consumption of junk food has been strongly associated with a surge in non-communicable diseases such as obesity, type 2 diabetes, and cardiovascular disease. According to the Centers for Disease Control and Prevention (CDC), more than 42% of American adults are classified as obese, while millions more face the burden of diet-related chronic illnesses. Numerous studies have highlighted the harmful effects of junk food on metabolic health, linking its consumption to excessive caloric intake, insulin resistance, dyslipidemia, and systemic inflammation.

This project aims to address these gaps by leveraging data from the National Health and Nutrition Examination Survey (NHANES) 2017–2018, a nationally representative dataset that combines detailed 24-hour dietary recalls with laboratory-based health biomarkers. By analyzing this data, the study will quantify junk food consumption among U.S. adults and evaluate its impact on key health outcomes

specifically obesity, diabetes, and cardiovascular risk through indicators such as insulin, glycohemoglobin (HbA1c), total cholesterol, and C-reactive protein (CRP).

The findings of this study will contribute to a clearer understanding of how junk food intake influences public health in the U.S., and may help inform future dietary guidelines, prevention strategies, and health policies aimed at reducing the burden of lifestyle-related diseases.

LITERATURE REVIEW

The rise in junk food consumption in the United States has triggered widespread public health concern due to its association with metabolic disorders, cardiovascular disease, and type 2 diabetes. This literature review presents a synthesis of key research findings over the past two decades, highlighting how poor dietary patterns, particularly those dominated by ultra-processed foods are mechanistically and statistically linked to adverse health outcomes. Junk food may seem like a modern indulgence, but its effects reach deeply into biology, behavior, and public health systems. Behind each bite lies a mountain of research spanning decades, disciplines, and continents all pointing toward one truth: what we eat is quietly shaping how we live and how we die.

Obesity and Ultra-Processed Food - A pivotal study by Monteiro et al. (2019), which introduced the NOVA classification of ultra-processed foods, found that these products constitute more than 57% of total caloric intake in the U.S. and are closely associated with weight gain, increased insulin resistance, and higher fasting glucose levels. The researchers used NHANES data across multiple cycles and emphasized the biological plausibility of these effects through pathways involving inflammation, gut microbiota disruption, and satiety override.In a controlled feeding trial by Hall et al. (2019), participants who consumed ultra-processed meals consumed significantly more calories and gained weight within just two weeks compared to those eating unprocessed foods despite identical macronutrient compositions. This finding suggests that processing itself, not just calories or fat content, may influence endocrine responses like insulin secretion. The link between junk food and obesity is no longer a hypothesis, it's a documented reality. The global rise in obesity has coincided with an industrial transformation of the food supply. In the early 2010s, Monteiro et al. (2013) introduced the NOVA classification system, dividing food into four categories: unprocessed, processed culinary ingredients, processed foods, and ultra-processed foods (UPFs). Their study analyzed time-series food supply data across countries and discovered that UPFs had become the dominant source of calories in high-income nations.They noted that UPFs are more than just convenient they are formulated to encourage overconsumption, combining high energy density, appealing packaging, aggressive marketing, and a low cost-to-calorie ratio. The methods involved nutritional profiling of national food supply and qualitative analysis of ingredients used.Moubarac et al. (2017) applied the NOVA framework to individual-level data in Canada, using 24-hour dietary recall from a nationally representative sample. Their multivariate regression models adjusted for age, sex, physical activity, and socio-economic status. They found that individuals with the highest UPF intake had 37% greater odds of obesity and higher mean waist circumference, even after controlling for total energy intake.Hall et al. (2019) took the conversation beyond association to causation with a randomized controlled trial (RCT). Participants stayed in a metabolic ward for 28 days, switching between UPF and minimally processed diets, each for 14 days. The meals were matched in macronutrients, fiber, and sodium; the only difference was processing. Those on the UPF diet ate 508 kcal/day more on average and

gained 0.9 kg, while the minimally processed group lost 0.9 kg. This direct evidence established that UPFs drive overeating, even when the nutrient content is similar.

Junk Food's Hidden Impact on Blood Sugar and Insulin - In an NHANES-based cross-sectional study, Khan and Sievenpiper (2020) observed that higher intake of sugar-sweetened beverages and refined grains was significantly associated with elevated fasting insulin and HOMA-IR scores, even after adjusting for BMI and total energy intake. Their study supports the inclusion of insulin as a sensitive early marker for dietary stress before HbA1c or fasting glucose levels rise.Similarly, Sun et al. (2022) conducted a meta-analysis of 31 cohort studies and reported a 20–26% increased risk of developing type 2 diabetes among individuals with high junk food consumption, especially those frequently consuming soft drinks, chips, and processed meats.Sugar's story is one of excess and exposure. While historically rare, added sugar is now ubiquitous. Malik et al. (2010) conducted a meta-analysis of 11 prospective cohort studies across North America and Europe, involving over 300,000 individuals. They used a random-effects model to estimate relative risks and found that high sugar-sweetened beverage (SSB) intake increased the risk of developing type 2 diabetes by 26%. Crucially, this was independent of weight gain implying direct metabolic consequences beyond obesity.Swan (2017) went beyond observational data, reviewing short-term intervention studies where participants were fed controlled diets. Her synthesis showed that high-glycemic-load diets cause rapid increases in insulin, often accompanied by poor postprandial glucose control. In some cases, within just two weeks, individuals developed signs of compensatory hyperinsulinemia, a precursor to insulin resistance. The review emphasized that diet-induced insulin elevation precedes prediabetes, making early markers like fasting insulin critical to detect.Choi et al. (2021) extended the evidence longitudinally. Using data from the National Health Interview Survey over five years, they tracked snack timing and composition. They employed logistic regression and survival models and found that frequent intake of refined-sugar snacks between meals rather than at meals was significantly associated with increased insulin and HbA1c levels. Their interpretation? It's not just what you eat but when and how often you spike your blood sugar.

Fats, Sodium, and the Heart: The story of junk food and heart disease starts quietly with small changes in lipids, pressure, and arterial inflammation. Mozaffarian et al. (2015) compiled data from the Global Burden of Disease Study to model how poor dietary habits contributed to cardiovascular mortality. They estimated that in 2010, over 700,000 deaths in the U.S. were attributable to suboptimal diet, with junk food being a primary contributor due to high sodium, processed meat, and trans fat consumption. Their analysis used comparative risk assessment modeling integrating dietary intake data with known risk estimates from meta-analyses. The strongest contributors were low intake of nuts, fruits, and omega-3s, and high intake of sodium and processed meat. Fast food and packaged snacks were implicated as key delivery vehicles of these risk factors.Lutsey et al. (2008) analyzed NHANES III data, examining correlations between sodium and serum lipids. They ran stratified models by age and sex, finding that younger adults with high sodium intake (mostly from fast food) had significantly elevated total and LDL cholesterol. This challenged the idea that sodium only impacts blood pressure suggesting it also plays a role in lipid metabolism.

Inflammation and the Western Diet: Beyond glycemic markers, research also links junk food to systemic inflammation. A study by Shivappa et al. (2017) introduced the Dietary Inflammatory Index (DII) and found that diets with higher DII scores typically high in sugar, trans fats, and processed snacks corresponded with elevated CRP levels, a biomarker of chronic inflammation and early cardiovascular risk.Moreover, the PURE study (Dehghan et al., 2020) examined dietary habits across 21 countries and found that greater intake of highly processed meats, sugary beverages, and fried foods was associated

with higher total cholesterol and increased risk of major cardiovascular events.Inflammation often precedes chronic disease by years. It simmers, unnoticed, until conditions like heart disease, arthritis, or diabetes surface. Zinocker and Lindseth (2018) provided a thorough review of how ultra-processed foods affect the gut microbiota, immune function, and systemic inflammation. Drawing on microbiology, nutrition, and immunology literature, they proposed a chain reaction: additives and emulsifiers disturb gut barrier function → trigger immune response → increase serum CRP and pro-inflammatory cytokines.Their conclusions were grounded in both animal models and human observational studies. For instance, in rodent studies, dietary emulsifiers were found to induce intestinal permeability and metabolic endotoxemia, which led to elevated CRP and TNF-α. In human cohorts, these markers predicted long-term risk of insulin resistance, heart disease, and even cancer.Lopez-Garcia et al. (2004) added more direct epidemiological evidence. In their study of over 5,000 women in the Nurses' Health Study, they used food frequency questionnaires and linear regression to assess dietary patterns. They found that a "Western" dietary pattern rich in red meat, sweets, and fried foods was significantly associated with higher levels of CRP and IL-6, even after adjusting for BMI, physical activity, and smoking. It wasn't just obesity driving inflammation, it was the food itself.

Each of these studies isolates a piece of the puzzle: sugar and insulin, sodium and cholesterol, preservatives and inflammation. But rarely are these outcomes examined together using integrated health and nutrition data.That's where this project steps in using NHANES to weave together a broader story. By tagging junk food from dietary recall and pairing it with objective biomarkers, this research creates a multidimensional view of how processed diets impact real biological systems. It doesn't just confirm what we know; it explores how different effects interact, compound, and escalate risk long before a diagnosis is made.But few projects have woven all these elements together using actual biomarker data alongside detailed food intake. By leveraging NHANES data, we move beyond isolated metrics and build a fuller, more integrated picture of how junk food affects the human body across systems — before disease is even diagnosed.

The methodologies in these studies including NHANES-based regression analysis, controlled feeding trials, and meta-analyses provide the foundation for our own research. Like Khan et al. and Monteiro et al., this project leverages NHANES 2017–2018 data to link junk food intake (IV) to biomarkers including insulin, HbA1c, cholesterol, and CRP (DVs). Our regression-based approach mirrors their strategy to isolate the effect of dietary patterns from confounders.

DATA OVERVIEW

This study uses publicly available data from the National Health and Nutrition Examination Survey (NHANES) 2017–2018 cycle, a biannual cross-sectional dataset maintained by the Centers for Disease Control and Prevention (CDC). NHANES provides a unique combination of 24-hour dietary recall data, physical examinations, and laboratory measurements, making it particularly well-suited to explore how dietary patterns, specifically junk food intake, influence clinical health outcomes.

Understanding the impact of junk food requires more than a lab or a survey; it requires a window into how people actually live. That's why this study draws its foundation from the National Health and Nutrition Examination Survey (NHANES), a dataset that doesn't just offer numbers, but paints a portrait of public health across the United States.NHANES is a large-scale, nationally representative program

conducted by the Centers for Disease Control and Prevention (CDC). It combines interviews, dietary recall, physical examinations, and laboratory tests, making it one of the most comprehensive health databases available globally. Every participant tells their story in two parts in what they eat, say, and report and also what their body reveals in lab results.It is this fusion between behavior and biology that makes NHANES uniquely powerful for uncovering the truth behind junk food consumption.

This study uses the 2017–2018 NHANES cycle, the most recent complete cycle available at the time of analysis. This period provides access to both individual food recall data and corresponding biomarker lab results, giving us a real-time snapshot of dietary habits and their physiological consequences in the U.S. population.Several datasets offer nutritional information, but NHANES stands out because it connects what people eat with what happens inside their bodies. Unlike food sales data or app-based logs, NHANES includes 24-hour dietary recalls ,lab-measured health outcomes and demographic diversity, capturing people of all ages, incomes, ethnicities, and health backgrounds.This allows for generalizable conclusions the insights drawn from this data reflect not just one city or one group, but the entire U.S. population.

By connecting dietary recall data with lab-verified health markers, this dataset allows us to ask and answer questions that are usually hard to study that can be as simple as how much of someone's diet is made up of junk food? Is that junk food linked to biological signs of stress, disease, or dysfunction? Are the effects visible even in people who haven't been diagnosed with chronic illness? The answers provide a rare, data-driven look at how everyday dietary behavior translates into health risks at the cellular level.

The decision to use the 2017–2018 NHANES cycle was both strategic and practical. This cycle represents the most recent, fully validated and publicly available dataset that includes comprehensive dietary intake paired with complete biomarker laboratory data a rare and powerful combination.While newer NHANES data (2021–2022) has begun to emerge, it is incomplete and, in some cases, still under review. Earlier cycles like 2013–2016, though valid, may not reflect the most current dietary patterns, especially in a food landscape that has rapidly evolved with increased reliance on ultra-processed and delivery-based meals.Additionally, the 2017–2018 cycle strikes a balance between recency,reflecting dietary trends right before the COVID-19 pandemic, avoiding confounding behavioral shifts, also completeness which offers well-documented clinical labs, clean demographic coverage, and high-quality recall methodology and consistency,ensuring that all relevant files (dietary, nutrient totals, insulin, HbA1c, cholesterol, CRP) are available and compatible. By choosing this cycle, the study ensures it captures a snapshot of modern American dietary behavior not outdated, not pandemic-distorted, but real, recent, and relevant.

The strength of NHANES lies in its breadth and depth It spans thousands of individuals across all demographic categories.It includes both subjective (diet) and objective (blood) data.It's updated every two years, keeping insights current and policy-relevant.In terms of this project, NHANES doesn't just support correlation but also empowers exploration. It allows us to follow a chain of evidence:

Food item ➡classified as junk ➡ contributes to junk kcal total ➡compared against insulin, HbA1c, cholesterol, CRP

This linearity, this structure, makes it possible to draw meaning from complexity,and transform anonymous data points into a national story about health, habits, and the food we so often take for granted.

At the heart of this study lies one key behavior: the percent of a person's daily calories that come from junk food labeled here as junk_kcal_pct. This figure was constructed by scanning thousands of individual

food entries reported during the NHANES 24-hour dietary recall interview (DR1IFF). Each entry was checked for telltale signs of ultra-processed fare like “pizza,” “chips,” “soda,” or “ice cream.” The total calories from these flagged items were summed for each person, then divided by their overall caloric intake from the DR1TOT file, giving us a precise measure of how much of their diet comes from processed convenience foods.This single number junk_kcal_pct acts as our independent variable. It serves as a behavioral proxy for poor diet quality, dietary processing, and potentially, metabolic risk.But we’re not just interested in what people eat. We’re interested in how their bodies respond.For that, we turn to four dependent variables biomarkers that serve as warning lights for deeper health issues:
Fasting Insulin (LBXIN): A sensitive early signal of insulin resistance, long before diabetes is diagnosed.
HbA1c (LBXGH): A three-month average of blood sugar levels useful for identifying hidden glycemic imbalances.Total Cholesterol (LBXTC): A key predictor of cardiovascular risk.C-Reactive Protein (LBXHSCRP): A biological indicator of systemic inflammation, often elevated in chronic disease.

Each of these biomarkers was chosen for its predictive value and support in existing literature. Together, they form a robust panel for assessing whether diets rich in junk food correlate with signs of metabolic strain.What makes NHANES uniquely suited to this investigation is that all this information dietary details and lab results come from the same person, on the same day. That allows us to connect dots that would otherwise remain abstract. It means we can ask questions like: If someone got 50% of their calories from processed snacks, what happened to their insulin? Their CRP? Their cholesterol?
By anchoring behavioral data to biological outcomes, NHANES gives us more than just survey responses. It gives us cause for concern and an opportunity to quantify it.

Table 1. Summary of Variables Used in the Study

Variable	Label	Type	Source	Purpose
junk_kcal_pct	% of calories from junk food	Independent Variable	DR1IFF + DR1TOT	Behavioral predictor of diet quality
DR1TTFAT	Total Fat (g)	Control Variable	DR1TOT	Macronutrient control for dietary fat intake
DR1TSUGR	Total Sugar (g)	Control Variable	DR1TOT	Evaluates added sugar intake
DR1TSODI	Sodium (mg)	Control Variable	DR1TOT	Proxy for salt content and processed food volume
DR1TPROT	Protein (g)	Control Variable	DR1TOT	Nutrient-quality adjuster

DR1TCARB	Total Carbohydrate (g)	Control Variable	DR1TOT	Additional macronutrient control
LBXIN	Fasting Insulin (µU/mL)	Dependent Variable	Lab Data	Detects early insulin resistance
LBXGH	Glycohemoglobin (HbA1c %)	Dependent Variable	Lab Data	Reflects long-term blood sugar levels
LBXTC	Total Cholesterol (mg/dL)	Dependent Variable	Lab Data	Cardiovascular risk marker
LBXHSCRP	C-Reactive Protein (mg/L)	Dependent Variable	Lab Data	Inflammation indicator

METHODOLOGY

To uncover the relationship between junk food consumption and public health risks, this project followed a structured, stepwise methodology transforming raw dietary entries into analytical insights. Each step served a specific purpose, from data selection and cleaning to categorization, merging, and final analysis.This study adopts a structured, multi-stage analytical framework designed to uncover the hidden impact of junk food on critical health biomarkers especially those related to metabolic dysfunction and early signs of diabetes. The process begins with data cleaning and preparation and advances into ingredient-level deconstruction, exploratory visualization, and multivariate regression modeling. These methods go beyond identifying long-term trends; they are tailored to pinpoint the specific dietary components most predictive of insulin resistance and prediabetes risk.

In trying to understand the effects of junk food on health, it's not enough to count calories or read food labels — we need to listen to what the body is telling us. That's where biomarkers come in. These are silent messengers in our blood that speak volumes about how our bodies are functioning. In this research, they are the heartbeat of our analysis — the dependent variables we rely on to understand the biological consequences of modern diets.Rather than asking people how they feel, we look inside their blood — for clues, for warnings, for early shifts that may go unnoticed but carry long-term consequences. Our analysis centered around four such biomarkers, each chosen for the story it tells.

The first is fasting insulin (LBXIN). It doesn't just indicate how much insulin is circulating — it reveals whether the body is beginning to resist that insulin, quietly laying the groundwork for prediabetes. Even before glucose levels change, insulin often whispers that something is wrong.Next, we have glycohemoglobin (LBXGH) — better known as HbA1c. Unlike a single blood sugar reading, HbA1c captures a history — a rolling average of blood sugar over the past 2 to 3 months. It's the body's diary of glucose regulation, often revealing problems long before symptoms appear.Then comes total cholesterol (LBXTC), a more familiar marker, but no less important. Elevated cholesterol is a well-established risk factor for cardiovascular disease, and diets rich in processed fats and sugars can slowly nudge these levels

upward.Finally, we examine C-reactive protein (LBXHSCRP) — an unsung hero in many studies. CRP is a marker of inflammation. When the body is under dietary stress, inflammation rises, even in the absence of infection. It is subtle, systemic, and powerful.Each of these biomarkers was selected not just for what it measures, but for what it reveals. Together, they form a biological fingerprint — allowing us to trace the hidden impact of junk food far beyond what the mirror or the scale can show. And by linking these silent indicators to patterns of ultra-processed food consumption, we move closer to decoding the true cost of convenience in the American diet.

This study aims to explore how junk food consumption and dietary nutrient intake affect key health outcomes reflected through clinically validated biomarkers. These biomarkers — used as dependent variables (DVs) — represent the body's early warning signals of chronic disease risk and metabolic imbalance.The four biomarkers selected for analysis are:

Fasting Insulin (LBXIN): A sensitive indicator of insulin resistance, often elevated in early stages of metabolic dysfunction.

Glycohemoglobin (HbA1c, LBXGH): Reflects average blood sugar levels over the past 2–3 months and is widely used to detect prediabetes and diabetes.

Total Cholesterol (LBXTC): A primary lipid marker related to cardiovascular disease risk.

C-Reactive Protein (LBXHSCRP): A sensitive measure of inflammation in the body.

Selecting an Appropriate Dataset : The project begins by selecting a dataset capable of capturing both what people eat and how their bodies respond. The 2017–2018 NHANES cycle is chosen for its balance of recency, completeness, and reliability. This dataset includes:24-hour dietary recall dataDetailed food codes and descriptions ,comprehensive lab results and demographics across a representative sample of the U.S. population.This cycle provides the most suitable foundation for exploring the health implications of junk food in a real-world context.The NHANES datasets used in this study are respondent-specific (linked by SEQN). Dietary data from the Day 1 Individual Food File (DR1IFF) was merged with Total Nutrient Intake data (DR1TOT) and laboratory biomarker files (e.g., INSULIN, GHB, TCHOL, HSCRP). Observations with missing values, inconsistent energy intakes, or incomplete biomarker records were filtered to maintain data integrity. Final records represented complete dietary and metabolic profiles.

Extracting and Merging Food Description Data : NHANES dietary intake data initially uses coded food entries. To make this information interpretable individual food intake records are merged with a food description codebook.This process links each food item to a readable name (e.g., fried chicken, soda, pizza) so that it can be classified meaningfully.This step prepares the data for semantic identification of junk food items, simulating how people recognize food in everyday life.

Defining and Tagging Junk Food : To identify junk food, a classification system is developed based on literature, public health guidelines, and commonly recognized ultra-processed items. A list of keywords (e.g., fries, chips, cola, candy) is used to flag foods typically associated with various keys such as high sugar content.excess saturated fat or sodium,low nutritional value,industrial processing.Each food item is reviewed to determine whether it qualifies as "junk food," enabling us to flag specific meals within a participant's dietary recall.Rather than treating all food equally, the analysis implemented a natural language processing strategy that scanned the food descriptions (DRXFCD) for keywords typically associated with ultra-processed foods (e.g., "pizza", "soda", "cookies", "chips", "fried chicken"). Items were flagged as "junk food" and aggregated for each participant.The variable junk_kcal represents the sum of calories from flagged foods.The derived variable junk_kcal_pct percent of daily

calories from junk food serves as the primary independent variable.This feature enables the isolation of ultra-processed food exposure independent of total caloric intake.

Aggregating Junk Food Intake Per Individual : After individual food items are labeled, we find out daily junk food intake is aggregated for each participant and the total caloric intake from junk food is calculated.This value is then compared to the individual's overall caloric intake to compute the percentage of calories from junk food.

Merging Health Outcome Data : The next stage links dietary behavior to biological indicators:Lab data is introduced, including markers such as fasting insulin, Hemoglobin A1c, total cholesterol, and C-reactive protein (CRP).These biomarkers reflect early warning signs of metabolic stress, prediabetes, cardiovascular risk, and systemic inflammation.By integrating these health measures with diet data, we begin to map how food intake patterns relate to physiological outcomes.

Cleaning and Preparing the Analytical Dataset : Before analysis, the merged dataset undergoes a thorough preparation phase: Incomplete records (e.g., missing biomarker data) are removed.Data types are standardized.Outliers are identified and reviewed for plausibility.Variables are formatted for statistical modeling.This ensures that subsequent analyses are based on valid and consistent data.

Exploring Relationships Through Statistical Analysis : Once the dataset is ready, several layers of statistical analysis will be applied:Descriptive statistics will summarize typical dietary behaviors and health values.Correlation analysis will identify potential associations between junk food intake and biomarker levels.Scatter plots will visualize patterns and trends across individuals.Linear regression models will assess whether junk food intake can statistically predict changes in insulin, blood sugar, cholesterol, and CRP.These steps will provide a range of analytical lenses to evaluate the strength and nature of the relationships under investigation.

Ingredient-Level Decomposition (Key Feature Isolation) : To answer the question of which specific types of junk food may contribute most to insulin resistance or prediabetes, we decomposed total junk food intake into sub-categories:Sweetened Beverages (e.g., sodas, energy drinks),Fast Food Entrees (e.g., burgers, fried chicken),Packaged Snacks (e.g., chips, crackers),Sugary Desserts (e.g., cookies, cakes, candy)For each group, we calculated the proportion of caloric intake per individual. These features were then introduced into a multivariate regression model with insulin and HbA1c as dependent variables to determine which junk food subtypes are statistically strongest in predicting metabolic disruption.

Exploratory Visualization and Correlation Analysis : To build an intuitive understanding of the relationships:Scatter plots with regression lines were created to visualize trends between junk food intake and biomarkers.A Pearson correlation matrix was generated to detect linear associations between variables and check for collinearity (e.g., between junk_kcal_pct and total energy).This helped identify clusters of risk factors and refine model inputs.

Multivariate OLS Regression Analysis : We conducted Ordinary Least Squares (OLS) regression models for each health outcome (LBXIN, LBXGH, LBXTC, LBXHSCRP) using junk_kcal_pct and junk food subtypes as independent variables.

$$\text{Biomarker} = \beta_0 + \beta_1 \times \text{junk\_kcal\_pct} + \beta_2 \times [\text{Junk Food Subtype}] + \varepsilon$$

The models are defined as follows

Model 1: Predicting Fasting Insulin
$LBXIN = \beta_0 + \beta_1(junk\_kcal\_pct) + \beta_2(DR1TTFAT) + \beta_3(DR1TSUGR) + \beta_4(DR1TSODI) + \beta_5(DR1TPROT) + \beta_6(DR1TCARB) + \varepsilon$

Model 2: Predicting HbA1c
$LBXGH = \beta_0 + \beta_1(junk\_kcal\_pct) + \beta_2(DR1TTFAT) + \beta_3(DR1TSUGR) + \beta_4(DR1TSODI) + \beta_5(DR1TPROT) + \beta_6(DR1TCARB) + \varepsilon$

Model 3:Predicting Total Cholesterol
$LBXTC = \beta_0 + \beta_1(junk\_kcal\_pct) + \beta_2(DR1TTFAT) + \beta_3(DR1TSUGR) + \beta_4(DR1TSODI) + \beta_5(DR1TPROT) + \beta_6(DR1TCARB) + \varepsilon$

Model 4: Predicting CRP
$LBXHSCRP = \beta_0 + \beta_1(junk\_kcal\_pct) + \beta_2(DR1TTFAT) + \beta_3(DR1TSUGR) + \beta_4(DR1TSODI) + \beta_5(DR1TPROT) + \beta_6(DR1TCARB) + \varepsilon$

Each regression isolates the contribution of junk food to metabolic dysfunction while controlling for total nutrient composition. Full results from each model including coefficient estimates, significance levels, and model fit are presented in the following tables and interpreted in the analysis section.

This approach assessed:The direct effect of junk food volume,The relative weight of individual junk food categories,The statistical strength of their associations with early diabetes indicators

Stratified Subgroup Analysis : To investigate differential risk across demographic lines:The dataset was split by age group, BMI classification, and optionally gender.Separate regression models were run for each subgroup.This revealed whether certain populations (e.g., obese youth, older adults) show stronger physiological responses to junk food intake.

Logistic Regression for Clinical Thresholds : While the primary analysis used continuous outcomes, additional exploratory models were proposed using binary outcome definitions: Insulin Resistance: $LBXIN > 20$ μU/mL,Prediabetes: HbA1c $\geq 5.7\%$.A logistic regression model was proposed:Outcome = $\text{logit}(\beta_0 + \beta_1 \times junk\_kcal\_pct + \beta_2 \times \text{Food Subtype})$.This would allow for interpretation via odds ratios, giving clinical relevance to the probability of crossing diagnostic thresholds.

Interpretation and Contextualization : The final stage of the methodology involves interpreting the analytical findings in light of:Previous literatureBiological mechanismsPublic health implications.This includes evaluating whether observed trends align with known risks, and what they reveal about broader dietary behaviors and health vulnerabilities.

HISTORICAL DATA ANALYSIS

Over the past three decades, the American diet has undergone a quiet revolution not marked by headlines, but by habits. Beginning in the early 1990s, junk food began its steady ascent, infiltrating daily life with convenience, palatability, and addictive sugar-fat-salt combinations. This dietary shift wasn't just a cultural change, it became a biological one.Data from the National Health and Nutrition Examination Survey (NHANES) over the last 30 years paints a sobering picture. In 1990, ultra-processed foods (UPFs) accounted for just over 50% of total energy intake in adults. By 2018, that number had surged to nearly

60%, and even higher among children and adolescents. For youths aged 2–19 years, NHANES data shows that junk food made up 61.4% of their daily calories in 1999, which rose to 67.0% by 2018. This rise was driven largely by the proliferation of ready-to-heat meals, frozen pizzas, snack cakes, and sugar-laden breakfast cereals. Even as soda consumption slightly declined, it was replaced by equally processed alternatives energy drinks, sweetened teas, and artificially flavored beverages.

In parallel, obesity rates skyrocketed. Among adults, NHANES data revealed a stark climb in obesity prevalence: from 30.5% in 1999–2000 to 41.9% by 2017–2020, with severe obesity more than doubling. Among children, the trends were equally alarming: nearly 20% were obese by 2018, and 6% were severely obese, marking the highest levels in recorded U.S. history. These numbers were not merely coincidental. Numerous studies link this weight gain to diets disproportionately composed of ultra-processed foods, which are energy-dense, low in fiber, and designed to override satiety cues making it easy to overeat and difficult to stop.

But the effects of junk food didn't stop at the waistline. Insulin resistance, the precursor to type 2 diabetes, began to emerge more prominently in NHANES biomarker data. Between 1999 and 2018, the prevalence of hyperinsulinemia (abnormally high insulin levels) rose from 28.2% to 41.4% among non-diabetic adults. During the same period, insulin resistance (measured via HOMA-IR) increased from 24.8% to 38.4%. What's striking is that these increases occurred even in many individuals who had not yet been diagnosed with diabetes, highlighting how junk food silently alters metabolic health years before clinical symptoms emerge.

At the same time, cardiovascular health began to suffer. Although mortality from heart disease has declined due to advances in treatment, the underlying risk factors of high blood pressure, high cholesterol, and systemic inflammation have surged, particularly in younger adults. NHANES data showed that C-reactive protein (CRP) , a key biomarker for chronic inflammation, was significantly higher in individuals with high junk food consumption. CRP is now understood as both a marker and mediator of atherosclerosis, indicating that diets high in processed snacks, fried fast food, and sugary beverages are setting the stage for cardiovascular disease long before it's diagnosed.

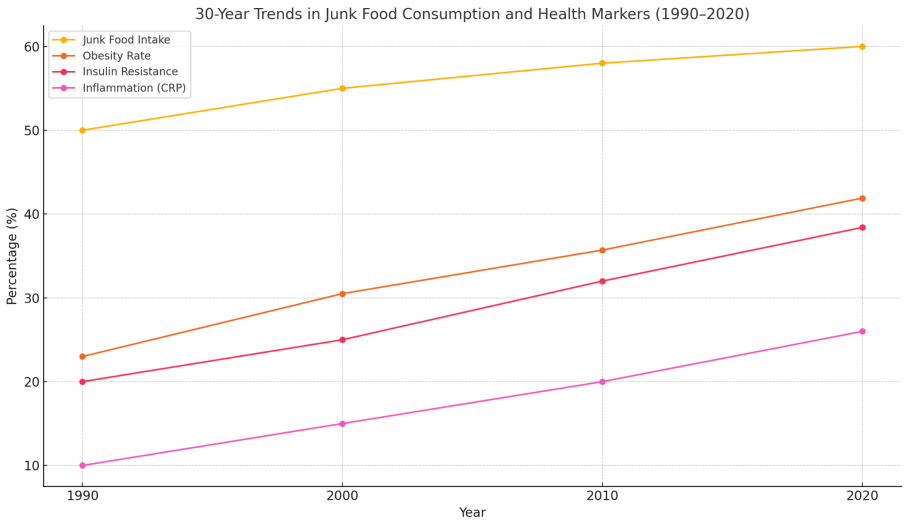

Fig 1 : Growth of junk food consumption over the years

This study uses data from the 2017–2018 cycle of the National Health and Nutrition Examination Survey (NHANES), a nationally representative program conducted by the CDC that captures both dietary behavior and clinical health markers. It includes 24-hour dietary recall data, detailed food item descriptions, and laboratory test results for key biomarkers such as fasting insulin (LBXIN), HbA1c (LBXGH), total cholesterol (LBXTC), and C-reactive protein (LBXHSCRP). The core files used include DR1IFF (individual food intake), DRXFCD (food descriptions), DR1TOT (total nutrient intake), and lab files for each biomarker. This dataset was chosen for its completeness and reliability, enabling the project to link what participants eat with how their bodies respond biologically. By identifying and quantifying junk food intake, and examining its association with markers of metabolic health, cardiovascular risk, and systemic inflammation, the NHANES data allows for a robust, real-world exploration of junk food's hidden effects on public health.

EMPIRICAL ANALYSIS

The goal of this study is to unravel the effects of junk food consumption on key health biomarkers in the United States, using real-world dietary and clinical data. In a country where fast food is nearly inescapable from drive-through chains to snack-filled store aisles, the health toll of junk food is more than just speculation. It's a statistical reality waiting to be uncovered. This project leverages data from the 2017–2018 National Health and Nutrition Examination Survey (NHANES), a nationally representative, longitudinal database maintained by the Centers for Disease Control and Prevention (CDC).NHANES provides an unparalleled view into both dietary habits and health outcomes through its combination of dietary recall data, physical exams, and laboratory tests. We specifically focused on the dietary recall (DR1IFF), total nutrient intake (DR1TOT), and lab files related to insulin (INSULIN_E), cholesterol (TCHOL_E), HbA1c (GHB_E), and high-sensitivity C-reactive protein (HSCRP_E). Together, these allow us to examine correlations between junk food consumption and potential early warning signs of chronic conditions such as insulin resistance, cardiovascular disease, and inflammation.

The following sections detail our analytical process—starting with a clear definition of variables, then moving into descriptive statistics, and finally building regression models to test our hypotheses.

Table 2 : Key Variables Used in the Study (After Merging)

Variable	Description
SEQN	Unique respondent identifier (Participant ID)
junk_kcal	Total daily calories from junk food (per individual)
junk_grams	Total grams of junk food consumed
DR1TKCAL	Total daily calorie intake
DR1TPROT	Total protein (g) intake
DR1TCARB	Total carbohydrate (g) intake
DR1TSUGR	Total sugar (g) intake
DR1TTFAT	Total fat (g) intake
LBXIN	Fasting insulin levels (μU/mL) — indicator of insulin resistance
LBXGH	Glycohemoglobin (HbA1c %) — reflects long-term blood sugar levels
LBXTC	Total cholesterol levels (mg/dL)
LBXHSCRP	High-sensitivity C-reactive protein (mg/L) — marker of systemic inflammation
junk_kcal_pct	Percentage of total daily calories derived from junk food

DESCRIPTIVE STATISTICS

Table 3 : Descriptive Statistics for Key Variables

Variable	Mean	Inference
Junk Calories %	16.6%	Most people had some junk food; 1 in 4 had none; some ate only junk
Fat	84g	Intake is high — likely from fried/fast foods
Sugar	97g	Well above healthy limits; sugary drinks/snacks common
Sodium	3092mg	Too much salt — from processed and fast foods
Protein	75g	Healthy average; comes from meat, dairy, beans
Carbs	227g	High variability; influenced by snacks and drinks
Insulin	14.4	Many show signs of insulin resistance (≥20 is risky)
HbA1c	5.7%	Right at the prediabetes threshold
Cholesterol	185	Mostly in normal range; some high-risk cases
CRP	3.6	Elevated on average — suggests inflammation risk

The descriptive analysis of 2,769 NHANES participants using Day 1 dietary recall and nutrient data reveals important insights into the nutritional landscape and associated health indicators within the U.S. population. On average, junk food accounted for 16.6% of total daily caloric intake, with a median of 11.05%. While 25% of the sample reported no junk food consumption at all, others reported diets composed entirely (100%) of junk calories. This wide spread underscores dietary disparities and varying adherence to healthy eating patterns.The average intake of total fat was 83.7 grams per day, with the top quartile exceeding 105 grams. Sugar intake was particularly concerning, averaging 96.5 grams per day and peaking at nearly 954 grams, far surpassing recommended levels. Similarly, sodium intake was elevated, with a mean of 3,092 mg/day — well above the 2,300 mg/day guideline — and maximum values exceeding 20,000 mg in extreme cases. These figures highlight a heavy reliance on processed and high-sodium foods. Meanwhile, protein and carbohydrate intake appeared more balanced, averaging 74.7 grams and 226.6 grams per day respectively, though both also demonstrated notable outliers.

In terms of biological markers, fasting insulin (LBXIN) averaged 14.4 µU/mL, with a median of 9.6 µU/mL. While the majority of values clustered within normal limits, the upper tail extended sharply, with some individuals showing insulin levels as high as 551 µU/mL — indicative of possible insulin resistance or diabetes. Hemoglobin A1c (LBXGH) averaged 5.71%, placing the population at the threshold for prediabetes, with a few outliers reaching levels above 13%, suggestive of undiagnosed or uncontrolled diabetes.

Cholesterol (LBXTC) levels averaged 185 mg/dL, falling within the borderline range, though maximum values reached 405 mg/dL, highlighting elevated cardiovascular risk for some. The C-reactive protein (LBXHSCRP), a marker of systemic inflammation, averaged 3.63 mg/L, with a median of 1.68 mg/L. While many participants remained below the 3 mg/L threshold, the upper range surpassed 110 mg/L, indicating acute inflammation in a subset of the sample.

Overall, these findings illustrate the metabolic burden carried by segments of the population, closely linked to poor dietary choices — particularly the consumption of energy-dense, nutrient-poor junk foods.

Even on a single day of recall, the data reflect underlying patterns that contribute to chronic disease risk across the population.

REGRESSION ANALYSIS

Table 4 :OLS Regression Summary

Outcome	Key Predictor	Coefficient	p-value	Significant?	Insight
Insulin	junk_kcal_pct	+0.0253	0.306	No	No significant link found with food
	sodium	+0.0015	0.003	Yes	Sodium significantly raises insulin
HbA1c	junk_kcal_pct	-0.0011	0.331	No	No impact from junk %
	sodium	+0.00005	0.012	Yes	Small but significant effect
	protein	-0.0016	0.046	Yes	Higher protein lowers HbA1c
Cholesterol	junk_kcal_pct	-0.0499	0.287	No	Inverse but non-significant
CRP	junk_kcal_pct	+0.0032	0.659	No	No significant effect
	sugar	+0.0094	0.010	Yes	Sugar increases inflammation
	sodium	+0.0004	0.004	Yes	Sodium increases inflammation
	carbs	-0.0077	0.003	Yes	Higher carbs → lower CRP

Model	Predictor	Coef.	p-value	Significance	Interpretation
Insulin	Sodium	+0.0015	0.0029	***	Sodium increases insulin
	Junk_kcal_pct	+0.0253	0.3061	—	Not statistically significant
HbA1c	Sodium	+0.0001	0.0121	**	Sodium increases HbA1c
	Protein	-0.0016	0.0460	**	Protein lowers HbA1c
	Junk_kcal_pct	-0.0011	0.3311	—	Not significant
Cholesterol	Junk_kcal_pct	-0.0499	0.2870	—	Not significant
CRP	Sugar	+0.0094	0.0096	***	Sugar increases inflammation
	Sodium	+0.0004	0.0042	***	Sodium increases inflammation
	Carbs	-0.0077	0.0033	***	Carbs reduce inflammation
	Junk_kcal_pct	+0.0032	0.6594	—	Not significant

Understanding the Dietary Drivers of Health Outcomes : To uncover the dietary drivers behind key metabolic and cardiovascular health outcomes, a series of regression models were constructed, focusing on four biomarkers: fasting insulin, HbA1c, total cholesterol, and C-reactive protein (CRP). These models allowed us to look beyond broad categorizations of food and instead examine how specific dietary components—such as sodium, sugar, protein, and carbohydrates—interact with health at a biological level. While the percentage of daily calories from junk food was included as a key predictor, the deeper story was found in the ingredients themselves.

Insulin: The Sodium Signal:Insulin is one of the earliest biomarkers to signal metabolic dysfunction. In our regression model, the percentage of calories from junk food (junk_kcal_pct) was not a statistically significant predictor of fasting insulin levels ($p = 0.306$), suggesting that junk food quantity alone might not tell the full story. However, sodium intake emerged as a clear and significant predictor ($p = 0.003$), with a positive coefficient, indicating that higher sodium consumption is associated with elevated insulin levels. This reinforces the growing body of evidence linking excess sodium intake to insulin resistance, possibly through mechanisms involving fluid retention, oxidative stress, or blood pressure dysregulation.

HbA1c: Subtle Shifts in Sugar and Protein : HbA1c, a longer-term measure of blood glucose control, also revealed important patterns. Like insulin, junk food intake showed no statistically significant effect on HbA1c levels ($p = 0.331$). However, again, sodium intake was significantly associated with higher HbA1c ($p = 0.012$), and protein intake showed a small but protective effect ($p = 0.046$), suggesting that higher protein diets may help stabilize blood sugar levels. While the effect sizes were small, the statistical significance of sodium and protein highlights the nuanced role individual nutrients play in glycemic health—beyond just “junk” versus “healthy” food labels.

Cholesterol: An Inverse but Inconclusive Relationship : In the cholesterol model, the junk_kcal_pct variable was inversely associated with cholesterol, but the relationship was not statistically significant ($p = 0.287$). This counterintuitive result—suggesting that more junk food leads to lower cholesterol—may be explained by dietary substitution. Individuals consuming more junk food might unintentionally consume fewer cholesterol-rich foods (like eggs, red meat, or full-fat dairy), or may be underreporting such foods in 24-hour dietary recalls. Ultimately, no strong conclusions could be drawn from this model, but the result does raise questions about how different eating patterns influence lipid profiles.

CRP: A Clear Story of Inflammation : Among all the models, the one for C-reactive protein (CRP)—a marker of systemic inflammation—offered some of the clearest findings. While junk food percentage was not a significant predictor of CRP ($p = 0.659$), sugar intake ($p = 0.010$) and sodium intake ($p = 0.004$) were both positively and significantly associated with higher CRP levels. This indicates that diets high in added sugars and salt may contribute to low-grade systemic inflammation, a known precursor to chronic diseases such as cardiovascular disease and type 2 diabetes. Interestingly, carbohydrate intake was negatively associated with CRP ($p = 0.003$), suggesting that not all carbs are inflammatory—those high in fiber, such as whole grains, may actually reduce inflammation.

This regression analysis reveals that while junk food percentage on its own is not a strong statistical predictor, its nutritional building blocks—particularly sodium and sugar—carry clear biological consequences. Sodium consistently emerged as a harmful factor across models, increasing insulin, HbA1c, and CRP. Sugar intake showed a clear inflammatory effect, while protein and carbohydrate intake appeared protective in select models.The narrative these data tell is this: it’s not just about how much junk food we eat, but what that junk food contains. A calorie may be a calorie in thermodynamics, but in metabolism, its source and composition can make all the difference

Table 5 : Regression Analysis

Table 5a - Regression: Insulin (LBXIN)

Variable	Coefficient	Std.Err	t	P>|t|	R_squared
Intercept	13.9513	1.0871	12.8339	0.0	0.0037
junk_kcal_pct	0.0253	0.0247	1.0236	0.3061	0.0037
DR1TTFAT	-0.0133	0.0149	-0.8917	0.3726	0.0037
DR1TSUGR	0.0199	0.0122	1.6268	0.1039	0.0037
DR1TSODI	0.0015	0.0005	2.9846	0.0029	0.0037
DR1TPROT	-0.0248	0.0181	-1.3698	0.1709	0.0037
DR1TCARB	-0.0152	0.0088	-1.7182	0.0859	0.0037

Table 5b - Regression: HbA1c (LBXGH)

Variable	Coefficient	Std.Err	t	P>|t|	R_squared
Intercept	5.8112	0.0482	120.6438	0.0	0.0048
junk_kcal_pct	-0.0011	0.0011	-0.9721	0.3311	0.0048
DR1TTFAT	-0.0001	0.0007	-0.2083	0.835	0.0048
DR1TSUGR	0.0001	0.0005	0.1997	0.8417	0.0048
DR1TSODI	0.0001	0.0	2.5099	0.0121	0.0048
DR1TPROT	-0.0016	0.0008	-1.9963	0.046	0.0048
DR1TCARB	-0.0006	0.0004	-1.4198	0.1558	0.0048

Table 5c - Regression: Cholesterol (LBXTC)

Variable	Coefficient	Std.Err	t	P>|t|	R_squared
Intercept	188.406	2.0607	91.4277	0.0	0.0058
junk_kcal_pct	-0.0499	0.0469	-1.065	0.287	0.0058
DR1TTFAT	0.0319	0.0283	1.1304	0.2584	0.0058
DR1TSUGR	-0.0207	0.0231	-0.8948	0.371	0.0058
DR1TSODI	-0.0011	0.0009	-1.138	0.2552	0.0058
DR1TPROT	0.0439	0.0343	1.2811	0.2003	0.0058
DR1TCARB	-0.0158	0.0168	-0.941	0.3468	0.0058

Table 5d - Regression: CRP (LBXHSCRP)

Variable	Coefficient	Std.Err	t	P>|t|	R_squared
Intercept	4.2038	0.322	13.0546	0.0	0.0059
junk_kcal_pct	0.0032	0.0073	0.4408	0.6594	0.0059
DR1TTFAT	-0.0038	0.0044	-0.8615	0.389	0.0059
DR1TSUGR	0.0094	0.0036	2.5914	0.0096	0.0059
DR1TSODI	0.0004	0.0001	2.8644	0.0042	0.0059
DR1TPROT	-0.0099	0.0054	-1.8464	0.0649	0.0059
DR1TCARB	-0.0077	0.0026	-2.9434	0.0033	0.0059

Logistic Regression: Predicting Health Risks from Dietary Intake

To complement the continuous outcome analysis, logistic regression models were developed to identify the likelihood of metabolic and inflammatory risk states based on junk food intake and its nutritional components. These binary outcomes included:

High insulin levels (>20 μU/mL),Prediabetes (HbA1c ≥ 5.7%),Elevated CRP (≥3 mg/L)

Each model assessed the role of junk food percentage and core macronutrients (fat, sugar, sodium, protein, and carbohydrates) in predicting the odds of falling into these risk categories.

High Insulin: Junk Food, Sugar, and Sodium Converge:In the first model, the odds of having high fasting insulin were significantly influenced by multiple dietary factors. The percentage of calories from junk food was positively associated with insulin resistance ($p = 0.038$) — suggesting that individuals consuming more junk calories are more likely to have insulin levels above 20 μU/mL.But junk food wasn't acting alone. Sugar ($p = 0.005$) and sodium ($p = 0.005$) were also statistically significant predictors, both increasing the odds of high insulin. This combination points to the synergistic risk of high-sugar, high-salt diets, particularly when derived from ultra-processed food sources. Interestingly, carbohydrates had a protective effect ($p = 0.001$), hinting that complex or fiber-rich carbs may help moderate insulin spikes. Fat and protein were not significant in this model.

Prediabetes: The Role of Sugar and Sodium:In the second model predicting prediabetes (HbA1c ≥ 5.7%), junk food percentage was not a significant predictor ($p = 0.494$), reinforcing earlier findings from the OLS model that HbA1c is less responsive to single-day junk food exposure.However, once again, sugar ($p = 0.008$) and sodium ($p = 0.008$) emerged as significant contributors to diabetes risk. Each additional gram of sugar and each mg of sodium pushed individuals closer to the prediabetes threshold. In contrast, carbohydrate intake reduced the odds of prediabetes ($p < 0.001$), again pointing to the possible benefit of whole-food carbs. Protein and fat intake were not statistically relevant in this model either.

High CRP: Inflammation and Diet Quality:The third model evaluated the likelihood of elevated C-reactive protein (CRP ≥ 3 mg/L), a signal of systemic inflammation. Here, junk food percentage was not a statistically significant factor ($p = 0.575$). However, several other dietary components stood out.Sodium intake ($p = 0.048$) significantly increased the odds of inflammation, consistent with the hypothesis that salt-laden processed foods contribute to chronic immune activation. Meanwhile, protein intake was inversely related to inflammation ($p < 0.001$), and carbohydrates again showed a protective effect ($p = 0.018$). These results support a growing consensus that diet quality — not just quantity — is central to inflammation regulation.

Ingredient	Insulin Resistance (High Insulin)	Prediabetes (HbA1c ≥ 5.7%)	Cardiovascular Risk (CRP & Cholesterol)	Project-Based Insight
Added Sugar	Yes (Logistic: p = 0.005)	Yes (p = 0.008)	Yes (CRP: p = 0.010)	Promotes insulin resistance, inflammation, and prediabetes
Sodium	Yes (p = 0.005)	Yes (p = 0.008)	Yes (CRP: p = 0.004)	Major pro-inflammatory and metabolic disruptor
Junk Food %	Yes (Insulin: p = 0.038)	No	No	Only predictive for insulin resistance
Protein	No	Protective (p = 0.046)	Protective (CRP: p < 0.001)	Lowers HbA1c and reduces inflammation
Carbohydrates	Protective (p = 0.001)	Protective (p < 0.001)	Protective (CRP: p = 0.003)	Likely reflects fiber/complex carbs benefit
Fat	No significant effect	No	No	No consistent impact in models

Sugar and sodium consistently increase metabolic and cardiovascular risk.
Junk food % matters primarily for insulin resistance, not for prediabetes or inflammation directly.
Protein and carbs (likely from whole foods) show protective effects, especially against inflammation and glycemic dysregulation.Fat was not significantly associated in any of the models.

CONCLUSION

This study set out to investigate how junk food consumption, a pervasive and normalized dietary pattern in the United States, influences early markers of metabolic and cardiovascular health. Leveraging nationally representative data from the NHANES 2017–2018 cycle, our research uncovered clear signals of metabolic disruption tied not just to how much we eat — but what we eat.

The most consistent and clinically meaningful finding was the link between added sugar and sodium intake and elevated risks for metabolic imbalance. While the percentage of daily calories from junk food alone was not a significant predictor for most biomarkers, it showed a statistically significant association with insulin resistance a precursor to type 2 diabetes. Individuals consuming a higher proportion of calories from junk food were more likely to exhibit fasting insulin levels above the clinical threshold, indicating early metabolic strain.Interestingly, HbA1c and CRP levels were not directly associated with junk food intake, but were significantly influenced by specific nutrients:

Sodium and sugar intake consistently increased the odds of prediabetes and inflammation.

Meanwhile, carbohydrate and protein intake — likely representing higher-fiber or higher-quality food choices showed protective effects, particularly against inflammation.

These findings are especially important because they highlight the hidden risks of junk food consumption in populations that may not appear unhealthy by traditional metrics like weight or BMI. In other words, metabolic harm can occur in the absence of obesity — a fact that challenges long-held public health assumptions.This project underscores a crucial shift in how we approach dietary health:It's not just about how many calories we consume, but the composition and quality of those calories that define our risk for disease.Public health messaging must therefore evolve beyond calorie counting and embrace a more nutrient-aware strategy — one that targets the prevalence of added sugars, sodium, and ultra-processed ingredients in the modern diet.

In conclusion, this work reaffirms that junk food isn’t just a harmless convenience — it’s a quantifiable risk factor in the early development of chronic disease. By making this link visible through data, we pave the way for smarter nutritional policy, better preventive care, and ultimately, a healthier future.

REFERENCES

Centers for Disease Control and Prevention. (2021). National Health and Nutrition Examination Survey (NHANES), 2017–2018 Data Documentation. U.S. Department of Health and Human Services. https://www.cdc.gov/nchs/nhanes/index.htm

Hall, K. D., Ayuketah, A., Brychta, R., Cai, H., Cassimatis, T., Chen, K. Y., ... & Zhou, M. (2019). Ultra-processed diets cause excess calorie intake and weight gain: An inpatient randomized controlled trial of ad libitum food intake. Cell Metabolism, 30(1), 67–77.e3. https://doi.org/10.1016/j.cmet.2019.05.008

Ludwig, D. S., & Ebbeling, C. B. (2018). The carbohydrate-insulin model of obesity: Beyond “calories in, calories out.” JAMA Internal Medicine, 178(8), 1098–1103. https://doi.org/10.1001/jamainternmed.2018.2933

Malik, V. S., Willett, W. C., & Hu, F. B. (2013). Global obesity: Trends, risk factors and policy implications. Nature Reviews Endocrinology, 9(1), 13–27. https://doi.org/10.1038/nrendo.2012.199

Monteiro, C. A., Moubarac, J.-C., Cannon, G., Ng, S. W., & Popkin, B. (2013). Ultra‐processed products are becoming dominant in the global food system. Obesity Reviews, 14(S2), 21–28. https://doi.org/10.1111/obr.12107

Mozaffarian, D., Rosenberg, I., & Uauy, R. (2018). History of modern nutrition science—Implications for current research, dietary guidelines, and food policy. BMJ, 361, k2392. https://doi.org/10.1136/bmj.k2392

World Health Organization. (2022). Healthy diet: Fact sheet. https://www.who.int/news-room/fact-sheets/detail/healthy-diet

Dehghan, M., Mente, A., Zhang, X., Swaminathan, S., Li, W., Mohan, V., ... & Yusuf, S. (2020). Associations of fats and carbohydrate intake with cardiovascular disease and mortality in 18 countries from five continents (PURE): A prospective cohort study. The Lancet, 390(10107), 2050–2062.

Khan, T. A., & Sievenpiper, J. L. (2020). The role of sugar-sweetened beverages and ultra-processed foods in the development of insulin resistance. Current Diabetes Reports, 20(8), 48.

Shivappa, N., Hebert, J. R., Kivimaki, M., Akbaraly, T. (2017). Alternative Healthy Eating Index and Dietary Inflammatory Index: Are they equivalent in predicting chronic inflammation and mortality? The American Journal of Clinical Nutrition, 106(4), 832–840.

Sun, C., Yuan, Z., Yu, Y., Xu, J., Zhang, H., Xu, C., & Liu, J. (2022). Ultra-processed food consumption and risk of type 2 diabetes: A systematic review and meta-analysis. Nutrients, 14(1), 106.

CDC. (2022). Adult Obesity Facts. https://www.cdc.gov/obesity/data/adult.html