

Burdens of type 2 diabetes and cardiovascular disease attributable to sugar-sweetened beverages in 184 countries

Received: 23 May 2024

Accepted: 7 October 2024

Published online: 06 January 2025

 Check for updates

Laura Lara-Castor^{1,2}✉, Meghan O’Hearn^{1,3}, Frederick Cudhea¹, Victoria Miller^{1,4,5}, Peilin Shi¹, Jianyi Zhang⁶, Julia R. Sharib¹, Sean B. Cash⁷, Simon Barquera⁸, Renata Micha^{1,9}, Dariush Mozaffarian^{1,10,11}✉ & Global Dietary Database*

The consumption of sugar-sweetened beverages (SSBs) is associated with type 2 diabetes (T2D) and cardiovascular diseases (CVD). However, an updated and comprehensive assessment of the global burden attributable to SSBs remains scarce. Here we estimated SSB-attributable T2D and CVD burdens across 184 countries in 1990 and 2020 globally, regionally and nationally, incorporating data from the Global Dietary Database, jointly stratified by age, sex, educational attainment and urbanicity. In 2020, 2.2 million (95% uncertainty interval 2.0–2.3) new T2D cases and 1.2 million (95% uncertainty interval 1.1–1.3) new CVD cases were attributable to SSBs worldwide, representing 9.8% and 3.1%, respectively, of all incident cases. Globally, proportional SSB-attributable burdens were higher among men versus women, younger versus older adults, higher- versus lower-educated adults, and adults in urban versus rural areas. By world region, the highest SSB-attributable percentage burdens were in Latin America and the Caribbean (T2D: 24.4%; CVD: 11.3%) and sub-Saharan Africa (T2D: 21.5%; CVD: 10.5%). From 1990 to 2020, the largest proportional increases in SSB-attributable incident T2D and CVD cases were in sub-Saharan Africa (+8.8% and +4.4%, respectively). Our study highlights the countries and subpopulations most affected by cardiometabolic disease associated with SSB consumption, assisting in shaping effective policies and interventions to reduce these burdens globally.

Sugar-sweetened beverages (SSBs) contribute to excess weight gain and cardiometabolic diseases such as type 2 diabetes (T2D) and cardiovascular disease (CVD), both directly and mediated by weight gain^{1,2}. Despite progress in elucidating the role of SSBs in health, an updated and comprehensive assessment of the global disease burden attributed to SSBs remains scarce. Our previous study estimated that, in 2010, intake of SSBs was responsible for 184,000 global deaths³. More recent analyses looking at 87 different risk factors in 2019, including

SSB intake⁴, relied primarily on national per capita estimates of added sugar availability or sales data⁵, rather than individual-level dietary data^{6,7}, limiting the validity and precision of estimates across population subgroups.

Due to their liquid form, SSBs are rapidly consumed and digested, resulting in lower satiety, higher caloric intake and weight gain⁸. High doses of rapidly digested glucose also activate insulin and other regulatory pathways, which can result in visceral fat production, hepatic and

Table 1 | Policy summary

Background	An updated and comprehensive assessment of the global, regional and national disease burdens attributable to SSBs remains scarce, particularly by key demographics such as education and urban or rural residence.
Main findings and limitations	In 2020, 2.2 million (95% UI 2.0–2.3) new T2D cases and 1.2 million (1.1–1.3) new CVD cases were attributable to SSBs worldwide, representing 9.8% and 3.1%, respectively, of all incident cases. By world region, the highest absolute cases per million adults (20+ years) were in Latin America and the Caribbean (T2D: 1,263; CVD: 522) and Middle East and North Africa (T2D: 1,001; CVD: 815), while the highest proportional SSB-attributable cases were in Latin America and the Caribbean (T2D: 24.4%; CVD: 11.3%) and sub-Saharan Africa (T2D: 21.5%; CVD: 10.5%). In 2020, the greatest absolute number of new T2D cases attributable to SSBs per million adults among the 30 most populous countries were in Mexico (2,007 per million adults; 30% of all T2D incidence cases), Colombia (1,971; 48.1%) and South Africa (1,258; 27.6%). For CVD, the greatest numbers were in Colombia (1,084; 23.0%), South Africa (828; 14.6%) and Mexico (721; 13.5%). Jointly considering education, urbanicity and world region, the highest proportions of incident T2D attributable to SSBs were among high-educated (31.9%) and mid-educated (34.2%) adults in urban sub-Saharan Africa, followed by high- and medium-educated adults in both urban and rural Latin America and the Caribbean (~26% each). Findings were similar for CVD, with the largest SSB-attributable proportions among higher-educated (19.5%) and mid-educated (17.6%) adults from urban areas in sub-Saharan Africa, but also among higher-educated and mid-educated adults from rural areas in sub-Saharan Africa and both urban and rural areas in Latin America and the Caribbean (~12–13% each). From 1990 to 2020, sub-Saharan Africa had the largest proportional increases in incident T2D and CVD (12.7–21.5% and 6.1–10.5%, respectively), while Latin America and the Caribbean experienced slight decreases. The findings represent estimates based on available data and reasoned assumptions and do not prove cause and effect.
Policy implications	Our study offers a comprehensive analysis of the global burden of SSB-attributable T2D and CVD burdens, incorporating sociodemographic disparities and regional nuances. While some policies to curb SSB intakes are currently in place in some countries, our study suggests that more work is needed. In Latin America and the Caribbean, for instance, several nations have implemented policies to curb SSB intakes, yet this region had the largest SSB-attributable cardiometabolic burdens in 2020. Mid- and high-educated adults in rural and urban Latin America and the Caribbean and rural sub-Saharan Africa should be given particular attention. The increasing burdens in sub-Saharan Africa shed light on the necessity of acting quickly in this region. By highlighting the countries and subpopulations most affected, our research can assist in shaping effective policies and interventions to efficiently reduce the burden of cardiometabolic diseases attributed to SSB consumption globally.

skeletal muscle insulin resistance and weight gain. High doses of rapidly digested fructose directly activate hepatic fat synthesis, leading to ectopic fat deposition and metabolic dysfunction in liver and muscle⁹. SSBs may also replace other healthier foods in the diet, contributing to harms through their absence. Excess adiposity and metabolic dysfunction activate inflammatory cytokines and increase risk of hypertension, dyslipidemia and diabetes¹⁰. All these risk factors accelerate atherosclerosis and plaque instability, contributing to ischemic cardiovascular events¹¹. Hence, both direct and adiposity-mediated effects of SSBs are relevant to assessing their health effects.

Both SSB intake and cardiometabolic risk also can vary substantially by key demographic factors within nations. For example, we recently reported that SSB intakes were higher among more versus less educated adults in sub-Saharan Africa, South Asia and Latin America and the Caribbean, while the inverse pattern was observed in Middle East and North Africa¹². By area of residence, intakes were higher in urban versus rural areas in sub-Saharan Africa and South Asia, whereas the inverse was true in the Middle East and North Africa. Yet, assessments of the global disease burden attributable to SSBs by key demographics such as educational attainment and urban versus rural residence have yet to be reported at a global scale.

This study aims to estimate the burdens of cardiometabolic diseases attributable to SSBs and the changes over time, in nations worldwide as well as subnationally, by key sociodemographic factors, as highlighted by the substantial public health challenge of SSB intake in most world regions¹². The findings would inform national, subnational and multinational surveillance and policy actions to address SSBs and their disease burdens, including inequities across nations and population subgroups (Table 1).

Results

Distributions of SSB intakes

SSBs were defined as any beverage with added sugars and ≥ 50 kcal per 8 oz serving, including commercial or homemade beverages, soft drinks, energy drinks, fruit drinks, punch, lemonade and aguas frescas. This definition excluded 100% fruit and vegetable juices, noncaloric artificially sweetened drinks and sweetened milk. We derived SSB intakes from the Global Dietary Database (GDD)^{7,12,13}, including 450 surveys with data on SSBs, totaling 2.9 million individuals from 118 countries representing 87.1% of the global population (Supplementary

Tables 1 and 2) and Bayesian hierarchical modeling. Globally in 2020, consistent with findings reported in 2018¹², adults consumed an average of 2.68 oz (248 g) servings per week (95% uncertainty interval (UI) 2.4–2.8). This ranged regionally from 0.7 (95% UI 0.5–1.1) in South Asia to 7.3 (95% UI 6.7–8.1) in Latin America and the Caribbean (Supplementary Table 3), and nationally among the 30 most populous countries from 17.4 in Colombia (95% UI 13.2–22.7), 9.6 in South Africa (95% UI 7.5–12.5), 8.5 in Mexico (95% UI 7.8–9.4) and 6.9 in Ethiopia (95% UI 5.5–8.7) to 0.2 in India, China and Bangladesh (Supplementary Table 4).

Globally, regionally and nationally, men had modestly higher energy-adjusted SSB intake than women. By age, SSB intakes were higher at younger compared with older ages in all world regions, though with varying absolute magnitudes of intakes and differences by region (Supplementary Table 3). The largest variations by age were observed in Latin America and the Caribbean, and the lowest in South Asia. By region and education, intakes were higher among more versus less educated adults in sub-Saharan Africa, South Asia and Latin America and the Caribbean, but lower among more versus less educated adults in the Middle East and North Africa, with smaller differences by education in other regions¹².

Global T2D and CVD burdens attributable to SSBs

SSB intakes and cardiometabolic disease rates were incorporated into a comparative risk assessment (CRA) model to assess risk. The CRA framework does not use ecologic correlations but is based on independent lines of evidence¹⁴, including age-adjusted etiologic effects of SSBs on T2D, ischemic heart disease and ischemic stroke, both directly and mediated by body mass index (BMI), from previous meta-analyses and pooled analyses of prospective cohorts, supported by evidence from randomized controlled trials^{2,15–17}. For each model, we ran 1,000 Monte Carlo simulations and report the median and 2.5th and 97.5th values (95% UI).

In 2020, an estimated 2.2 million (95% UI 2.0–2.3) new T2D cases and 1.2 million (95% UI 1.1–1.3) new CVD cases were attributable to intake of SSBs globally, corresponding to 9.8% (95% UI 9.1–10.5) and 3.1% (95% UI 2.8–3.4) of total incident cases, respectively (Supplementary Data 1 and Extended Data Fig. 1). SSBs contributed to 12.5 million cardiometabolic disability-adjusted life years (DALYs), including 5.0 million (95% UI 4.6–5.4) from T2D (6.9% (6.4–7.4) of all T2D DALYs) and 7.6 million (95% UI 6.9–8.3) from CVD (3.0% (2.7–3.3) of all CVD DALYs). SSBs were estimated to cause 80,278 (72,297–88,824) deaths from T2D

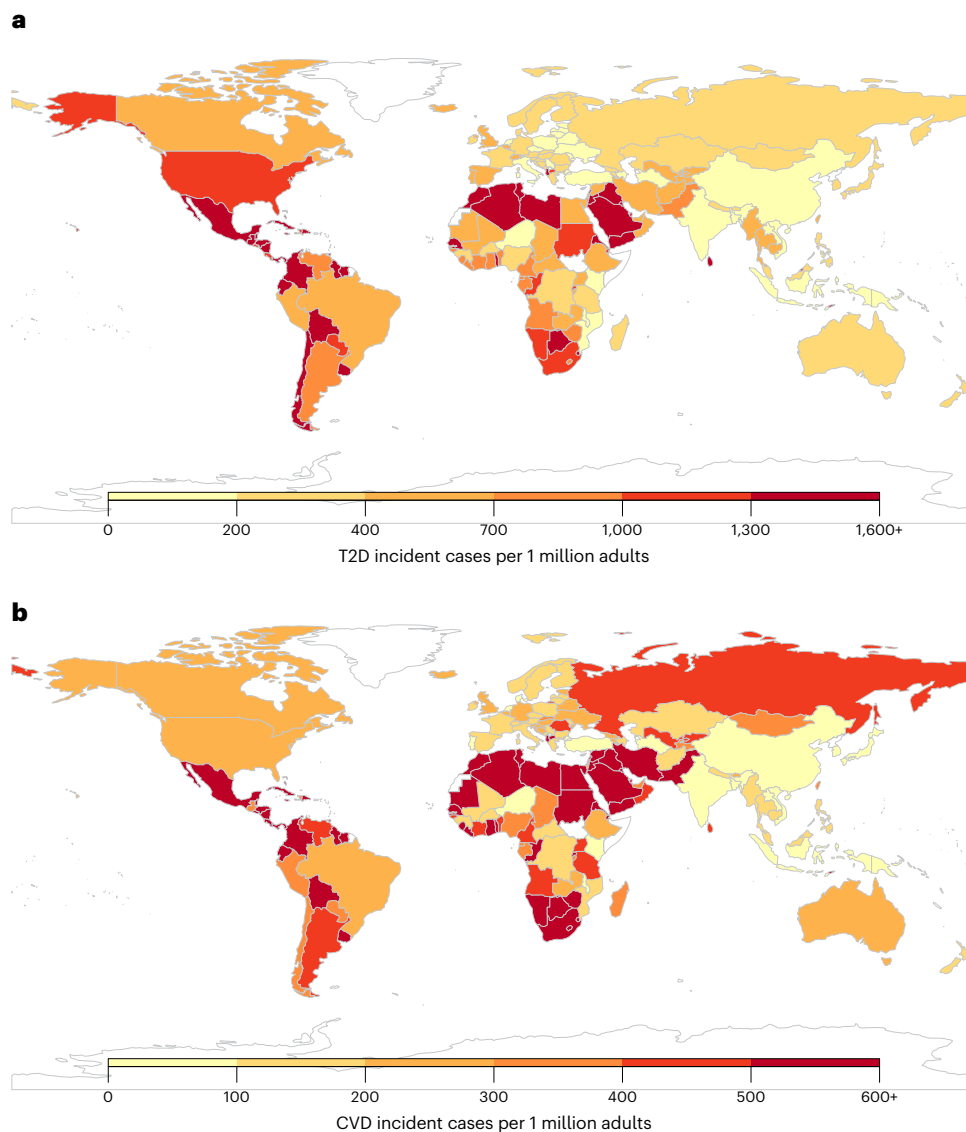


Fig. 1 | Incidence of T2D and CVD per 1 million adults attributable to SSB intake among adults (20+ years) in 184 countries in 2020. a, b. Absolute SSB-attributable T2D incidence (**a**) and absolute SSB-attributable CVD incidence (**b**). The SSB-attributable absolute burden per 1 million adults was calculated by dividing the country absolute number of SSB-attributable cases by the country

adult population (20+ years) in that same year and multiplying by 1 million. Values were truncated at 1,600 for **a** and at 600 for **b** to better reflect the absolute case distribution globally for T2D and CVD. The analysis of the data was done using the worldmap package (v1.3-6). Source data are provided in Source Data Fig. 1.

(5.1% (4.6–5.7) of all T2D deaths) and 257,962 (235,059–283,798) deaths from CVD (2.1% (1.9–2.3) of all CVD deaths).

Among the 30 most populous countries, the greatest absolute numbers of new T2D cases attributable to SSBs were in Mexico (2,007 per million adults (1,754–2,338)), Colombia (1,971 (1,612–2,354)) and South Africa (1,258 (1,005–1,575)) (Fig. 1 and Supplementary Data 1). For CVD, the greatest numbers were in Colombia (1,084 (832–1,381)), South Africa (828 (645–1,083)) and Mexico (721 (612–889)). As a proportion of all new cases, the highest SSB burdens of T2D were in Colombia (48.1% (39.3–57.3)), Mexico (30.0% (26.4–35.0)) and South Africa (27.6% (22.1–34.6)), and those of CVD were in Colombia (23.0% (18.0–29.2)), South Africa (14.6% (11.4–19.2)) and Mexico (13.5% (11.5–16.7)). Findings on cardiometabolic deaths and DALYs attributable to SSBs in 184 nations are presented in Supplementary Figs. 1 and 2 and Supplementary Data 1.

Regional T2D and CVD burdens attributable to SSBs

By world region, Latin America and the Caribbean had the highest absolute and proportional T2D incidence due to SSBs (1,263 new cases

per 1 million (1,146–1,400); 24.4% (22.3.0–26.9)), and Southeast and East Asia had the lowest (119 new cases per 1 million (103–145); 3.1% (2.7–3.8)) in 2020 (Fig. 2 and Supplementary Table 5). SSB-attributable CVD incidence ranged from 815 new cases per 1 million (674–980) in the Middle East and North Africa to 46.8 new cases per 1 million (41.0–57.1) in Southeast and East Asia (Fig. 2 and Supplementary Table 6). SSBs were estimated to have caused more than 1 in 10 new CVD cases in Latin America and the Caribbean (11.3% (10.1–12.8)) and sub-Saharan Africa (10.5% (8.1–13.3)), compared with less than 1 in 100 cases in South Asia (0.60% (0.6–0.8)). Cardiometabolic mortality and DALYs from SSBs in different world regions are presented in Supplementary Figs. 3–6 and Supplementary Tables 5 and 6.

Global heterogeneity by age, sex, education and urbanicity

Globally, SSBs were estimated to cause more T2D cases in men (447 per 1 million adults (413–491); 10.1% (9.3–11.10) of total cases) versus women (388 per 1 million adults (358–422); 9.5% (8.8–10.4)), higher-educated (531 per 1 million adults (486–620); 11.1% (10.1–12.9))

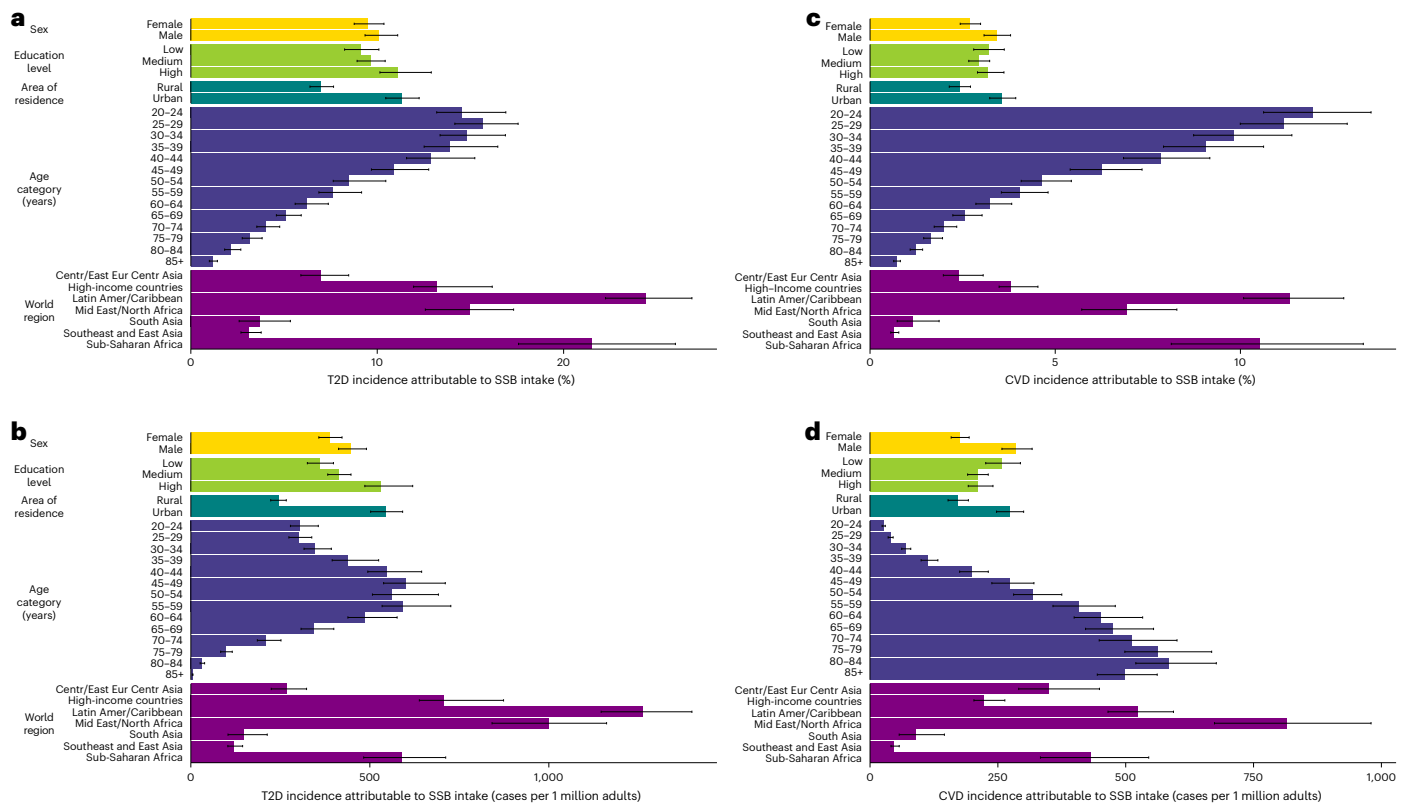


Fig. 2 | Incidence of T2D and CVD attributable to SSB intake by key sociodemographic factors at the global level and by world region in 2020. **a–d.** The bars represent the central estimate (median) of the proportional SSB-attributable T2D incidence (**a**), the absolute SSB-attributable T2D incidence per 1 million adults (**b**), the proportional SSB-attributable CVD incidence (**c**) and the absolute SSB-attributable CVD incidence per 1 million adults (**d**). The error bars represent the 95% UI derived from the 2.5th and 97.5th percentiles of 1,000 multiway probabilistic Monte Carlo model simulations. The SSB-attributable absolute burden per 1 million adults was calculated by dividing the stratum

absolute number of SSB-attributable cases by the stratum adult population (20+ years) in that same year and multiplying by 1 million. In previous GDD reports, the region ‘Central and Eastern Europe and Central Asia’ was referred to as ‘Former Soviet Union’, and ‘Southeast and East Asia’ was referred to as ‘Asia’. See Supplementary Table 2 for a list of countries included in each world region. Source data are provided in Source Data Fig. 2. Centr/East Eur Centr Asia, Central or Eastern Europe and Central Asia; Latin Amer/Caribbean, Latin America and the Caribbean; Mid East/North Africa, Middle East and North Africa.

versus lower-educated adults (360 per 1 million adults (326–398); 9.1% (8.2–10.1)) and urban (543 per 1 million adults (502–592); 11.3% (10.5–12.2)) versus rural adults (244 per 1 million adults (233–267); 7.0% (6.4–7.7)) in 2020 (Fig. 2 and Supplementary Table 5). By age, absolute burdens of SSB-attributable T2D cases were highest at ages 45–49 years (601 per 1 million adults (539–711); 10.9% (9.7–12.8)), while proportional risk was highest at ages 25–29 years (301 per 1 million adults (274–338); 15.6% (14.2–17.6)).

For CVD, absolute SSB-attributable incident cases per 1 million adults were higher among men (285 per 1 million adults (258–317)) than women (176 per 1 million adults (159–194)) in 2020, owing to both higher SSB intake and higher baseline CVD risk (Fig. 2 and Supplementary Table 6). SSB-attributable CVD incidence was also higher in urban adults (273 per 1 million adults (247–300)) than in rural adults (172 per 1 million adults (153–192)), for similar reasons. By contrast, global SSB-attributable CVD incidence was similar across education levels. Absolute incidence of SSB-attributable CVD increased with age, while proportion risk decreased with age. For example, SSBs were estimated to contribute to 585 new CVD cases (520–677) per 1 million adults among adults aged 80–84 years (1.2% (1.1–1.4) of total incident CVD in this age group) versus 26.2 new CVD cases (23.6–29.8) per 1 million among adults aged 20–24 years (12.0% (10.6–13.5) of total incident CVD in this age group). T2D and CVD mortality and DALYs attributable to SSBs followed similar patterns as for T2D and CVD incidence (Supplementary Figs. 3–6 and Supplementary Tables 5 and 6).

Regional and national heterogeneity by age, education and urbanicity

In all world regions, the proportion of SSB-attributable T2D and CVD cases was highest at the youngest ages (Extended Data Fig. 2), with most pronounced variations by age in high-income countries, Latin America and the Caribbean, Middle East and North Africa, and sub-Saharan Africa. By world region and age, the highest proportional incidence due to SSBs was seen among younger adults in Latin America and the Caribbean: 43.7% (39.0–50.4) in 20–24-year-olds and 41.2% (36.1–48.8) in 25–30-year-olds. Patterns were similar for proportions of CVD attributable to SSBs (Extended Data Fig. 2). Cardiometabolic deaths and DALYs by age and world region are shown in Supplementary Figs. 7 and 8.

Cardiometabolic burdens due to SSB varied by education and urban or rural residence across world regions (Fig. 3). When education, urbanicity and world region were jointly considered, it was revealed that the highest proportions of incident T2D attributable to SSBs were among high-educated (34.2% (26.8–42.7)) and mid-educated (31.9% (25.7–38.4)) adults in urban sub-Saharan Africa, followed by high- and medium-educated adults in both urban and rural Latin America and the Caribbean (~26% each). Our findings were similar for CVD, with the largest SSB-attributable proportions among higher-educated (19.5% (14.6–26.1)) and mid-educated (17.6% (13.5–22.4)) adults from urban areas in sub-Saharan Africa, but also among higher-educated and mid-educated adults from rural areas in sub-Saharan Africa and both urban and rural areas in Latin America and the Caribbean (~12–13%

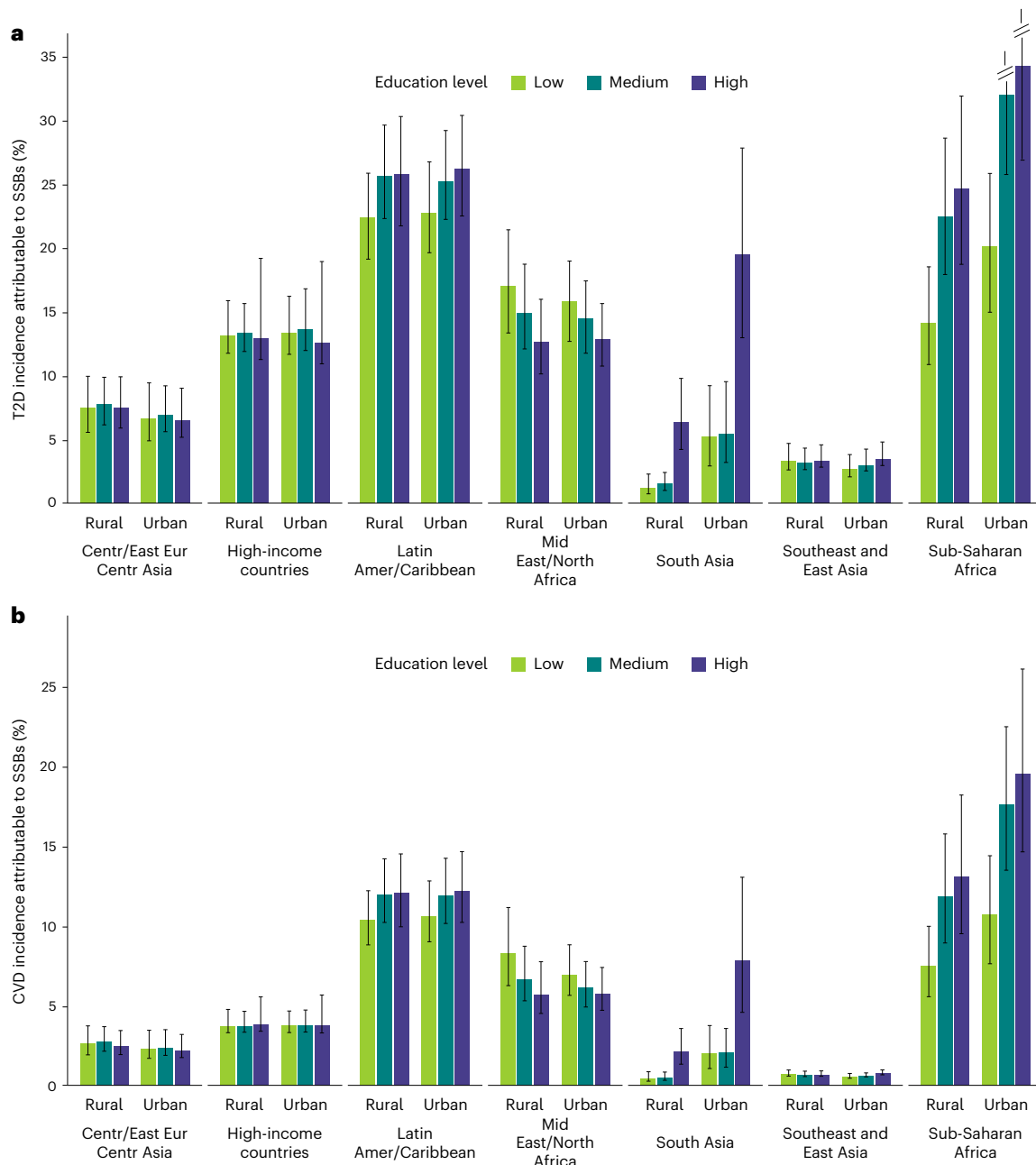


Fig. 3 | Proportional incidence of T2D and CVD attributable to SSB intake among adults (20+ years) jointly stratified by world region, area of residence and education level in 2020. a, b. The bars represent the central estimate (median) of the proportional SSB-attributable T2D incidence (a) and

CVD incidence (b). The error bars represent the 95% UI derived from the 2.5th and 97.5th percentiles of 1,000 multiway probabilistic Monte Carlo model simulations. Values were truncated at 35 for a, and 95% UIs above 35 are shown with diagonal lines. Source data are provided in Source Data Fig. 3.

each). Patterns for SSB-attributable deaths and DALYs were consistent with these results (Supplementary Figs. 9 and 10).

Changes over time in T2D and CVD attributable to SSBs

Globally from 1990 to 2020, the proportion of T2D incidence attributable to SSBs increased by 1.3% absolute percentage points (0.9–1.7), and that of CVD decreased by –0.1% (–0.3 to 0.0). By region, the greatest percentage increase in T2D and CVD burdens due to SSBs was in sub-Saharan Africa (Fig. 4 and Supplementary Data 2), where SSB-attributable T2D increased by 8.8 percentage points (6.8–11.0) and CVD by 4.4 percentage points (3.1–5.8). More moderate increases were also identified in the Middle East and North Africa and Central or Eastern Europe and Central Asia. By contrast, T2D and CVD proportional

burdens were generally stable over time in other regions, while Latin America and the Caribbean and high-income countries experienced a slight decrease. Similar patterns were identified in SSB-attributable deaths and DALYs (Supplementary Fig. 11).

Among the 30 most populous countries, the largest increase over time in SSB-attributable new T2D cases per 1 million adults was in Colombia with 793 more cases (627–972), followed by the United States (671 (576–985)), Argentina (544 (432–682)), Myanmar (522 (364–772)) and Thailand (512 (249–982)) (Fig. 5). Incident SSB-attributable CVD cases increased most in Nigeria (291 (188–464)), Russia (274 (213–414)), Colombia (216 (125–335)) and Thailand (166 (79.6–340)). By contrast, Turkey (–156 (–234, –59.7)) experienced the largest decrease in incident T2D due to SSBs, while the largest reductions in incident CVD due to

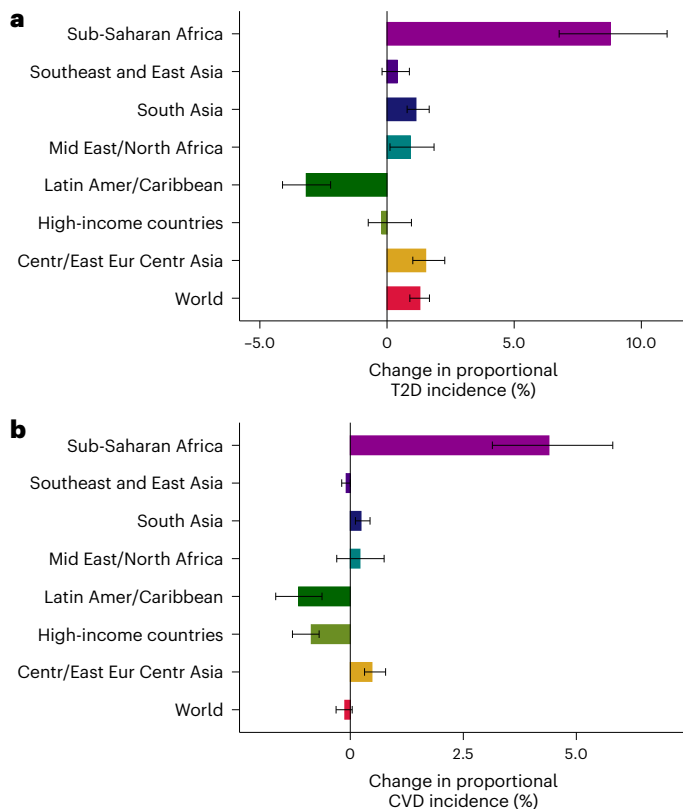


Fig. 4 | Change in proportional incidence of T2D and CVD attributable to intake of SSBs among adults (20+ years) from 1990 to 2020 by world region. **a,b**, The bars represent the central estimate (median) of the difference between 1990 and 2020 for the proportional T2D incidence (**a**) and CVD incidence (**b**) attributable to SSB. The error bars represent the 95% UI derived from the 2.5th and 97.5th percentiles of 1,000 multiway probabilistic Monte Carlo model simulations. Source data are provided in Source Data Fig. 4.

SSBs were in Turkey (−541 (−789, −378)), the United States (−382 (−536, −322)), South Africa (−202 (−304, −111)) and the United Kingdom (−172 (−245, −133)). Changes over time in SSB-attributable T2D and CVD deaths and DALYs among the 30 most populous countries are presented in Supplementary Figs. 12 and 13.

To better understand relationships of these trends with national development, we investigated findings by sociodemographic index (SDI), a composite measure of national development based on fertility rate under age 25, mean education level among people aged 15+ years and income per capita¹⁸. In 1990, the national SDI was not correlated with SSB-attributable T2D or CVD disease burdens. However, by 2020, the national SDI was inversely correlated with both incident T2D ($r = -0.30$, $P < 0.01$) and CVD ($r = -0.33$, $P < 0.01$), with higher SSB-attributable health burdens occurring among nations with lower SDI (Extended Data Figs. 3 and 4, Supplementary Figs. 14–17 and Supplementary Data 3).

Discussion

This investigation provides updated estimates of T2D and CVD health burdens attributable to SSBs worldwide, with stratification by age, sex, education, and rural or urban residence³. Globally, we found that 2.2 million new cases of T2D and 1.2 million new cases of CVD in 2020 were attributable to SSBs—representing about 1 in 10 new T2D and 1 in 30 new CVD cases. In addition, we estimated that about 340,000 people died in 2020 from SSB-related T2D and CVD. Important trends over time were identified by world region and demographic subgroups.

These SSB-attributable health burdens are informed by independently derived etiologic effects incorporated into our CRA model,

identified from published meta-analyses of prospective cohort studies and randomized controlled trials of SSBs and cardiometabolic outcomes. A larger global effect on T2D than on CVD can be explained by the greater impact of adiposity on T2D versus CVD, as well as T2D onset generally occurring at younger ages (when SSB intakes are much higher) compared with CVD. Also, competing risk factors for CVD, such as smoking, LDL cholesterol and blood pressure, lower the relative impact of SSBs on CVD compared with T2D¹⁹.

We identified important heterogeneity in these cardiometabolic disease burdens. Regionally, sub-Saharan Africa experienced the largest increases in SSB-attributable burdens between 1990 and 2020, while Latin America and the Caribbean, despite modest declines over this period, retained significant burdens compared with other regions. We also identified relatively high SSB-related health burdens among individuals with higher educational attainment in Latin America and the Caribbean, South Asia and sub-Saharan Africa; lower educational attainment in the Middle East and North Africa; and both urban and rural residence in South Asia and sub-Saharan Africa. Proportional burdens were largest among younger versus older adults in most world regions, whereas absolute burdens were more substantial among middle-aged and older adults. While proportional SSB-attributable burdens were lower among older adults owing to other competing risk factors for cardiometabolic disease later in life, absolute numbers of cases and deaths were substantial.

The large SSB-attributable burdens in sub-Saharan Africa are consistent with reports of substantial increases in SSB intakes and, separately, in cardiometabolic disease rates in the region between 1990 and 2020^{12,20}. Our findings provide a clarion call that the ‘nutrition transition’ from traditional toward Western diets has already occurred in much of the region, requiring urgent policy and public health attention. Yet, many sub-Saharan African nations have not implemented any measures to curb SSB intakes, perhaps owing to both industry opposition and previous lack of credible country-specific data^{21,22}. As SSB intakes have leveled or started to decline in high-income nations¹², the beverage industry has turned to emerging markets where populations are highly susceptible to marketing appeal of aspirational ‘Western’ lifestyles^{23,24}. Critically, we found in many such regions that SSB-attributable health burdens are now largest among more educated adults, in line with regional economic growth and burgeoning middle classes^{23,25}. In South Africa, for example, SSB advertising is mainly directed at wealthy males under 45 years old, who also show the highest SSB consumption in the country²⁶. Although South Africa has implemented an SSB tax, the beverage industry increased their advertising to offset negative effects on sales²⁷. Independent advertising regulations—that is, not designed or policed by industry—have been largely missing^{26,27}. Our findings highlight the need to address high and increasing health burdens from SSBs in these nations. Our results further suggest, given higher intakes and health burdens among more versus less educated adults in many regions, that general education alone is unlikely to effectively reduce SSB intakes (and could even augment intakes and health burdens).

The modest decreases in SSB-related cardiometabolic burdens identified in Latin America and the Caribbean are consistent with slowly decreasing consumption of SSBs in this region¹². Nations in this region have implemented several policy efforts targeting SSBs, including taxes, marketing regulations, front-of-package warnings and education campaigns^{28,29}. Yet, SSB-attributable health burdens remain high in the region, and absolute burdens per million adults continue to rise owing to continuing increased rates in obesity, T2D and CVD. While declining SSB intake may have slowed regional growth in obesity, SSB intake remains high, while other risks such as high refined grain intake and physical inactivity remain prevalent³⁰. Given the lag between obesity and development of T2D and CVD, as well as the persistent impact of early life influences on later cardiometabolic risk, concerted multigenerational efforts over many years may be needed to reverse these challenges.

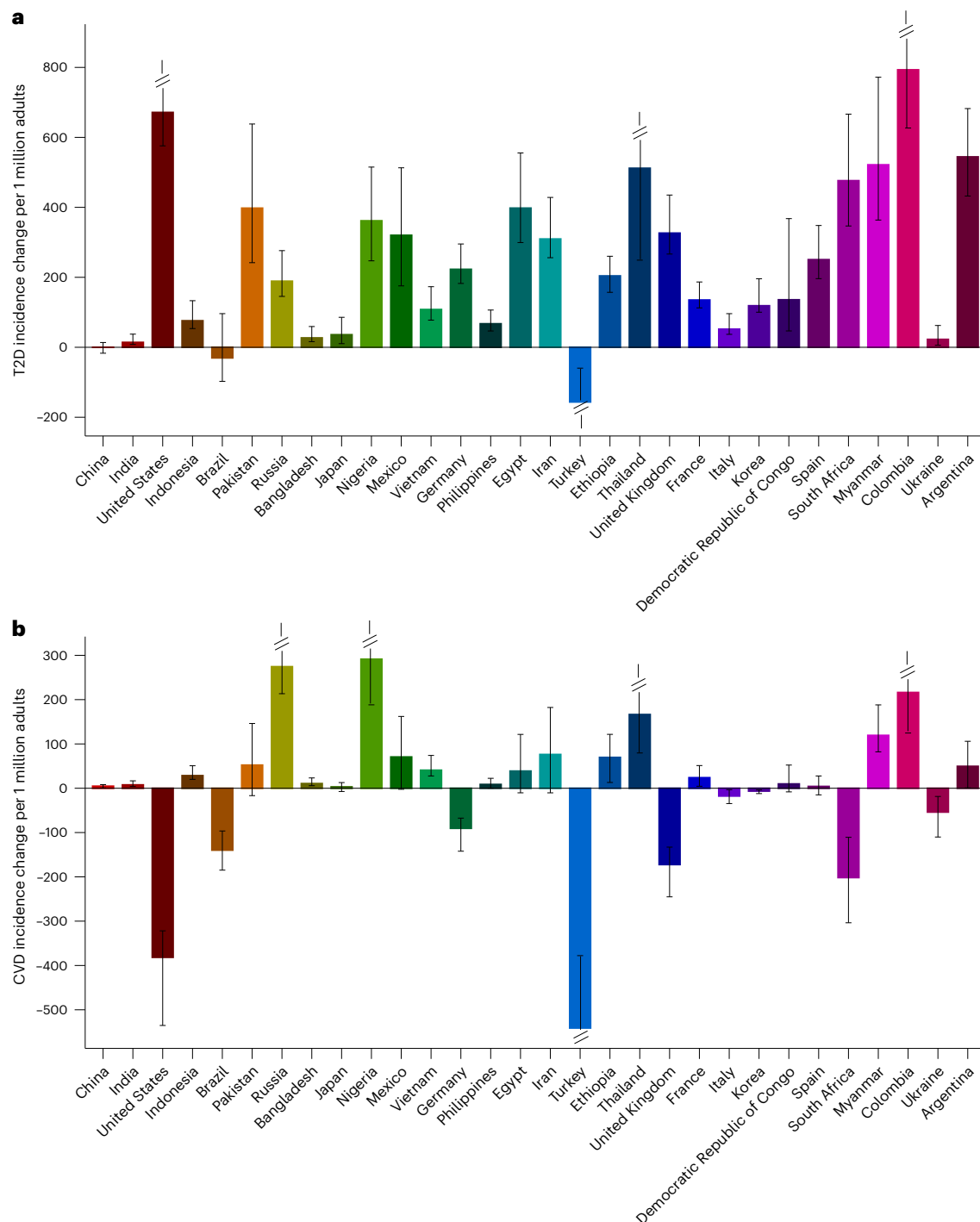


Fig. 5 | Change in incident cases per 1 million adults of T2D and CVD attributable to SSB intake among adults (20+ years) from 1990 to 2020 among the 30 most populous countries. a, b, The bars represent the central estimate (median) of the difference between 1990 and 2020 of the absolute SSB-attributable T2D incidence (**a**) and CVD incidence (**b**) per 1 million adults. The error bars represent the 95% UI derived from the 2.5th and 97.5th percentiles of 1,000 multiway probabilistic Monte Carlo model simulations. Values were truncated from -200 to 870 for T2D (**a**) and from -545 to 300 for CVD (**b**). The 95% UIs above or below these values are shown with diagonal lines.

The SSB-attributable absolute burden per 1 million adults was calculated by dividing the country absolute number of SSB-attributable cases by the country adult population (20+ years) in that same year and multiplying by 1 million. The difference in the absolute burden per 1 million adults was calculated by subtracting per 1 million adult burdens in 1990 from per 1 million adult burdens in 2020. From left to right, the countries are ordered from most to least populous based on 2020 adult (20+ years) population data. Source data are provided in Source Data Fig. 5.

Compared with other regions, South Asia and Southeast Asia had the lowest SSB-attributable cardiometabolic burdens, consistent with their much lower intakes¹². Other competing dietary risks, such as large T2D burdens attributable to refined rice³¹,

may also lower relative SSB-attributable burdens. In line with these findings, a set of prospective cohorts across 21 countries found that incident T2D associated with refined rice intake was highest in South Asia³². Our lack of data on sweetened teas could also partly

underestimate burdens attributable to these beverages, given their expanding market in Asia³³.

In nations with high SSB intakes and T2D rates among younger adults, resulting attributable burdens will be greater. Too large or too low nation-specific findings should be interpreted within their estimated uncertainty ranges. For example, Colombia had high central estimates of SSB-attributable burdens (48.1% for T2D incidence; 23.0% for CVD incidence), and these estimates' lower bounds (39.3% and 18.0%, respectively) are also plausible SSB-attributable burdens in this country. Given imperfections of global modeling, a reasonable conclusion is that Colombia's SSB-attributable T2D harms are at least as high as in other high-burden nations such as Mexico (30.0%) or South Africa (27.6%).

Between 1990 and 2020, SSB-attributable T2D incidence increased 1.3% globally while SSB-attributable CVD incidence remained stable, although with wide variations by region and nation. For instance, sub-Saharan Africa experienced an increase in T2D incidence due to SSBs of 8.8 percentage points, while Latin America and the Caribbean experienced a decrease of -3.2 percentage points. Among large nations, the largest increases in SSB-related T2D burdens were in Mexico, Thailand and the United Kingdom, and in CVD burdens, Colombia, Nigeria, Thailand and Russia. These changes align with rises in SSB consumption in these nations¹². Similarly, declining SSB-related cardiometabolic burdens in Brazil, the United States and the United Kingdom (for CVD) are consistent with their decreasing SSB consumption from 1990 to 2020¹².

Commercial interests of multinational and local SSB manufacturers, who invest in supply chain and marketing efforts to increase sales while opposing policy efforts to curb intakes, are probably drivers of increased SSB intakes and attributable cardiometabolic burdens in many nations³⁴. Mexico faces industry opposition to its soda tax, including industry-supported reports questioning the efficacy of the tax to reduce intakes and suggesting harms to jobs and the national economy, as well as amplified marketing through advertising, price reductions and bonus products³⁴. Colombia's 2016 efforts to pass an SSB tax were blocked by industry opposition, although renewed efforts were successful in 2022³⁵. In Nigeria, increased SSB intake has motivated studies looking into potential SSB taxation and other policies to curb intakes³⁶. However, increasing SSB sales in Nigeria have likewise motivated the expansion of the soda industry, supported by the vulnerability of the youthful population and the availability of relevant natural resources for SSB production³⁷. The case of Nigeria highlights a reality that growth in SSB sales frequently greatly outpaces consideration, development and implementation of countering policies. When Thailand implemented an SSB tax in 2017 in response to high intakes, promising effects were seen among older and lower-income individuals, but not in youth³⁸, which may relate to heavy advertising targeting youth in Thailand³⁹. In the United Kingdom, a 2018 SSB graduated tax encouraged industry reformulation to lower sugar content, but the total volume of SSBs purchased went up⁴⁰, potentially owing to counteracting industry marketing strategies⁴¹. Beyond commercial interests, insufficient availability of safe drinking water can contribute to rising SSB-related health burdens in many nations, such as among rural areas in Colombia, Mexico and Thailand⁴²⁻⁴⁴. Clean water scarcity and commercial interests can go hand in hand: in one Mexican town, water scarcity was partly attributable to water concessions for soda companies⁴⁵. Our findings show the downstream health consequences, which are high and often rising, of these realities.

Addressing cardiometabolic diseases is crucial to several United Nations 2030 Sustainable Development Goals, including promoting good health and well-being, reducing inequalities, promoting responsible consumption and reducing poverty⁴⁶. Our research contributes to these endeavors by quantifying SSB-attributable cardiometabolic burdens, including population subgroups with the largest risk and increases over time, to more accurately inform national strategies. Policies targeting the food environment, including taxes and food labeling,

are gaining traction and can influence consumer behavior⁴⁷⁻⁴⁹. Taxation based on sugar density also influences product reformulation, promoting the availability of lower-sugar options⁵⁰. Currently, a higher proportion of people in low- and lower-middle-income countries is exposed to SSB taxes compared with those living in upper-middle-income and high-income countries⁵¹. Yet, these taxes are generally new (many implemented after 2018) and often authorized at low levels to raise revenue, rather than higher levels needed to reduce consumption. Our findings also reveal that lower-SDI nations are more likely to have higher T2D and CVD SSB-attributable burdens than higher-SDI nations—a worrisome change since 1990, and consistent with other recent reports⁵². Lower-SDI countries can face significant constraints in policy implementation, including lower tax rates, implementation oversight and administration capacity⁵³, as well as challenging population-wide access to clean water, which may hinder the declines in SSB consumption⁵⁴. Our findings suggest the need for national and multilateral design, implementation and evaluation of policy to reduce SSBs, such as taxes, front-of-package labeling, marketing regulations, school food regulations and water sanitation efforts^{55,56}, with particular focus on nations and subgroups with the largest SSB-attributable cardiometabolic burdens.

Some industry segments have initiated sugar reduction in certain products and regions^{57,58}. In Europe, Coca-Cola and PepsiCo have cut sugar contents by 30–50% in some products including Sprite and 7-Up⁵⁸. SSB marketing and distribution are often replaced by those for beverages with nonnutritive sweeteners. While short-term weight effects of the latter may be less problematic, growing evidence suggests that such compounds are not innocuous and may have harms on the microbiome and glucose tolerance⁵⁹⁻⁶¹. Given the widespread and increasing use of nonnutritive sweeteners, additional research on their health effects is critically needed, as well as public health messaging for avoidance whenever possible.

The GBD recently estimated, among other risk factors, SSB-attributable cardiometabolic burdens globally^{52,62,63}, although incidence was not reported, which is the primary focus of our study. Considering deaths, the GBD estimated 52,882 SSB-attributable T2D deaths in 2020, representing 3.4% of total T2D deaths⁶³. In comparison, we estimated 80,278 SSB-attributable T2D deaths, or 5.1% of the total. The GBD estimated only 13,691 SSB-attributable CVD deaths in 2020, compared with 257,962 in our analysis. Important methodologic differences between these estimates include (1) our use of a much larger number of individual-level dietary surveys with data on SSBs (GDD: 450; GBD: 44)⁶⁴; (2) individual dietary data derived from more countries (GDD: 118; GBD: 17); (3) our use of individual-level surveys as the primary input to estimate global SSB intakes, rather than sales data and Food and Agriculture Organization (FAO) food balance estimates of national per capita added sugar availability used by GBD^{12,65}; and (4) our incorporation of updated findings on both direct and BMI-mediated etiologic effects of SSBs, compared with GBD that did not include mediation by BMI⁶⁵. We consider the latter to be important given effects of SSBs on adiposity, contributions of overweight and obesity to cardiometabolic risk, and the fact that SSB-disease estimates in the literature generally adjust for BMI^{2,17}. Ultimately, differing careful global investigations such as these are critical to help triangulate the reality on the ground in the absence of uniformly available data in all nations, while similarities and differences between studies can help identify how differences in reasoned assumptions and methods might influence results.

Our study has several strengths. To our knowledge, previous investigations have not reported global, regional and national estimates of T2D and CVD SSB-attributable burdens jointly stratified by age, sex, education, and urban or rural residence. About 85% of dietary inputs were derived from individual-level dietary surveys (that is, 24 h recalls, food records and food frequency questionnaires), and uncertainty in the individual data sources, their methods and representativeness

is incorporated into our hierarchical Bayesian model to help address potential bias from less reliable dietary collection methods, variability in representativeness and survey level stratification, and sampling and model uncertainties. Compared with previous studies, our final GDD dietary estimates are stratified by education and area of residence, in addition to sex and age. Our CRA incorporated etiologic effects of SSBs from meta-analyses of prospective cohorts or randomized trials that facilitate inference of independence and temporality, rather than performing cross-sectional correlational analyses that can be strongly limited by reverse causation and cross-national confounding. The cohort-derived effects were multivariable adjusted for sociodemographic, lifestyle and other dietary factors and are consistent with randomized trials of SSBs, representing reasonable estimates of long-term health effects. Studies of etiologic effects with increased risk of bias, such as retrospective or cross-sectional studies, were excluded². We incorporated sampling and modeling uncertainty from each of our model inputs, providing central estimates and measures of uncertainty representing the 95% most probable values¹².

Limitations should be considered. Our estimates are based on best available data and reasoned assumptions, and do not prove cause and effect. The CRA framework is not a microsimulation estimating the impact of a specific intervention to reduce future SSB intakes, but a counterfactual approach that estimates the health effects of current SSB intakes compared with the scenario in which such an exposure was not present. While etiologic effects of SSBs on weight gain, T2D and CVD were obtained from multiple cohorts across world regions, these were mostly from high-income countries especially for weight gain that used pooled US studies, which could imperfectly represent other populations if health effects of SSBs are in the future shown to biologically vary by world region or over time. While we varied such estimates by age and baseline BMI, current evidence is likewise insufficient to vary such risk estimates by other population characteristics. Although etiologic effects were obtained from multivariable-adjusted studies and, where available, were consistent with findings from trials, measurement error, residual confounding and publication bias cannot be ruled out, which could alter findings in unpredictable directions. We did not incorporate other likely SSB-related health harms, such as dental caries, other effects of adiposity, hepatic steatosis or microbiome dysfunction; thus, our findings probably underestimate the full health burdens of SSBs. Despite extensive efforts of the GDD, dietary data were limited for several time periods and countries, particularly lower-income nations¹². Accordingly, estimated burdens in countries with fewer individual-level surveys have higher uncertainty. All dietary assessments include some error; however, validated methods such as multiple 24 h recalls, food records and food frequency questionnaires included in the GDD are considered realistic and reliable tools for individual-level dietary collection in large-scale demographic studies⁶⁶. Our SSB definition did not include 100% fruit juices or sugar-sweetened milk, which have shown inconsistent evidence for cardiometabolic effects^{2,9}. Global dietary surveys often did not collect information on sugar-sweetened tea or coffee, highlighting a future surveillance need, particularly in Asia³³.

In summary, our study offers a comprehensive analysis of the global burden of SSB-attributable T2D and CVD, incorporating sociodemographic disparities and regional nuance. The largest proportional T2D and CVD attributable burdens in 2020 were in Latin America and the Caribbean and sub-Saharan Africa, and the largest increases from 1990 to 2020 were in sub-Saharan Africa. These findings emphasize the need for targeted interventions, accounting for social inequities and aligned with global health objectives. While some policies to curb SSB intakes are currently in place in some countries, our study suggests that more work is needed. By highlighting the countries and subpopulations most affected, our research can assist in shaping effective policies and interventions to ultimately reduce the cardiometabolic health burdens of SSBs globally.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-024-03345-4>.

References

1. Malik, V. S. & Hu, F. B. The role of sugar-sweetened beverages in the global epidemics of obesity and chronic diseases. *Nat. Rev. Endocrinol.* **18**, 205–218 (2022).
2. Miller, V. et al. Evaluation of the quality of evidence of the association of foods and nutrients with cardiovascular disease and diabetes: a systematic review. *JAMA Netw. Open* **5**, e2146705 (2022).
3. Singh, G. M. et al. Estimated global, regional, and national disease burdens related to sugar-sweetened beverage consumption in 2010. *Circulation* **132**, 639–666 (2015).
4. Murray, C. J. L. et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1223–1249 (2020).
5. Beal, T., Herforth, A., Sundberg, S., Hess, S. Y. & Neufeld, L. M. Differences in modelled estimates of global dietary intake. *Lancet* **397**, 1708–1709 (2021).
6. Del Gobbo, L. C. et al. Assessing global dietary habits: a comparison of national estimates from the FAO and the Global Dietary Database. *Am. J. Clin. Nutr.* **101**, 1038–1046 (2015).
7. Miller, V. et al. Global Dietary Database 2017: data availability and gaps on 54 major foods, beverages and nutrients among 5.6 million children and adults from 1220 surveys worldwide. *BMJ Glob. Health* **6**, e003585 (2021).
8. Pan, A. & Hu, F. B. Effects of carbohydrates on satiety: differences between liquid and solid food. *Curr. Opin. Clin. Nutr. Metab. Care* **14**, 385–390 (2011).
9. Imamura, F. et al. Consumption of sugar sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: systematic review, meta-analysis, and estimation of population attributable fraction. *BMJ* **351**, h3576 (2015).
10. Khanna, D., Welch, B. & Rehman, A. *Pathophysiology of Obesity* (StatPearls Publishing, 2022).
11. Powell-Wiley, T. M. et al. Obesity and cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* **143**, e984–e1010 (2021).
12. Lara-Castor, L. et al. Sugar-sweetened beverage intakes among adults between 1990 and 2018 in 185 countries. *Nat. Commun.* **14**, 5957 (2023).
13. Global Nutrition and Policy Consortium. *Global Dietary Database* <https://globaldietarydatabase.org> (Tufts, 2020).
14. Murray, C. J. L., Ezzati, M., Lopez, A. D., Rodgers, A. & Vander Hoorn, S. Comparative quantification of health risks: conceptual framework and methodological issues. *Popul. Health Metr.* **1**, 1 (2003).
15. Singh, G. M. et al. The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. *PLoS ONE* **8**, e65174 (2013).
16. Lu, Y. et al. Metabolic mediators of the effects of body-mass index, overweight, and obesity on coronary heart disease and stroke: a pooled analysis of 97 prospective cohorts with 1.8 million participants. *Lancet* **383**, 970–983 (2014).
17. Mozaffarian, D., Hao, T., Rimm, E. B., Willett, W. C. & Hu, F. B. Changes in diet and lifestyle and long-term weight gain in women and men. *N. Engl. J. Med.* **364**, 2392–2404 (2011).
18. Global Burden of Disease Collaborative Network *Global Burden of Disease Study 2019 (GBD 2019) Socio-Demographic Index (SDI) 1950–2019* (Global Health Data Exchange, 2022).

19. Razo, C. et al. Effects of elevated systolic blood pressure on ischemic heart disease: a burden of proof study. *Nat. Med.* **28**, 2056–2065 (2022).
20. Ong, K. L. et al. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **402**, 203–234 (2023).
21. Erzse, A. et al. The data availability landscape in seven sub-Saharan African countries and its role in strengthening sugar-sweetened beverage taxation. *Glob. Health Action* **14**, 1871189 (2021).
22. Essman, M., Stoltze, F. M., Carpentier, F. D., Swart, E. C. & Taillie, L. S. Examining the news media reaction to a national sugary beverage tax in South Africa: a quantitative content analysis. *BMC Public Health* **21**, 454 (2021).
23. Audain, K., Levy, L. & Ellahi, B. Sugar-sweetened beverage consumption in the early years and implications for type-2 diabetes: a sub-Saharan Africa context. *Proc. Nutr. Soc.* **78**, 547–553 (2019).
24. Azzam, A. Is the world converging to a ‘Western diet’? *Public Health Nutr.* **24**, 309–317 (2021).
25. Ndlovu, N. & Swinburn, B. Readiness for sugar sweetened beverage taxation in sub-Saharan Africa. *Glob. Health Action* **14**, 1917801 (2021).
26. Stacey, N., van Walbeek, C., Maboshe, M., Tugendhaft, A. & Hofman, K. Energy drink consumption and marketing in South Africa. *Prev. Med.* **105**, S32–S36 (2017).
27. Boachie, M. K. et al. Beverage industry’s advertising expenditures and airtimes in South Africa from 2013 to 2019 target children and families. *J. Public Health Res.* **12**, 22799036231168207 (2023).
28. Bergallo, P., Castagnari, V., Fernández, A. & Mejía, R. Regulatory initiatives to reduce sugar-sweetened beverages (SSBs) in Latin America. *PLoS ONE* **13**, e0205694 (2018).
29. Melo, G. et al. Structural responses to the obesity epidemic in Latin America: what are the next steps for food and physical activity policies? *Lancet Reg. Health Am.* **21**, 100486 (2023).
30. (NCD-RisC), N.R.F.C. Worldwide trends in underweight and obesity from 1990 to 2022: a pooled analysis of 3663 population-representative studies with 222 million children, adolescents, and adults. *Lancet* **403**, 1027–1050 (2024).
31. O’Hearn, M. et al. Incident type 2 diabetes attributable to suboptimal diet in 184 countries. *Nat. Med.* **29**, 982–995 (2023).
32. Bhavadharini, B. et al. White rice intake and incident diabetes: a study of 132,373 participants in 21 countries. *Diabetes Care* **43**, 2643–2650 (2020).
33. Statista. Ready-to-drink (RDT) tea—Southeast Asia. *Statista Marketing Insights* <https://www.statista.com/outlook/cmo/non-alcoholic-drinks/ready-to-drink-rt-d-coffee-tea/ready-to-drink-rt-d-tea/southeast-asia> (2024).
34. Velasco Bernal, A., Hernández Chavez, P. & Aguilar Menéndez, P. *Estrategias de mercadotecnia de la industria de bebidas azucaradas* (Instituto Nacional de Salud Pública, 2020).
35. Lowery, C. M., Mora-Plazas, M., Gómez, L. F., Popkin, B. & Taillie, L. S. Reformulation of packaged foods and beverages in the Colombian food supply. *Nutrients* **12**, 3260 (2020).
36. Darsamo, V. & Walbeek, C. Effect of price and income on the demand for sugar-sweetened beverages in Nigeria: an analysis of household consumption data using an almost ideal demand system (AIDS). *BMJ Open* **13**, e072538 (2023).
37. Juma, F. BBC Nigeria: Investing in a new bottling line to grow Nigeria’s soft drinks industry. *Food Business Africa* (March 2021).
38. Phulkerd, S., Thongcharoenchupong, N., Chamrathirong, A., Soottipong Gray, R. & Prasertsom, P. Changes in population-level consumption of taxed and non-taxed sugar-sweetened beverages (SSB) after implementation of SSB excise tax in Thailand: a prospective cohort study. *Nutrients* **12**, 3294 (2020).
39. Jindarattanaporn, N., Kelly, B. & Phulkerd, S. A comparative analysis of unhealthy food and beverage television advertising to children in Thailand, between 2014 and 2022. *Global Health* **20**, 2 (2024).
40. Rogers, N. T. et al. Changes in soft drinks purchased by British households associated with the UK soft drinks industry levy: a controlled interrupted time series analysis. *BMJ Open* **13**, e077059 (2023).
41. Lauber, K. Framing marketing responses to national regulation: the four Ps in transnational corporate political discourse comment on “Understanding Marketing Responses to a Tax on Sugary Drinks: A Qualitative Interview Study in the United Kingdom, 2019”. *Int. J. Health Policy Manag.* **12**, 7618 (2023).
42. Lugo-Arias, J., Burgos-Vergara, J., Lugo-Arias, E., Gould, A. & Ovallos-Gazabon, D. Evaluation of low-cost alternatives for water purification in the stilt house villages of Santa Marta’s Ciénaga Grande. *Heliyon* **6**, e03062 (2020).
43. Smith, C. D., Jackson, K., Peters, H. & Herrera Lima, S. Lack of safe drinking water for Lake Chapala basin communities in Mexico inhibits progress toward sustainable development goals 3 and 6. *Int. J. Environ. Res. Public Health* **17**, 8328 (2020).
44. Sudsandee, S., Fakkaew, K., Siriratruengsuk, W., Worakhunpiset, S. & Apidechkul, T. Quality of sources of drinking water and health among the hill tribe people of northern Thailand. *Environ. Geochem. Health* **44**, 873–892 (2022).
45. Lopez, O. & Jabobs, A. In town with little water, Coca-Cola is everywhere. So is diabetes. *The New York Times* (14 Jul 2018).
46. UN General Assembly. *Transforming Our World: The 2030 Agenda for Sustainable Development* (UN, 2015).
47. Hammond, D. et al. The conceptual framework for the International Food Policy Study: evaluating the population-level impact of food policy. *J. Nutr.* **152**, 1s–12s (2022).
48. Elliott, L. M., Dalglis, S. L. & Topp, S. M. Health taxes on tobacco, alcohol, food and drinks in low- and middle-income countries: a scoping review of policy content, actors, process and context. *Int. J. Health Policy Manag.* **11**, 414–428 (2022).
49. Hansen, K. L., Golubovic, S., Eriksen, C. U., Jørgensen, T. & Toft, U. Effectiveness of food environment policies in improving population diets: a review of systematic reviews. *Eur. J. Clin. Nutr.* **76**, 637–646 (2022).
50. Popkin, B. M. & Ng, S. W. Sugar-sweetened beverage taxes: lessons to date and the future of taxation. *PLoS Med.* **18**, e1003412 (2021).
51. Hattersley, L. & Mandeville, K. L. Global coverage and design of sugar-sweetened beverage taxes. *JAMA Netw. Open* **6**, e231412 (2023).
52. Sun, H. et al. Global disease burden attributed to high sugar-sweetened beverages in 204 countries and territories from 1990 to 2019. *Prev. Med.* **175**, 107690 (2023).
53. Dom, R., Custers, A., Davenport, S. & Prichard, W. *Innovations in Tax Compliance: Building Trust, Navigating Politics, and Tailoring Reform* (World Bank, 2022).
54. Cha, S., Mankadi, P. M., Elhag, M. S., Lee, Y. & Jin, Y. Trends of improved water and sanitation coverage around the globe between 1990 and 2010: inequality among countries and performance of official development assistance. *Glob. Health Action* **10**, 1327170 (2017).
55. von Philipsborn, P. et al. Environmental interventions to reduce the consumption of sugar-sweetened beverages and their effects on health. *Cochrane Database Syst. Rev.* **6**, CD012292 (2019).
56. *The CDC Guide to Strategies for Reducing the Consumption of Sugar-Sweetened Beverages* (Centers for Disease Control and Prevention (CDC), 2010).
57. Keller, K. O. & Guyt, J. Y. A war on sugar? Effects of reduced sugar content and package size in the soda category. *J. Mark.* **87**, 698–718 (2023).

58. Macdonald, C. Soft drink industry beats EU sugar reduction targets—but challenges remain. *Global Insights* (15 Nov 2021).
59. Suez, J. et al. Personalized microbiome-driven effects of non-nutritive sweeteners on human glucose tolerance. *Cell* **185**, 3307–3328.e3319 (2022).
60. del Pozo, S. et al. Potential effects of sucralose and saccharin on gut microbiota: a review. *Nutrients* **14**, (2022).
61. Gauthier, E., Milagro, F. I. & Navas-Carretero, S. Effect of low- and non-calorie sweeteners on the gut microbiota: a review of clinical trials and cross-sectional studies. *Nutrition* **117**, 112237 (2024).
62. Shi, D. et al. The burden of cardiovascular diseases attributed to diet high in sugar-sweetened beverages in 204 countries and territories from 1990 to 2019. *Curr. Probl. Cardiol.* **49**, 102043 (2024).
63. *GBD Compare Data Visualization* (IHME, University of Washington, 2024); <http://vizhub.healthdata.org/gbd-compare>
64. *Global Burden of Disease Study 2021 (GBD 2021) Sources Tool* (IHME, University of Washington, 2024); <https://ghdx.healthdata.org/gbd-2021/sources?components=25&risks=118&locations=1>
65. Brauer, M. et al. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **403**, 2162–2203 (2024).
66. Willett, W. *Nutritional Epidemiology* (Oxford Scholarship Online, 2012).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025

¹Food Is Medicine Institute, Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA, USA. ²Institute of Health Metrics and Evaluation, University of Washington, Seattle, WA, USA. ³Food Systems for the Future Institute, Chicago, IL, USA. ⁴Department of Medicine, McMaster University, Hamilton, Ontario, Canada. ⁵Population Health Research Institute, Hamilton, Ontario, Canada. ⁶Center for Surgery and Public Health, Brigham and Women's Hospital, Boston, MA, USA. ⁷Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA, USA. ⁸Research Center on Nutrition and Health, National Institute of Public Health, Cuernavaca, Mexico. ⁹University of Thessaly, Volos, Greece. ¹⁰Tufts University School of Medicine, Boston, MA, USA. ¹¹Department of Medicine, Tufts Medical Center, Boston, MA, USA. *A list of authors and their affiliations appears at the end of the paper.

✉e-mail: lauralac@uw.edu; dariush.mozaffarian@tufts.edu

Global Dietary Database

Antonia Trichopoulou¹², Murat Bas¹³, Jemal Haidar Ali¹⁴, Tatyana El-Kour¹⁵, Anand Krishnan¹⁶, Puneet Misra¹⁶, Nahla Hwalla¹⁷, Chandrashekar Janakiram¹⁸, Nur Indrawaty Lipoeto¹⁹, Abdulrahman Musaiger²⁰, Farhad Pourfarzi²¹, Iftikhar Alam²², Celine Termote²³, Anjum Memon²⁴, Marieke Vossenaar²⁵, Paramita Mazumdar²⁶, Ingrid Rached²⁷, Alicia Roviroso²⁸, María Elisa Zapata²⁸, Roya Kelishadi²⁹, Tamene Taye Asayehu³⁰, Francis Oduor³¹, Julia Boedecker³¹, Lilian Aluso³¹, Emanuele Marconi³², Laura D'Addezio³², Raffaella Piccinelli³², Stefania Sette³², Johana Ortiz-Ulloa³³, J. V. Meenakshi³⁴, Giuseppe Grosso³⁵, Anna Waskiewicz³⁶, Umber S. Khan³⁷, Kenneth Brown³⁸, Lene Frost Andersen³⁹, Anastasia Thanopoulou⁴⁰, Reza Malekzadeh⁴¹, Neville Calleja⁴², Anca Ioana Nicolau⁴³, Cornelia Tudorie⁴³, Marga Ocke⁴⁴, Zohreh Etemad⁴⁴, Mohannad Al Nsour⁴⁵, Lydiah M. Waswa⁴⁶, Maryam Hashemian⁴⁷, Eha Nurk⁴⁸, Joanne Arsenault⁴⁹, Patricio Lopez-Jaramillo⁵⁰, Abba Mehio Sibai⁵¹, Albertino Damasceno⁵², Pulani Lanerolle⁵³, Carukshi Arambepola⁵³, Carla Lopes⁵⁴, Milton Severo⁵⁴, Nuno Lunet⁵⁴, Duarte Torres⁵⁵, Heli Tapanainen⁵⁶, Jaana Lindstrom⁵⁶, Suvi Virtanen⁵⁶, Cristina Palacios⁵⁷, Noel Barengo⁵⁷, Eva Roos⁵⁸, Irmgard Jordan⁵⁹, Charmaine Duante⁶⁰, Corazon Cerdena⁶⁰, Imelda Angeles-Agdeppa⁶⁰, Josie Desnacio⁶⁰, Mario Capanzana⁶¹, Anoop Misra^{62,63}, Ilse Khouw⁶⁴, Swee Ai Ng⁶⁴, Edna Gamboa Delgado⁶⁵, Mauricio T. Caballero⁶⁶, Johanna Otero⁶⁷, Hae-Jeung Lee⁶⁸, Eda Koksali⁶⁹, Idris Guessous⁷⁰, Carl Lachat⁷¹, Stefaan De Henauw⁷¹, Ali Reza Rahbar⁷², Alison Tedstone⁷², Annie Ling⁷², Beth Hopping⁷², Catherine Leclercq⁷², Christian Haerpfner⁷², Christine Hotz⁷², Christos Pitsavos⁷², Coline van Oosterhout⁷², Debbie Bradshaw⁷², Dimitrios Trichopoulos⁷², Dorothy Gauci⁷², Dulitha Fernando⁷², Elzbieta Sygnowska⁷², Erkki Vartiainen⁷², Farshad Farzadfar⁷², Gabor Zajkas⁷², Gillian Swan⁷², Guansheng Ma⁷², Hajah Masni Ibrahim⁷², Harri Sinkko⁷², Isabelle Sioen⁷², Jean-Michel Gaspoz⁷², Jillian Odenkirk⁷², Kanitta Bundhamcharoen⁷², Keiu Nelis⁷², Khairul Zarina⁷², Lajos Biro⁷², Lars Johansson⁷², Leanne Riley⁷², Mabel Yap⁷², Manami Inoue⁷², Maria Szabo⁷², Marja-Leena Ovaskainen⁷², Meei-Shyuan Lee⁷², Mei Fen Chan⁷², Melanie Cowan⁷², Mirnalini Kandiah⁷², Ola Kally⁷², Olof Jonsdottir⁷², Pam Palmer⁷², Philippos Orfanos⁷², Renzo Asciac⁷², Robert Templeton⁷², Rokiah Don⁷², Roseyati Yaakub⁷², Rusidah Selamat⁷², Safiah Yusof⁷², Sameer Al-Zenki⁷², Shu-Yi Hung⁷², Sigrid Beer-Borst⁷², Suh Wu⁷², Widjaja Lukito⁷², Wilbur Hadden⁷², Xia Cao⁷², Yi Ma⁷², Yuen Lai⁷², Zaiton Hjadaud⁷², Jennifer Ali⁷³, Ron Gravel⁷³, Tina Tao⁷³, Jacob Lennert Veerman⁷⁴, Mustafa Arici⁷⁵, Demosthenes Panagiotakos⁷⁶, Yanping Li⁷⁷, Gulden Pekcan⁷⁸, Karim Anzid⁷⁹, Anuradha Khadilkar⁸⁰, Veena Ekbote⁸⁰, Irina Kovalskys⁸¹, Arlappa Nimmathota⁸², Avula Laxmaiah⁸², Balakrishna Nagalla⁸², Brahmam Ginnela⁸², Hemalatha Rajkumar⁸², Indrapal Meshram⁸², Kalpagam Polasa⁸², Licia Iacoviello^{83,84}, Marialaura Bonaccio⁸³, Simona Costanzo⁸³, Yves Martin-Prevel⁸⁵, Nattinee Jitnarin⁸⁶, Wen-Harn Pan⁸⁷, Yao-Te Hsieh⁸⁷, Sonia Olivares⁸⁸, Gabriela Tejada⁸⁹, Aida Hadziomeragic⁹⁰, Le Tran Ngoan^{91,92}, Amanda de Moura Souza⁹³, Daniel Illescas-Zarate^{94,95}, Inge Huybrechts⁹⁶, Alan de Brauw⁹⁷, Mourad Moursi⁹⁷, Augustin Nawidimbassa Zeba⁹⁸, Maryam Maghroun⁹⁹, Nizal Sarrafzadegan⁹⁹, Noushin Mohammadi⁹⁹, Lital Keinan-Boker¹⁰⁰, Rebecca Goldsmith¹⁰⁰, Tal Shimony¹⁰⁰, Gudrun B. Keding¹⁰¹, Shivanand C. Mastiholi¹⁰², Moses Mwangi¹⁰³, Yeri Kombe¹⁰³, Zipporah Bukania¹⁰³, Eman Alissa¹⁰⁴, Nasser Al-Daghri¹⁰⁵, Shaun Sabico¹⁰⁵, Rajesh Jeewon^{105,106}, Martin Gulliford¹⁰⁷, Tshilenge S. Diba¹⁰⁸, Kyungwon Oh¹⁰⁹, Sihyun Park¹⁰⁹, Sungha Yun¹⁰⁹,

Yoonsu Cho¹¹⁰, Suad Al-Hooti¹¹¹, Chanthaly Luangphaxay¹¹², Daovieng Douangvichit¹¹², Latsamy Siengsounthone¹¹², Pedro Marques-Vidal¹¹³, Peter Vollenweider¹¹³, Constance Rybak¹¹⁴, Amy Luke¹¹⁵, Nipa Rojroongwasinkul¹¹⁶, Noppawan Piaseu¹¹⁶, Kalyana Sundram¹¹⁷, Jeremy Koster¹¹⁸, Donka Baykova¹¹⁹, Parvin Abedi¹²⁰, Sandjaja Sandjaja¹²¹, Fariza Fadzil¹²², Norikilil Bukhary Ismail Bukhary¹²³, Pascal Bovet^{124,125}, Yu Chen¹²⁶, Norie Sawada¹²⁷, Shoichiro Tsugane¹²⁷, Lalka Rangelova¹²⁸, Stefka Petrova¹²⁸, Vesselka Duleva¹²⁸, Ward Siamusantu¹²⁹, Lucjan Szponar¹³⁰, Hsing-Yi Chang¹³¹, Makiko Sekiyama¹³², Khanh Le Nguyen Bao¹³³, Sesikeran Boindala¹³⁴, Jalila El Ati¹³⁵, Ivonne Ramirez Silva¹³⁶, Juan Rivera Dommarco¹³⁶, Luz Maria Sanchez-Romero¹³⁶, Simon Barquera¹³⁶, Sonia Rodríguez-Ramírez¹³⁶, Nayu Ikeda¹³⁷, Sahar Zaghloul¹³⁸, Anahita Houshiar-rad¹³⁹, Fatemeh Mohammadi-Nasrabadi¹³⁹, Morteza Abdollahi¹³⁹, Khun-Aik Chuah¹⁴⁰, Zaleha Abdullah Mahdy¹⁴⁰, Alison Eldridge¹⁴¹, Eric L. Ding¹⁴², Herculina Kruger¹⁴³, Sigrun Henjum¹⁴⁴, Milton Fabian Suarez-Ortegon¹⁴⁵, Nawal Al-Hamad¹⁴⁶, Veronika Janská¹⁴⁷, Reema Tayyem¹⁴⁸, Bemnet Tedla¹⁴⁹, Parvin Mirmiran¹⁵⁰, Almut Richter¹⁵¹, Gert Mensink¹⁵¹, Lothar Wieler¹⁵¹, Daniel Hoffman¹⁵², Benoit Salanave¹⁵³, Shashi Chiplonkar¹⁵⁴, Anne Fernandez¹⁵⁵, Androniki Naska¹⁵⁶, Karin De Ridder¹⁵⁷, Cho-il Kim¹⁵⁸, Rebecca Kuriyan¹⁵⁹, Sumathi Swaminathan¹⁵⁹, Didier Garriguet¹⁶⁰, Anna Karin Lindroos¹⁶¹, Eva Warensjo Lemming¹⁶¹, Jessica Petrelius Sipinen¹⁶¹, Lotta Moraeus¹⁶¹, Saeed Dastgiri¹⁶², Sirje Vaask¹⁶³, Tilakavati Karupaiah¹⁶⁴, Fatemeh Vida Zohoori¹⁶⁵, Alireza Esteghamati¹⁶⁶, Sina Noshad¹⁶⁶, Suhad Abumweis¹⁶⁷, Elizabeth Mwaniki¹⁶⁸, Simon G. Anderson¹⁶⁹, Justin Chileshe¹⁷⁰, Sydney Mwanza¹⁷⁰, Lydia Lera Marques¹⁷¹, Samuel Duran Aguero¹⁷², Mariana Oleas¹⁷³, Luz Posada¹⁷⁴, Angelica Ochoa¹⁷⁵, Alan Martin Preston¹⁷⁶, Khadijah Shamsuddin¹⁷⁷, Zalilah Mohd Shariff¹⁷⁸, Hamid Jan Bin Jan Mohamed¹⁷⁹, Wan Manan¹⁷⁹, Bee Koon Poh¹⁸⁰, Pamela Abbott¹⁸¹, Mohammadreza Pakseresht¹⁸², Sangita Sharma¹⁸², Tor Strand¹⁸³, Ute Alexy¹⁸⁴, Ute Nöthlings¹⁸⁴, Indu Waidyatilaka¹⁸⁵, Ranil Jayawardena¹⁸⁵, Julie M. Long¹⁸⁶, K. Michael Hambidge^{186,206}, Nancy F. Krebs¹⁸⁶, Aminul Haque¹⁸⁷, Liisa Korkalo¹⁸⁸, Majjaliisa Erkkola¹⁸⁸, Riitta Freese¹⁸⁸, Laila Eleraky¹⁸⁹, Wolfgang Stuetz¹⁸⁹, Laufey Steingrimsdottir¹⁹⁰, Inga Thorsdottir¹⁹⁰, Ingibjorg Gunnarsdottir¹⁹⁰, Lluís Serra-Majem¹⁹¹, Foong Ming Moy¹⁹², Corina Aurelia Zugravu¹⁹³, Elizabeth Yakes Jimenez¹⁹⁴, Linda Adair¹⁹⁵, Shu Wen Ng¹⁹⁵, Sheila Skeaff¹⁹⁶, Regina Fisberg¹⁹⁷, Carol Henry¹⁹⁸, Getahun Ersino¹⁹⁸, Gordon Zello¹⁹⁸, Alexa Meyer¹⁹⁹, Ibrahim Elmadfa¹⁹⁹, Claudette Mitchell²⁰⁰, David Balfour²⁰⁰, Johanna M. Geleijnse²⁰¹, Mark Manary²⁰², Laetitia Nikiema²⁰³, Masoud Mirzaei²⁰⁴ & Rubina Hakeem²⁰⁵

¹²Academy of Athens, Athens, Greece. ¹³Acibadem University, Istanbul, Turkey. ¹⁴Addis Ababa University, Addis Ababa, Ethiopia. ¹⁵Aga Khan Foundation, Geneva, Switzerland. ¹⁶All India Institute of Medical Sciences, New Delhi, India. ¹⁷American University of Beirut, Beirut, Lebanon. ¹⁸Amrita School of Dentistry, Eranakulam, India. ¹⁹Andalas University, Padang, Indonesia. ²⁰Arab Center for Nutrition, Manama, Bahrain. ²¹Ardabil University of Medical Sciences, Ardabil, Iran. ²²Bacha Khan University, Charsadda, Pakistan. ²³Biodiversity International, Nairobi, Kenya. ²⁴Brighton and Sussex Medical School, Brighton, UK. ²⁵Center for Studies of Sensory Impairment Aging and Metabolism (CeSIAM), Guatemala City, Guatemala. ²⁶Centre For Media Studies, New Delhi, India. ²⁷Centro de Atención Nutricional Antimano (CANIA), Miami, FL, USA. ²⁸Centro de Estudios sobre Nutrición Infantil (CESNI), Buenos Aires, Argentina. ²⁹Child Growth and Development Research Center, Isfahan University of Medical Sciences, Isfahan, Iran. ³⁰College of Applied Sciences, Department of Food Science and Applied Nutrition, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia. ³¹Consultative Group on International Agricultural Research (CGIAR), Montpellier, France. ³²Council for Agricultural Research and Economics, Research Centre for Food and Nutrition, Rome, Italy. ³³Cuenca University, Cuenca, Ecuador. ³⁴Delhi School of Economics, University of Delhi, Delhi, India. ³⁵Department of Biomedical and Biotechnological Sciences, University of Catania, Catania, Italy. ³⁶Department of CVD Epidemiology, Prevention and Health Promotion, Institute of Cardiology, Warsaw, Poland. ³⁷Department of Community Health Sciences, Aga Khan University, Karachi, Pakistan. ³⁸Department of Nutrition and Institute for Global Nutrition, University of California Davis, Davis, CA, USA. ³⁹Department of Nutrition, University of Oslo, Oslo, Norway. ⁴⁰Diabetes Center, 2nd Department of Internal Medicine, Athens University, Athens, Greece. ⁴¹Digestive Disease Research Institute, Tehran University of Medical Sciences, Tehran, Iran. ⁴²Directorate for Health Information and Research, Pietà, Malta. ⁴³Dunarea de Jos University of Galati, Galati, Romania. ⁴⁴Dutch National Institute for Public Health and the Environment (RIVM), Bilthoven, Netherlands. ⁴⁵Eastern Mediterranean Public Health Network (EMPHNET), Amman, Jordan. ⁴⁶Egerton University, Njoro, Kenya. ⁴⁷Epidemiology and Community Health Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA. ⁴⁸Estonian National Institute for Health Development, Tallinn, Estonia. ⁴⁹FHI360, Washington DC, USA. ⁵⁰FOSCAL and UDES, Bucaramanga, Colombia. ⁵¹Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon. ⁵²Faculty of Medicine, Eduardo Mondlane University, Maputo, Mozambique. ⁵³Faculty of Medicine, University of Colombo, Colombo, Sri Lanka. ⁵⁴Faculty of Medicine/ Institute of Public Health, University of Porto, Porto, Portugal. ⁵⁵Faculty of Nutrition and Food Sciences, University of Porto, Porto, Portugal. ⁵⁶Finnish Institute for Health and Welfare, Helsinki, Finland. ⁵⁷Florida International University, Miami, FL, USA. ⁵⁸Folkhälsan Research Center, Helsinki, Finland. ⁵⁹Food Environment Consumer Behaviour Lever of the Alliance Bioversity International and CIAT, Nairobi, Kenya. ⁶⁰Food and Nutrition Research Institute (DOST-FNRI), Manila, Philippines. ⁶¹Food and Nutrition Research Institute (DOST-FNRI), Taguig City, Philippines. ⁶²Fortis CDOC Center for Excellence for Diabetes, New Delhi, India. ⁶³National Diabetes, Obesity and Cholesterol Foundation (N-DOC), Diabetes Foundation (India), Fortis CDOC Hospital for Diabetes and Allied Sciences, New Delhi, India. ⁶⁴FrieslandCampina, Amersfoort, The Netherlands. ⁶⁵Fundacion Cardiovascular de Colombia, Bucaramanga, Colombia. ⁶⁶Fundacion INFANT and Consejo Nacional De Investigaciones Cientificas y Tecnicas (CONICET), Buenos Aires, Argentina. ⁶⁷Fundacion Oftalmologica de Santander (FOSCAL), Floridablanca, Colombia. ⁶⁸Gachon University, Seongnam-si, South Korea. ⁶⁹Gazi University, Ankara, Turkey. ⁷⁰Geneva University Hospitals, Geneva, Switzerland. ⁷¹Ghent University, Ghent, Belgium. ⁷²Global Dietary Database Consortium, Boston, MA, USA. ⁷³Statistics Canada, Government of Canada, Ottawa, Ontario, Canada. ⁷⁴Griffith University, Gold Coast, Queensland, Australia. ⁷⁵Faculty of Medicine, Hacettepe University, Ankara, Turkey. ⁷⁶Harokopio University, Athens, Greece. ⁷⁷Harvard School of Public Health, Cambridge, MA, USA. ⁷⁸Department of Nutrition and Dietetics, Hasan Kalyoncu University, Gaziantep, Turkey. ⁷⁹Higher Institute of Nursing Professions and Health Techniques, Marrakesh, Morocco. ⁸⁰Hirabai Cowasji Jehangir Medical Research Institute, Pune, India. ⁸¹ICCAS (Instituto para la Cooperacion Cientifica en Ambiente y Salud), Buenos Aires, Argentina. ⁸²ICMR-National Institute of Nutrition, Hyderabad, India. ⁸³IRCCS Neuromed, Pozzilli, Italy. ⁸⁴LUM University "Giuseppe Degennaro", Casamassima, Italy. ⁸⁵Institut de Recherche pour le Développement, Montpellier, France. ⁸⁶Institute for International Investigation, NDRI-USA, New York, NY, USA. ⁸⁷Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. ⁸⁸Institute of Nutrition and Food Technology (INTA), University of Chile, Santiago, Chile. ⁸⁹Institute of Nutrition in Central America and Panama (INCAP), Guatemala City, Guatemala. ⁹⁰Institute of Public Health of Federation of Bosnia and Herzegovina, Sarajevo, Bosnia and Herzegovina. ⁹¹Institute of Research and Development, Duy Tan University, Da Nang City, Vietnam. ⁹²School of Preventive Medicine and Public Health, Hanoi City, Vietnam. ⁹³Institute of Studies in Public Health, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil. ⁹⁴Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico. ⁹⁵Tecnologico de Monterrey, , Escuela de Medicina y Ciencias de la Salud, Mexico

City, Mexico. ⁹⁶International Agency for Research on Cancer, Lyon, France. ⁹⁷International Food Policy Research Institute (IFPRI), Washington DC, USA. ⁹⁸Intitut de Recherche en Sciences de la Sante, Bobo Dioulasso, Burkina Faso. ⁹⁹Isfahan Cardiovascular Research Center, Cardiovascular Research Institute, Isfahan University of Medical Sciences, Isfahan, Iran. ¹⁰⁰Israel Center for Disease Control, Tel-Hashomer, Israel. ¹⁰¹Institute of Nutritional Sciences, Justus Liebig University of Giessen, Giessen, Germany. ¹⁰²KLE Academy of Higher Education and Research (Deemed-to-be-University) Jawaharlal Nehru Medical College, Belagavi, India. ¹⁰³Kenya Medical Research Institute, Nairobi, Kenya. ¹⁰⁴King Abdulaziz University, Jeddah, Saudi Arabia. ¹⁰⁵King Saud University, Riyadh, Saudi Arabia. ¹⁰⁶University of Mauritius, Reduit, Mauritius. ¹⁰⁷King's College London, London, UK. ¹⁰⁸Kinshasa School of Public Health, Kinshasa, Democratic Republic of Congo. ¹⁰⁹Korea Disease Control and Prevention Agency (KDCA), Cheongju-si, South Korea. ¹¹⁰Korea University, Seoul, South Korea. ¹¹¹Kuwait Institute for Scientific Research, Safat, Kuwait. ¹¹²Lao Tropical and Public Health Institute, Vientiane, Lao People's Democratic Republic. ¹¹³Lausanne University Hospital (CHUV) and University of Lausanne, Lausanne, Switzerland. ¹¹⁴Leibniz Centre for Agricultural Landscape Research, Müncheberg, Germany. ¹¹⁵Loyola University Chicago, Chicago, IL, USA. ¹¹⁶Mahidol University, Pathom, Thailand. ¹¹⁷Malaysian Palm Oil Council (MPOC), Kelana Jaya, Malaysia. ¹¹⁸Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. ¹¹⁹Medical Center Markovs, Sofia, Bulgaria. ¹²⁰Menopause Andropause Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran. ¹²¹Ministry of Health, Jakarta, Indonesia. ¹²²Ministry of Health, Kuala Lumpur, Malaysia. ¹²³Ministry of Health, Sungai Besar, Malaysia. ¹²⁴Ministry of Health, Victoria, Seychelles. ¹²⁵University Center for Primary Care and Public Health (Unisanté), Lausanne, Switzerland. ¹²⁶NYU School of Medicine, New York, NY, USA. ¹²⁷National Cancer Center Institute for Cancer Control, Tokyo, Japan. ¹²⁸National Centre of Public Health and Analyses (NCPHA), Sofia, Bulgaria. ¹²⁹National Food and Nutrition Commission, Lusaka, Zambia. ¹³⁰National Food and Nutrition Institute, Warsaw, Poland. ¹³¹National Health Research Institutes, Zhunan Township, Taiwan, Republic of China. ¹³²Health and Environmental Risk Division, National Institute for Environmental Studies, Tsukuba, Japan. ¹³³National Institute of Nutrition, Hanoi, Vietnam. ¹³⁴National Institute of Nutrition, Hyderabad, India. ¹³⁵National Institute of Nutrition and Food Technology & SURVEN RL, Tunis, Tunisia. ¹³⁶National Institute of Public Health (INSP), Cuernavaca, Mexico. ¹³⁷National Institutes of Biomedical Innovation Health and Nutrition, Osaka, Japan. ¹³⁸National Nutrition Institute, Cairo, Egypt. ¹³⁹National Nutrition and Food Technology Research Institute (NNFTRI): SBMU, Tehran, Iran. ¹⁴⁰National University of Malaysia (UKM), Kuala Lumpur, Malaysia. ¹⁴¹Nestlé Research, Lausanne, Switzerland. ¹⁴²New England Complex Systems Institute, Cambridge, MA, USA. ¹⁴³North-West University, Potchefstroom South Africa, Potchefstroom, South Africa. ¹⁴⁴Oslo Metropolitan University (OsloMet), Oslo, Norway. ¹⁴⁵Pontificia Universidad Javeriana Seccional Cali, Cali, Colombia. ¹⁴⁶Public Authority For Food and Nutrition, Sabah Al Salem, Kuwait. ¹⁴⁷Public Health Authority of the Slovak Republic, Bratislava, Slovak Republic. ¹⁴⁸Qatar University and University of Jordan, Doha, Qatar. ¹⁴⁹Queensland University of Technology, Brisbane, Queensland, Australia. ¹⁵⁰Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ¹⁵¹Robert Koch Institute, Berlin, Germany. ¹⁵²Rutgers University, New Brunswick, NJ, USA. ¹⁵³Santé publique France, the French Public Health Agency, Saint Maurice, France. ¹⁵⁴Savitribai Phule Pune University (SPPU), Pune, India. ¹⁵⁵School of Medicine, Deakin University, Geelong, Victoria, Australia. ¹⁵⁶School of Medicine, National and Kapodistrian University of Athens, Athens, Greece. ¹⁵⁷Sciensano (Belgian Public Health Institute), Brussels, Belgium. ¹⁵⁸Seoul National University, Seoul, South Korea. ¹⁵⁹St John's Research Institute, Bangalore, India. ¹⁶⁰Statistics Canada, Ottawa, Ontario, Canada. ¹⁶¹Swedish Food Agency, Uppsala, Sweden. ¹⁶²Tabriz University of Medical Sciences, Tabriz, Iran. ¹⁶³Tallinn University, Tallinn, Estonia. ¹⁶⁴Taylor's University, Subang Jaya, Malaysia. ¹⁶⁵Teesside University, Middlesbrough, UK. ¹⁶⁶Tehran University of Medical Sciences, Tehran, Iran. ¹⁶⁷The Hashemite University, Az Zarqa, Jordan. ¹⁶⁸The Technical University of Kenya, Nairobi, Kenya. ¹⁶⁹The University of the West Indies, Kingston, Jamaica. ¹⁷⁰Tropical Diseases Research Centre, Ndola, Zambia. ¹⁷¹Unidad de Nutricion Publica, Macul, Chile. ¹⁷²Universidad San Sebastian, Providencia, Chile. ¹⁷³Universidad Tecnica del Norte, Ibarra, Ecuador. ¹⁷⁴Universidad de Antioquia, Medellin, Colombia. ¹⁷⁵Universidad de Cuenca, Cuenca, Ecuador. ¹⁷⁶Medical Sciences Department of Biochemistry, Universidad de Puerto Rico, San Juan, Puerto Rico. ¹⁷⁷Universiti Kebangsaan Malaysia Medical Centre, Kuala Lumpur, Malaysia. ¹⁷⁸Universiti Putra Malaysia, Serdang, Malaysia. ¹⁷⁹Universiti Sains Malaysia, Kubang Kerian, Malaysia. ¹⁸⁰University Kebangsaan Malaysia, Bangi, Malaysia. ¹⁸¹University of Aberdeen, Aberdeen, UK. ¹⁸²University of Alberta, Edmonton, Alberta, Canada. ¹⁸³University of Bergen, Bergen, Norway. ¹⁸⁴Department of Nutrition and Food Sciences, University of Bonn, Bonn, Germany. ¹⁸⁵University of Colombo, Colombo, Sri Lanka. ¹⁸⁶University of Colorado School of Medicine, Aurora, CO, USA. ¹⁸⁷University of Dhaka, Dhaka, Bangladesh. ¹⁸⁸Department of Food and Nutrition, University of Helsinki, Helsinki, Finland. ¹⁸⁹University of Hohenheim, Stuttgart, Germany. ¹⁹⁰University of Iceland, Reykjavik, Iceland. ¹⁹¹University of Las Palmas de Gran Canaria (ULPGC), Las Palmas, Spain. ¹⁹²University of Malaya, Kuala Lumpur, Malaysia. ¹⁹³University of Medicine and Pharmacy Carol Davila, Bucharest, Romania. ¹⁹⁴University of New Mexico Health Sciences Center, Albuquerque, NM, USA. ¹⁹⁵University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ¹⁹⁶University of Otago, Dunedin, New Zealand. ¹⁹⁷University of Sao Paulo, Sao Paulo, Brazil. ¹⁹⁸University of Saskatchewan, Saskatoon, Saskatchewan, Canada. ¹⁹⁹University of Vienna, Vienna, Austria. ²⁰⁰University of the Southern Caribbean, Port of Spain, Trinidad and Tobago. ²⁰¹Wageningen University, Wageningen, Netherlands. ²⁰²Washington University in St. Louis, St. Louis, MO, USA. ²⁰³World Health Organization (WHO), Geneva, Switzerland. ²⁰⁴Yazd Cardiovascular Research Centre, Shahid Sadoughi University of Medical Sciences, Yazd, Iran. ²⁰⁵Ziauddin University Karachi, Karachi City, Pakistan. ²⁰⁶Deceased: K. M. Hambidge.

Methods

Inclusion and ethics statement

Data informing the GDD modeling estimates for this study, including from LMICs (low- and middle-income countries), were collected between 1980 and 2018 from GDD consortium members and publicly available sources in the form of dietary intake surveys. If nationally representative surveys were not available for a country, we also considered national surveys without representative sampling, followed by regional, urban or rural surveys, and finally large local cohorts, provided that selection and measurement biases were not apparent limitations (for example, excluding studies focused on a selected population with a specific disease, a certain profession or following a particular dietary pattern). For countries with no surveys identified, other sources of potential data were considered, including the WHO Infobase, the STEP database and household budget survey data. As of August 2021, we identified and retrieved 1,634 eligible survey years of data from public and private sources. Of these, 1,224 were checked, standardized and included in the GDD model, including 450 surveys informing SSB intake estimates¹².

Most surveys identified were privately held or, if public, not available in relevant format for GDD modeling (for example, not jointly stratified by age, sex, education, and urban or rural status). We thus relied almost entirely on direct consortium member contacts for each survey to provide us with exposure data directly. Roles and responsibilities of GDD consortium members were determined and agreed upon before data sharing as part of a standardized data sharing agreement. The draft manuscript was shared with all GDD consortium members before submission for peer review, and all members are included as coauthors of this work. We endorse the Nature Portfolio journals' guidance on LMIC authorship and inclusion and are committed to the inclusion of researchers from LMICs in publications from the GDD. We share the GDD data with the entire consortium, encourage authors from LMICs to take the lead on analyses and papers, and can provide technical and writing support to LMIC authors. For more details on the collaborative GDD data collection process, please visit our website at <https://www.globaldietarydatabase.org/methods/summary-methods-and-data-collection>.

This research is locally relevant to all countries included, given that it disaggregates findings nationally and subnationally by key demographic factors such as age, sex, education level and urbanicity, providing decision-makers with the CVD and diabetes risk associated with SSB intakes over time.

This modeling investigation was exempt from ethical review board approval because it was based on published data and nationally representative, de-identified datasets, without personally identifiable information. Individual surveys underwent ethical review board approval required for the applicable local context.

Study design

A CRA model¹⁴ estimated the numbers, proportions and uncertainty of global T2D and CVD incidence, DALYs and mortality attributable to intake of SSBs among adults aged 20+ years. Importantly, the CRA framework does not use ecologic correlations to estimate risk, but incorporates independently derived input parameters and their uncertainties on sociodemographics, population size, risk factor (that is, SSBs) their multivariable-adjusted estimated etiologic effects on disease based on external studies, and background disease incidence, mortality and DALYs¹⁴. These parameters are entered into the model to estimate the disease burdens and their uncertainties. Specifically for this investigation, we leveraged input data and corresponding uncertainty in 184 countries including (1) population SSB intake distributions based on individual-level survey data from the GDD (<https://www.globaldietarydatabase.org/>)^{7,12,13}; (2) optimal SSB intake levels from previous analyses⁶⁷; (3) direct age-adjusted etiologic effects of SSBs on T2D, ischemic heart disease and ischemic stroke adjusted for BMI^{2,68–70},

and of weight gain on T2D¹⁵, ischemic heart disease¹⁶ and ischemic stroke¹⁵ from previous meta-analyses and pooled analyses of prospective cohorts, as well as linear, BMI-stratified effects of SSBs on weight gain or loss¹⁷; (4) population overweight (BMI ≥ 25 kg m⁻²) and underweight (BMI < 18.5 kg m⁻²) distributions from the (non-communicable disease) NCD Risk Factor Collaboration (NCD-RisC)⁷¹; (5) total T2D, ischemic heart disease, and ischemic stroke incidence, DALYs and mortality estimate distributions from the GBD study^{72,73}; and (6) population demographic data from the United Nations Population Division^{74,75} and the Barro and Lee Educational Attainment Dataset 2013⁷⁶, as previously reported³¹ (Supplementary Table 7).

Bias and reliability were addressed in each of the independent data sources used in our model. The GDD selected national and subnational dietary surveys without apparent measurement or selection biases⁷, and leveraged a Bayesian model to incorporate differences in data comparability and sampling uncertainty. In GBD, bias adjustment of underlying rates of T2D and CVD not specifically meeting the gold-standard definition of these causes was done using network meta-regression before estimation in DisMod, while implausible or unspecified causes of death were redistributed to valid underlying causes of death using reclassification algorithms^{73,77}. Etiologic effects were obtained from published meta-analyses or pooled analyses of prospective cohorts and randomized control trials including multivariable adjustment for age, sex, BMI and other risk factors to reduce bias from confounding^{2,68–70}. Studies with increased risk of bias such as retrospective or cross-sectional studies were excluded². Underlying adiposity rates were obtained from the NCD-RisC, which used national or subnational surveys that collected measured height and weight data to avoid bias in self-reported data⁷¹.

The GBD study uses a different approach to dietary assessment, primarily relying on adjusted United Nations (UN) and FAO national per capita availability of sugar as primary data to estimate SSBs, with a limited set of individual-level dietary surveys ($N = 44$). In comparison, the GDD uses a much more comprehensive database of largely individual-level dietary surveys to estimate SSB intake ($N = 450$), with other data (such as UN FAO sugar) used as covariates rather than as primary data. Thus, in addition to novel stratification by educational level and area of residence, the GDD dietary estimates may be more valid and informative. Our investigation leverages published diet-disease etiologic effects from extensive meta-analyses identified through reviews conducted by our team, includes both direct and BMI-mediated effects, and incorporates new data on prevalence of overweight and obesity from the NCD-RisC. Our study also estimates incident cases, which is not a measure reported in previous global studies.

Compared with our previous 2010 estimates³, our present investigation includes major expansion of individual-level dietary surveys and global coverage through 2018; updated modeling methods, covariates and validation to improve estimates of stratum-specific mean intakes and uncertainty; inclusion of updated dietary and disease data that are jointly stratified subnationally by age, sex, education level, and urban or rural residence; and updated SSB etiologic effect estimates on T2D, ischemic stroke and ischemic heart disease. This present analysis focused on adults aged 20+ years given the low rates of T2D and CVD globally at younger ages.

Global distributions of SSB intakes

The GDD systematically searched for and compiled representative data on individual-level dietary intakes from national surveys and subnational surveys^{7,12}. The final GDD model incorporated 1,224 dietary surveys representing 185 countries from 7 world regions and 99.0% of the global population in 2020. Of these, 450 surveys reported data on SSBs, totaling 2.9 million individuals from 118 countries representing 86.8% of the global population. Most surveys were nationally or subnationally representative (94.2%), collected at the individual level (84.7%), and included estimates in both urban and rural area of

residence (61.6%). Further details on characteristics of surveys with data on SSBs, including availability of surveys per world region, are available in Supplementary Table 1. The world region classification used in our study was based on groupings that are likely to have consistent exposures to disease risk and rates of disease outcomes, and this or similar classifications have been previously used by our team and others⁷³. Countries included in each world region are listed in Supplementary Table 2. Global, regional and national estimates among the 30 most populous countries, by population characteristics in 2020, are available in Supplementary Tables 3 and 4.

SSBs were defined as any beverages with added sugars and ≥ 50 kcal per 8 oz serving, including commercial or homemade beverages, soft drinks, energy drinks, fruit drinks, punch, lemonade and aguas frescas. This definition excluded 100% fruit and vegetable juices, noncaloric artificially sweetened drinks and sweetened milk. All included surveys used this definition. We used an average sugar content per SSB serving, an assumption that probably has little influence on large-scale demographic estimates such as these but could be a problem for more focused local studies. Home-sweetened teas and coffees (which often would have less than 50 kcal per serving) were not explicitly excluded from the SSB definition at the time of data collection, but total tea and coffee intake were separately collected in the dietary surveys and by the GDD as separate variables. Compared with soda and other industrial SSBs, 100% fruit juices and sugar-sweetened milk, coffee and tea have shown inconsistent evidence for health effects, and were therefore excluded from our definition of SSBs^{2,9}. Differences in health effects may relate to additional nutrients in those drinks, such as calcium, vitamin D, fats, and protein in milk, caffeine and polyphenols in coffee and tea, and fiber and vitamins in 100% juice, or to differences in the rapidity of consumption and/or drinking patterns of these beverages. Notably, each of these other beverages is also generally excluded in policy and surveillance efforts around SSBs¹². At high intakes, alcoholic beverages have been associated with T2D and CVD in prospective cohorts and genome-wide association studies^{78,79}. However, the effect of alcoholic beverages on T2D and CVD differs from the effect of SSBs on these diseases, and thus, alcohol and SSB should be analyzed separately^{2,79,80}. Moreover, the exclusion of alcoholic beverages ensures comparability across diverse populations, given variations in alcohol consumption due to religious and cultural factors⁸¹. Regulatory shortcomings in labeling 100% fruit and vegetable juices may have led to underestimations in SSB intake and attributable burdens for certain populations^{82,83}.

For our present analysis, we updated SSB intake estimates for 2020 using similar methodology as previously reported¹², but with updated food availability data released by FAO for 2014–2020 as covariates. Because FAO updated its methodology for these new estimates, the FAO estimates from this period versus their estimates from earlier years are not directly comparable (for example, a ‘step change’ in FAO estimates was noted comparing 2013 versus 2014 data for most countries). To account for this and retain the relative ranking between nations, we calculated a nation-specific adjustment factor for each FAO covariate, based on the ratio of that nation’s 2013 versus 2014 data, and applied this to each nation’s FAO estimates from 2014 to 2020.

A Bayesian model with a nested hierarchical structure (with random effects by country and region) was used to estimate the mean consumption level of SSBs and its statistical uncertainty for each of 264 population strata across 185 countries from 1990 through 2020, incorporating and addressing differences in data comparability and sampling uncertainty^{12,84}. The model then estimated intakes jointly stratified by age (22 age categories from 0 to 6 months through 95+ years), sex (female, male), education (≤ 6 years of education, >6 years to 12 years, >12 years) and urbanicity (urban, rural). Although this analysis focuses only on adults aged 20+ years, the model used all age data to generate the strata estimates.

Of the 188 countries with survey data, 3 were dropped from the GDD estimation model owing to unavailability of FAO food availability

data (Andorra, Democratic People’s Republic of Korea and Somalia), an important covariate in the estimation model. Uncertainty of each stratum-specific estimate was quantified using 4,000 iterations to determine posterior distributions of mean intake jointly by country, year and sociodemographic subgroup. The median intake and the 95% UI for each stratum were computed at the 50th, 2.5th and 97.5th percentiles of the 4,000 draws, respectively.

Global, regional, national and within-country population subgroup intakes of SSBs and their uncertainty were calculated as population-weighted averages using all 4,000 posterior estimates for each of the 264 demographic strata in each country–year. Population weights for each year were derived from the United Nations Population Division^{74,75}, supplemented with data for education and urban or rural status from a previous study⁸⁵. Intakes were calculated as 8 oz (248 g) servings per week. For our present analysis, GDD SSB estimates were collapsed for adults aged 85+ years using the 4,000 simulations corresponding to the stratum-level intake data derived from the Bayesian model. In this study, regression-based methods were used to estimate the standard deviation corresponding to each estimated, stratum-specific mean from the dietary survey input data. These mean–standard deviation pairs were then used to generate gamma distribution parameters for usual dietary intake as detailed in the following section.

Estimation of gamma parameters for the distribution of usual intake

Dietary intakes cannot be negative, and the usual intake distributions tend to be skewed to the right^{86,87}. Gamma distributions were shown to be more appropriate than normal distributions for SSBs based on the analysis of GDD input data (for example, NHANES data) in a previous study⁸⁸ and other research on assessment of population dietary intake^{89,90}, as it is nonnegative and includes a wide range of shapes with varying degrees of skewness⁹¹. Standard deviation (s.d.) needed to be obtained to construct the gamma distribution of intakes. Parameters for gamma distribution were generated using the mean estimate from the GDD estimation model and the estimated s.d. for the mean estimate from 1,000 simulations.

Standard deviation estimates for the distribution of usual dietary intake. Stratum-level GDD input survey data were used to fit a linear regression of the s.d. of intake on mean intake (both adjusted for total energy). To determine the appropriate transformation of the input data used for fitting the linear regression, scatter plots of energy-adjusted means versus energy-adjusted s.d. were created. Using this approach, we concluded that a natural log transformation for both mean and s.d. was most appropriate. We also explored excluding demographic and health surveys, household surveys and outlier data owing to the potential unreliability of such surveys for estimating s.d., but determined that no one dietary assessment method contributed unevenly to the observed linear trend. Thus, all available data were included, allowing for the largest possible sample size and greatest generalizability. We also investigated whether the log mean and log s.d. relationship differed by world region, but did not find strong evidence for such heterogeneity. A regression model was used for each individual diet factor to calculate the s.d.:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

where i refers to each survey stratum, Y_i is the natural log of the s.d. of stratum-specific intake, x_i is the natural log of the mean of stratum-specific intake and ε_i is the random error that follows $N(0, \sigma^2)$.

Monte Carlo simulations for generating standard deviation distributions. Estimates for β_0 and β_1 were used to predict 1,000 $\ln(\text{s.d.})$ values

corresponding to 1,000 iterations (k) of the predicted mean intake for each population stratum (j) using Monte Carlo simulations.

$$\hat{Y}_{jk} = \hat{\beta}_0 + \hat{\beta}_1 \hat{X}_{jk}$$

in which \hat{X}_{jk} is the k th sample draw of the posterior distribution for mean intake for population stratum j . We propagated uncertainty from the model estimates, as well as variation within the sampling data itself, by randomly drawing from a t -distribution with $n - 1$ degrees of freedom using the following equation:

$$\ln(\widehat{\text{s.d.}}_{jk}) = \hat{Y}_{jk} + \hat{\sigma} \sqrt{1 + \left(\frac{1}{n}\right) \times t_k^{n-1}},$$

in which $\hat{\sigma}$ is the estimate for σ , n is the number of survey strata, t_k^{n-1} is the k th sample drawn from a t -distribution with $n - 1$ degrees of freedom and $\widehat{\text{s.d.}}_{jk}$ is the k th sample draw of the predicted s.d. distribution for population stratum j .

Estimation of gamma parameters for the distribution of usual intake. The posterior distributions for each stratum-specific s.d. were used to generate 1,000 corresponding shape and rate gamma parameters for the distribution of usual intake, a primary input in the CRA model, using the following equations:

$$\widehat{\text{Shape}}_{jk} = \left(\frac{\hat{X}_{jk}}{\widehat{\text{s.d.}}_{jk}} \right)^2$$

$$\widehat{\text{Rate}}_{jk} = \frac{\hat{X}_{jk}}{\widehat{\text{s.d.}}_{jk}^2}$$

Estimated SSB–disease relationships

The direct risk estimates between SSB intake and T2D, ischemic heart disease and ischemic stroke were obtained from published systematic reviews and evidence grading, based on published meta-analyses of prospective cohort studies and randomized controlled trials including multivariable adjustment for age, sex, BMI and other risk factors to reduce bias from confounding (Supplementary Table 8)^{2,68–70}. The methods and results for the review, identification and assessment of evidence for the SSB–disease relationships have been described^{2,67}. Briefly, evidence for each SSB–disease relationship was first evaluated by grading the quality of evidence according to nine different Bradford Hill criteria for causation: strength, consistency, temporality, coherence, specificity, analogy, plausibility, biological gradient and experiment. This evidence grading was completed independently and in duplicate by two expert investigators. Based on these assessments, probable or convincing evidence was determined independently and in duplicate, in accordance with the criteria of the FAO and World Health Organization⁹² and with consideration of consistency with similar criteria of the World Cancer Research Fund and the American Institute for Cancer Research⁹³. SSBs had at least probable association for direct etiologic effects (BMI independent) on T2D, ischemic heart disease and ischemic stroke risk, as well as on weight gain. See Supplementary Table 9 for further details on the evidence grading criteria and results of this evaluation. All SSB–disease estimates were standardized from the originally reported 250 ml serving size to 8 oz servings (248 g), the unit used in our analysis.

Given that these studies adjusted for BMI, we separately assessed the BMI-mediated effects (BMI change in kg m^{-2}) based on pooled analyses from long-term prospective cohort studies of changes in diet and changes in BMI (Supplementary Table 8)¹⁷. Specifically, we used data from three separate prospective cohort studies: the Nurses' Health Study (1986–2006), involving 50,422 women with 20 years of follow-up; the Nurses' Health Study II (1991–2003), including 47,898 women with 12 years of follow-up; and the Health Professionals Follow-up

Study (1986–2006) with 22,557 men with 20 years of follow-up. Participants included in these analyses were initially free of obesity (that is, $\text{BMI} < 30 \text{ kg m}^{-2}$) or chronic diseases and had complete baseline data on weight and lifestyle habits. The associations between SSBs and weight gain were estimated separately for overweight and obese ($\text{BMI} \geq 25 \text{ kg m}^{-2}$) and non-overweight adults ($\text{BMI} < 25 \text{ kg m}^{-2}$), given observed effect modification by baseline BMI status¹⁷. We used linear regression with robust variance, accounting for within-person repeated measures, to assess the independent relationships between changes in SSB intake and changes in BMI over 4 year periods. Women who became pregnant during follow-up were excluded from the analysis. BMI-mediated effects did not specifically differentiate between overweight and obesity, which could have led to an underestimation in the BMI-mediated effects among adults with obesity.

To examine the BMI-mediated associations, we assessed the impact of differences in BMI on the risk of T2D, ischemic heart disease and ischemic stroke (Supplementary Table 8)^{15,16}. These relationships were obtained from pooled analyses of multiple cohort studies investigating the quantitative effects of BMI on T2D¹⁵, ischemic heart disease¹⁶ and ischemic stroke¹⁵. The risk estimates were transformed from the originally reported 5 kg m^{-2} to 1 kg m^{-2} .

Heterogeneity in diet–disease relationships using age-specific relative risks

Age-specific relative risks were calculated for each SSB–disease etiologic relationship based on evidence showing decreasing proportional effects of metabolic risk factors on cardiometabolic disease incidence at older ages (for example, due to other competing risk factors)^{15,67}. The age-specific relative risks were calculated based on the age at event and were assumed to have a log-linear age association, although the true age relationship may differ.

To calculate the age at event for each SSB–disease pair, we obtained relevant data from the original studies. This included the average age at baseline in years, the follow-up time in years, the type of follow-up time reported (for example, maximum, median or mean) and the study weight for each study in each meta-analysis (Supplementary Tables 10–12 and Supplementary Data 4). In cases in which the age at baseline was reported as a range rather than as the average, we used the central value to estimate the mean. If follow-up time to events was not reported, we estimated it based on the duration of the study. For studies that reported maximum follow-up time, we estimated the mean time to event as half of the maximum follow-up, and for studies that reported mean or median follow-up times, as two-thirds of the mean or median follow-up. The unweighted mean age at event for each study was calculated by summing the mean age at baseline and the appropriate mean time to event, and the weighted mean age at event for the meta-analysis as the weighted age at event across all studies. In cases in which specific studies were excluded from the meta-analysis owing to limitations in study quality, or when the meta-analysis was conducted for multiple outcomes, the weights were adjusted accordingly. When study weights were not reported, we assigned equal weights to each study when calculating the mean overall age at event.

Given limited evidence of significant effect modification by sex, we incorporated similar proportional effects of risk factors by sex⁶⁷. In previous research, we evaluated the proportional differences in relative risk for key diet-related cardiometabolic risk factors, including systolic blood pressure, BMI, fasting plasma glucose and total cholesterol, across six 10 year age groups from 25–34 years to 75+ years⁶⁷. Given similarities across these four risk factors, the mean proportional differences in relative risk across all risk factors were applied to the SSB–disease relative risks. In this study, we disaggregated the mean proportional differences into 14 5 year age groups from 20–24 years to 85+ years. This was achieved by linearly scaling between each 10 year mean proportional difference in log relative risk, anchoring at the calculated mean age at event for each SSB–disease.

We used Monte Carlo simulations to estimate the uncertainty in the age-distributed log relative risk, sampling from the distribution of log relative risk at the age at event. On the basis of 1,000 simulations, we used the 2.5th and 97.5th percentiles to derive the 95% UI.

Global distributions of adiposity

Prevalence of overweight (BMI ≥ 25 kg m⁻²) and underweight (BMI < 18.5 kg m⁻²) in each country–year–age–sex–urbanicity stratum and their uncertainty was obtained from the NCD-RisC. The NCD-RisC collected data from 1,820 population-based studies encompassing national, regional and global trends in mean BMI, with measurements of height and weight taken from over 97 million adults^{71,94}. Surveys were excluded if they relied solely on self-report data, focused on specific subsets of the population or involved pregnancy. The NCD-RisC used a Bayesian hierarchical model to estimate age-specific mean BMI and prevalence of overweight and obesity by country, year and sex. The model incorporated data-driven fixed effects to account for differences in BMI by rural and urban area of residence. A Gibbs sampling Markov Chain Monte Carlo algorithm was used to fit the model, producing 5,000 samples of the posterior distributions of the model parameters. These samples were then used to generate posterior distributions of mean BMI and prevalence of overweight and obesity for each stratum. Estimates were age standardized using age weights from the WHO standard population. Weighting was also used at the global, regional and national levels, taking into account the respective age-specific population proportions by country, year and sex. The estimates of mean BMI and overweight and obesity prevalence were presented along with their respective 95% credible intervals, representing the uncertainty around the estimates. To further stratify the NCD-RisC overweight and obesity prevalence estimates by education level and urbanicity, we assumed that the prevalence did not vary across different education levels or between urban and rural residences. In addition, it was assumed that these estimates remained constant between 2016 and 2020 (as NCD-RisC reports only through 2016, but this CRA analysis assesses estimates for 2020), a conservative assumption that probably underestimates the prevalence of overweight and obesity and, thus, SSB-attributable burdens.

Characterization of optimal intake

The optimal intake level of SSBs served as the counterfactual in our CRA modeling analysis, allowing the quantification of impacts of SSBs on disease risk at the population level. We determined the optimal intake level based on probable or convincing evidence for effects of SSBs on cardiometabolic outcomes. The methodology for defining the optimal intake level has been described⁶⁷. Briefly, it was determined primarily based on disease risk (observed consumption levels associated with lowest disease risk in meta-analyses) with further considerations of feasibility (observed national mean consumption levels in nationally representative surveys worldwide)^{95,96}, and consistency with existing major food-based dietary guidelines^{97,98}. The term ‘optimal intake’ can be considered analytically analogous to what has been referred to as the ‘theoretical minimum risk exposure level’ in other analyses^{99,100}. We prefer the former term as it is more relevant to dietary risks, which can serve as a benchmark for quantifying disease risk, informing dietary guidance and informing policy priorities.

Global distributions of T2D, ischemic heart disease and ischemic stroke

The estimates of underlying cardiometabolic disease burdens at global, regional and national levels were obtained from the GBD 2021. The GBD collected data from censuses, household surveys, civil registration, vital statistics and other relevant records to estimate incidence, prevalence, mortality, years lived with disability (YLDs), years of life lost (YLLs) and DALYs for 371 diseases and injuries⁷³. These estimates were stratified by 204 countries and territories, 23 age groups and sex,

yearly from 1990 to 2021. For this analysis, we used GBD estimates of incidence, mortality and DALYs for T2D, ischemic heart disease and ischemic stroke for 1990 and 2020. The GBD defined T2D as fasting plasma glucose greater than or equal to 126 mg dl⁻¹ (7 mmol l⁻¹) or reporting the use of diabetes medication⁷³. Estimated cases of type 1 diabetes were subtracted from the overall diabetes cases at the most stratified level of age, sex, location and year to estimate T2D cases. Ischemic heart disease was estimated in the GBD as the aggregate of myocardial infarction (heart attack), angina (chest pain) or ischemic cardiomyopathy (heart failure due to ischemic heart disease). Ischemic stroke was defined as rapidly developing clinical signs of (usually focal) cerebral function disturbance lasting over 24 h, or leading to death, according to the WHO criterion of sudden occlusion of arteries supplying the brain due to a thrombus¹⁰¹.

GBD mortality estimates were generated using the Cause of Death Ensemble Model framework, which incorporated various models including different sets of covariates testing the predictive validity, and generating cause-specific mortality estimates^{73,102,103}. Cause of Death Ensemble Model estimates were scaled among all causes such that the sum of cause-specific deaths did not exceed all-cause mortality. YLLs were calculated as the product of the number of deaths for each cause by age, sex, location and year times the standard life expectancy. Life expectancy was first decomposed by cause of death, location and year to represent the cause-specific effects on life expectancy¹⁰². Then, the sum across age groups was taken to estimate the impact of a given cause on the at-birth life expectancy from 1990 to 2021. Incidence was modeled using DisMod, a meta-regression tool that used epidemiologic data to estimate the occurrence disease within a population and determines whether cases remain prevalent, go into remission or result in death. YLDs were calculated by splitting the prevalence of each cause into mutually exclusive sequela, each defined by a health state; each health state was then weighted by the corresponding disability weight⁷³. Finally, DALYs were calculated as the sum of YLLs and YLDs.

Disaggregation of T2D and CVD burdens by education level and urbanicity

The GBD provides underlying disease estimates at global, regional and national levels for 1990 to 2021, jointly stratified by age and sex. Extensive previous evidence shows that T2D and CVD outcomes vary by educational attainment and urbanicity^{104–122}. We further stratified the 1990 and 2020 GBD estimates by education level (low, medium, high) and area of residency (urban, rural) to examine potential variations in risk within these subpopulations and to align with the demographic and GDD dietary data stratifications available. This approach required assumptions on distributions of disease burdens by these demographic factors and potentially underestimated uncertainty in our results stratified by these factors.

To stratify the GBD estimates, we conducted a search of scientific literature to identify recent meta-analysis, pooled analyses and large surveys evaluating the association between educational attainment and urbanicity with the risk of T2D and CVD. Because we hypothesized that country income level was a potential effect modifier for the relationships of educational attainment and urbanicity with T2D and CVD risk, we further collected and collated risk estimates stratified by country income level. We limited our analysis to studies adjusting only for age and sex, when possible, to avoid the attenuating effects of adjusting for additional covariates^{104–122}.

We conducted fixed-effects meta-analysis of collated effect sizes (associations between education or urbanicity and disease rates), stratified by country income level. Published estimates were standardized to high versus low education level, matched as closely as possible to the GDD definitions (low: 0–6 years of education; high: >12 years of education), as well as to urban versus rural residence. We pooled estimates within studies when (1) multiple estimates were reported for different CVD outcomes, (2) separate estimates were provided for men

and women, (3) estimates were reported for different locations (except by country income) or (4) an intermediate category matched our definitions for education level or area of residence. The characteristics of the studies used to calculate the effect estimates, including their original and calculated effect sizes, can be found in Supplementary Data 5 and 6 for education level and area of residence, respectively.

We conducted a separate fixed-effect meta-analysis for the relationship of education or urbanicity to T2D and CVD, stratified by country income level. We distributed the central estimate of our meta-analyzed risk estimate equally for high versus low education, and urban versus rural residence, by taking its square root and inverse square root (Supplementary Table 13). This approach assumed similar differences from high to medium education as from medium to low education. We also explored distributing the central estimate by incorporating information on the actual distance (for example, grade years) from high to medium education and medium to low education, when such information was available. As the results did not appreciably differ, we used the square root and inverse square root approach to maintain consistency across studies, particularly given heterogeneity in categorizations of education levels. The final calculated effect estimates for the association between education level and area of residence with T2D and CVD, by income country level, can be found in Supplementary Table 13.

The T2D, ischemic heart disease and ischemic stroke estimates for each year–country–age–sex stratum (mean and 95% UI) were multiplied by their respective population proportion, education effect and urban effect. This process created six de novo strata with the raw (unscaled) fully proportioned burden estimates and their uncertainty. The global population proportions for each year were derived from the United Nations Population Division⁷⁵, supplemented with data on education attainment from a previous study⁷⁶. Finally, to prevent under- or overestimation of the absolute number of T2D, ischemic heart disease and ischemic stroke cases globally, the raw fully proportioned burden estimates were scaled to match the total burden estimate for each stratum. This scaling ensured that the overall burden estimates remained consistent. Supplementary Table 14 provides a fictitious, illustrative example of how 1,000 T2D cases in a single age–sex population stratum (low-income country) in a given year were disaggregated into the 6 finer education–urbanicity strata using the central estimate of the meta-analyzed education and urban effects. The population proportioned only burden estimates is also provided as a comparison. While uncertainty was incorporated in all the modeling parameters, we were unable to include uncertainty in the stratification of T2D and CVD cases by educational attainment and urban or rural residence as rigorous data to do so were not available.

Statistical analysis: CRA analysis

The CRA framework incorporated the data inputs and their uncertainty to estimate the absolute number, rate (per million adults 20+ years) and proportion of T2D, ischemic heart disease and ischemic stroke cases attributable to intake of SSBs in 1990 and 2020 (Supplementary Fig. 18). For each stratum, the model calculated the percentage (population attributable fraction (PAF)) of total T2D, ischemic heart disease and ischemic stroke incidence, mortality and DALYs attributable to intake of SSBs. For BMI-mediated effects, the model considered the associations between observed SSB intakes and changes in BMI at the stratum level. This association was weighted by the prevalence of overweight (BMI ≥ 25), normal weight (BMI >18.5 to <25) and underweight (BMI <18.5 ; assumed to have no effect) in each stratum. The resulting weighted BMI change was combined with the relative risk (RR) of BMI change and T2D or CVD using the same continuous PAF formula. Further details on each calculation for the PAF can be found in the sections below.

Given that summing direct and BMI-mediated PAFs would overestimate the combined effect, for each disease stratum (that is,

country–year–age–sex–education–residence), the PAF was calculated using proportional multiplication of the direct and BMI-mediated PAFs as follows:

$$\text{PAF} = 1 - ((1 - \text{direct PAF}) \times (1 - \text{BMI-mediated PAF}))$$

The resulting PAF was then multiplied by the corresponding number of disease cases to calculate the attributable burden in each stratum. Findings were evaluated globally, regionally and nationally, and by specific population subgroups of age, sex, education and urbanicity. The results are presented as proportional attributable burden (percentage of cases) and attributable rate (per one million adults). This representation of the proportional multiplication for a single risk factor (that is, SSBs) is equivalent to the formula commonly reported for several risk factors: $\text{PAF} = 1 - \prod_{i=1}^n (1 - \text{PAF}_i)$

Direct-effect PAF. The PAF formula is used to quantify the burden of disease attributable to a particular exposure. It involves comparing the disease cases associated with the observed exposure levels in the population to a counterfactual scenario with an optimal intake distribution, given a known etiologic exposure–disease risk relationship.

In this analysis, we aimed to estimate the burden of incidence, mortality and DALYs for T2D, ischemic heart disease and ischemic stroke attributable to intake of SSBs.

$$\text{PAF} = \frac{\int_{x=0}^m \text{RR}(x) P(x) dx - 1}{\int_{x=0}^m \text{RR}(x) P(x) dx},$$

The PAF formula used is as follows:

$$\text{PAF} = \frac{\int_{x=0}^m \text{RR}(x) P(x) dx - 1}{\int_{x=0}^m \text{RR}(x) P(x) dx},$$

where $P(x)$ is the usual SSB intake distribution in a specific population stratum, assumed to follow a gamma distribution as used in previous analyses^{3,31,88}; $\text{RR}(x)$ is the age-specific relative risk function for T2D or CVD risk; and m is the maximum exposure level.

$\text{RR}(x)$ is defined as:

$$\begin{cases} \exp(\beta(x - y(x))) & : x - y(x) \geq 0 \\ 1 & : x - y(x) < 0 \end{cases}$$

where β is the stratum-specific change in log relative risk per unit of exposure, x is the current exposure level and $y(x)$ is the optimal exposure level. $y(x)$ is defined to be $F_{\text{optimal}}(F_x^{-1}(x))$, where F_{optimal} is the cumulative distribution function of the optimal intake and F_x^{-1} is the inverse cumulative distribution function of the current exposure distribution. Implicit in how we characterize the relative risk function are certain assumptions, including a linear relationship between the log relative risk (beta) and the unit of exposure. This model assumes that no further risk is associated with exposure beyond the optimal intake level, and that both x and the optimal intake level for an individual at exposure level x are the q th quantile of their respective distributions (the observed exposure distribution and the optimal intake distribution, respectively).

PAF calculation. In practice, simple numerical integration using Riemann sums can be used to compute the integrals in the PAF formula⁸⁸.

$$\text{PAF} = \frac{\sum_{i=1}^n P_i(\text{RR}_i - 1)}{\sum_{i=1}^n P_i(\text{RR}_i - 1) + 1}$$

$$\text{PAF} = \frac{\sum_{i=1}^n P_i(\text{RR}_i - 1)}{\sum_{i=1}^n P_i(\text{RR}_i - 1) + 1}$$

n categories are determined by dividing the exposure range (chosen here to be $0, F_x^{-1}(\Phi(-6))$) into 121 intervals, each of length 0.1 when converted to the standard normal scale (except for the first one). Φ is defined as the cumulative distribution function of the standard normal distribution $N(0,1)$. More precisely, the range of exposure groups i can be described as:

$$\begin{aligned} & (0, F_x^{-1}(\Phi(-6))) & : i = 1 \\ & (F_x^{-1}(\Phi(-6 + 0.1(i-2))), F_x^{-1}(\Phi(-6 + 0.1(i-1)))) & : i = 1 \end{aligned}$$

BMI-mediated effects PAF. The association of change in BMI with change in SSB intake was assessed in three pooled US cohorts using multivariate linear regression accounting for within-person repeated measures, as described in an earlier study¹⁷. Separate linear relationships were estimated for underweight (BMI < 18.5), normal weight (BMI > 18.5 to < 25) and overweight (BMI ≥ 25 to < 30), given observed effect modification by baseline BMI status¹⁷. Because individuals with obesity were excluded in these previous analyses, we used the risk estimate for individuals with overweight for individuals with obesity, which could underestimate the full effects of SSB on weight change.

To assess the BMI-mediated effects of SSB intake on incidence, mortality and DALYs of T2D, ischemic heart disease and ischemic stroke, we first calculated the monotonic effect of SSB intake on BMI change for each population stratum by weighting the baseline BMI-specific effect by the respective prevalence of underweight, normal weight and overweight (including obesity) within each stratum. We obtained overweight and underweight population distributions from the NCD-RisC⁷¹ and calculated the prevalence of normal weight as 1 minus the sum of these prevalences⁷¹. The NCD-RisC estimates go up to 2016, and thus, for our 2020 analysis, we used data from 2016 as a proxy for 2020. Given increasing adiposity globally, this assumption could result in underestimation of disease burdens due to SSBs in 2020. We assumed that individuals with underweight did not experience increased risk of T2D, ischemic heart disease or ischemic stroke with increased consumption of SSBs. As such, the monotonic effect for this population segment was set at 0:

$$\begin{aligned} \text{SSB-to-BMI effect} &= \beta_{\text{BMI} \geq 25} \times (\text{overweight prevalence}) + \beta_{\text{BMI} 18.5-25} \\ &\times (\text{normal weight prevalence}) + 0 \\ &\times (\text{underweight prevalence}) \end{aligned}$$

We then estimated the BMI-mediated log(RR) by multiplying the log(RR) per unit increase in BMI and the SSB-to-BMI effect (associated increase in BMI per one-unit-associated increase in SSB intake).

Quantification of uncertainty using Monte Carlo simulations. We used Monte Carlo simulations to quantify the uncertainty around the PAF estimate. In this calculation, we incorporated uncertainty of multiple key parameters, including the usual intake distribution of SSBs in each stratum; underlying T2D, CVD and DALY burden estimates in each stratum; the etiologic estimates (RR) for SSB-BMI, SSB-T2D and SSB-CVD relationships; and the prevalence of individuals with underweight, normal weight or overweight in each stratum. For each SSB-disease combination and stratum, we drew randomly 1,000 times from the respective probability distributions. This included drawing randomly from the normal distribution of the estimate of disease-specific changes in the log(RR) of BMI-mediated and direct etiologic effects for a one-unit increase in SSB intake, the posterior distributions for shape and rate parameters for usual dietary intake and the normal distribution of the estimate for the prevalence of underweight, normal weight and overweight. Draws of proportions that were less than 0 or greater than 1 were truncated at 0 or 1, respectively, and draws of mean intake

that were 0 or less were truncated at 0.00001. Each set of random draws was used to calculate the PAFs and, multiplied by the stratum-specific disease rates, the associated absolute attributable disease burden. Corresponding 95% UIs were derived from the 2.5th and 97.5th percentiles of 1,000 estimated models.

Sociodemographic development index

We assessed national-level findings by SDI in 1990 and 2020. The SDI is a composite measure of a nation's development based on factors such as income per capita, educational attainment and fertility rates¹⁸.

Changes in SSB-attributable burdens over time

To compare estimates across different years (1990 and 2020), we calculated differences for absolute and proportional burdens from 1990 to 2020 (that is, 2020 minus 1990). We performed this calculation for each simulation resulting in a distribution of differences, and we report the median and 95% UIs for each difference. We did not formally standardize comparisons over time by age or sex. This decision was made to ensure that findings would reflect the actual population differences in attributable burdens that are relevant to policy decisions. However, we also performed analyses stratified by age and sex, taking into account changes in these demographics over time. All analyses were conducted using R statistical software, R version 4.4.0 (ref. 123) on the Tufts High Performance Cluster.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data used in this analysis are publicly available from the following sources: (1) population SSB intake distributions based on individual-level survey data from the GDD (<https://www.globaldietarydatabase.org/data-download>)^{71,72}; (2) optimal SSB intake levels from previous analyses⁶⁷; (3) direct age-adjusted etiologic effects of SSBs on T2D, ischemic heart disease and ischemic stroke adjusted for BMI², and of weight gain on T2D¹⁵, ischemic heart disease¹⁶ and ischemic stroke¹⁵ from previous meta-analyses and pooled analyses of prospective cohorts, as well as linear, BMI-stratified effects of SSBs on weight gain or loss¹⁷; (4) population overweight (BMI ≥ 25 kg m⁻²) and underweight (BMI < 18.5 kg m⁻²) distributions from the NCD-RisC (<https://ncdrisc.org/data-downloads.html>)⁷¹; (5) total T2D, ischemic heart disease and ischemic stroke incidence, DALYs, and mortality estimate distributions from the GBD study (<https://vizhub.healthdata.org/gbd-results/>)^{72,73}; and (6) population demographic data from the United Nations Population Division (UN, <https://population.un.org/wpp/>)^{74,75}, the Barro and Lee Educational Attainment Dataset 2013 (<https://doi.org/10.3386/w15902>)⁷⁶ and SDI data (Global Health Data Exchange: GBD, <https://ghdx.healthdata.org/record/ihme-data/gbd-2019-socio-demographic-index-sdi-1950-2019>). The GDD SSB intake data were collapsed for 85+ years using the 4,000 simulations corresponding to the stratum-level intake data derived from the Bayesian model. These data were used to obtain the gamma parameters of the SSB intake distribution used in the model. The 4,000 simulation files can be made available to researchers upon request. Eligibility criteria for such requests include utilization for nonprofit purposes only, for appropriate scientific use based on a robust research plan and by investigators from an academic institution. If you are interested in requesting access to the data, please submit the following documents: (1) proposed research plan (please download and complete the proposed research plan form: https://www.globaldietarydatabase.org/sites/default/files/manual_upload/research-proposal-template.pdf), (2) data sharing agreement (please download this form: https://www.globaldietarydatabase.org/sites/default/files/manual_upload/tufts-gdd-data-sharing-agreement.docx) and complete the highlighted

fields; have someone who is authorized to enter your institution into a binding legal agreement with outside institutions sign the document; note that this agreement does not apply when protected health information or personally identifiable information is shared), and (3) email items 1 and 2 to info@globaldietarydatabase.org. Please use the subject line 'GDD Data Access Request'. Once all documents have been received, the GDD team will be in contact with you within 2–4 weeks regarding subsequent steps. Data will be shared as .csv or .xlsx files, using a compressed format when appropriate. Source data are provided with this paper.

Code availability

Custom code was developed using R (version 4.4.0) for this research including calculation of age-adjusted relative risks, SSB intake gamma parameters, CRA analysis (including the PAF and attributable mortality for each stratum) and data visualizations. Given the computational size, calculation of the SSB intake gamma parameters, CRA modeling, absolute and relative differences, and summary statistics were run on the Tufts University High Performance Computing Cluster (<https://it.tufts.edu/high-performance-computing>), supported by the National Science Foundation (grant: 2018149, https://www.nsf.gov/awardsearch/showAward?AWD_ID=2018149&HistoricalAwards=false) under active development by Research Technology, Tufts Technology Services (<https://it.tufts.edu/researchtechnology.tufts.edu>). The statistical code can be made available to researchers upon request. Eligibility criteria for such requests include utilization for nonprofit purposes only, for appropriate scientific use based on a robust research plan and by investigators from an academic institution. GDD will nominate coauthors to be included in any papers generated using GDD-generated statistical code. If you are interested in requesting access to the statistical code, please submit the following documents: (1) proposed research plan (please download and complete the proposed research plan form: https://www.globaldietarydatabase.org/sites/default/files/manual_upload/research-proposal-template.pdf), (2) data sharing agreement (please download this form: https://www.globaldietarydatabase.org/sites/default/files/manual_upload/tufts-gdd-data-sharing-agreement.docx and complete the highlighted fields; have someone who is authorized to enter your institution into a binding legal agreement with outside institutions sign the document; note that this agreement does not apply when protected health information or personally identifiable information is shared) and (3) email items 1 and 2 to info@globaldietarydatabase.org. Please use the subject line 'GDD Code Access Request'. Once all documents have been received, the GDD team will be in contact with you within 2–4 weeks regarding subsequent steps. Data will be shared as .csv or .xlsx files, using a compressed format when appropriate.

References

67. Micha, R. et al. Etiologic effects and optimal intakes of foods and nutrients for risk of cardiovascular diseases and diabetes: systematic reviews and meta-analyses from the Nutrition and Chronic Diseases Expert Group (NutriCoDE). *PLoS ONE* **12**, e0175149 (2017).
68. Yin, J. et al. Intake of sugar-sweetened and low-calorie sweetened beverages and risk of cardiovascular disease: a meta-analysis and systematic review. *Adv. Nutr.* **12**, 89–101 (2021).
69. Bechthold, A. et al. Food groups and risk of coronary heart disease, stroke and heart failure: a systematic review and dose-response meta-analysis of prospective studies. *Crit. Rev. Food Sci. Nutr.* **59**, 1071–1090 (2019).
70. Qin, P. et al. Sugar and artificially sweetened beverages and risk of obesity, type 2 diabetes mellitus, hypertension, and all-cause mortality: a dose-response meta-analysis of prospective cohort studies. *Eur. J. Epidemiol.* **35**, 655–671 (2020).
71. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. *Lancet* **390**, 2627–2642 (2017).
72. Global Burden of Disease Collaborative Network *Global Burden of Disease Study 2021 (GBD 2021) Results* (IHME, 2024).
73. Ferrari, A. J. et al. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **403**, 2133–2161 (2024).
74. United Nations Population Division. *Urban Population (% of Total Population)* (UN, 2018); <https://population.un.org/wup/Download/>
75. United Nations Population Division. *Total Population by Sex (Thousands)* (UN, 2019); <https://population.un.org/wpp/>
76. Barro, R. J. & Lee, J. W. A new data set of educational attainment in the world, 1950–2010. *J. Dev. Econ.* **104**, 184–198 (2013).
77. Vos, T. et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1204–1222 (2020).
78. Rosoff, D. B., Davey Smith, G., Mehta, N., Clarke, T. K. & Lohoff, F. W. Evaluating the relationship between alcohol consumption, tobacco use, and cardiovascular disease: a multivariable Mendelian randomization study. *PLoS Med.* **17**, e1003410 (2020).
79. Llamosas-Falcón, L. et al. The relationship between alcohol consumption, BMI, and type 2 diabetes: a systematic review and dose-response meta-analysis. *Diabetes Care* **46**, 2076–2083 (2023).
80. Zhang, C. et al. Alcohol intake and risk of stroke: a dose–response meta-analysis of prospective studies. *Int. J. Cardiol.* **174**, 669–677 (2014).
81. Tuck, A., Robinson, M., Agic, B., Ialomiteanu, A. R. & Mann, R. E. Religion, alcohol use and risk drinking among Canadian adults living in Ontario. *J. Relig. Health* **56**, 2023–2038 (2017).
82. El jugoso engaño con bebidas de «fruta». *El Poder del Consumidor* (2009).
83. Rellinger, D. Is your fruit juice really 100 percent fruit juice? *Michigan State University* (Michigan State Univ., 2013); https://www.canr.msu.edu/news/is_your_fruit_juice_really_100_percent_fruit_juice
84. Miller, V. et al. Global, regional, and national consumption of animal-source foods between 1990 and 2018: findings from the Global Dietary Database. *Lancet Planet. Health* **6**, e243–e256 (2022).
85. Barro, R. & Lee, J. A new data set of educational attainment in the world, 1950–2010. *J. Dev. Econ.* **104**, 184–198 (2013).
86. National Research Council *Nutrient Adequacy: Assessment Using Food Consumption Surveys* (National Academy of Sciences, 1986).
87. Emrich, L. J., Dennison, D. & Dennison, K. F. Distributional shape of nutrition data. *J. Am. Diet. Assoc.* **89**, 665–670 (1989).
88. Micha, R. et al. Association between dietary factors and mortality from heart disease, stroke, and type 2 diabetes in the United States. *JAMA* **317**, 912–924 (2017).
89. Chang, H. Y., Suchindran, C. M. & Pan, W. H. Using the overdispersed exponential family to estimate the distribution of usual daily intakes of people aged between 18 and 28 in Taiwan. *Stat. Med.* **20**, 2337–2350 (2001).
90. Yokoi, K. Simplified population data analysis using gamma distribution for nutritional requirements and its application to the estimation of iron requirements for women of child-bearing age. *J. Trace Elem. Med. Biol.* **62**, 126597 (2020).

91. Jensen, H. H., Nusser, S. M., Riddick, H. & Sands, L. A critique of two methods for assessing the nutrient adequacy of diets. *J. Nutr. Educ.* **24**, 123–129 (1992).
92. *Diet, Nutrition and the Prevention of Chronic Diseases: Report of a Joint WHO/FAO Expert Consultation* (World Health Organization, 2003).
93. World Cancer Research Fund/American Institute for Cancer Research. *Diet, Nutrition, Physical Activity and Cancer: A Global Perspective* <https://www.wcrf.org/wp-content/uploads/2021/02/Summary-of-Third-Expert-Report-2018.pdf> (World Cancer Research Fund, 2020).
94. Finucane, M. M. et al. National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet* **377**, 557–567 (2011).
95. Singh, G. M. et al. Global, regional, and national consumption of sugar-sweetened beverages, fruit juices, and milk: a systematic assessment of beverage intake in 187 countries. *PLoS ONE* **10**, e0124845 (2015).
96. Khatibzadeh, S. et al. A global database of food and nutrient consumption. *Bull. World Health Organ.* **94**, 931–934 (2016).
97. Lloyd-Jones, D. M. et al. Defining and setting national goals for cardiovascular health promotion and disease reduction. *Circulation* **121**, 586–613 (2010).
98. *2015–2020 Dietary Guidelines for Americans* (US Department of Agriculture and US Department of Health and Human Services, 2015).
99. Lin, X. et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Sci. Rep.* **10**, 14790 (2020).
100. Afshin, A. et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **393**, 1958–1972 (2019).
101. *WHO STEPS Stroke Manual: The WHO STEPwise Approach to Stroke Surveillance/Noncommunicable Diseases and Mental Health* (World Health Organization, 2005).
102. Naghavi, M. et al. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **403**, 2100–2132 (2024).
103. James, S. L. et al. Estimating global injuries morbidity and mortality: methods and data used in the Global Burden of Disease 2017 study. *Inj. Prev.* **26**, i125–i153 (2020).
104. Dagenais, G. R. et al. Variations in diabetes prevalence in low-, middle-, and high-income countries: results from the prospective urban and rural epidemiological study. *Diabetes Care* **39**, 780–787 (2016).
105. Agardh, E., Allebeck, P., Hallqvist, J., Moradi, T. & Sidorchuk, A. Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. *Int. J. Epidemiol.* **40**, 804–818 (2011).
106. den Braver, N. R. et al. Built environmental characteristics and diabetes: a systematic review and meta-analysis. *BMC Med.* **16**, 12 (2018).
107. O'Connor, A. & Wellenius, G. Rural–urban disparities in the prevalence of diabetes and coronary heart disease. *Public Health* **126**, 813–820 (2012).
108. Corsi, D. J. & Subramanian, S. V. Socioeconomic gradients and distribution of diabetes, hypertension, and obesity in India. *JAMA Netw. Open* **2**, e190411 (2019).
109. Seiglie, J. A. et al. Diabetes prevalence and its relationship with education, wealth, and BMI in 29 low- and middle-income countries. *Diabetes Care* **43**, 767–775 (2020).
110. Yusuf, S. et al. Modifiable risk factors, cardiovascular disease, and mortality in 155722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. *Lancet* **395**, 795–808 (2020).
111. Jackson, C. A., Sudlow, C. L. M. & Mishra, G. D. Education, sex and risk of stroke: a prospective cohort study in New South Wales, Australia. *BMJ Open* **8**, e024070 (2018).
112. Hassen, H. Y., Bastiaens, H., Van Royen, K. & Abrams, S. Socioeconomic and behavioral determinants of cardiovascular diseases among older adults in Belgium and France: a longitudinal analysis from the SHARE study. *PLoS ONE* **15**, e0243422 (2020).
113. Min, Y. I. et al. Cardiovascular disease burden and socioeconomic correlates: findings from the Jackson Heart Study. *J. Am. Heart Assoc.* **6**, e004416 (2017).
114. Khaing, W., Vallibhakara, S. A., Attia, J., McEvoy, M. & Thakkinstant, A. Effects of education and income on cardiovascular outcomes: a systematic review and meta-analysis. *Eur. J. Prev. Cardiol.* **24**, 1032–1042 (2017).
115. Liao, L. Z., Zhuang, X. D., Zhang, S. Z., Liao, X. X. & Li, W. D. Education and heart failure: new insights from the atherosclerosis risk in communities study and mendelian randomization study. *Int. J. Cardiol.* **324**, 115–121 (2021).
116. Reshetnyak, E. et al. Impact of multiple social determinants of health on incident stroke. *Stroke* **51**, 2445–2453 (2020).
117. Li, S. et al. Modifiable risk factors associated with cardiovascular disease and mortality in China: a PURE substudy. *Eur. Heart J.* **43**, 2852–2863 (2022).
118. Joseph, P. et al. Cardiovascular disease, mortality, and their associations with modifiable risk factors in a multi-national South Asia cohort: a PURE substudy. *Eur. Heart J.* **43**, 2831–2840 (2022).
119. Lopez-Jaramillo, P. et al. Risk factors, cardiovascular disease, and mortality in South America: a PURE substudy. *Eur. Heart J.* **43**, 2841–2851 (2022).
120. Kapral, M. K. et al. Rural–urban differences in stroke risk factors, incidence, and mortality in people with and without prior stroke. *Circ. Cardiovasc. Qual. Outcomes* **12**, e004973 (2019).
121. Bernabe-Ortiz, A. & Carrillo-Larco, R. M. Urbanization, altitude and cardiovascular risk. *Glob. Heart* **17**, 42 (2022).
122. Lu, J. et al. Characteristics of high cardiovascular risk in 1.7 million Chinese adults. *Ann. Intern. Med.* **170**, 298–308 (2019).
123. R Core Team. *R: A Language and Environment for Statistical Computing*. <http://www.r-project.org/> (R Foundation for Statistical Computing, 2020).

Acknowledgements

The GDD estimates that underlie this research were supported by the Gates Foundation (grant OPP1176682 to D.M.). This analysis was further supported by the American Heart Association (grant 903679 to L.L.-C.) and Consejo Nacional de Ciencia y Tecnología in Mexico (to L.L.-C.). We acknowledge the Tufts University High Performance Computing Cluster (<https://it.tufts.edu/high-performance-computing>), which was used for the research reported in this paper. This cluster is based on work supported by the National Science Foundation (grant 2018149) and is under active development by Research Technology, Tufts Technology Services. The funding agencies had no role in the design of the study; collection, management, analysis or interpretation of the data; preparation, review or approval of the paper; or decision to submit for publication.

Author contributions

Conceptualization and design: L.L.-C., R.M. and D.M.; methodology: L.L.-C., R.M., D.M., F.C. and M.O.; data acquisition: L.L.-C., F.C., M.O., P.S., J.Z., J.R.S. and V.M.; statistical analysis and data curation: L.L.-C., M.O., F.C. and V.M.; validation: L.L.-C., M.O., R.M. and D.M.; data

interpretation: L.L.-C., R.M., D.M, F.C., S.B. and S.B.C.; visualization: L.L.-C., M.O., R.M., D.M., V.M., S.B. and S.B.C.; writing—original draft preparation: L.L.-C.; writing, review and editing: all authors provided critical revision to the paper; supervision: D.M., R.M., S.B. and S.B.C.; project administration: D.M.; funding acquisition: D.M., R.M. and L.L.-C. The GDD Consortium members provided survey intake data to inform the SSB intake predictions used in this analysis.

Competing interests

The investigators did not receive funding from a pharmaceutical company or other agency to write this report. L.L.-C. reports research funding from the Gates Foundation, the American Heart Association and Consejo Nacional de Ciencia y Tecnología in Mexico (CONACyT), outside of the submitted work. M.O. reports previous research funding from the Gates Foundation, as well as the National Institutes of Health and employment with Food Systems for the Future, outside of the submitted work. R.M. reports research funding from the Gates Foundation and (ended) the US National Institutes of Health, Danone and Nestle. She also reports consulting from Development Initiatives and as IEG chair for the Global Nutrition Report, outside of the submitted work. F.C., J.Z. and P.S. report research funding from the Gates Foundation, as well as the National Institutes of Health, outside of the submitted work. V.M. reports research funding from the Canadian Institutes of Health Research and from the American Heart Association, outside the submitted work. J.R.S. reports research funding from the Gates Foundation, as well as the National Institutes of Health, Nestlé, Rockefeller Foundation and Kaiser Permanent Fund at East Bay Community Foundation, outside of the submitted work. S.B. reports funding from Bloomberg Philanthropies, CONACyT, United Nations International Children's Emergency Fund (UNICEF) and Fundación Rio Arronte, all outside the submitted work. S.B.C.

reports research funding from the US National Institutes of Health, US Department of Agriculture, the Rockefeller Foundation, US Agency for International Development and the Kaiser Permanente Fund at East Bay Community Foundation, outside the submitted work. D.M. reports research funding from the US National Institutes of Health, the Gates Foundation, the Rockefeller Foundation, the Kaiser Permanente Fund at East Bay Community Foundation and the National Association of Chain Drug Stores Foundation; scientific advisory board membership for Beren Therapeutics, Brightseed, Calibrate, Elysium Health, HumanCo, Instacart, January Inc., Season Health and the Validation Institute; equity in Calibrate and HumanCo; and chapter royalties from UpToDate, all outside the submitted work.

Additional information

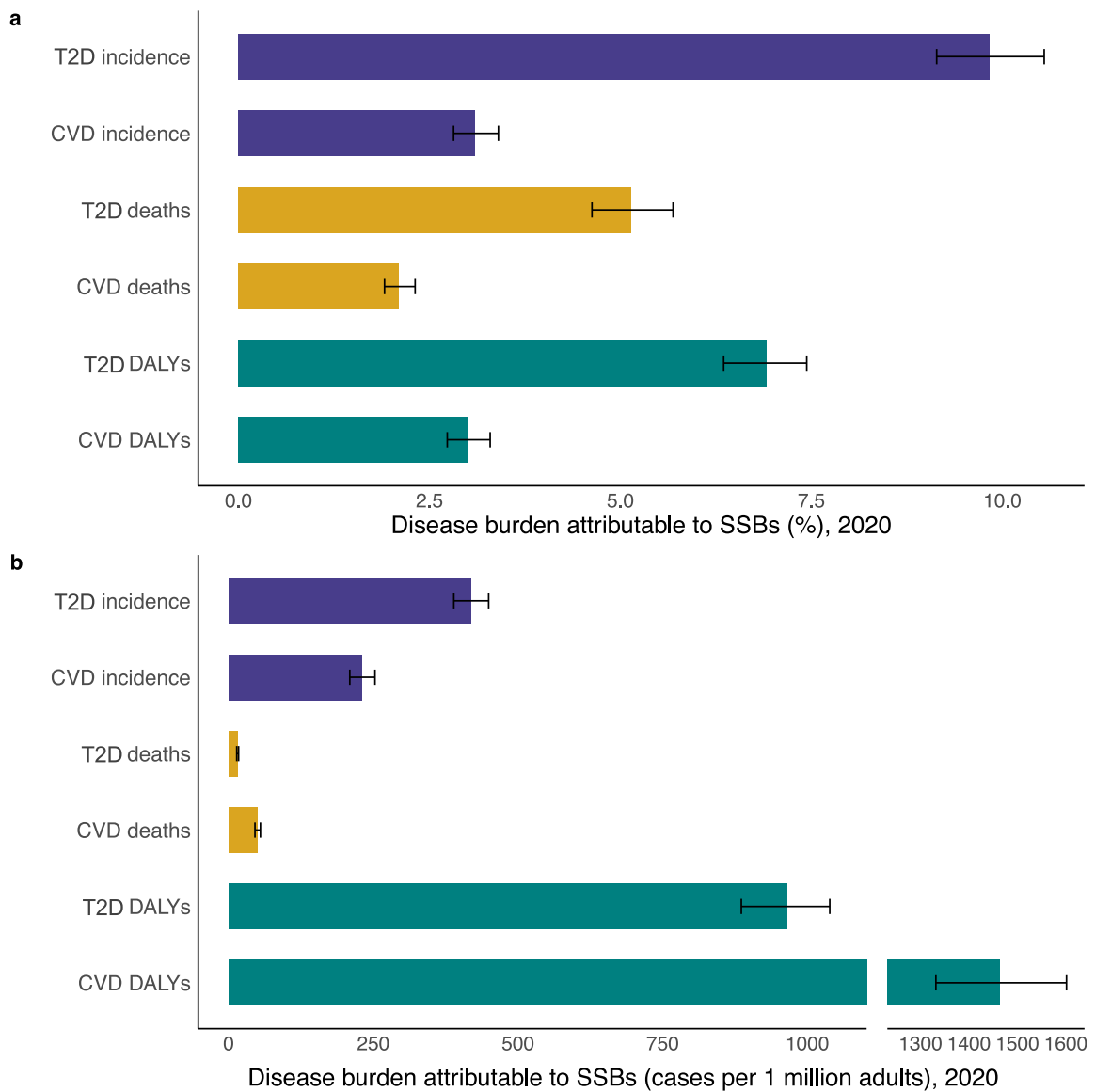
Extended data is available for this paper at <https://doi.org/10.1038/s41591-024-03345-4>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41591-024-03345-4>.

Correspondence and requests for materials should be addressed to Laura Lara-Castor or Dariush Mozaffarian.

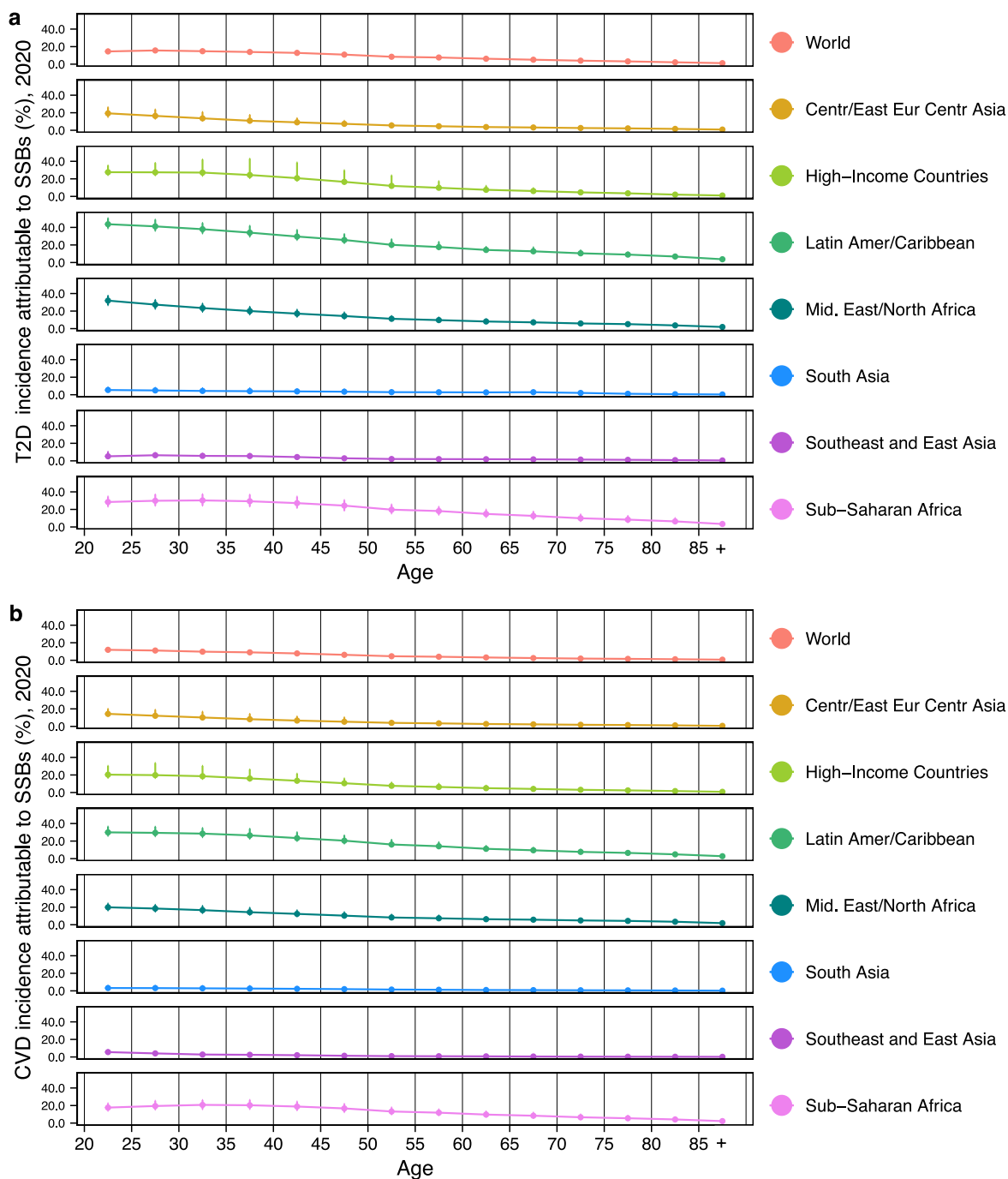
Peer review information *Nature Medicine* thanks Bo Xi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Ming Yang, in collaboration with the *Nature Medicine* team.

Reprints and permissions information is available at www.nature.com/reprints.



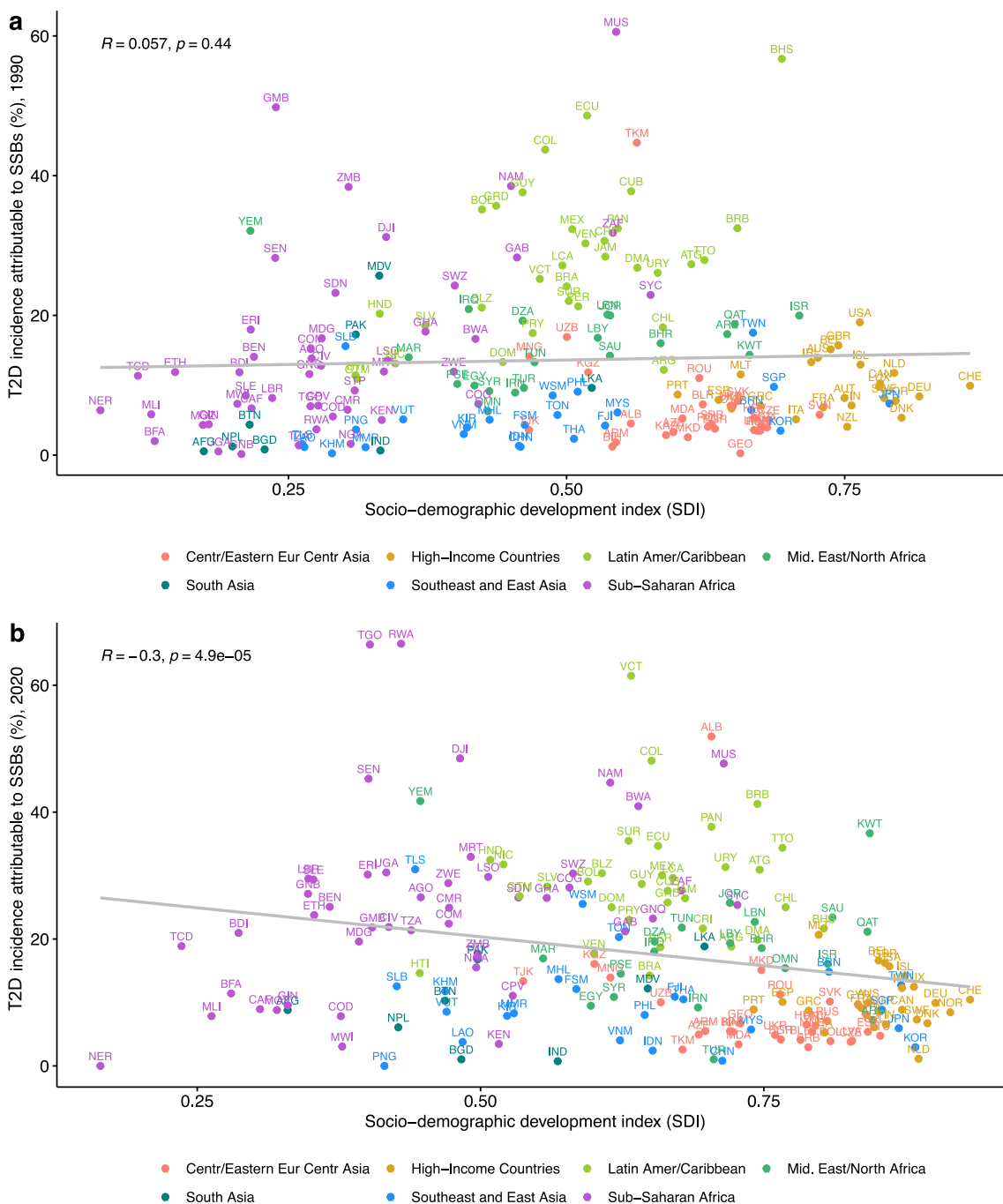
Extended Data Fig. 1 | Global disease burden attributable to intake of sugar-sweetened beverages among adults (20+ years) in 2020. Bars represent the estimated proportional SSB-attributable burden in the top panel (a) and the estimated SSB-attributable absolute cases per 1 million adults in the bottom panel (b) of T2D and CVD incidence, deaths, and DALYs. Data are presented as central estimate (median) and the corresponding 95% UI, derived from the

2.5th and 97.5th percentiles of 1,000 multiway probabilistic Monte Carlo model simulations. The SSB-attributable absolute burden per 1 million adults was calculated by dividing the absolute number of SSB-attributable cases by the total adult population (20+ years) in that year and multiplying it by 1 million. Source data are provided in Source Data file 6. CVD, cardiovascular disease; SSBs, sugar sweetened beverages; T2D, type 2 diabetes; UIs, uncertainty intervals.



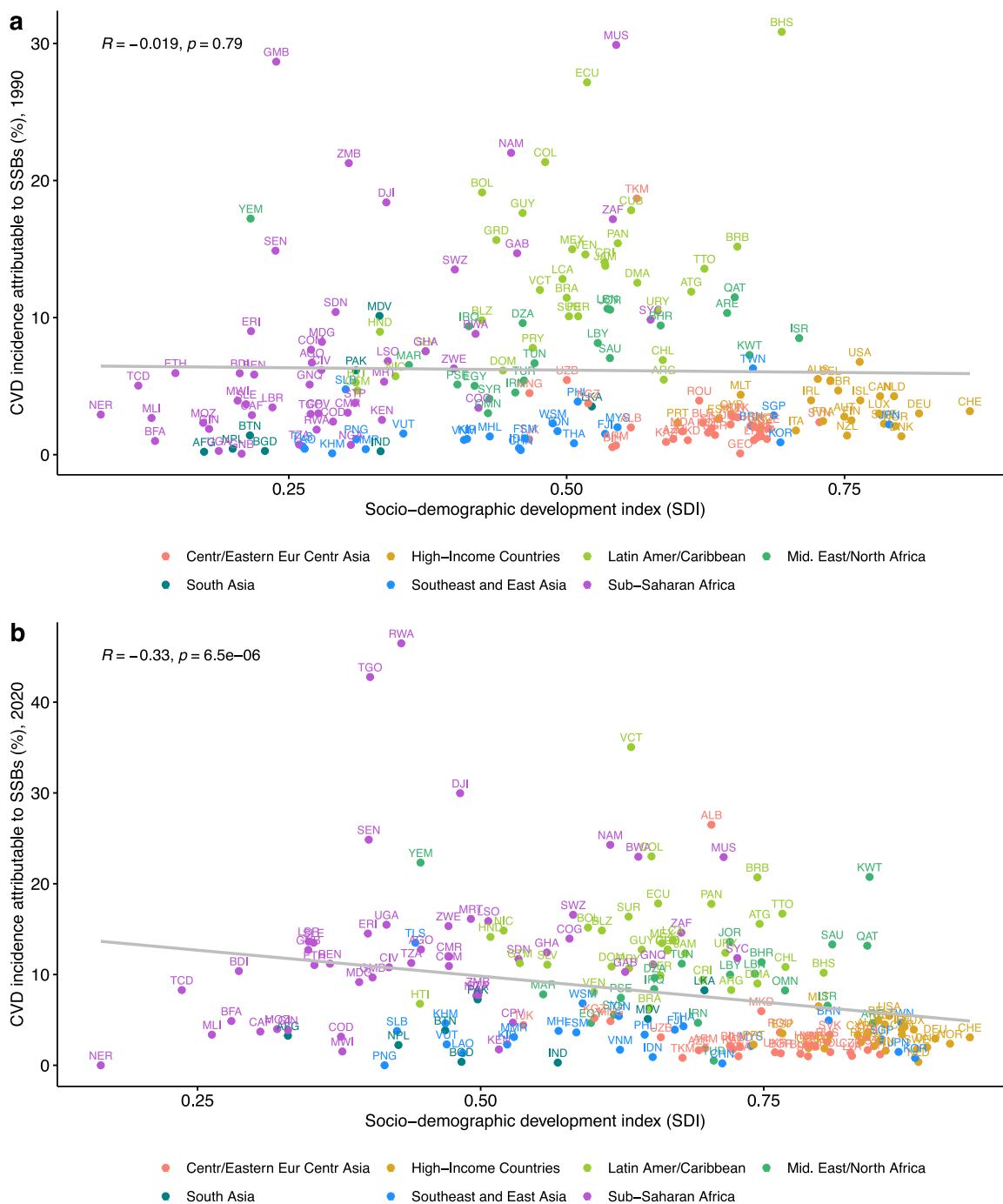
Extended Data Fig. 2 | Proportional incidence of T2D and CVD attributable to SSBs intake among adults (20+ years) jointly stratified by world region and age in 2020. The filled circles represent the central estimate (median) of the proportional SSB-attributable diabetes incidence in the top panel (a) CVD incidence in the bottom panel (b). The error bars represent the 95% UI derived from the 2.5th and 97.5th percentiles of 1,000 multiway probabilistic Monte Carlo model simulations. The age groups are 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+ years. In prior GDD reports, the

region Central/ Eastern Europe and Central Asia was referred as Former Soviet Union, and Southeast and East Asia was referred as Asia. See Supplementary Table 1 for a list of countries included in each world region. Source data are provided in Source Data file 7. Centr/East Eur Centr Asia, Central/Eastern Europe and Central Asia; CVD, cardiovascular disease; GDD, Global Dietary Database; Latin Amer/Caribbean, Latin America/Caribbean; SSBs, sugar sweetened beverages; T2D, type 2 diabetes; UIs, uncertainty intervals.



Extended Data Fig. 3 | National correlation of proportional SSB-attributable T2D incidence and SDI in 1990 and 2020. (a) 1990 and (b) 2020. Points represent the 184 countries included in this analysis (labeled with their ISO3 code and colored based on world region). The gray line represents the overall linear association, with Spearman correlation coefficient and associated P value (two-tailed) provided. No adjustments were made for multiple comparisons. SDI is a measure of a nation’s development expressed on a scale of 0 to 1 sourced

from the Global Burden of Disease study, based on a composite average of the rankings of income per capita, average educational attainment and fertility rates. In prior GDD reports, the region Central/ Eastern Europe and Central Asia was referred to as Former Soviet Union, and Southeast and East Asia was referred to as Asia. See Supplementary Table 1 for a list of countries included in each world region. Source data are provided in Source Data file 8.



Extended Data Fig. 4 | National correlation of proportional SSB-attributable CVD incidence and SDI at the national level in 1990 and 2020. (a) 1990 and (b) 2020. Points represent the 184 countries included in this analysis (labeled with their ISO3 code and colored based on world region). The gray line represents the overall linear association, with Spearman correlation coefficient and associated p value (two-tailed) provided. No adjustments were made for multiple comparisons. SDI is a measure of a nation’s development expressed on a scale of 0 to 1 sourced from the Global Burden of Disease study, based on a composite average of the rankings of income per capita, average educational attainment

and fertility rates. In prior GDD reports, the region Central/ Eastern Europe and Central Asia was referred as Former Soviet Union, and Southeast and East Asia was referred to as Asia. See Supplementary Table 1 for a list of countries included in each world region. Source data are provided in Source Data file 8. Cent/Eastern Eur Centr Asia, Central/Eastern Europe and Central Asia; CVD, cardiovascular disease; GDD, Global Dietary Database; Latin Amer/Caribbean, Latin America/Caribbean; SDI, sociodemographic development index; SSBs, sugar sweetened beverages; UIs, uncertainty intervals.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	No software was used for data collection for this analysis.
Data analysis	<p>Custom code was developed using R (Version 4.4.0) for this research including, calculation of age-adjusted relative risks, SSB intake gamma parameters, comparative risk assessment analysis (including the PAF and attributable mortality for each stratum), and data visualizations. Packages used include rworldmap package (v1.3-6). Given the computational size, calculation of the SSB intake gamma parameters, comparative risk assessment modeling, absolute and relative differences, and summary statistics were run on the Tufts University High Performance Computing Cluster (https://it.tufts.edu/high-performance-computing), supported by the National Science Foundation (grant:2018149, https://www.nsf.gov/awardsearch/showAward?AWD_ID=2018149&HistoricalAwards=false) under active development by Research Technology, Tufts Technology Services (https://it.tufts.edu/researchtechnology.tufts.edu).</p> <p>The statistical code can be made available to researchers upon request. Eligibility criteria for such requests include utilization for nonprofit purposes only, for appropriate scientific use based on a robust research plan, and by investigators from an academic institution. GDD will nominate co-authors to be included on any papers generated using GDD-generated statistical code. If you are interested in requesting access to the statistical code, please submit the following documents: (1) proposed research plan (please download and complete the proposed research plan form: https://www.globaldietarydatabase.org/sites/default/files/manual_upload/research-proposal-template.pdf), (2) data-sharing agreement (please download this form https://www.globaldietarydatabase.org/sites/default/files/manual_upload/tufts-gdd-data-sharing-agreement.docx and complete the highlighted fields, have someone who is authorized to enter your institution into a binding legal agreement with outside institutions sign the document. Note that this agreement does not apply when protected health information or personally identifiable information are shared), (3) email items (1) and (2) to info@globaldietarydatabase.org.</p>

Please use the subject line "GDD Code Access Request". Once all documents have been received, the GDD team will be in contact with you within 2-4 weeks regarding subsequent steps. Data will be shared as .csv or .xlsx files, using a compressed format when appropriate.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data used in this analysis are publicly available from the following sources: (1) population SSB intake distributions based on individual-level survey data from the GDD (<https://www.globaldietarydatabase.org/data-download>); (2) optimal SSB intake levels from previous analyses; (3) direct age-adjusted etiologic effects of SSBs on diabetes, ischemic heart disease, and ischemic stroke adjusted for BMI, and of weight gain on diabetes, ischemic heart disease, and ischemic stroke from previous meta-analyses and pooled analyses of prospective cohorts; as well as linear, BMI-stratified effects of SSBs on weight gain or loss; (4) population overweight (BMI ≥ 25 kg/m²) and underweight (BMI < 18.5 kg/m²) distributions from the NCD Risk Factor Collaboration (NCD-RisC: <https://ncdrisc.org/data-downloads.html>); (5) total diabetes, ischemic heart disease, and ischemic stroke incidence, DALYs, and mortality estimate distributions from the GBD study (<https://vizhub.healthdata.org/gbd-results/>); and (6) population demographic data from the United Nations Population Division (UN, <https://population.un.org/wpp/>), the Barro and Lee Educational Attainment Dataset 2013 (Barro and Lee, DOI: 10.3386/w15902), and Socio-demographic Index (SDI) data (Global Health Data Exchange: GBD, <https://ghdx.healthdata.org/record/ihme-data/gbd-2019-socio-demographic-index-sdi-1950-2019>).

The GDD SSB intake data was collapsed for 85+ years using the 4,000 simulations corresponding to the stratum level intake data derived from the Bayesian model. These data were used to obtain the gamma parameters of the SSB intake distribution used in the model. The 4,000 simulations files can be made available to researchers upon request. Eligibility criteria for such requests include utilization for nonprofit purposes only, for appropriate scientific use based on a robust research plan, and by investigators from an academic institution. If you are interested in requesting access to the data, please submit the following documents: (1) proposed research plan (please download and complete the proposed research plan form: https://www.globaldietarydatabase.org/sites/default/files/manual_upload/research-proposal-template.pdf), (2) data-sharing agreement (please download this form https://www.globaldietarydatabase.org/sites/default/files/manual_upload/tufts-gdd-data-sharing-agreement.docx and complete the highlighted fields, have someone who is authorized to enter your institution into a binding legal agreement with outside institutions sign the document. Note that this agreement does not apply when protected health information or personally identifiable information are shared), (3) email items (1) and (2) to info@globaldietarydatabase.org.

Please use the subject line "GDD Data Access Request". Once all documents have been received, the GDD team will be in contact with you within 2-4 weeks regarding subsequent steps. Data will be shared as .csv or .xlsx files, using a compressed format when appropriate.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

This study involved secondary data analysis from publicly available databases. No human research participants were involved in this original analysis. Sex was defined as female or male and this was extracted from the individual surveys as provided. No distinction between sex vs. gender was possible for this study.

Reporting on race, ethnicity, or other socially relevant groupings

This analysis is based in multiple inputs from published de-identified studies. The population characteristics for each study are available within each study publication/report of each input. In our findings we report estimates stratified by socially relevant groups including education level and area of residence.

Population characteristics

This analysis is based in multiple inputs from published de-identified studies. The population characteristics for each study are available within each study publication/report of each input.

Recruitment

The recruitment methodology for each study is available within each study publication/report of each input.

Ethics oversight

This investigation was exempt from ethical review board approval because it was based on published de-identified nationally representative data, without personally identifiable information.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Quantitative, comparative risk assessment (CRA) modeling analysis incorporating independently derived inputs and parameters on demographics, risk factors, their etiologic effects, and disease burdens to model attributable SSB cardiometabolic burdens at the global, regional, and national levels, and by key sociodemographic factors.
Research sample	This study involved a secondary data analysis from publicly available databases. The model estimated SSB-attributable cardiometabolic disease burdens for the global adult population (20+ years) stratified by sex (men and women), age (14 age groups), educational attainment (low, medium, high), and urbanicity (rural or urban), yielding 30,912 population strata across 184 countries in each year (1990 and 2020). No distinction between sex vs. gender was possible for this study. Sex was defined as female or male and this was extracted from the datasets as provided.
Sampling strategy	Given the nature of the study, using independently derived datasets, the sampling procedures varied across datasets including: dietary intake surveys, T2D incidence data, BMI surveys, estimated etiologic effects.
Data collection	<p>This study consisted on a secondary data analysis, thus no individual-level data was collected. Data used in this analysis are publicly available from the following sources: (1) population SSB intake distributions based on individual-level survey data from the GDD (https://www.globaldietarydatabase.org/data-download); (2) optimal SSB intake levels from previous analyses; (3) direct age-adjusted etiologic effects of SSBs on diabetes, ischemic heart disease, and ischemic stroke adjusted for BMI, and of weight gain on diabetes, ischemic heart disease, and ischemic stroke from previous meta-analyses and pooled analyses of prospective cohorts; as well as linear, BMI-stratified effects of SSBs on weight gain or loss; (4) population overweight (BMI ≥ 25 kg/m²) and underweight (BMI < 18.5 kg/m²) distributions from the NCD Risk Factor Collaboration (NCD-RisC: https://ncdrisc.org/data-downloads.html); (5) total diabetes, ischemic heart disease, and ischemic stroke incidence, DALYs, and mortality estimate distributions from the GBD study (https://vizhub.healthdata.org/gbd-results/); and (6) population demographic data from the United Nations Population Division (UN, https://population.un.org/wpp/), the Barro and Lee Educational Attainment Dataset 2013 (Barro and Lee, DOI: 10.3386/w15902), and Socio-demographic Index (SDI) data (Global Health Data Exchange: GBD, https://ghdx.healthdata.org/record/ihme-data/gbd-2019-socio-demographic-index-sdi-1950-2019).</p> <p>Our primary dietary input, SSB intakes from the GDD, is unique to this study and was modeled by our team. The GDD systematically searched for and compiled representative data on individual-level dietary intakes from national surveys and sub-national surveys. The final GDD model incorporated 1,224 dietary surveys representing 185 countries from seven world regions and 99.0% of the global population in 2020. Of these, 450 surveys reported data on SSBs, totaling 2.9 million individuals from 118 countries representing 86.8% of the global population. Most surveys were nationally or sub-nationally representative (94.2%), collected at the individual level (84.7%), and included estimates in, both urban and rural area of residence (61.6%). The global sample size consisted of 2,941,704 participants: 44.3% female and 55.7% male; 70% urban areas and 30% from rural areas; 16% low, 37.6% medium, and 46.4% high education; and 53% adults and 47% children/adolescent. Further details on characteristics of surveys with data on SSBs, including availability of surveys per world region, is available in Supplementary Table 1.</p>
Timing	The estimates reported in this study correspond to years 1990 and 2020.
Data exclusions	Of 188 countries with dietary survey data, three were dropped from the GDD estimation model due to unavailability of FAO food availability data (Andorra, Democratic People's Republic of Korea, and Somalia); and 1 additional country (South Sudan) was excluded as it lacked overweight/obesity estimates from the NCD Risk Factor Collaboration study.
Non-participation	No human research participants were involved in this original analysis.
Randomization	This study is a comparative risk assessment analysis and thus randomization is not applicable

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

Methods

- n/a Involved in the study
- Antibodies
 - Eukaryotic cell lines
 - Palaeontology and archaeology
 - Animals and other organisms
 - Clinical data
 - Dual use research of concern
 - Plants

- n/a Involved in the study
- ChIP-seq
 - Flow cytometry
 - MRI-based neuroimaging