G-Computation Quantifying Caries Reduction by World Health Organization Sugar Limits in Children

A.T.M. Dao[1,2,3] (iD), L.G. Do[1] (iD), N. Stormon[1,4] (iD),
H.V. Nguyen[2,3,5,6], and D.H. Ha[1]

Abstract

The World Health Organization (WHO) recommends reducing free sugar intake (FSI) to below 10%, and ideally below 5%, of the estimated energy requirement (EER) to prevent noncommunicable diseases, including dental caries. Little progress has been made in lowering FSI to reduce early childhood caries (ECC) over the past few decades. Although sugar's impact on health is well established, no studies have quantified the extent to which reducing FSI to these thresholds would reduce decayed, missing, and filled surfaces (dmfs) scores. Data from 2,182 Australian children in the SMILE birth cohort were analyzed using G-computation analysis to estimate counterfactual dmfs if FSI at age 2 y had been reduced below 10% or 5% of the EER. A dose-response analysis using restricted cubic splines was also conducted to empirically assess thresholds by modeling dmfs at age 5 y as a function of continuous FSI. The G-computation results indicate that, in the general preschool population, reducing FSI to less than 10% or 5% of the EER would result in an absolute reduction (AR) in dmfs of 1.3 and 1.5, respectively, corresponding to attributable fractions among the exposed (AFE) of 84% and 97%. Among high-risk children, the estimation impact is even greater, with ARs in dmfs of 4.4 to 4.5 and AFEs ranging from 75% to 99%. The dose-response analysis identified an empirical threshold of approximately 6.25 g/d, equivalent to 2.5% of the EER, where dmfs began to increase most steeply, which is lower than the WHO cutoffs. These findings underscore the importance of reducing FSI to below 10% of the EER for all children and suggest a target below 5% for high-risk groups. The results offer evidence to support clinical guidance and population-level interventions to lower FSI in early childhood. Future research should test these findings in diverse sociocultural settings, including children and adults in low- and middle-income countries, to strengthen the evidence for global dietary sugar-reduction policies.

Keywords: dietary sugars, dental caries, preschool child, primary dentition, counterfactual analysis, causal inference

Introduction

Early childhood caries (ECC) represents a significant global public health challenge, affecting approximately 48% of children worldwide (Uribe et al 2021). Free sugar intake (FSI), a critical factor causing dental caries (Moynihan and Kelly 2014), has been identified as an important mediator in the socioeconomic status (SES)–ECC pathway (Dao et al 2025b), accounting for 52% of the total effect. A 1-unit SES increase corresponded to a 6% reduction in ECC risk via FSI (Dao et al 2025b). Reducing FSI is recommended to reduce not only levels of dental caries (World Health Organization [WHO] 2015) but also diminish the level of inequality in dental health (Dao et al 2025a).

The WHO recommends limiting FSI to less than 10% of the total estimated energy requirement (EER), with further reductions below 5% providing additional health benefits such as reducing levels of noncommunicable diseases, including dental caries (WHO 2015). However, robust evidence supporting these thresholds remains limited (Yan and Louie 2023). The <10% EER recommendation was based on a systematic review of 8 longitudinal studies that mainly assessed associations rather than causal effects (WHO 2015; Moynihan 2016; Dao

et al 2024), while the <5% EER guideline relied on very-low-quality evidence from ecological studies (WHO 2015). None of these studies quantified the extent of reduction, a critical component for clear guidance, nor provided tangible estimates of FSI reduction benefits (Yan and Louie 2023). Furthermore,

[1]School of Dentistry, Faculty of Health and Behavioural Sciences, The University of Queensland, Brisbane, Queensland, Australia
[2]University of Cuu Long, Vinh Long, Vietnam
[3]Division of Epidemiology, Department of Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, Worcester, MA, USA
[4]Community and Oral Health, Metro North Hospital and Health Services, Queensland, Australia
[5]Health Innovation and Transformation Centre (HITC), Federation University, Australia
[6]School of Public Health, The University of Queensland, Australia

A supplemental appendix to this article is available online.

Corresponding Author:
A.T.M. Dao, School of Dentistry, Faculty of Health and Behavioural Sciences, The University of Queensland, 288 Herston Road, Brisbane, Queensland 4006, Australia.
Email: t.dao@uq.edu.au

most evidence focuses on sugar-sweetened beverages (SSBs), with limited data on FSI from all food sources (Yan and Louie 2023).

In the Australian context, little progress has been made in reducing FSI levels. Between 2007 and 2012, the proportion of children consuming FSI exceeding the WHO-recommended thresholds of <10% and <5% of total energy intake declined only slightly, from more than 68.0% and 98.0% in 2007 (Louie et al 2016) to 60.6% and 93.4% in 2011 to 2012, respectively (Lei et al 2016). Similarly, caries prevalence among 5- to 6-y-olds declined by less than 10.0% during this period (Mejia et al 2012), with 34.3% of Australian preschoolers still experiencing caries (Do et al 2016).

To strengthen global evidence supporting WHO's FSI thresholds and to inform strategies for reducing FSI among Australian children, it is essential to generate robust causal evidence addressing existing knowledge gaps. Using data from the Study of Mothers' and Infants' Life Events Affecting Oral Health (SMILE) birth cohort, this study applies a causal inference framework to evaluate the potential impact of targeting FSI reduction on ECC.

Specifically, G-computation was used to estimate the potential effect of reducing FSI at age 2 y on the occurrence of ECC at age 5 y in the overall population of preschool children. In addition, an exploratory subgroup analysis was conducted to examine the effect within a high-risk population, defined as children who exceeded the recommended FSI threshold by age 2 and developed ECC by age 5.

Methods

This study conformed to the guidelines for Strengthening the Reporting of cohort studies in Epidemiology (von Elm et al 2007). It also incorporated the target trial framework for causal inference using observational data (Hernan et al 2025).

Data Source

The SMILE is a longitudinal observational study that recruited mother–child dyads from 3 main hospitals in Adelaide between 2013 and 2014. It was funded by National Health and Medical Research Council (NHMRC) Project Grant #1046219 (2013 to 2016) and conducted with the approval of ethical research committees (Do et al 2014). A total of 2,182 children were recruited at baseline and followed up at 3, 6, 12, 24, and 60 mo with informed consent. For this study, FSI at age 2, dental caries at age 2 and 5, and plaque index at age 2 and 5 were analyzed using only completed relevant data.

Study Design

The analysis addresses the research question by quantifying the reduction in dental caries resulting from lowering FSI to below 10% and 5% of the EER through 2 populations: (1) general preschoolers and (2) the high-risk subgroup, children who exceeded the recommended FSI threshold by age 2 and developed ECC by age 5. For the general population, it

estimates the change in dmfs scores at age 5 if the entire preschool population had reduced their FSI intake from ≥10% or ≥5% to below these thresholds. For the high-risk subgroup, it assesses how their dmfs scores would have differed had their FSI been reduced below 10% or 5%.

A causal inference approach using g-computation was applied in 3 stages (Snowden et al 2011). In the first stage, a regression model was fitted to predict the effect of exposure (FSI at age 2) on the outcome (ECC at age 5). The second stage estimates 2 counterfactual outcome values, derived by conditioning on the presence or absence of the risk factor/exposure across the entire study population, based on the model's parameters fit in the first stage. In the third stage, the change in the outcome was estimated by comparing these 2 counterfactual outcome values. The absolute reduction (AR) represented the magnitude of change, while the attributable fraction among the exposed (AFE) was calculated by dividing the AR by the dmfs scores of the population exposed to the risk factor at age 2 (Fig. 1).

Measurements

Outcome: ECC. ECC is defined as the presence of 1 or more carious lesions, white spot lesions, tooth loss due to caries, or filled surfaces due to caries in the primary teeth of a child younger than 6 y (WHO). ECC was evaluated through oral epidemiological examinations conducted by trained dental examiners following a standardized protocol when the children turned 2 and 5 y of age. For each child, any tooth surface affected by decay (cavitated or noncavitated), missing due to caries, or filled was recorded and assigned a single score. The cumulative score, known as the score of decayed, missing, and filled surfaces (dmfs score) (Do et al 2020), was used to create a continuous ECC variable at age 5 y.

Exposure: FSI. FSI is defined as the consumption of any type of sugar added to foods and beverages. This includes Sugar-Sweetened Beverages (SSBs) such as fruit-based and milk-based sweetened drinks and 100% fruit juices, as well as confectionery, cakes, biscuits, sweetened cereals, sweet desserts, sucrose, honey, syrups, and preserves. It does not include naturally occurring sugars found in fruits, vegetables, and milk (WHO 2015).

At age 2 y, FSI was assessed using 89-item Food Frequency Questionnaire (FFQ), whose repeatability and validity were evaluated against 24-h dietary recalls (Devenish et al 2017; Devenish et al 2019). FSI values were extracted from these food items and converted into milligrams per day (mg/day) as a continuous variable, and outliers exceeding the 99th percentile were excluded due to the highly right-skewed distribution.

For G-computation analysis, the continuous FSI variable was categorized into a binary variable, 1 (≥10% or ≥5% of EER) and 0 (<10% or <5% of EER). Based on the Nutrient Reference Values for Australia and New Zealand Including Recommended Dietary Intakes Australia, the daily total EER is 4,200 kJ for girls and 4,400 kJ for boys (NHMRC et al 2006). Given that 17 kJ is equivalent to 1 g of free sugar, 10% of EER
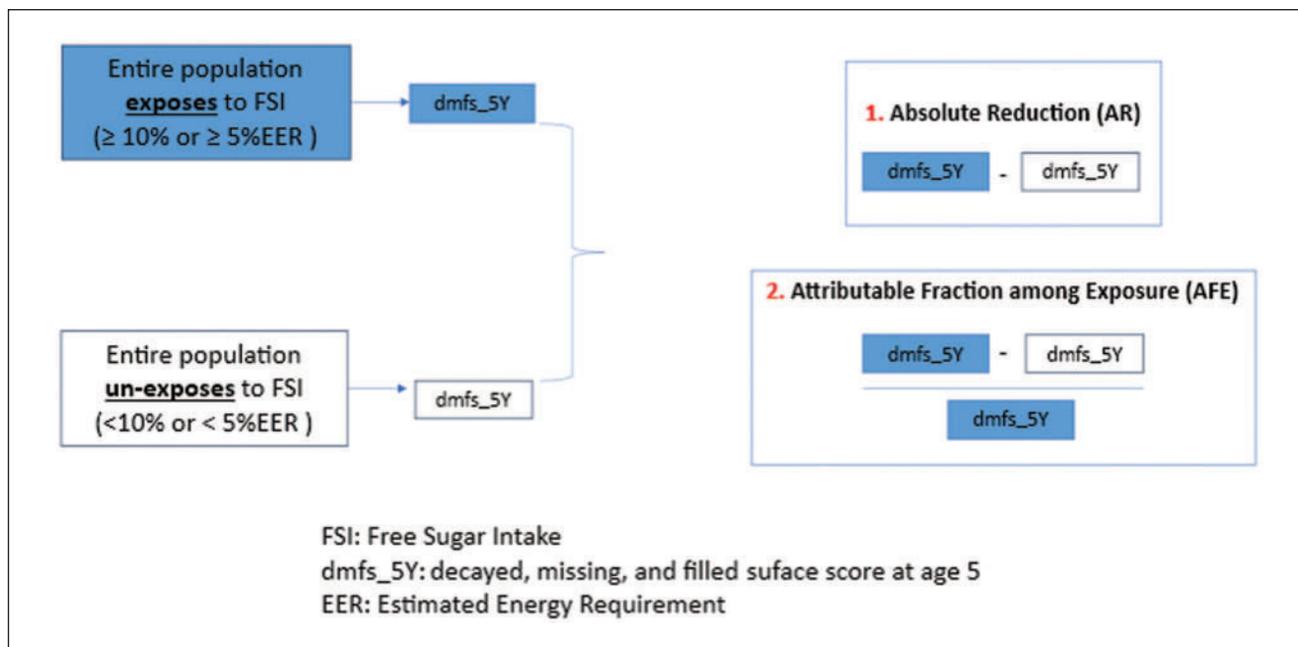
Figure 1. Counterfactual framework for G-computation analysis.

corresponds to 24.7 g/d for girls and 25.9 g/d for boys. Similarly, the 5% threshold corresponds to 12.4 g/d for girls and 12.9 g/d for boys. These thresholds were used to create the binary FSI variables for the analysis.

Confounders. SES at baseline (Dao et al 2024; Dao et al 2025b), ECC at age 2 y (Lin et al 2021), gender (Ferraro et al 2010; Ra and Park 2022), dental visits pattern (Andrew et al 2021) during the first 5 y, and plaque index at ages 2 and 5 y (Marsh 2010) were identified as potential confounding variables. These 6 factors were therefore included in the analysis to control for confounding (Fig. 2).

SES is defined as one's access to social-economic resources across various dimensions (Antonoplis 2023). It is measured through multiple dimensions and considered a covariate for the causal relationship between FSI and ECC. In this study, SES was represented using a composite created by Dao et al through a hybrid approach (Dao et al 2025a), described in detail elsewhere. In brief detail, at baseline, mothers or caregivers answered 21 SES-related questions, focusing on 12 SES indicators. Of these, 11 were potentially relevant and were used as inputs for principal component analysis (PCA) to create the SES composite. This composite was generated by summing the standardized values of the 5 significant loading indicators,
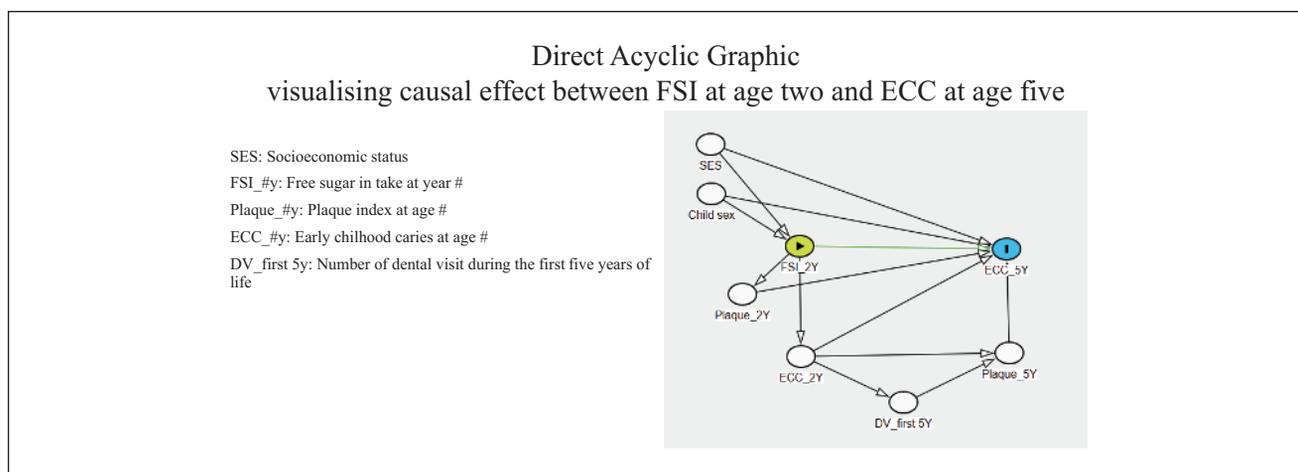

Figure 2. Direct acyclic graphic: visualizing the causal effect of free sugar intake (FSI) at age 2 y on early childhood caries (ECC) at age 5. DV, dental visit; SES, socioeconomic status.

family education, work, occupation, income, and health insurance, from the first PCA component, each weighted for its loading value. The composite was used as a continuous SES variable measured by the standardized score.

ECC experience at age 2 y was obtained by a clinical exam and described above. Plaque presence was assessed during the oral clinical examinations at ages 2 and 5 y and was classified into 4 categories: 0, no plaque; 1, thin plaque detected with dental probing; 2, visible plaque without probing; and 3, abundant plaque (Do et al 2020). The plaque variable was then treated as a binary variable, with the first 2 categories defined as "invisible plaque" and the latter two as "visible plaque." Dental visit pattern was assessed through a survey administered when the child turned 5 y. Parents were asked to list all dental visits during their child's first 5 y. The total number of dental visits was calculated by summing all reported visits (Do et al 2020).

Statistical Analyses

The model was fitted using a negative binomial regression in Stata (nbreg) (UCLA 2024), as dmfs is count data and exhibits overdispersion (variance greater than the mean) (equation 1). Based on the model coefficients, counterfactual dmfs values at age 5 y corresponding to FSI levels of ≥10% versus <10% and ≥5% versus <5% were estimated as log-transformed values, following equations 2 and 3, respectively. The comparison reduction or change in dmfs score at age 5 y was first estimated by using log-transformed dmfs values. The AR in the observed dmfs scores was then calculated by exponentiating the changes in the log values given the real values have skew distributions (equation 4), while the proportion of reduction or the AFE was determined by dividing the AR by the dmfs score for individuals exposed to FSI ≥10% or 5% of the EER, as presented in equation 5.

A command nbreg was first fitted without adjusting for confounding factors. It was then adjusted for SES and, finally, for all 6 potential confounders. The final model was selected based on the values of the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The AIC penalizes the number of parameters to prevent overfitting and is based on information theory. BIC also penalizes model complexity but does so more strongly, as it incorporates sample size into the penalty term. The model with the lowest AIC and BIC was considered the best-fitting model and was selected for estimating counterfactual values in the G-computation approach, following the general rule of thumb (Pham 2019).

$$
\begin{aligned}
logECC_{5Y} = \beta_0 &+ \beta_1.FSI + \beta_2.SES \\
&+ \beta_3.SES + \beta_4.DV_{5Y} + \beta_5.ECC_{2Y} + \beta_6.Sex \\
&+ \beta_7.plaque\_2Y + \beta_8.plaque\_5Y
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
logECC_{5Y} \mid FSI_{2Y} \geq WHO_{thresholds} &= \beta_0 + \beta_1 FSI \\
+ \beta_2 SES + \beta_3 DV_{5Y} + \beta_4 ECC_{2Y} &+ \beta_5 Sex \\
+ \beta_6 plaque\_2Y + \beta_7 plaque\_5Y
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
logECC_{5Y} \mid FSI_{2Y} &< WHO_{thresholds} \\
= \beta_0 + \beta_2 SES &+ \beta_3 DV_{5Y} + \beta_4 ECC_{2Y} \\
+ \beta_5 Sex + \beta_6 plaque\_2Y &+ \beta_7 plaque\_5Y
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
AR = logECC_{5Y} \mid FSI_{2Y} &\geq WHO_{thresholds} \\
- logECC_{5Y} \mid FSI_{2Y} &< WHO_{thresholds}
\end{aligned} \tag{4}
$$

$$
AFE = AR / logECC_{5Y} \mid FSI_{2Y} \geq WHO_{thresholds} \tag{5}
$$

Converting continuous FSI into a binary variable for G-computation results in some loss of granularity. However, this tradeoff reflects real-world public health recommendations (<10% EER) and facilitates the translation of our findings into actionable intervention scenarios. To balance this limitation and empirically assess WHO recommended thresholds, we conducted a dose–response analysis (Steenland and Deddens 2004) using restricted cubic splines functions with 4 knots at placed at the 5th, 35th, 65th, and 95th percentiles of the FSI distribution, modeling dmfs at age 5 as a function of continuous FSI.

Results

Of 2,182 children recruited at baseline, 879 completed clinical examination at age 5 y (Appendix Fig. 1). The completers were more likely have higher SES than those who dropped out before age 5 (Appendix Table 1).

The distributions of study variables including exposure FSI at age 2 y, dmfs at age 5 y, covariate SES at baseline, and 5 confounders are presented in Table 1. Of the 2,182 recruited children, response rates varied across these variables: 1,047 provided data for 89 FFQ items asking about FSI (excluding 10 cases provided values exceeded 99th percentile), 1,878 responded to all 21 SES items at baseline, 824 had information on the number of dental visits in the first 5 y, and 1,394 had data on gender. For clinical examinations, dmfs scores were recorded for 1,040 children at age 2 y and 879 at age 5 y, while plaque index measurements were available for 1,041 and 866 children, respectively.

Of 1,047 children having FSI measured at age 2 y via the 89 FFQ items, their FSI distribution showed a right-skewed distribution (Appendix Fig. 2a) with a median intake of 22.6 g/d. Of these children, 464 (44.3%) and 787 (75.2%) reported having an FSI exceeding the 10% and 5% of the EER, respectively (Table 1), with no significant FSI difference between boys and girls (Appendix Table 2). The ECC score distribution at age 5 y was also right skewed (Appendix Fig. 2b), with a median of 0 and a range of 0 to 60.0, with 32.5% of children experiencing ECC age 5 y. In contrast, SES distribution at baseline was approximately normal (Appendix Fig. 2c).

A total of 541 children had completed data for FSI at age 2 y and dmfs at age 5 y, SES at baseline, and 5 potential confounders. Within this subgroup, the high-risk population was defined as follows: 78 children who consumed FSI exceeding

Table 1. Distribution of the Study Variables.

	ECC			FSI at Age 2 y					Plaque	
	Age 2 y	Age5 y			EER				Age 2 y	Age 5 y
			SES Baseline (Standardized Score)	Grams per Day	≥10%	≥5%	DV_5Y (Times)	Sex (Male: 1; Female: 0)		
Statistics	(dmfs Score)				(Yes: 1 or No: 0)				(Categories: 0, 1, 2, 3)	
Among recruited participants										
p50	0.0	0.0	0.3	22.6	0.0	1.0	1.0	0.0	0.0	1.0
Mean/%	0.5	1.7	0.2	29.7	44.3	75.2	1.8	0.5	0.4	1.2
SD	2.7	4.8	1.5	24.8	—	—	1.7	—	0.6	0.6
Min	0.0	0.0	−4.2	0.3	0.0	0.0	0.0	0.0	0.0	0.0
Max	72.0	60.0	3.2	155.1	1.0	1.0	9.0	1.0	2.0	3.0
Skewness	19.2	5.8	−0.4	2.0	—	—	0.9	—	0.9	0.2
Kurtosis	489.9	48.3	2.7	8.1	—	—	3.1	—	2.8	3.1
n	1,040	879	1,878	1,047			824	1,394	1,041	866
Among those completing information of FSI, ECC_5Y, and confounders										
p50	0.0	0.0	0.8	20.4	0.0	1.0	1.0	0.0	0.0	1.0
Mean	0.5	1.4	0.7	27.4	39.9	70.8	1.8	0.5	0.4	1.2
SD	3.4	4.5	1.3	23.0	—	—	1.7	—	0.5	0.6
Min	0.0	0.0	−2.7	0.7	0.0	0.0	0.0	0.0	0.0	0.0
Max	72.0	60.0	3.0	153.7	1.0	1.0	6.0	1.0	2.0	3.0
Skewness	18.1	7.3	−0.5	2.0	—	—	0.9	—	0.8	0.3
Kurtosis	379.4	73.6	2.6	8.5	—	—	3.0	—	2.5	3.4
n	541									
Among high-risk group (FSI ≥ 10% EER and ECC_5Y > 0) with completed confounders' information										
p50	0.0	2.5	0.5	41.2	1.0	1.0	2.0	0.0	0.0	1.0
Mean	1.5	5.9	0.3	50.5	1.0	1.0	1.9	0.4	0.5	1.4
SD	8.2	9.8	1.4	26.1	—	—	1.7	—	0.6	0.7
Min	0.0	1.0	−2.6	24.8	1.0	1.0	0.0	0.0	0.0	0.0
Max	72.0	60.0	2.8	133.1	1.0	1.0	6.0	1.0	2.0	3.0
Skewness	8.2	3.3	−0.4	1.7	—	—	0.6	—	0.7	0.3
Kurtosis	70.5	15.4	2.4	5.4	—	—	2.4	—	2.4	3.0
n	78									
Among high-risk group (FSI ≥ 5% EER and ECC_5Y > 0) with completed confounders' information										
p50	0.0	2.0	0.7	30.0	1.0	1.0	2.0	0.0	0.0	1.0
Mean	1.1	4.8	0.5	37.2	0.6	1.0	2.0	0.4	0.5	1.4
SD	6.4	7.8	1.3	25.6	—	—	1.7	—	0.6	0.7
Min	0.0	1.0	−2.6	13.0	0.0	1.0	0.0	0.0	0.0	0.0
Max	72.0	60.0	2.9	133.1	1.0	1.0	6.0	1.0	2.0	3.0
Skewness	10.5	4.2	−0.4	1.9	—	—	0.6	—	0.5	0.2
Kurtosis	116.1	24.6	2.6	6.7	—	—	2.4	—	2.1	2.9
n						133				

—, not applicable; DV: dental visit; ECC, early childhood caries; ECC_5Y, ECC at age 5 y; EER: estimated energy requirement at age 2 y (girls: 4,200 kJ/d; boys: 4,400 kJ/d); FSI: free sugar intake; SES: socioeconomic status; n, sample size.

10% of EER and experienced ECC at age 5 y or 133 children who consumed FSI exceeding 5% of EER and experienced ECC at age 5 y.

FSI levels at age 2 and dmfs levels at age 5 among the 541 children were comparable with those of the overall sample. However, among the high-risk population, FSI levels were nearly double those of the overall study sample, with a median of 41.2 g/d for children consuming ≥10% of EER and 1½ times higher, with a median of 30 g/d, for those consuming ≥5% of EER. Similarly, dmfs levels of the high-risk group were substantially higher than those observed in the overall study sample.

Negative bionomial regression (nbreg), estimating the effects of FSI at age 2 y on dmfs at age 5 y, both with and

Table 2. Change in dmfs Score among the General Population when FSI Changed from ≥10% to Less than 10% or from ≥5% to Less than 5%.

ECC at Age 5 y (dmfs Score)		p50	Mean	SD	Min	Max	Skewness	Kurtosis	P value (Observed vs Counterfactual)
FSI ≥ 10% of the EER (*n* = 541)									
Log values	Counterfactual[a]	−0.5	−0.3	0.8	−1.7	2.6	1.2	4.5	0.00001 (*t* test)
	Counterfactual[b]	−0.3	−0.1	0.9	−1.7	3.1	1.1	4.3	
Real values	Counterfactual[c]	0.6	1.2	1.8	0.2	13.8	4.0	20.5	0.00001 (signed-rank test)
	Counterfactual[d]	0.8	1.5	2.4	0.2	21.6	4.3	24.4	
Change of dmfs score at age 5 y when FSI reduces from ≥10% to less than 10%									Proportion of reduction
Log value change		0.0	0.2	0.2	0.0	0.5	0.4	1.2	
Absolute difference		1.0	1.3	0.3	1.0	1.6	0.4	1.2	(1.25 × 100)/ 1.49 = 83.9

dmfs score at age 5 y		p50	Mean	SD	Min	Max	Skewness	Kurtosis	P value (Observed vs Counterfactual)
FSI ≥ 5% of the EER (*n* = 541)									
Log values	Counterfactual[a]	−0.7	−0.5	0.9	−1.9	2.5	1.2	4.5	0.00001 (*t* test)
	Counterfactual[b]	−0.3	−0.1	0.9	−1.9	3.0	1.0	4.2	
Real values	Counterfactual[c]	0.5	1.0	1.6	0.2	12.4	4.0	20.4	0.00001 (signed-rank test)
	Counterfactual[d]	0.8	1.5	2.4	0.2	19.5	4.1	21.9	
Change of dmfs score at age 5 y when FSI reduces from ≥ 5% to under 5%									Proportion of reduction
Log value change		0.5	0.4	0.2	0.0	0.5	−0.9	1.8	
Absolute difference		1.6	1.5	0.3	1.0	1.6	−0.9	1.8	(1.45 × 100)/ 1.50 = 96.7
FSI ≥ 2.5% of the EER (*n* = 541)									
Log values	Counterfactual[a]	−2.1	−1.9	0.8	−3.2	1.1	1.2	4.6	0.00001 (*t* test)
	Counterfactual[b]	−0.3	−0.2	1.0	−3.2	2.9	0.3	4.4	
Real values	Counterfactual[c]	0.1	0.2	0.4	0.0	2.9	3.9	19.9	0.00001 (signed rank test)
	Counterfactual[d]	0.8	1.5	2.5	0.0	18.4	3.9	20.2	
Change of dmfs score at age 5 y when FSI reduces from ≥2.5% to less than 2.5%									not applicable
Log value change		1.9	1.7	0.5	0.0	1.9	−3.3	11.9	
Absolute difference		6.4	6.0	1.4	1.0	6.4	−3.3	11.9	

Dmfs, decayed, missing, and filled surfaces of primary tooth/teeth; ECC, early childhood caries; EER, estimated energy requirement; FSI, free sugar intake; *n*, sample size.

EER for 2-y-olds: girls: 4,200 kJ/d; boys: 4,400 kJ/d.

Counterfactual[a] = log values of ECC at age 5 y estimated from the developed model, conditioning that FSI level either less than 10% or 5% of the EER= $logECC_{5Y} \mid FSI_{2Y} < WHO_{thresholds} = \beta_0 + \beta_2.SES + \beta_3 DV_{5Y} + \beta_4 ECC_{2Y} + \beta_5 Sex + \beta_6 plaque\_2Y + \beta_7 plaque\_5Y$ (3).

Counterfactual[b] = log values of ECC at age 5 y estimated from the developed model, conditioning that FSI level either from and greater than 10% or 5% of the EER= $logECC_{5Y} \mid FSI_{2Y} \geq WHO_{thresholds} = \beta_0 + \beta_1.FSI + \beta_2.SES + \beta_3 DV_{5Y} + \beta_4 ECC_{2Y} + \beta_5 Sex + \beta_6 plaque\_2Y + \beta_7 plaque\_5Y$ (2).

Counterfactual[c] = exp (log [counterfactual[a]]).

Counterfactual[d] = exp (log [counterfactual[b]]).

Log value change = counterfactual[b] − counterfactual[a].

Absolute difference = exp (log value change).

without adjustment for confounders, were conducted. The fully adjusted model, which included all 6 confounders, showed the best fit, as indicated by the lowest AIC and BIC values (Appendix Table 3). Consuming FSI at levels ≥10% or ≥5% of EER significantly increased dmfs at age 5 y. Specifically, children consuming FSI ≥10% or ≥5% of EER had dmfs scores nearly 0.5 units higher compared with those consuming FSI below these thresholds (Appendix Table 4).

The results of dmfs score reduction for the general and high-risk populations are presented in Tables 2 and Table 3, respectively. Findings consistently show that if either the general population or the high-risk population reduced their FSI intake from ≥10% of EER to below this threshold, the mean difference in dmfs score at age 5 y would decrease by 1.3 or 4.4, representing approximately an 84% or 75% dmfs

reduction among exposure group, respectively, compared with when the population consumes ≥10% of EER. Notably, if FSI intake were reduced to below 5% of EER, the dmfs score at age 5 y would decrease by 1.5 or 4.7, corresponding to nearly a 97% or 99% reduction, respectively, compared with when the population consumes ≥5% of EER. The dose-response analysis revealed an empirical threshold of FSI at ~6.25 g/d (≈2.5% of EER) where dmfs began to increase most steeply, which is lower than the WHO 5% cutoff (Appendix Fig. 3).

Discussion

This study provides robust evidence quantifying the potential change in dmfs scores among the preschool population if their FSI levels were reduced to below either 10% or 5% of the

Table 3. Change in dmfs Score among the High-Risk Population when FSI Changed from ≥10% to Less than 10% or from ≥5% to Less than 5%.

dmfs Score at Age 5 y		p50	Mean	SD	Min	Max	Skewness	Kurtosis	P Value (Observed vs Counterfactual)
High-risk population who both consumed FSI ≥ 10% of the EER and experienced ECC at age 5 y ($n = 78$)									
Log values	Counterfactual[a]	0.9	1.1	1.1	0.0	4.1	0.9	3.2	0.00001 (*t* test)
	Counterfactual[b]	−0.1	0.1	1.0	−1.5	2.6	0.8	2.9	
Real values	Counterfactual[c]	2.5	5.9	9.8	1.0	60.0	3.3	15.4	0.00001 (signed-rank test)
	Counterfactual[d]	0.9	1.9	2.6	0.2	13.3	2.5	9.2	
Change of dmfs score at age 5 y when FSI reduces from ≥ 10% to less than 10%									Proportion of reduction
Log value change		1.3	1.0	1.1	−2.4	2.9	−0.5	2.9	
Absolute difference		3.5	4.4	4.4	0.1	18.4	1.5	4.7	(4.44 × 100)/5.92 = 75.0
High-risk population who both consumed FSI ≥ 5% of the EER and experienced ECC at age 5 y ($n = 133$)									
Log values	Counterfactual[a]	0.7	1.0	0.9	0.0	4.1	1.0	3.6	0.00001 (*t* test)
	Counterfactual[b]	−0.3	−0.1	1.0	−1.7	2.5	0.7	2.8	
Real values	Counterfactual[c]	2.0	4.8	7.8	1.0	60.0	4.2	24.6	0.00001 (signed-rank test)
	Counterfactual[d]	0.7	1.6	2.2	0.2	11.9	2.6	9.7	
Change of dmfs score at age 5 y when FSI reduces from ≥ 5% to less than 5%									Proportion of reduction
Log value change		1.2	1.1	1.1	−2.3	3.0	−0.5	3.1	
Absolute difference		3.3	4.7	4.7	0.1	20.8	1.8	5.6	(4.71 × 100)/4.78 = 98.5

dmfs, decayed, missing, and filled surfaces of primary tooth/teeth; ECC, early childhood caries; EER: estimated energy requirement; FSI: free sugar intake; *n*, sample size.
High-risk population: children who consumed FSI from and over either 10% or 5% of the EER and experienced dental caries at age 5 y. EER for 2-y-olds: girls: 4,200 kJ/d; boys: 4,400 kJ/d.
Counterfactual[a] = log values of ECC at age 5 y estimated from the developed model, conditioning that FSI level either from and greater than 10% or 5% of the EER = $logECC_{5Y} \mid FSI_{2Y} \geq WHO_{thresholds} = \beta_0 + \beta_1 FSI + \beta_2 SES + \beta_3 DV_{5Y} + \beta_4 ECC_{2Y} + \beta_5 Sex + \beta_6 plaque\_2Y + \beta_7 plaque\_5Y$ (2).
Counterfactual[b] = log values of ECC at age 5 y estimated from the developed model, conditioning that FSI level either less than 10% or 5% of the EER= $logECC_{5Y} \mid FSI_{2Y} < WHO_{thresholds} = \beta_0 + \beta_2 SES + \beta_3 DV_{5Y} + \beta_4 ECC_{2Y} + \beta_5 Sex + \beta_6 plaque\_2Y + \beta_7 plaque\_5Y$ (3).
Counterfactual[c] exp (log (counterfactual[a])).
Counterfactual[d] exp (log (counterfactual[b])).
Log value change = counterfactual[a] − counterfactual[b].
Absolute difference = exp (log value change).

EER, addressing the research question: "How would dmfs scores among preschoolers have changed if FSI levels had been reduced to below either 10% or 5% of the EER?" The findings indicate that reducing FSI from ≥10% to below this threshold would decrease dmfs scores at age 5 y by nearly 84% within the exposed of the general population and 75% in the high-risk population. Further reducing FSI to less than 5% of the EER would result in a nearly 97% and 99% reduction, respectively.

These findings are robust evidence supporting the WHO's recommended FSI thresholds. First, the study captured FSI from all food sources, not just SSBs, which have been the narrow focus of much previous research (Yan and Louie 2023). Second, using the G-computation framework, an advanced causal approach (Hernán and Robins 2011), and birth cohort data, it estimated both absolute and relative reductions in dmfs scores. These findings add important specificity to the otherwise general WHO recommendations, particularly the <5% of EER threshold (WHO 2015). Specifically, the study quantified the percentage reductions for each threshold, supporting the application of the 10% threshold for the general preschool population and suggesting a stricter 5% threshold for high-risk preschool children. Furthermore, a dose-response analysis, a data-driven validation approach, identified an empirical threshold of 6.25 g/d.

This value, lower than the WHO recommendation, may reflect the greater susceptibility of preschool children to caries, further reinforcing the WHO's guidance.

The current study also represents a paradigm shift in caries prevention by integrating stages of translational science (Nazarian 2023), building on evidence that SES is a fundamental driver of dental health inequalities (Fisher-Owens et al 2007; Do et al 2022) and that FSI is a key mediator linking SES to ECC (Dao et al 2025b), our study advances to estimate the potential causal effect of reducing FSI on the burden of ECC in both general and high-risk populations. Overall, the study offers actionable insights for a range of stakeholders: health care providers can incorporate these results into routine pediatric dental care and nutritional counseling, policymakers can implement WHO's recommended thresholds to different populations and settings, and schools, daycare centers, and health care providers can collaborate to promote low-FSI diets among young children.

The study faced several challenges, such as attrition and recall bias. Attrition, common in cohort studies, was addressed by oversampling low-SES mothers, as they were more likely to be lost to follow-up (Do et al 2014). This helped maintain sample representativeness, as reflected in the similar dmfs scores between the SMILE cohort (1.34, 95% CI: 1.0–1.6; 23.4%) and the broader South Australian population (1.40, 95%

CI: 1.0–1.6; 25.3%) (Do et al 2016). Recall bias in dietary intake was reduced using a validated 89-item dietary checklist, developed from 3-d food diaries at 12 mo and tested for repeatability and accuracy (Do et al 2020), while that in number of dental visits was controlled by not collecting as a single summary measure; instead, parents were asked to provide the dates of the first, second, and subsequent visits. Consistent results across both general and high-risk groups strengthen the robustness of our findings. Although generated for the Australian preschool population, these results have not yet been tested across diverse sociocultural contexts, including low- and middle-income countries.

In addition, some other unmeasured confounders, such as breastfeeding, caregivers' caries status, dietary diversity, and parental oral health behaviors, were not controlled. Sensitivity analysis indicated an E-value for the observed effect of approximately 2.62 to 2.64, with a lower-bound E-value of 1.24 to 1.34 (Appendix A.6), suggesting that an unmeasured confounder would need to be moderately strongly associated with both the exposure and the outcome to fully explain the observed effect. Although breastfeeding and parental dietary factors could theoretically act as confounders, existing evidence suggests that their effect sizes are relatively small. For example, Shrestha et al (2024) reported effect sizes of 1.54 (lower CI = 1.16) for children breastfed for 12 mo or longer and 2.44 (lower CI = 1.09) for those breastfed for 18 mo or longer. Similarly, Ha et al (2022) found effect sizes ranging from 1.24 (95% CI 0.96–1.60) to 1.37 (95% CI 1.01–1.67) for maternal high-sugar intake trajectories. These estimates are smaller than the calculated E-value, indicating that the results appear moderately robust to potential residual confounding from these unmeasured factors.

Finally, the subgroup analysis of high-risk children was conducted on a small sample and intended for descriptive purposes only. Because it conditions on the outcome, it may introduce selection bias, limiting the causal interpretation. While the estimates highlight those with the worst outcomes, they should be interpreted with caution.

Conclusion

This study used the advanced G-computation approach to estimate the causal impact of reducing FSI in accordance with WHO guidelines on the prevalence of ECC. The findings underscore the importance of reducing FSI to less than 10% of the EER for the general population of children. The results also suggest additional benefits of less than 5% among high-risk children who consumed excessive FSI and experienced dental caries.

The corresponding reductions in dmfs at these thresholds provide tangible evidence to inform both clinical practice and community-level interventions. Clinically, these findings can support dental practitioners and pediatric health providers in counseling parents about optimal dietary sugar limits and incorporating sugar reduction guidance into preventive oral health programs. At the community level, the results can inform public health initiatives such as nutrition education campaigns, reformulation of children's foods and beverages, and integration of sugar reduction goals into school and childcare nutrition policies. Future research should build on these results by applying causal inference methods to broader populations, including adults, to further strengthen the evidence base for dietary sugar reduction policies.

Author Contributions

A.T.M. Dao, contributed to conception and design, data analysis and interpretation, drafted and critically revised manuscript; L.G. Do, D.H. Ha, contributed to conception and design, data acquisition and interpretation, critically revised manuscript; N. Stormon, contributed to conception and design, data interpretation, critically revised the manuscript; H.V. Nguyen, contributed to conception and design, data analysis and interpretation, critically revised the manuscript. All authors gave their final approval and agreed to be accountable for all aspects of the work.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iDs

A.T.M. Dao [iD] https://orcid.org/0000-0002-1942-3009
L. G. Do [iD] https://orcid.org/0000-0003-3684-9949
N. Stormon [iD] https://orcid.org/0000-0003-0758-1605

Data Availability Statement

The data underlying this article cannot be shared publicly due to the privacy of individuals who consented to participate in the SMILE study. The data will be shared on reasonable request to Prof. Loc Do (l.do@uq.edu.au), the principal investigator of the SMILE study.

References

Andrew L, Wallace R, Wickens N, Patel J. 2021. Early childhood caries, primary caregiver oral health knowledge and behaviours and associated sociological factors in Australia: a systematic scoping review. BMC Oral Health. 21(1):521. https://doi.org/10.1186/s12903-021-01887-4

Antonoplis S. 2023. Studying socioeconomic status: conceptual problems and an alternative path forward. Perspect Psychol Sci. 18(2):275–292. https://doi.org/10.1177/17456916221093615

Dao ATM, Do LG, Stormon N, Dhanapriyanka M, Ha DH. 2024. Causal analyses in longitudinal observational studies in oral health: a scoping review. Community Dent Oral Epidemiol. 53(1):7–16. https://doi.org/10.1111/cdoe.13002

Dao ATM, Do LG, Stormon N, Nguyen HV, Ha DH. 2025a. Enhancing socioeconomic status prediction for cavities: a hybrid method. J Dent Res. 104(9):947–954. https://doi.org/10.1177/00220345251324494

Dao ATM, Do LG, Stormon N, Nguyen HV, Ha DH. 2025b. Key mediators reducing socioeconomic inequality in early childhood caries. JDR Clin Trans Res. [published online ahead of print 2025 Sep 10]. https://doi.org/10.1177/23800844251365536

Devenish G et al. 2019. Free sugars intake, sources and determinants of high consumption among Australian 2-year-olds in the SMILE cohort. Nutrients. 11(1):161. https://doi.org/10.3390/nu11010161

Devenish G, Mukhtar A, Begley A, Do L, Scott J. 2017. Development and relative validity of a food frequency questionnaire to assess intakes of total and free sugars in australian toddlers. Int J Environ Res Public Health. 14(11):1361. https://doi.org/10.3390/ijerph14111361

Do LG et al. 2014. Common risk factor approach to address socioeconomic inequality in the oral health of preschool children—a prospective cohort study. BMC Public Health. 14:429. https://doi.org/10.1186/1471-2458-14-429

Do LG et al. 2016. Trends in child oral health in Australia. In: Do LG, Spencer AJ, editors. Oral health of Australian children: the National Child Oral Health Study 2012-14. University of Adelaide Press. p 288–305.

Do LG et al. 2020. Study of mothers' and infants' life events affecting oral health (SMILE) birth cohort study: cohort profile. BMJ Open. 10(10):e041185. https://doi.org/10.1136/bmjopen-2020-041185

Do LG, Song YH, Du M, Spencer AJ, Ha DH. 2022. Socioecological determinants of child oral health—a scoping review. Community Dent Oral Epidemiol. 51(5):1024–1036. https://doi.org/10.1111/cdoe.12819

Ferraro M, Vieira AR, Seymen F. 2010. Explaining gender differences in caries: a multifactorial approach to a multifactorial disease. Int J Dent. 2010:649643. https://doi.org/10.1155/2010/649643

Fisher-Owens SA et al. 2007. Influences on children's oral health: a conceptual model. Pediatrics. 120(3):e510–520. https://doi.org/10.1542/peds.2006-3084

Ha DH et al. 2022. Group-based trajectories of maternal intake of sugar-sweetened beverage and offspring oral health from a prospective birth cohort study. J Dent. 122:104113. https://doi.org/10.1016/j.jdent.2022.104113

Hernan MA, Dahabreh IJ, Dickerman BA, Swanson SA. 2025. The target trial framework for causal inference from observational data: why and when is it helpful? Ann Intern Med. 178(3):402–407. https://doi.org/10.7326/annals-24-01871

Hernán MA, Robins JM. 2011. Causal inference: what if. 1st ed. Chapman & Hall/CRC.

Lei L, Rangan A, Flood VM, Louie JCY. 2016. Dietary intake and food sources of added sugar in the Australian population. Br J Nutr. 115(5):868–877. https://doi.org/10.1017/s0007114515005255

Lin Y-T, Chou C-C, Lin Y-TJ. 2021. Caries experience between primary teeth at 3–5 years of age and future caries in the permanent first molars. J Dent Sci. 16(3):899–904. https://doi.org/10.1016/j.jds.2020.11.014

Louie JCY, Moshtaghian H, Rangan AM, Flood VM, Gill TP. 2016. Intake and sources of added sugars among Australian children and adolescents. Eur J Nutr. 55(8):2347–2355. https://doi.org/10.1007/s00394-015-1041-8

Marsh PDP. 2010. Microbiology of dental plaque biofilms and their role in oral health and caries. Dent Clin North Am. 54(3):441–454. https://doi.org/10.1016/j.cden.2010.03.002

Mejia GC, Amarasena N, Ha DH, Roberts-Thomson KF, Ellershaw AC. 2012. Child Dental Health Survey Australia 2007: 30-year trends in child oral health. Dental statistics and research series no. 60. Cat. no. DEN 217. Australian Institute of Health and Welfare (AIHW).

Moynihan P. 2016. Sugars and dental caries: evidence for setting a recommended threshold for intake Adv Nutr. 7(1):149–156. https://doi.org/10.3945/an.115.009365

Moynihan PJ, Kelly SA. 2014. Effect on caries of restricting sugars intake: systematic review to inform WHO guidelines. J Dent Res. 93(1):8–18. https://doi.org/10.1177/0022034513508954

National Health and Medical Research Council, Australian Government Department of Health and Ageing, New Zealand Ministry of Health. 2006. Nutrient reference values for Australia and New Zealand including recommended dietary intakes. National Health and Medical Research Council (NHMRC); [accessed 2025 Nov 20]. https://www.nhmrc.gov.au/about-us/publications/nutrient-reference-values-australia-and-new-zealand-including-recommended-dietary-intakes

Nazarian A. 2023. Translational science research: bridging the gap between lab and life for a better tomorrow. J Med Org Chem. 6(5):119–120.

Pham H. 2019. A new criterion for model selection. Mathematics (Basel). 7(12):1215. https://doi.org/10.3390/math7121215

Ra JS, Park M. 2022. Sex-based differences in factors associated with sugar-sweetened beverage consumption among Korean high school students. Fronti Nutr. 9:907922. https://doi.org/10.3389/fnut.2022.907922

Shrestha SK, Arora A, Manohar N, Ekanayake K, Foster J. 2024. Association of breastfeeding and early childhood caries: a systematic review and meta-analysis. Nutrients. 16(9):1355. https://doi.org/10.3390/nu16091355

Snowden JM, Rose S, Mortimer KM. 2011. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. Am J Epidemiol. 173(7):731–738. https://doi.org/10.1093/aje/kwq472

Steenland K, Deddens JA. 2004. A practical guide to dose-response analyses and risk assessment in occupational epidemiology. Epidemiology. 15(1):63–70. https://doi.org/10.1097/01.ede.0000100287.45004.e7

UCLA Statistical Consulting Group. 2024. Negative binomial regression| Stata data analysis examples. [accessed 2025 Mar 10]. https://stats.oarc.ucla.edu/stata/dae/negative-binomial-regression/

Uribe SE, Innes N, Maldupa I. 2021. The global prevalence of early childhood caries: a systematic review with meta-analysis using the who diagnostic criteria. Int J Paediatr Dent. 31(6):817–830. https://doi.org/10.1111/ipd.12783

von Elm E et al. 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 370(9596):1453–1457. https://doi.org/10.1016/s0140-6736(07)61602-x

World Health Organization. 2015. Guideline: sugars intake for adults and children. World Health Organization.

Yan RR, Louie JCY. 2023. Sugar guidelines should be evidence-based and contain simple and easily actionable messages. Front Nutr. 10:1227377. https://doi.org/10.3389/fnut.2023.1227377