

## Development and internal validation of the SMILE-FSS - A Free Sugars Screener for Australian children aged 2 and 5 years

Lucinda K Bell<sup>1\*</sup>, Shalem Leemaqz<sup>2</sup>, Gemma Devenish-Coleman<sup>3</sup>, Loc G Do<sup>4</sup>, Diep Ha<sup>4</sup>, Jane A Scott<sup>3</sup>, Rebecca K Golley<sup>1</sup>

<sup>1</sup>Caring Futures Institute, College of Nursing and Health Sciences, Flinders University, Adelaide 5001, Australia; lucinda.bell@flinders.edu.au (L.K.B.); rebecca.golley@flinders.edu.au (R.K.G.)

<sup>2</sup>South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia; [shalem.leemaqz@sahmri.com](mailto:shalem.leemaqz@sahmri.com)

<sup>3</sup>School of Population Health, Curtin University, Perth, WA, 6102, Australia; gemma.devenish@curtin.edu.au (G.D.C.); jane.scott@curtin.edu.au (J.A.S.)

<sup>4</sup>School of Dentistry, The University of Queensland, Brisbane, QLD, 4072 Australia; d.ha@uq.edu.au (D.H.); l.do@uq.edu.au (L.G.D.)

**\*Corresponding author:** Lucinda Bell, Caring Futures Institute, College of Nursing and Health Sciences, Flinders University, GPO Box 2100, Adelaide, SA, 5001, [lucy.bell@flinders.edu.au](mailto:lucy.bell@flinders.edu.au)



**The  
Nutrition  
Society**

This is an Accepted Manuscript for Public Health Nutrition. This peer-reviewed article has been accepted for publication but not yet copyedited or typeset, and so may be subject to change during the production process. The article is considered published and may be cited using its DOI 10.1017/S1368980023002380

Public Health Nutrition is published by Cambridge University Press on behalf of The Nutrition Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

## Abstract

*Objective:* To develop and internally validate a Free Sugars Screener (FSS) for Australian children aged 2 and 5 years.

*Design:* Using data collected from a 99-item (2-year-olds) and 98-item (5-year-olds) Food Frequency Questionnaire in the Study of Mother's and Infant's Life Events affecting oral health (SMILE-FFQ), a regression-based prediction modelling approach was employed to identify a subset of items that accurately estimate total free sugars intake (FSI). The predictors were grams of free sugars (FSg) for individual items in the SMILE-FFQ and child's age and sex. The outcome variable was total FSI per person. To internally validate the SMILE-FSS items, the estimated FSg was converted to percent energy from free sugars (%EFS) for comparison to the WHO free sugars guideline categories (<5%, 5-<10%,  $\geq$ 10%EFS) using cross-classification analysis.

*Setting:* Australia.

*Participants:* 858 and 652 2- and 5-year-old children, respectively, with complete dietary (<5% missing) and sociodemographic data.

*Results:* 22- and 26-items were important in predicting FSI at 2 and 5 years, respectively. Items were similar between ages with more discretionary beverage items (e.g., sugar-sweetened beverages) at 5 years. %EFS was overestimated by 4.4% and 2.6%. Most children (75% and 82%) were categorised into the same WHO free sugars category with most (87% and 95%) correctly identified as having <10%EFS in line with the WHO recommendation.

*Conclusions:* The SMILE-FSS has good internal validity and can be used in research and practice to estimate young Australian children's free sugars intake and compare to the WHO free sugars guidelines to identify those 'at risk'.

**Keywords:** Early childhood; dietary assessment; short screener; dental screener; free sugars; food frequency questionnaire; oral health; pre-school; short screener

## 1. Introduction

High consumption of dietary sugars is a major public health concern, having a negative impact on oral health<sup>(1)</sup> and contributing to an increased risk of chronic diseases such as obesity<sup>(2)</sup>. Excessive intakes of added sugars is of particular concern in children<sup>(3)</sup>, with approximately one quarter of Australian children having overweight or obesity<sup>(4)</sup> and one-third having caries by the age of five<sup>(5)</sup>. Frequent exposure to added sugars in the first years of life may reinforce the innate preference that infants have for sweet foods<sup>(6)</sup> with taste preferences persisting throughout life and influencing dietary patterns in adulthood<sup>(7)</sup>. It is therefore important to understand patterns of sugars intakes in early life to inform early prevention efforts to reduce dental caries and overweight prevalence in children<sup>(8,9)</sup>.

Dietary free sugars, defined as sugar added to foods and beverages by the manufacturer, cook or consumer plus those naturally present in honey, syrups, and fruit juices, are the most important risk factor for dental caries<sup>(1, 10, 11)</sup> and can contribute to excess energy intake with little nutritional benefit<sup>(12)</sup>. The World Health Organization (WHO) recommends that less than 10% of total energy intake come from free sugars, with a conditional recommendation to reduce intake to less than 5% of energy intake for additional health benefits<sup>(1)</sup>. However, recent analysis of data from the Australian SMILE birth cohort has shown that some children (2.4%) as young as 1 year old were exceeding the <10%EFS recommendation (23% exceeding <5%EFS recommendation)<sup>(13)</sup>, increasing to over a third at 2 years (38% exceeding <10%EFS and 71% exceeding <5%EFS)<sup>(14)</sup>, with similar findings at 5 years (37% exceeding <10%EFS and 75% exceeding <5%EFS, unpublished data). Thus, there is an urgent need to address excess intakes of free sugars in early childhood.

Accurate measurement of free sugars intakes in early childhood supports the monitoring of population intakes to inform policy, the evaluation of interventions in practice and at scale, and for screening intakes within the primary health care setting to identify 'at risk' children requiring referral and intervention<sup>(15, 16)</sup>. However, an ongoing limitation in research and practice is the lack of consistency and precision in the assessment of dental-specific dietary factors, leading to a call within the WHO Sugars Intake Guidelines for studies with improved dietary assessment methodology<sup>(1)</sup>. In response to this, the Study of Mothers' and Infants' Life Events affecting oral health Food Frequency Questionnaire (SMILE-FFQ) was developed<sup>(17)</sup>. The SMILE-FFQ is a dental-specific, semi-quantitative 89-item FFQ (generating a list of 99 foods) designed to capture the leading dietary contributors to dental caries risk in toddlers

aged 18 – 30 months<sup>(17)</sup>. Although the SMILE-FFQ was determined to be a valid tool for assessing free sugars intake of Australian toddlers using parental proxy report<sup>(17)</sup>, the length limits its use in both the research and practice settings.

Short dietary questionnaires or screeners, typically completed in less than 15 minutes<sup>(18)</sup>, provide an alternative to longer FFQs in settings where brief tools are needed<sup>(15, 16)</sup>. Short questionnaires can be administered quickly in a variety of formats, tailored to outcomes of interest, are associated with a low respondent burden, and allow for easy comparison of intake against guidelines to provide a quick assessment of ‘risk’<sup>(15, 16)</sup>. Although short questionnaires have been developed and validated in Australian children for assessing various aspects of diet, predominately dietary intake and diet quality<sup>(15, 16, 19)</sup>, none specifically assess free sugars intake. Short dental-specific questionnaires that assess free sugars intakes of young children and that allow for comparison against the WHO free sugars guidelines are therefore required.

Thus, the aim of this study was to develop and internally validate a short-form Free Sugars Screener (FSS) for Australian children aged 2- and 5-years using data derived from the long-form SMILE-FFQ. The objectives were to: 1) identify a subset of SMILE-FFQ questions that accurately estimate free sugars intake (FSI) from the complete set of survey questions, 2) internally validate the items by comparing *measured* FSI with *predicted* FSI, and 3) internally validate the items by classifying participants’ FSI into the WHO free sugars guidelines categories (<5%, 5-<10%, ≥10%EFS)<sup>(1)</sup> and comparing measured categorization with *predicted* categorization.

## 2. Methods

### 2.1. *The Study of Mother’s and Infant’s Life Events affecting oral health (SMILE)*

SMILE is a birth cohort study conducted in Adelaide, South Australia. The objectives, methods and cohort profile have been described elsewhere<sup>(20, 21)</sup>. Briefly, 2,147 mothers and 2,181 newborns were recruited from the three major maternity hospitals in Adelaide between July 2013 and August 2014. All new mothers with sufficient English competency and with no intention of moving out of the greater Adelaide area within the following year were invited to participate. Mothers in hospitals that service lower socio-economic areas were oversampled to compensate for anticipated higher attrition rates<sup>(21)</sup>. Participants were invited to complete questionnaires on dental and dietary habits at recruitment, when their child reached 3 and 6

months of age, and again at 1, 2, and 5 years of age. This included the SMILE-FFQ, administered at 2 and 5 years of age. The SMILE study was approved by the Southern Adelaide Clinical Human Research Ethics Committee (HREC/50.13, approval date: 28 Feb 2013) and the South Australian Women and Children Health Network (HREC/13/WCHN/69, approval date: 7 August 2013).

## 2.2. Data collection

### 2.2.1. The SMILE-FFQ

Development and validation of the SMILE-FFQ has been described in detail elsewhere<sup>(17)</sup>. Briefly, the 89-item SMILE-FFQ was constructed to capture dietary factors associated with dental caries risk in children aged 1-3 years. This included major food and beverage sources of total and free sugars, and dental-protective foods such as milk products<sup>(17)</sup>. Participants report the frequency (from “never or rarely” to “3 or more times per day”) and quantity of consumption for each food item in the questionnaire. Portion size options were tailored to the food in question and described using a combination of household measures (teaspoon, tablespoon, cup) and common portion sizes (piece, tub, pouch), each with a corresponding weight (g) and/or volume (ml). Four supplementary questions within the SMILE-FFQ split the coding of 6 food items into 16 items, so that the 89-item FFQ generates a list of 99 foods. For example, in the SMILE-FFQ a supplementary question “If your child eats yoghurt, do you choose reduced fat versions?” is used to split four yoghurt line-items differentiated by flavour into eight foods, sub-divided by both flavour and fat type ([supplementary table S1](#)). The SMILE-FFQ was validated against repeat 24hr recalls in an external cohort of toddlers aged 18-30 months, performing similarly at ranking individuals’ total and free sugars intakes, with a tendency to underestimate intakes in participants with lower reported 24hr recall values and overestimate intakes in those with higher reported values<sup>(17)</sup>, highlighting its suitability for observational studies wanting to use ranked total and free sugars intakes.

Minor adjustments were made to the SMILE-FFQ ahead of data collection when children were approximately 5 years of age. Checks were undertaken to ensure: (1) the item response wording was appropriate for the target age group, (2) the portion sizes listed reflected current supermarket products for older children; and (3) the portion sizes listed were distributed around the median intakes of 4–8-year-old Australians<sup>(22)</sup>. Examples of wording changes included: changing “infant formula” to “junior formula”, removing “baby jars” and “sippee cups” from the portion size descriptors for some items, and removing the smallest portion

size response option for some questions. No changes were made to the frequency response options. One item was added (coffee e.g., ready-to-drink coffee beverages such as iced coffee, as coffee is a source of fluoride<sup>(23)</sup>) and two items removed (infant fruit purees and infant vegetable purees due to not being age-appropriate), resulting in 88 items. As above, the additional four supplementary questions are used to split the coding of six food items into 16 items, generating a list of 98 foods at 5 years.

### 2.2.2. Sociodemographic data

Sociodemographic data were collected from mothers at recruitment via self-completed questionnaires. These included measures of maternal age, country of birth and educational attainment, in addition to household income, number of children at home and whether their child lived in a one- or two-parent household (Table 1). Mothers' pre-pregnancy Body Mass Index (BMI, kg/m<sup>2</sup>) was calculated from self-reported height and weight and classified into weight status categories (healthy weight <25 kg/m<sup>2</sup>, overweight 25–29.9 kg/m<sup>2</sup>, obesity ≥30 kg/m<sup>2</sup>)<sup>(24)</sup>. Postcode was used to derive a measure of socioeconomic status using the Index of Relative Socio-Economic Advantage and Disadvantage (IRSAD)<sup>(25)</sup>, one of the four Socio-Economic Indexes for Areas (SEIFA) indices that ranks areas across Australia on a continuum of social disadvantage to advantage<sup>(20)</sup>. Child age was calculated from date of birth and date of completion of the SMILE-FFQ at both 2 and 5 years of age.

### 2.2.3. Child anthropometric data

Child weight and height data were collected at the 2- and 5-year physical examinations using standardised methodology and equipment, including calibrated electronic scales and medical stadiometers<sup>(26)</sup>. Age and sex-specific BMI z-scores were calculated for children and classified into weight status categories (healthy weight ≥-2 and ≤+2SD, overweight >+2 and ≤+3SD, and obesity >+3) using the WHO reference<sup>(27)</sup>.

## 2.3. Development of the SMILE-Free Sugars Screener

### 2.3.1 Data source and variables

The SMILE Free Sugars Screener (SMILE-FSS), a short-form version of the SMILE-FFQ, was developed using data obtained from completion of the SMILE-FFQ at 2 and 5 years of the SMILE study. Data reduction techniques were employed to identify a subset of items that accurately estimate FSI in Australian children aged 2 and 5 years, derived from the long-form

SMILE-FFQ. The predictor variable was measured FSI in grams calculated for each questionnaire item of the long-form SMILE-FFQ based on quantity and frequency response options. The outcome variable was total usual FSI derived from the long-form SMILE-FFQ calculated per person. A specific number of items was not identified as the target, however, to substantially reduce participant burden, fewer than 50, and ideally around 20-30 items was anticipated, consistent with other short screeners<sup>(28-34)</sup>.

### 2.3.2 Data cleaning and preparation

On return of the completed SMILE-FFQ at 2 and 5 years, data were cleaned and questionnaires with more than 5% of line items left blank were excluded<sup>(17)</sup>. Microsoft Access version 15 (Microsoft Corporation, 2013, Washington, DC, USA) was used to link the questionnaire responses to an accompanying database developed specifically for the SMILE study which contained lookup tables of free sugars amounts (g) for the foods, quantity options and frequencies (per day) listed in the SMILE-FFQ<sup>(17)</sup>. This database linked FFQ responses to nutrient values for total sugars and free sugars using the AUSNUT 2011-2013 food consumption database<sup>(35)</sup>, from which total usual FSI in grams per day was calculated. Implausible FSI intakes were excluded, defined as values greater than 3SD above the mean for both total and free sugars<sup>(17)</sup>.

### 2.4 Statistical analysis

To identify a subset of FFQ items that can reasonably estimate total FSI of each child calculated from the complete set of survey questions, a regularised regression-based prediction modelling approach was employed at both 2 and 5 years. Participants with complete dietary (<5% missing FFQ data) and complete sociodemographic data were included (i.e., “complete-case analysis”). The predictors included grams of free sugars for each food in the SMILE-FFQ and child’s age and sex. Child’s weight and height were also included as predictors in the 5-year regression model to improve model fit, although not at 2 years due to the number of missing values ( $n = 279$ , 26.7%). A sensitivity analysis was performed with height and weight *included* at 2 years, and height and weight *excluded* at 5 years, to assess the consistency of variables selected.

Survey items (herein referred to as predictors) with near-zero variance (for example, foods such as porridge, milk and milk alternatives, fruit [fresh, tinned, puree], cheese and custard, and drinks such as tea, coffee, cordial, vegetable juice and water), defined as less than 10%



unique values or the ratio of most common values is greater than 19, were removed prior to analyses (i.e., survey items which were rarely consumed and made limited contribution to grams of free sugars intake (FSg) in participants) (See Supplementary table S2). A 10-fold cross-validation was performed, in which the sample was randomly split into a training (70%) and testing (30%) sample with the testing sample reserved for use only in the final internal validation step. In the first step of the analysis, we used regularised linear regression with Elastic Net (a combination of L1 Lasso and L2 Ridge) to select variables which are ‘important’ in predicting the outcome (i.e., total usual FSI per person). Lasso and Ridge regression works by estimating the regression parameters under a constraint which shrinks the estimates towards zero with the aim of reducing prediction variance (determined by the minimum root mean squared error (RMSE)), and in the process ‘removes’ (i.e., shrinks to zero) unimportant predictors. The optimal model was chosen on the basis of minimum differences between training and testing RMSE across cross-validation runs.

To internally validate the items in the optimal model, the predicted free sugars in grams (FSg) was converted to percent energy from free sugars (%EFS) for comparison to the WHO free sugars guidelines categories (<5%, 5-<10%, ≥10%EFS)<sup>(1)</sup>. However, as the SMILE-FFQ was not designed to capture total energy intake, Estimated Energy Requirements (EER) were used as an alternative for calculating the percentage of EER coming from FSI. The method for calculating EER at 2 years has been described previously<sup>(14)</sup>. That is, EER was determined from age and weight data using the Australian Nutrient Reference Values<sup>(36)</sup> equation for children aged 13 – 35 months. At 5 years, the equation for children aged 3-8 year was used and EER determined from child age, weight and height data and applying a standard Physical Activity level (PAL) of 1.6 (“light activity”) due to a lack of physical activity data for the current sample. The percent of EER coming from FSI was then calculated using the following equation:  $\%EERFreeSug = [(Sugars\_g \times 16.7) \div EER\_kJ] \times 100$ . The *measured* and *predicted* %EFS were then compared using the root mean square error (RMSE), and participants in the testing sample categorised into the WHO categories (<5%EFS, 5-<10%EFS, ≥10%EFS)<sup>(1)</sup> and compared descriptively. Sensitivity analysis was also conducted to compare the measured and predicted FSg in the optimal model using the testing sample at 2 and 5 years.



### 3. Results

#### 3.1. Population studied

Of the 2181 mothers-infant dyads recruited for the SMILE study,  $n = 1043/1195$  and  $n = 716/825$  SMILE-FFQs were returned complete (<5% missing) and plausible at 2 and 5 years, respectively ([Figure 1](#)). Complete-case analysis was conducted on  $n = 885/1043$  (2 years) and  $n = 652/716$  (5 years). Sample characteristics are presented in [Table 1](#). At the 2-year data collection wave, the mean age of children was 2.1 years ( $SD \pm 0.1$  years, range 1.9-3.2 years), with sex almost evenly split (females 47%, males 53%) and most being from a two-parent household (94%). The majority of mothers (mean age  $30.5 \pm 5$  years) were born in Australia, New Zealand (NZ) or United Kingdom (UK) (78%), with pre-pregnancy BMI less than  $25 \text{ kg/m}^2$  (54%), and were university educated or higher (57%). Similarly, nearly half of the children at the 5-year data collection wave (mean age  $5.3 \pm 0.4$  years, range 4.7-6.5 years) were female (45%) and approximately two-thirds from two-parent households (65%). The majority of mothers were born in Australia, NZ or UK (74%), with a pre-pregnancy BMI less than  $25 \text{ kg/m}^2$  (48%), and were university educated or higher (61%).

[Supplementary table S3](#) presents the sample characteristics for the complete-case samples compared to the excluded cases (i.e., those with complete dietary data but incomplete sociodemographic data). Participants excluded were largely similar to the analysed samples at 2 and 5 years; however, children from the analysed sample were slightly younger at 2 years ( $2.1 \pm 0.1$  years vs  $2.2 \pm 0.2$  years,  $p = 0.03$ ) and slightly older at 5 years ( $5.3 \pm 0.4$  vs  $5.2 \pm 0.4$  years,  $p = 0.004$ ). Mothers who provided complete data at 2 years were more likely to be born in Australia/NZ/UK (78% vs 68%,  $p = 0.006$ ) while at 5 years they were less likely to be university educated (61% vs 75%,  $p = 0.04$ ).

#### 3.2. Development of the SMILE-FSS

##### 3.2.1 SMILE-FFQ item reduction at 2 years of age

[Figure 2](#) shows the reduction of the 2-year SMILE-FFQ from 99 items to 22 items using data from the training sample ( $n = 622$ ). First, the 99 FFQ items were reduced by collapsing 25 FFQ items into 8 ([Supplementary table S1](#)) and subsequently removing 60 items with near-zero variance ([Supplementary table S2](#)). A total of 24 predictors (22 FFQ items plus child age and sex) were entered into the regularized linear regression prediction model. [Table 2](#) shows the model coefficients after variable shrinkage across 10 cross-validation runs. The RMSE

across the 10 runs was  $2275.6 \pm 28.2$  (%EFS,  $4.06 \pm 0.26\%$ ). The best model was R8, with an RMSE of 4.4%EFS, meaning that on average a child's predicted %EFS was overestimated by 4.4%.

Of the 24 predictors entered into the regression model, all remained after variable shrinkage. The 22 items included discretionary items (i.e. those that do not fit into the five food groups due to their high saturated fat and/or added sugars or salt content<sup>(37)</sup>) such as: icy poles and sorbets; ice creams and frozen yoghurts; flavoured breakfast cereals (i.e. chocolate, vanilla or malt-flavoured with added sugar content); snack foods (i.e. potato crisps), sweet biscuits; muesli and cake-type bars; cakes and puddings; sweet breads and pancakes; spreads such as jam and honey; chocolate and lollies; fruit juice (no added sugar); and sugar. Sugar-sweetened beverages (such as soft drinks, cordials and fruit drinks) were not a strong predictor of FSI in the 2-year-old sample. Core ('healthy')<sup>(37)</sup> food items such as custard, nut paste, wheat biscuits and cereal flakes, flavoured yoghurts (i.e. fruit flavoured), and plain savoury biscuits also remained as strong predictors of FSI. The secondary analysis showed that for the model including child height and weight ([supplementary table S3](#)), the resultant FFQ items ( $n = 21$ ) were similar, however without sweet style sauces and marinades, and savoury biscuits and snack foods.

### 3.2.2. SMILE-FFQ item reduction at 5 years of age

[Figure 2](#) shows the reduction of the 5-year SMILE-FFQ from 98 items to 26 items using data from the training sample ( $n = 460$ ). First, the 98 FFQ items were reduced by collapsing 25 FFQ items into 8 ([Supplementary table S1](#)) and removing 51 items with near-zero variance ([Supplementary table S2](#)). A total of 34 items (30 FFQ items plus child age, sex, weight and height) were entered into the linear regression prediction model. [Table 3](#) shows the model coefficients resulting after variable shrinkage across 10 cross-validation runs. The RMSE across the 10 runs was  $2193.3 \pm 37.1$  (%EFS,  $2.5 \pm 0.4\%$ ). The best model was R8, with an RMSE of 2.6%EFS, meaning that on average a child's predicted %EFS was overestimated by 2.6%.

Of the 34 predictors entered into the regression model, 26 remained after variable shrinkage. That was, four SMILE-FFQ items (nut paste, wheat biscuits, cereal flakes and mayonnaise) were removed via the regression based predictive analysis model, in addition to the four child predictors. Items remaining as strong predictors were those as per the 2-year analysis, with additional items including discretionary drinks such as flavoured milk, fruit drink, soft drink

and powdered milo/ovaltine (added to milk drinks) and discretionary foods such as chocolate spreads, sweet-style sauces/marinades, sugar syrups (e.g., maple syrup), and sweet pastries. The sensitivity analysis showed that for the model excluding child height and weight ([supplementary table S4](#)), the resultant FFQ items ( $n = 27$ ) were similar, however did not include wheat biscuits, cereal flakes and mayonnaise.

### 3.3. Internal validation of the SMILE-FSS

#### 3.3.1. Comparison of the measured and predicted %EFS

[Figure 3](#) shows a scatter plot of the *measured* versus *predicted* %EFS, including a unity line which represents perfect calibration (i.e., where predicted=measured), developed using the testing sample ( $n = 263$  at 2 years;  $n = 193$  at 5 years). The figure indicates that at both 2 and 5 years, the models performed worse at predicting larger values of %EFS. The horizontal and vertical dashed and dotted lines represent the cut-points for <5%EFS and <10%EFS respectively and indicate that no participant with a %EFS of less than 5% was predicted to have a %EFS greater than 10%, or vice versa. [Supplementary Figure S3](#) shows the scatter plot of the measured versus predicted free sugars intake in grams (FSg), also developed using the testing sample ( $n = 263$  at 2 years;  $n = 193$  at 5 years).

#### 3.3.2. Comparison to WHO free sugars guideline categories at 2 years of age

[Table 4](#) shows the cross-classification of measured versus predicted WHO categorization of %EFS, in which the column-wise percentages in the diagonal equal the proportion of the participants correctly predicted to be in a given category. At 2 years, a total of 74.9% were classified into the same category and 25.1% into an adjacent category (15.6% over classified and 9.5% under classified). Correct classification was highest for category 3 ( $\geq 10\%$ EFS, 84.5%), followed by category 1 (<5%EFS, 75.9%) and category 2 (5-<10%EFS, 63.2%). In regard to risk identification ([Supplementary table S6](#)), more than 4 out of 5 children ( $n = 144/166$ , 87%) with FSI <10%EFS were correctly identified as such, with less than one-quarter ( $n = 22/166$ , 13%) misclassified as ‘at risk’ (i.e.,  $\geq 10\%$ EFS). Conversely, few children ( $n = 15/263$ , 6%) with FSI  $\geq 10\%$ EFS were not identified as such by the SMILE-FSS (i.e., were predicted to have FSI <10%EFS).

### 3.3.3. Comparison to WHO free sugars guideline categories at 5 years of age

At 5 years, a total of 82.3% were classified into the same WHO free sugars guideline category and 17.7% classified into an adjacent category (11.9% over classified and 5.8% under classified) (Table 4). Correct classification was highest for category 3 ( $\geq 10\%$ EFS, 89.5%), followed by category 2 ( $5 < 10\%$ EFS, 88.3%) and category 1 ( $< 5\%$ EFS, 56.4%). In regard to risk identification (Supplementary table S6), nearly all children ( $n = 110/116$ , 95%) with FSI  $< 10\%$ EFS were correctly identified as such, with few ( $n = 6/116$ , 5%) misclassified as ‘at risk’ (i.e.,  $\geq 10\%$ EFS). Conversely, few ( $n = 8/192$ , 4%) children with FSI  $\geq 10\%$ EFS at 5 years were not identified as such by the SMILE-FSS (i.e., were predicted to have FSI  $< 10\%$ EFS).

## Discussion

This study employed a regression-based prediction modelling approach to reduce the long-form SMILE-FFQ at 2 and 5 years to a short-form SMILE Free Sugars Screener (SMILE-FSS). Food items that strongly predicted free sugars intake (FSI) and were included in the SMILE-FSS (22 and 26-items, respectively) were relatively consistent between age groups, with more discretionary food and beverage items included at 5 years. Internal validation of the SMILE-FSS showed that it performs well in comparison to the long-form SMILE-FFQ at both ages. Most children (75% and 82%, respectively) were categorised into the same WHO free sugars category with nearly all children (87% and 95%, respectively) correctly identified as having  $< 10\%$ EFS in line with the WHO recommendation<sup>(1)</sup>. Overall, the SMILE-FSS was shown to have good internal validity and can be used in future research and practice to estimate total free sugars intake in Australian children aged 2-3 and 5-6 years and compare to the WHO free sugars guidelines to determine those ‘at risk’.

Key predictors of free sugars intake included in the SMILE-FSS were similar at 2 (22-items) and 5 years (26-items) and included discretionary items such as ice-cream, sweet biscuits, cakes and puddings, sweet breads and pancakes, sugar-based sauces and spreads, chocolate and lollies. At 5 (but not 2 years) of age, additional discretionary beverage items such as sugar sweetened beverages (known as “soft drink” in Australia) and fruit drinks were key predictors of free sugars intake. This is not overly surprising given trends of increasing discretionary beverage intake in the first years of life in Australian children<sup>(38, 39)</sup>. For example, a study of Australian children’s beverage intake reported that the proportion of consumers of sweet beverages (including flavoured milk, 100% juice, fruit drink, cordial and

soft drink) increased from approximately one third (38%) at 2 years, to half at 3.7 (55%) and 5 (47%) years<sup>(38)</sup>. A similar trend was observed in a longitudinal study of beverage consumption in Australian children<sup>(39)</sup>, whereby soft drink and cordial consumers increased from 1% in the first year of life to one quarter (28%) at 2 years and nearly half (43%) at 10 years. Discretionary items such as chocolate spreads and sugar syrups were also additional predictors of FSI at 5 years (but not at 2 years), while core items such as nut paste, wheat (cereal) biscuits (e.g. Weet-Bix, Vita Brits etc), cereal flakes (e.g. Cornflakes, Wheaties, Sultana Bran, Light n Tasty etc.) were important in predicting free sugars intake at 2 years (but not at 5 years). Together these findings demonstrate a shift from healthy/core foods as key free sugar contributors in the first few years of life towards greater discretionary food contributors in school-aged children, consistent with population trends<sup>(40)</sup>. Overall, consistency and discrepancies in key free sugar contributors between 2 and 5 years are reflective of changes in patterns of consumption as children move from toddlerhood to childhood.

Importantly, in this study, 100% fruit juice and flavoured milk ((i.e. chocolate/strawberry milk, milkshakes) were key predictors of free sugars intake at 2 and 5 years, respectively, aligning with the downward trend by age group in fruit juice consumption (78% 2-3 year olds, 59% in 4-8 year olds)<sup>(41)</sup> and the upward trend in flavoured milk consumption<sup>(42)</sup>. However, although the primary goal of the SMILE-FSS is a reduction in free sugars intake for dental health, and thus identification of free sugar containing items is important, 100% fruit juice and flavoured milk are considered core beverage items within the Australian Dietary Guidelines<sup>(37)</sup>, providing essential nutrients and contributing to children's fruit and dairy serves<sup>(43)</sup> This is a particularly important consideration for the 5-year-old SMILE FFS with respect to flavoured milk consumption and dairy serve recommendations given that Australian children older than 4 years of age do not reach recommendations for dairy food intake, consuming  $\leq 2$  servings/d<sup>(42)</sup>. Thus, end-users of the SMILE-FSS should be cautious not to recommend avoidance of these items but rather balance recommendations to reduce free sugars intake with respect to their nutrient contribution.

Validation of the SMILE-FSS showed that the short-form screener has good internal validity. Although a child's predicted %EFS was overestimated by 4.4% in 2-3 year olds and by 2.6% in 5-6 year olds, the SMILE-FSS correctly classified most children (75% and 82%, respectively) into the WHO free sugars guidelines categories (<5, 5-<10,  $\geq 10\%$ EFS)<sup>(1)</sup>. Further, nearly all children (87% and 95%, respectively) were correctly identified as having

<10%EFS in line with the WHO free sugars recommendation with few children (6% at 2 years, 4% at 5 years) with intakes  $\geq 10\%$ EFS (i.e., ‘at risk’) not identified as such by the SMILE-FSS. In contrast, the proportion of children misclassified as ‘at risk’ (i.e.,  $\geq 10\%$ EFS) at 2 (13%) years was higher than that at 5 years (5%), although similar studies have shown misclassification levels as high as 25%<sup>(31, 32)</sup>. Overall, findings show that the SMILE-FSS performs well in identifying children ‘at risk’ ( $\geq 10\%$ EFS) and in need of referral and intervention and can be used in future research to estimate total free sugars intake in Australian children aged 2-3 and 5-6 years.

To our knowledge, no other short dietary-related screening tools exist within the dental context to assess FSI in Australian children. A short-form 8-16 item child oral health-related quality-of life tool has been developed in a sample of New Zealand 11–14-year-old children, while several short diet-related screening tools (7 – 28 items) that assess food intake<sup>(29, 31-34)</sup>, diet quality<sup>(30)</sup> and or obesity-related behaviours<sup>(28)</sup> have been developed and validated in Australian children ranging from 12 months to 16 years of age. These include, the 28 item Children’s Dietary Questionnaire (4-16 years)<sup>(34)</sup>, the 19 item Toddler (ages 12-36 months)<sup>(29, 31)</sup> and Preschooler (3-5 years) Dietary Questionnaires<sup>(32)</sup>, and the 7-15 item Early Prevention of Obesity in Childhood Dietary Questionnaire (EPOCH-DQ)<sup>(28)</sup>. None of these tools were designed to capture the leading dietary contributors to dental caries risk and thus none specifically assess total and/or free sugars intake from major food and beverages sources in young Australian children. The 2- and 5-year-old versions of the SMILE-FSS therefore add to the small yet increasing number of short tools to assess dietary intakes in young Australian children.

The SMILE-FSS is the first short screening tool to assess FSI across the first years of life that is suitable for dental research and practice, showing good internal validity compared with the longer SMILE-FFQ. The SMILE-FSS was developed in a large sample of children ( $n = 885$  at 2 years and  $n = 652$  at 5 years) through application of an innovative approach to item selection, a regression-based model utilising existing data of 2- and 5-year-olds dietary intakes. Other strengths include combining items a-priori to ensure the final screener would be fit-for-purpose and the 10-fold cross-validation that was undertaken to determine the best subset of SMILE-FFQ questions that accurately estimate children’s free sugars intake. Child-factors (age, sex, height and weight) were added to the model to improve model fit, and sensitivity analyses conducted (with or without height and weight and using both %EFS and FSg to internally validate items in the optimal model) with minimal differences seen between



models. Additionally, demographic characteristics of included participants ( $n = 885$  at 2 years;  $n = 652$  at 5 years) were compared to those excluded ( $n = 158$  at 2 years;  $n = 64$  at 5 years) with no meaningful differences found. However, despite the extensive analyses undertaken in this study to determine internal validity, the external validity of the SMILE-FSS is unknown. This could be determined through completion of the short-form SMILE-FSS and repeat 24-hour recalls (the gold standard in dietary assessment)<sup>(44)</sup> by the same sample to compare FSI of Australian children derived from the two tools. The SMILE-FSS is also limited by the fact that it was developed and internally validated in discrete samples of Australian children aged 2-3 years and 5-6 years and thus the applicability to children outside of these age ranges is unknown.

Despite these limitations, the SMILE-FSS fills an important gap in the literature, being highly useful in dental research where dietary assessment resources are limited. However, caution should be taken to ensure that core foods that provide nutritional benefits, such as 100% fruit juice and flavoured milks are not restricted to ensure young children's nutritional intake is not compromised. Further, if the SMILE-FSS is to be used in a population substantially different to the one described here, or the tool is modified in any way, further validation (internal or external) would be required to determine its validity within that context. It also has potential, following further work to translate it into a user-friendly online version, to be utilised within the primary health care dental setting to screen children's FSI and support appropriate intervention (i.e. behaviour change messages for oral health regarding reducing and/or limiting sugar consumption) to reduce caries risk<sup>(45)</sup>. Future research however is required to determine the feasibility and acceptability (by dental practitioners and caregivers) of this approach<sup>(45)</sup>.

This study used a regression-based prediction modelling approach to develop and internally validate a short 22-item (2 years) and 26-item (5 years) Free Sugars Screener (the SMILE-FSS) using data derived from a longer (99- and 98-items, respectively) Food Frequency Questionnaire (the SMILE-FFQ). The SMILE-FSS was shown to have good internal validity, accurately estimating free sugars intakes (FSI) and allowing for good comparison to the WHO free sugars guideline categories to determine those 'at risk' (i.e., not meeting the <10%EFS WHO recommendation). Food items in the SMILE-FSS at both 2 and 5 years are relatively consistent, with more discretionary foods and beverages featuring in the 5-year SMILE-FSS. In summary, the SMILE-FSS is a novel, age- appropriate, culturally appropriate tool that can be used in research and practice settings to assess FSI of Australian children



aged 2-3 and 5-6 years and provide quick assessment against the WHO free sugars guideline. Future research should determine the acceptability and feasibility of implementing the SMILE-FSS within the primary care dental setting.

### **Acknowledgements**

The authors wish to acknowledge and thank all participants and research support staff of the Study of Mothers and Infants life Events (SMILE) affecting oral health.

### **Funding / sources of support**

The SMILE birth cohort is funded by Australian National Health and Medical Research Council Grants #1046219 and #1144595. LKB is supported by this grant.

### **Conflict of Interest**

None declared

### **Authorship**

The study was conceived by GD-C, JAS, LKB, and RKG with funding obtained by LGD, DH, JAS and RKG. GD-C conducted the validation of the original SMILE-FFQ and LB and GD-C led the implementation of the present study with supervision from RKG and JAS. SL conducted the statistical analyses with contribution from LB and all authors contributed to the interpretation of results. LB drafted the manuscript and all authors have read and approved the final version.

### **Ethical Standards Disclosure**

This study was conducted according to the guidelines laid down in the Declaration of Helsinki and all procedures involving research study participants were approved by the Southern Adelaide Clinical Human Research Ethics Committee (HREC/50.13, approval date: 28 Feb 2013) and the South Australian Women and Children Health Network (HREC/13/WCHN/69, approval date: 7 August 2013). Written informed consent was obtained from all subjects/patients.

## Abbreviations list

BMI, Body Mass Index; %EFS, percent energy from free sugars; FFQ, Food Frequency Questionnaire; FS, Free Sugars; FSI, Free Sugars Intake; FSg, grams of Free Sugars; RMSE, Root Mean Square Error; SMILE, Study of Mothers' and Infants' Life Events affecting oral health; SMILE-FSS, SMILE Free Sugars Screener; SMILE-FFQ, SMILE Food Frequency Questionnaire; WHO, World Health Organization.

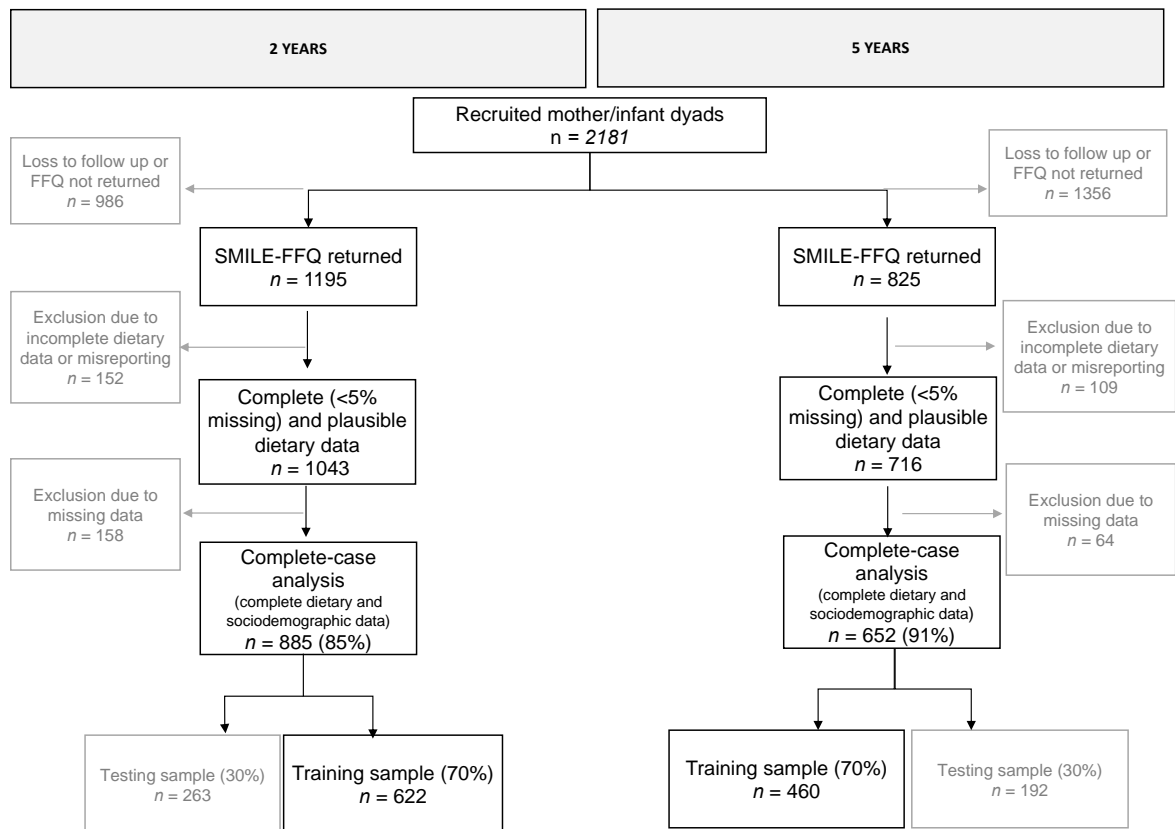
## References

1. World Health Organization (2017) *Sugars and dental caries*. WHO Technical Information Note. Geneva, Switzerland: WHO.
2. Delli Bovi AP, Di Michele L, Laino G *et al.* (2017) Obesity and Obesity Related Diseases, Sugar Consumption and Bad Oral Health: A Fatal Epidemic Mixtures: The Pediatric and Odontologist Point of View. *Transl Med UniSa* **16**, 11-16.
3. Vos MB, Kaar JL, Welsh JA *et al.* (2017) Added Sugars and Cardiovascular Disease Risk in Children: A Scientific Statement From the American Heart Association. *Circulation* **135**, e1017-e1034.
4. Australian Institute of Health and Welfare (AIHW) (2017) *A picture of overweight and obesity in Australia*. Canberra, Australia: Australian Institute of Health and Welfare (AIHW),.
5. Ha D, Roberts-Thomson K, Arrow P *et al.* (2016) Oral health of Australian Children: The National Child Oral Health Study. In *Children's oral health status in Australia, 2012-14*, pp. 86-152 [LG Doc and AJ Spencer, editors]. Adelaide: University of Adelaide Press.
6. Fidler Mis N, Braegger C, Bronsky J *et al.* (2017) Sugar in Infants, Children and Adolescents: A Position Paper of the European Society for Paediatric Gastroenterology, Hepatology and Nutrition Committee on Nutrition. *J Pediatr Gastroenterol Nutr* **65**, 681-696.
7. Birch L, Savage JS Ventura A (2007) Influences on the Development of Children's Eating Behaviours: From Infancy to Adolescence. *Can J Diet Pract Res* **68**, s1-s56.
8. COAG (Council of Australian Governments) Health Council (2015) *Healthy Mouths Healthy Lives - Australia's National Oral Health Plan 2015-2024*. Adelaide: South Australian Dental Service.
9. World Health Organization (2017) *Report of the commission on ending childhood obesity*. Geneva.

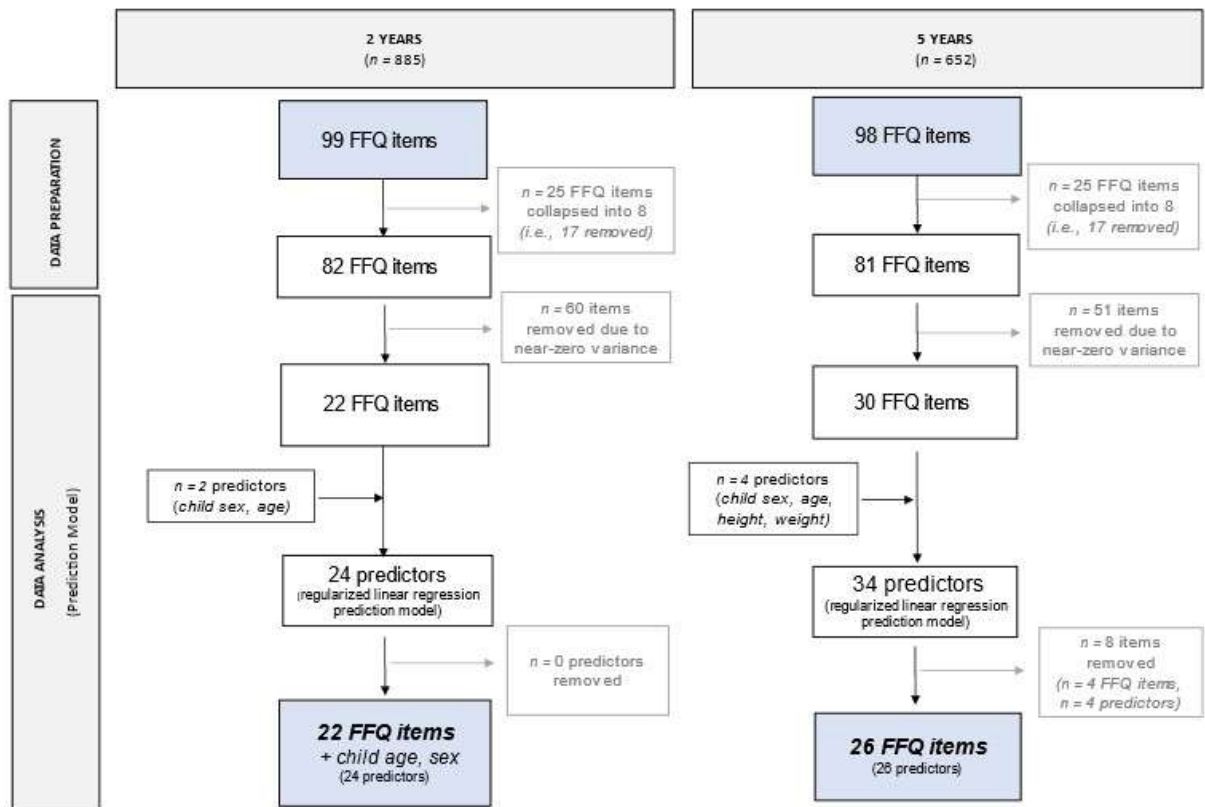
10. Mahboobi Z, Pakdaman A, Yazdani R *et al.* (2021) Dietary free sugar and dental caries in children: A systematic review on longitudinal studies. *Health Promot Perspect* **11**, 271-280.
11. Moynihan P (2016) Sugars and Dental Caries: Evidence for Setting a Recommended Threshold for Intake. *Adv Nutr* **7**, 149-156.
12. Crowe M, O'Sullivan M, Cassetti O *et al.* (2020) Estimation and consumption pattern of free sugar intake in 3-year-old Irish preschool children. *Eur J Nutr* **59**, 2065-2074.
13. Devenish G, Ytterstad E, Begley A *et al.* (2019) Intake, sources, and determinants of free sugars intake in Australian children aged 12-14 months. *Matern Child Nutr* **15**, e12692.
14. Devenish G, Golley RK, Mukhtar A *et al.* (2019) Free Sugars Intake, Sources and Determinants of High Consumption among Australian 2-Year-Olds in the SMILE Cohort. *Nutrients* **11**, 161.
15. Byrne R, Bell L, Taylor RW *et al.* (2019) Brief tools to measure obesity-related behaviours in children under 5 years of age: A systematic review. *Obes Rev* **20**, 432-447.
16. Golley RK, Bell LK, Hendrie GA *et al.* (2017) Validity of short food questionnaire items to measure intake in children and adolescents: a systematic review. *J Hum Nutr Diet* **30**, 36-50.
17. Devenish G, Mukhtar A, Begley A *et al.* (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**.
18. National Institutes of Health National Cancer Institute. Dietary Assessment Primer, Screeners at a Glance. <https://dietassessmentprimer.cancer.gov/profiles/screeners/> (accessed 21 June 2022)
19. Bell LK, Golley RK, Magarey AM (2013) Short tools to assess young children's dietary intake: a systematic review focusing on application to dietary index research. *J Obes* **709626**, Epublication 26 September 2013.
20. Do LG, Ha DH, Bell LK *et al.* (2020) Study of Mothers' and Infants' Life Events Affecting Oral Health (SMILE) birth cohort study: cohort profile. *BMJ Open* **10**, e041185.
21. Do LG, Scott JA, Thomson WM *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.
22. van der Bend D, Bucher T, Schumacher TL *et al.* (2017) Trends in Food and Beverage Portion Sizes in Australian Children; a Time-Series Analysis Comparing 2007 and 2011-2012 National Data. *Children (Basel)* **4**.

23. Alejandro-Vega S, Suarez-Marichal D, Niebla-Canelo D *et al.* (2022) Fluoride Exposure from Ready-To-Drink Coffee Consumption. *Life (Basel)* **12**.
24. World Health Organization (2000) *Obesity: preventing and managing the global epidemic. Report of a WHO consultation*. WHO Technical Report Series Geneva: WHO.
25. Australian Bureau of Statistics (2013) *Socio-Economic Indexes for Areas (SEIFA) 2011* [Australian Bureau of Statistics, editor]. Canberra, ACT, Australia.
26. World Health Organization (1995) *Physical Status: Use and Interpretation of Anthropometry*. Geneva, Switzerland: WHO Expert Committee on Physical Status.
27. World Health Organization (1995) *Weighing and measuring a child training course on child growth assessment*. Geneva, Switzerland: WHO.
28. Zarnowiecki D Development and validation of a short dietary questionnaire for assessment of obesity-related dietary behaviours in 6 month – 5-year-old children.
29. Bell LK, Golley RK Magarey AM (2015) Dietary risk scores of toddlers are associated with nutrient intakes and socio-demographic factors, but not weight status. *Nutrition and Dietetics* **73**, 73-80.
30. Hendrie GA, Viner Smith E Golley RK (2014) The reliability and relative validity of a diet index score for 4-11-year-old children derived from a parent-reported short food survey. *Public Health Nutr* **17**, 1486-1497.
31. Bell LK, Golley RK Magarey AM (2014) A short food-group based dietary questionnaire is reliable and valid for assessing toddlers' dietary risk in relatively advantaged samples. *British Journal of Nutrition* **112**, 627-637.
32. Bell LK, Golley RK, Mauch C *et al.* (2018) Validation testing of a short food-group based questionnaire to assess dietary risk in preschoolers aged 3-5 years. *Nutrition and Dietetics* **In press**.
33. Grady A, Fielding A, Golley RK *et al.* (2020) Adaptation, acceptability and feasibility of a Short Food Survey to assess the dietary intake of children during attendance at childcare. *Public Health Nutr* **23**, 1484-1494.
34. Magarey A, Golley RK, Spurrier N *et al.* (2009) Reliability and validity of the Children's Dietary Questionnaire; a new tool to measure children's dietary patterns. *Int J Pediatr Obes* **4**, 257-265.
35. Food Standards Australia and New Zealand Determining the Amount of Added Sugars and Free Sugars in Foods Listed in the AUSNUT 2011–13 Dataset. <https://www.foodstandards.gov.au/science/monitoringnutrients/ausnut/foodnutrient/Pages/Determining-the-amount-of-added-sugars-and-free-sugars-in-foods-listed-in-the-AUSNUT-201113-dataset.aspx> (accessed 19 February 2018)

36. National Health and Medical Research Council (2006) Nutrient Reference Values for Australia and New Zealand; [NHMRC, editor]. Canberra, Australia.
37. National Health and Medical Research Council (2013) Eat for Health Guidelines: Canberra, Australia.
38. Byrne R, Zhou Y, Perry R *et al.* (2018) Beverage intake of Australian children and relationship with intake of fruit, vegetables, milk and body weight at 2, 3.7 and 5 years of age. *Nutr Diet* **75**, 159-166.
39. Cockburn N, Lalloo R, Schubert L *et al.* (2018) Beverage consumption in Australian children. *Eur J Clin Nutr* **72**, 401-409.
40. Australian Bureau of Statistics (2011) Australian Health Survey: Nutrition First Results - Foods and Nutrients. <https://www.abs.gov.au/statistics/health/health-conditions-and-risks/australian-health-survey-nutrition-first-results-foods-and-nutrients/latest-release#discretionary-foods> (accessed 13 December 2022)
41. Australian Bureau of Statistics (2018) Australian Health Survey: Consumption of Food Groups from the Australian Dietary Guidelines, 2011-2012. <https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/4364.0.55.012~2011-12~Main%20Features~Fruit~15#2> (accessed August 2023)
42. Baird DL, Syrette J, Hendrie GA *et al.* (2012) Dairy food intake of Australian children and adolescents 2-16 years of age: 2007 Australian National Children's Nutrition and Physical Activity Survey. *Public Health Nutr* **15**, 2060-2073.
43. Fayet-Moore F, Cassettari T, McConnell A *et al.* (2019) Australian children and adolescents who were drinkers of plain and flavored milk had the highest intakes of milk, total dairy, and calcium. *Nutr Res* **66**, 68-81.
44. Burrows TL, Martin RJ Collins CE (2010) A systematic review of the validity of dietary assessment methods in children when compared with the method of doubly labeled water. *J Am Diet Assoc* **110**, 1501-1510.
45. Daley AJ (2022) Time to get our teeth into reducing obesity: should dentists screen and deliver interventions to reduce obesity in the population? *Br Dent J* **232**, 78-79.



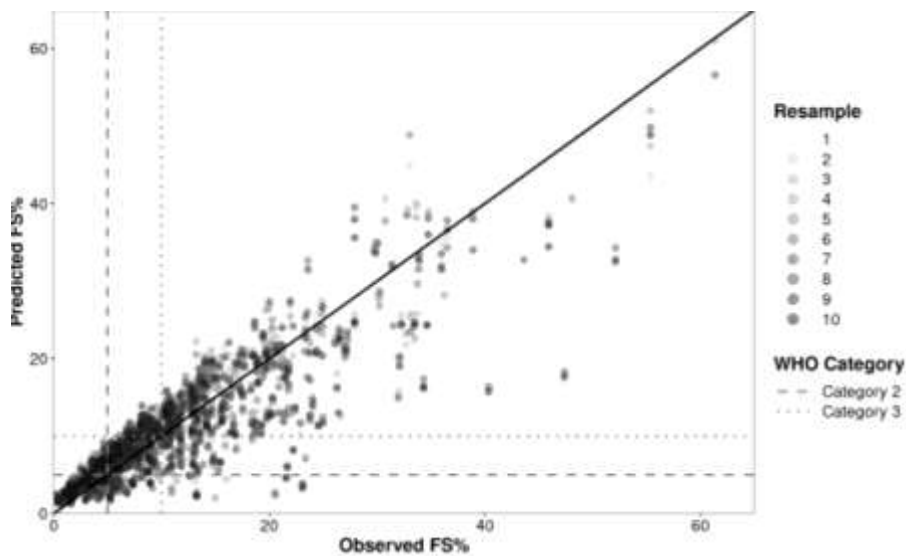
**Figure 1.** Participant flow through the study to develop a short-form version of the SMILE-Food Frequency Questionnaire (SMILE-FFQ)



**Figure 2.** Reduction of the 99-item and 98-item SMILE-FFQ into the 22-item and 26-item SMILE-FSS at 2 years and 5 years, respectively

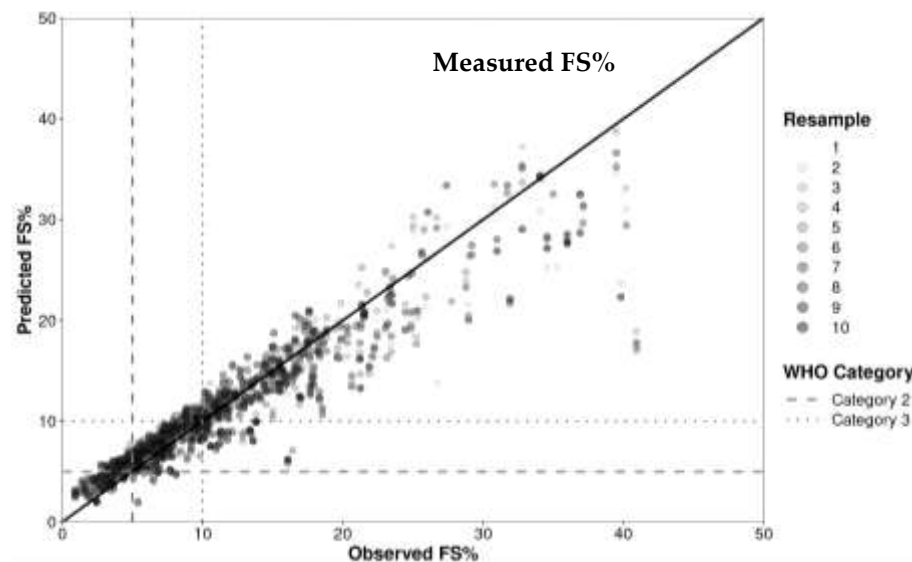


## 2 Years



&lt;5%EFS

## 5 years



&lt;5%EFS

**Figure 3:** Scatter plot of measured and predicted %EFS with the unity line (representing perfect calibration), using the testing sample ( $n = 263$  at 2 years;  $n = 192$  at 5 years). The horizontal and vertical dashed  $\epsilon$  Measured FS% represent the cut-off for <5% total energy from free sugars (EFS) and <10% total energy from free sugars (EFS), respectively.

**Table 1.** Characteristics of mother-child dyads in the samples with complete dietary and sociodemographic data at 2 and 5 years <sup>a</sup>

	<b>2 Years</b>		<b>5 Years</b>	
	(n = 885)		(n = 652)	
<b>Child Characteristics</b>				
Age at time of questionnaire completion (years) (mean, SD)	2.1	0.1	5.3	0.4
Child's sex				
Male	469	53.0	357	54.8
Female	416	47.0	295	45.2
Child's BMI category <sup>b</sup>				
Healthy weight	673	76.1	448	68.7
Overweight	71	8.0	143	21.9
Obesity	23	2.6	61	9.4
<b>Maternal and Household Characteristics</b>				
Mother's age at child's birth (years) (mean, SD)	30.5	5.0	30.8	5.1
Mother's pre-pregnancy BMI <sup>c, d</sup>				
<25kg/m <sup>2</sup>	476	53.8	315	48.3
25-29.9kg/m <sup>2</sup>	176	19.9	122	18.7
≥30kg/m <sup>2</sup>	174	19.7	90	13.8
Mother's country of birth				
Australia or New Zealand or UK	686	77.5	480	73.6
Other	191	21.6	165	25.3
Mother's education attainment				
High school / vocational	379	42.8	254	93.0
Some university and above	506	57.2	398	61.0
Annual household income, \$AUD				
<40,000	97	11.0	65	10.0
40,001-80,000	267	30.2	208	31.9

80,001-120,000	269	30.4	216	33.1
>120,001	245	27.7	156	23.9
<hr/>				
Index of Relative Socio-Economic Advantage and Disadvantage				
<hr/>				
Deciles 1-2 (most disadvantaged)	203	22.9	146	22.4
Deciles 3-4	231	26.1	161	24.7
Deciles 5-6	151	17.1	120	18.4
Deciles 7-8	220	24.9	166	25.5
Deciles 9-10 (most advantaged)	73	8.2	54	8.3
<hr/>				
Number of children at home				
<hr/>				
1	413	46.7	317	48.6
2	310	35.0	232	35.6
≥3	137	15.5	84	12.9
<hr/>				
Two parent household				
<hr/>				
Yes	834	94.2	616	94.5
<hr/>				

Abbreviations: BMI, Body Mass Index; NZ, New Zealand; UK, United Kingdom;

<sup>a</sup> Data presented as N (%) unless otherwise stated

<sup>b</sup> Age and sex-specific BMI z-scores were calculated for children and classified into weight status categories (healthy weight  $\geq -2$  and  $\leq +2SD$ , overweight  $> +2$  and  $\leq +3SD$ , and obesity  $> +3$ ) using the WHO reference<sup>(27)</sup>

<sup>c</sup> Weight status categories equivalent to body mass index (BMI) (kg/m<sup>2</sup>); Healthy weight  $< 25$  kg/m<sup>2</sup>, Overweight 25–29.9 kg/m<sup>2</sup>, Obesity  $\geq 30$  kg/m<sup>2</sup>.

<sup>d</sup> Missing data  $n = 118$  at 2 years.

**Table 2:** Model coefficients after variable shrinkage across 10 cross-validation runs of the regularised regression-based prediction model, using the training sample ( $n = 622$ ) at **2 years**<sup>a</sup>

Description	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
(Intercept)	-2.452	-5.189	3.227	-1.301	0.28	-8.584	2.151	<b>4.445</b>	-6.137	-4.717
<i>Child factors</i>										
Child age (at 2 years)	3.695	4.535	0.527	2.71	1.957	6.055	1.705	<b>0.04</b>	5.014	4.646
Child sex	-0.773	-0.667	-0.214	0	-0.073	-0.293	-0.317	<b>-0.223</b>	-0.575	-0.67
<i>Food frequency items</i>										
Custard: plain or vanilla	1.331	1.314	1.398	1.391	1.123	1.334	1.242	<b>1.335</b>	1.353	1.326
Icy-poles and sorbet	2.037	1.753	1.752	1.528	1.518	1.397	1.678	<b>1.565</b>	1.776	1.748
Nut paste	0	5.973	4.183	3.517	5.444	5.138	4.149	<b>6.578</b>	5.92	1.467
Jam, marmalade and other fruit spreads	0.904	0.87	1.208	1.082	1.338	1.264	1.207	<b>1.204</b>	1.366	1.219
Honey	0.909	0.975	0.597	0.704	0.918	0.9	0.859	<b>0.946</b>	0.905	0.942
Wheat biscuits: plain	-0.368	-0.12	0.245	0	0.158	0.408	0	<b>-0.353</b>	-0.026	0.844
Cereal flakes	1.787	3.216	3.653	3.421	2.643	3.182	2.247	<b>2.385</b>	2.57	1.904
Flavoured breakfast cereals	0.984	0.758	0.831	0.563	0.798	1.026	0.992	<b>1.181</b>	0.946	1.055
Sweet biscuits: plain	0.66	0.707	0.579	0.868	0.849	0.923	0.61	<b>0.727</b>	0.918	0.792
Sweet biscuits: not plain	2.375	2.342	2.324	2.03	1.605	2.094	1.997	<b>1.884</b>	2.282	1.792

Muesli and cake-type bars	0.927	1.122	1.088	0.986	0.994	0.879	0.671	<b>0.57</b>	0.687	0.889
Cakes and puddings	1.266	1.233	1.119	1.131	1.182	1.11	1.09	<b>1.212</b>	1.205	1.168
Sweet breads and pancakes	0.331	0.722	0.673	0.767	0.847	0.702	0.662	<b>0.719</b>	0.421	0.491
Juice: fruit (no added sugar/100% fruit juice)	0.879	0.912	0.923	0.889	0.925	0.906	0.821	<b>0.944</b>	0.921	0.949
Sauce: tomato or barbecue	1.478	1.133	1.044	1.184	0.981	1.161	1.169	<b>1.745</b>	1.472	1.05
Sauce and marinade: sweet style	-1.13	-0.676	1.257	0	0	0.313	0	<b>0.389</b>	0	0.281
Sugar: regular table sugar	1.499	2.075	1.491	1.646	1.466	2.368	1.426	<b>1.783</b>	1.564	1.451
Chocolate or carob	1.418	1.358	1.623	1.413	1.607	1.382	1.289	<b>1.131</b>	1.328	1.814
Ice-cream, frozen yoghurt, and non-dairy alternatives	1.89	1.865	1.601	1.643	1.535	1.363	1.86	<b>1.655</b>	1.5	1.802
Yoghurt and alternatives, flavoured	1.3	1.402	1.345	1.384	1.363	1.316	1.271	<b>1.374</b>	1.355	1.232
Savoury biscuits and snack foods	0	-1.021	0.91	0	0.401	0	0.055	<b>-1.165</b>	0.132	0.099
Lollies: regular (not sugarfree)	1.866	1.363	1.762	1.864	1.521	1.635	1.893	<b>1.85</b>	1.716	1.385

<sup>a</sup>Bold values indicate coefficients for best model with minimum Root Mean Square Error difference between training (2263.86) and testing (2263.59) data

**Table 3:** Model coefficients after variable shrinkage across 10 cross-validation runs of the regularised regression-based prediction model, using the training sample ( $n = 460$ ) at 5 years<sup>a</sup>

Description	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
(Intercept)	6.769	8.003	7.013	6.562	6.466	7.048	7.628	<b>6.894</b>	7.012	6.78
<i>Child factors</i>										
Child age (at 5 years)	0	0	0	0	0	0	0	<b>0</b>	0	0
Child sex	0	0	0	0	0	0	0	<b>0</b>	0	0
Child height	0	0	0	0	0	0	0	<b>0</b>	0	0
Child weight	0	0	0	0	0	0	0	<b>0</b>	0	0
<i>Food frequency items</i>										
Milk: flavoured	1.487	0.439	1.255	1.355	0.865	1.544	1.186	<b>1.153</b>	1.197	1.706
Custard: infant and toddler	2.411	0	2.08	0.888	1.241	0.589	2.253	<b>1.109</b>	1.932	1.145
Icy-poles and sorbet	1.286	1.014	1.253	1.229	1.107	1.08	1.308	<b>1.067</b>	1.238	1.149
Nut paste	0	0	0	0	0	0.07	0	<b>0</b>	0.389	0
Chocolate, choc-nut or carob spreads	1.038	1.13	1.191	0.942	0.958	0.973	0.758	<b>0.602</b>	0.754	0.926
Jam, marmalade and other fruit spreads	1.03	0.944	0.993	0.98	1.001	0.89	0.977	<b>1.065</b>	0.95	0.976
Honey	0.723	0.729	0.691	0.797	0.773	0.852	0.804	<b>0.82</b>	0.793	0.813
Wheat biscuits: plain	0	0	0	0	0	0	0	<b>0</b>	0	0
Cereal flakes	0	0	0	0	0	0	0	<b>0</b>	0	0
Flavoured breakfast cereals	0.758	1.093	1.011	0.722	1.097	0.781	0.735	<b>0.728</b>	0.837	0.934
Sweet biscuits: plain	0.756	0.628	0.885	0.886	1.226	1.062	0.493	<b>0.826</b>	0.991	0.781
Sweet biscuits: not plain	1.207	1.251	1.054	1.08	0.917	1.013	1.179	<b>1.072</b>	1.095	1.141
Muesli and cake-type bars	0.789	1.081	0.712	0.883	0.64	0.813	1.135	<b>1.041</b>	0.627	0.78
Cakes and puddings	0.859	1.034	0.86	0.916	0.877	0.834	0.904	<b>0.946</b>	0.848	0.895
Sweet breads and pancakes	0.663	0.8	0.647	0.61	0.554	0.632	0.61	<b>0.793</b>	0.567	0.74
Sweet pastry	1.049	0.965	1.112	1.184	1.137	1.425	1.113	<b>1.273</b>	1.058	1.212
Soft drink: regular (not sugarfree)	0.668	1.191	0.784	1.24	1.072	0.772	0.708	<b>0.773</b>	0.777	0.607

Description	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
Juice: fruit drink	1.153	0.829	1.262	1.162	1.137	1.101	1.131	<b>1.165</b>	1.13	0.733
Juice: fruit (no added sugar/100% fruit juice)	0.911	0.831	0.872	0.787	0.848	0.872	0.884	<b>0.771</b>	0.882	0.738
Milo or Ovaltine in milk drinks	0.037	0.768	0.36	0	0	0	0.456	<b>0.432</b>	0.014	0.258
Sauce: tomato or barbecue	0.817	1.128	0.839	1.109	0.878	0.79	0.782	<b>1.11</b>	0.868	1.09
Sauce and marinade: sweet style	1.426	0.138	1.475	1.299	0.947	1.092	1.141	<b>1.349</b>	0.943	0.262
Mayonnaise: regular	0.671	0	0	0	0	0	0	<b>0</b>	0	0
Sugar: regular table sugar	0.872	0.856	0.896	1.544	0.908	0.905	0.855	<b>0.894</b>	0.873	0.948
Sugar syrups	1.066	0.476	0.647	0.635	0.964	0.631	0.509	<b>0.454</b>	0.831	0.581
Chocolate or carob	0.792	0.792	0.908	0.847	0.76	0.836	0.811	<b>0.829</b>	1.123	0.837
Ice-cream, frozen yoghurt, and non-dairy	1.114	1.36	0.973	1.148	1.315	1.097	1.208	<b>1.218</b>	1.094	1.193
Yoghurt and alternatives, flavoured	1.162	0.942	1.13	1.141	1.061	1.133	0.982	<b>1.004</b>	1.022	1.105
Savoury biscuits and snack foods	3.572	0.148	3.065	2.381	3.118	3.316	2.32	<b>2.453</b>	3.686	3.839
Lollies: regular (not sugarfree)	1.162	1.243	1.213	1.11	1.244	1.183	1.333	<b>1.225</b>	1.329	1.137

<sup>a</sup>Bold values indicate coefficients for best model with minimum Root Mean Square Error difference between training (2199.41) and testing (2196.15) data



**Table 4:** Cross-classification table of measured and predicted WHO free sugars percentage categories<sup>a</sup>, using the testing sample ( $n = 263$  at 2 years;  $n = 192$  at 5 years)<sup>b</sup>

Age group	Predicted category	Measured category		
		1	2	3
2 years	1	<b>60 (75.9%)</b>	10 (11.5%)	1 (1.0%)
	2	19 (24.1%)	<b>55 (63.2%)</b>	14 (14.4%)
	3	0 (0%)	22 (25.3%)	<b>82 (84.5%)</b>
5 years	1	<b>22 (56.4%)</b>	3 (3.9%)	0 (0%)
	2	17 (43.6%)	<b>68 (88.3%)</b>	8 (10.5%)
	3	0 (0%)	6 (7.8%)	<b>68 (89.5%)</b>

<sup>a</sup>WHO categories: (1) <5%EFS, (2) 5-<10%EFS, (3) ≥10%EFS

<sup>b</sup>column-wise percentages in the diagonal equal the proportion of the participants **within each category** correctly predicted to be in a given category