Weighting strategy details

Three sets of weights were developed to make the participant sample comparable to the Swiss population. The first set was calculated for participants who attended the on-site visit (N = 1 852 participants), the second set was for those who consented to blood sample collection (N = 848 participants) and the third for all participants with a questionnaire, including those without an on-site visit (N = 1 935). The weight calculation process was similar for all sets of weights and can be summarized in four steps.

In the first step, sampling weights were calculated to reflect the sampling methodology. The sampling was conducted in four waves across 18 strata (six centres by three age groups), with population totals available for each stratum at each wave. To simplify the weight calculation, data from the four waves were pooled and an average sampling frame was calculated by averaging the population totals over the four waves. The sampling weights were defined as the inverse of the inclusion probabilities, which were assumed equal for all individuals within the same stratum. The inclusion probabilities were defined as the ratio of the total number of individuals invited to participate in a given stratum to the size of the average sampling frame for that stratum.

The second step involved correcting the sampling weights obtained at step one for non-response. This adjustment was achieved by modelling the response probability based on variables that were used to define the strata (i.e. age group and study centre) as well as sex, nationality, household size, and Swiss-SEP quintiles[1] using a logistic regression model. Initially, all predictors and their pairwise interactions were included in the model. A backwards elimination process based on the Bayesian Information Criterion (BIC)[2] was then applied to refine the model, retaining only those predictors and interactions whose removal would result in a higher BIC (indicating a poorer fit). The predicted response probabilities from the final model were then analysed using a hierarchical clustering algorithm (k-means) to build up to five homogeneous response groups. The common response probability in each group was calculated as the weighted average of the participation indicator (1 for participants, 0 for non-participants) for all individuals in the group, using the sampling weights obtained at the first step. The weights corrected for non-response were then computed as the ratio of the sampling weights obtained in step one to the common response probability for each individual's group (defined by k-means).

In the third step, the weights obtained after non-response correction were used to calibrate the sample to align with population totals for design and auxiliary variables through a raking procedure executed in two phases[3]. First, a preliminary calibration was performed to align the sample with population totals for the 7 major regions of Switzerland. The resulting weights were then adjusted for the remaining variables: sampling strata (18 categories which combine age groups and study centres), age (separate years), sex, nationality, household size, Swiss SEP quintiles, and season. The process was then repeated until convergence. This sequential calibration approach was adopted because the raking algorithm often had difficulty converging when both the 18 strata and the 7 major regions of Switzerland were included simultaneously, as the strata information (which includes the area around the study centre) is largely colinear with the regional information. This method ensured that the strata margins were matched while still achieving adequate calibration for the regions. Note that the four seasons were defined based on the month of the visit with winter covering the months December to February. The margins for each season were calculated by dividing the size of the average sampling frame by four. To account for difference in dietary intake during the week, the day of the week was used as an additional auxiliary variable in the calibration process for the first set of weights only (participants who attended the on-site visit and had 2x24HDR). The day of the week was categorized into three groups (week=Monday to Thursday, weekend=Friday to Sunday, mix of both) based on the days of the first and second 24HDRs. Participants with a single 24HDR were allocated to the "week" or "weekend" categories based on the day of their unique 24HDR. The margins for the day of the week were obtained by multiplying the size of the average sampling frame by 16/49 (week), 9/49 (weekend) or 24/49 (mix), corresponding to the combinations of ordered pairs of days of the two weeks falling into each category.

Finally, the last step involved trimming excessively large weights that could adversely affect the variance of the weighted estimator. A trimming limit was set at the median plus five times the interquartile range of the weights. Any weight exceeding this limit was capped at the limit, and the remaining weights were adjusted proportionally to ensure the sum of the weights still equals the size of the sampling frame. When weight trimming was necessary, the calibration process (step three) was repeated on the trimmed weights to ensure proper calibration. This process of weight trimming and calibration was carried out sequentially until no weight exceeded the trimming limit.

A rescaling bootstrap procedure based on the approach proposed by Rao and Wu (1988) was used to build 1000 replicates for each set of weights[4]. These replicates were used to calculate the variance of the weighted estimators while accounting for the sampling design.

1. Panczak, R., Berlin, C., Voorpostel, M., Zwahlen, M. & Egger, M. The Swiss neighbourhood index of socioeconomic position: update and re-validation. *Swiss Med Wkly* **153**, 40028 (2023).

2. Schwarz, G. Estimating the Dimension of a Model. *Ann. Statist.* **6**, (1978).

3. Deville, J.-C. & Särndal, C.-E. Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association* **87**, 376–382 (1992).

4. Rao, J. N. K. & Wu, C. F. J. Resampling Inference with Complex Survey Data. *Journal of the American Statistical Association* **83**, 231–241 (1988).