PERFORMANCE OF SUBSTITUTION MODELS IN NUTRITIONAL EPIDEMIOLOGY

Background Dietary guidelines often recommend substituting certain nutrients or foods with healthier alternatives, based on the available evidence from nutritional epidemiology. The effects of food substitutions can be examined by conducting isocaloric dietary interventions, but experimental studies are often not practical or sufficiently generalisable. Therefore, nutritional epidemiology is highly reliant on observational data, in which food substitutions can be explored using mathematical modelling. The two modelling approaches commonly used for estimating substitution effects are known as (1) the ‘leave-one-out’ model, in which total energy intake and all dietary components are included as covariates, excluding the nutrient(s) that the exposure should be substituted with; and (2) the energy partition model, in which all dietary components are included as covariates, without further adjustment for total energy intake. It remains underappreciated that these approaches do not perform equally well for estimating substitution effects, and that there is limited evidence on whether they produce unbiased estimates.

Methods Semi-parametric directed acyclic graphs and Monte Carlo data simulations were used to explore the performance of the two approaches for estimating the following estimands: 1) the average relative causal effect (i.e. the joint effect of increasing intake of the exposure and decreasing the intake of all other nutrients, while keeping total energy intake constant), 2) the relative effect of increasing the exposure nutrient and decreasing the intake of one other nutrient, and 3) the relative effect of increasing the exposure nutrient and decreasing the intake of a combination of other nutrients. The approaches were explored both in the absence and presence of confounding that acts through diet.

Results The ‘leave-one-out’ model produced a biased estimate of the average relative causal effect even in the absence of confounding. It robustly estimated substituting the exposure with another specific nutrient regardless of whether confounding was present but produced biased estimates of substituting the exposure for a combination of other nutrients even in the absence of confounding. The energy partition model robustly estimated all three estimands of interest, producing unbiased estimates regardless of whether confounding was present or not.

Conclusion Only the energy partition model produces unbiased estimates of different substitution effects in the context of nutritional epidemiology. It performs equally well even in the presence of confounding that acts through diet. Substitution analyses using the ‘leave-one-out’ approach might not be robust and any existing studies using this model might suffer from bias.