*total causal effect* (i.e. an 'additive' effect) but only provides unbiased estimates in the absence of confounding or when all individual nutrients have equal effects on the outcome. The 'nutrient density model' does not target a causally meaningful estimand but can provide extremely biased estimates of the average relative causal effect of the exposure rescaled as a percentage of total energy intake. Accurate estimates of both the total and average relative causal effects were obtained with the 'all-components model'.

**Conclusion** Only the 'all-components model' produces unbiased estimates of different causal effects. Lack of awareness of the estimand differences and accuracy of the different modelling approaches may explain some of the apparent heterogeneity among existing nutritional studies. Serious questions may be raised regarding the validity of meta-analyses where different strategies returning different estimands have been inappropriately pooled.

## OP82 PERFORMANCE OF SUBSTITUTION MODELS IN NUTRITIONAL EPIDEMIOLOGY

[1,2,3]Georgia Tomova*, [1,4]Kellyn Arnold, [1,2,3]Mark Gilthorpe, [1,2,3]Peter Tennant. [1]*Leeds Institute for Data Analytics, University of Leeds, Leeds, UK*; [2]*Faculty of Medicine and Health, University of Leeds, Leeds, UK*; [3]*The Alan Turing Institute, London, UK*; [4]*Faculty of Environment, University of Leeds, Leeds, UK*

10.1136/jech-2021-SSMabstracts.82

**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 any 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.

## OP83 DEVELOPING A MEASURE OF DIETARY QUALITY FOR THE UK BIOBANK

[1,2]Chloe Montague*, [1]Stefania D'Angelo, [1,3]Nicholas C Harvey, [1,3]Christina Vogel, [1,3]Janis Baird. [1]*MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, UK*; [2]*Health Education England – Wessex, Health Education England, Winchester, UK*; [3]*NIHR Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust, Southampton, UK*

10.1136/jech-2021-SSMabstracts.83

**Background** The UK Biobank (UKB) has assessed diet among participants using a food frequency questionnaire (FFQ). Studies of UKB have tended to examine the association of intake of single foods with health outcomes. We used the FFQ data to develop a diet quality score that could be used to examine the relationship of dietary patterns with markers of cardiometabolic health.

**Methods** We applied principal components analysis to the FFQ data for participants from the Foodscapes section of UKB (n=52,345). The first component was transformed using Fisher-Yates normal scores, mapping it onto a Normal distribution with a mean of 0 and a standard deviation of 1. The median value was used to categorise diet as healthy or unhealthy. We examined the relationship of the diet scores with markers of cardiometabolic health (blood pressure, serum lipids, BMI) in linear regression analyses, adjusting for confounding factors identified in a Directed Acyclic Graph: sex, age, ethnicity, educational attainment, smoking, Index of Multiple Deprivation.

**Results** Of the 52,345 potential participants, 33,149 (63%) had complete dietary data. Just under half (45%) were men. The average age was 56.2 years (SD=8.2). Most (84%) were white and almost half had a college/university degree (49.5%). The first component explained 14% of the variation in the diet data. It was characterised by high consumption of beef, any type of meat, white bread and low fibre cereals, and low intake of fresh/dried fruit, salad and cooked vegetables and was defined as an 'unhealthy diet score'. Coefficients were multiplied by -1, so a positive value of the score indicated a healthy diet. A higher score, indicative of healthier diet, was associated with lower systolic and diastolic blood pressure (β -0.81, 95% CI -1.0, -0.62; β -0.61, 95% CI -0.72, -0.05), lower BMI (β -0.67, 95% CI -0.73, -0.62) and with healthier lipid profile (lower levels of cholesterol β -0.05, 95% CI -0.06, -0.04, triglycerides β -0.05, 95% CI -0.06, -0.03, and higher HDL cholesterol β 0.01, 95% CI 0.0, 0.01). Participants with healthier diet scores (>1) were less likely to be current smokers and more likely to be physically active (>1/day vigorous activity) (both P<0.001).