Causal inference of the factors leading individual to choose protein rich foods using directed acyclic graphical models

Irène Demongeot, Nicolas Darcel, Antoine Cornuéjols

UMR MIA , AgroParisTech, INRA, Université Paris-Saclay, Paris, France

irene.demongeot@agroparistech.fr

A better understanding of the determinants of preferences for high-protein food would be a precious help in view of formulating efficient food recommendations, particularly in a general context of recommendation of reduction of meat consumption. The aim of this study was to identify causal factors of choices for protein rich sources. We conducted an online survey on food preferences consisting in collecting in adults subjects, individual informations, and current levels of hunger and thirst as well as preferences between pairs of pictures of food items at the time of the survey. Foods were selected among a representative subset of common dishes varying in protein content. We aimed to measure the causal effect of the different collected parameters on the preference for protein-rich foods. The preference for protein-rich foods was computed as the mean difference between the actual content in the chosen food and the content of rejected food. As causal links cannot be discovered with usual statistical methods (since correlation does not imply causality) we created a graph representing the relationships between all collected parameter. We used back-door criterion (Pearl, 1995) to determine which variable had an identifiable causal effect. We used several matching methods and propensity score methods to estimate the causal effects. Among all the considered variables, we identified the state of hunger and the gender as causal factors for choosing high-protein food. In future studies, this promising methodology should be applied to bigger datasets from large numbers of participants to yield precise information on factors leading individuals to choose protein rich foods.