Methodological Considerations in the Application of Multiverse Analyses for More Rigorous Inferences About Nutrition and Health

Colby Vorland,[1] Lauren O'Connor,[2] Beate Henschel,[1] Cuiqiong Huo,[1] James Shikany,[3] Keisuke Ejima,[1] Aurelian Bidulescu,[1] and Andrew Brown[1]

[1]Indiana University; [2]Purdue University; and [3]University of Alabama at Birmingham

Objectives: Evaluating associations between nutritional exposures and disease outcomes requires numerous analytical decisions. Among these decisions are how nutritional exposures are defined and configured, choices and configuration of covariates, participant exclusion criteria, among others. Each decision potentially affects the results and conclusions, yet only a small number of potential models are generally conducted as sensitivity analyses. One approach to evaluate the robustness of conclusions from these myriad decisions is to analyze all reasonable approaches in a so called 'multiverse analysis'. As part of a project to apply multiverse methodology in nutrition research, we report here some methodological considerations, including the magnitude of covariate selection possibilities.

Methods: We are conducting a multiverse analysis of the association between self-reported beef intake and coronary heart disease (CHD) using the Reasons for Geographic and Racial Differences in Stroke (REGARDS) cohort. We identified common approaches to defining the exposure, outcome, and covariate selection used in the literature to identify reasonable models to incorporate in our multiverse analysis.

Results: Covariates were pooled into 29 distinct concepts (e.g., 'Physical activity', 'Weight') each which contain a number of possible covariate configurations. For practicality, we limited our analysis to: use of Cox proportional hazards models; one operationalization of CHD; covariate and exposure definitions to a subset of possible configurations for food and nutrients; and excluding linear transformations (e.g., grams of beef versus an 80 gram serving of beef) that do not affect statistical significance. We then varied the number of covariates included and excluded within each model to ensure that covariate concepts were not duplicative, as well as covariate configuration (e.g., continuous, categorical). We estimated 338 quadrillion model specifications with our exposure and covariate lists.

Conclusions: Researchers are faced with innumerable choices during model specification in nutrition epidemiology. Multiverse approaches have potential to inform how these choices may influence conclusions from nutrition epidemiological research, but further methodological development is needed.

Funding Sources: Funded by the Beef Checkoff