

1 **Variations in the results of nutritional epidemiology studies due to analytic**  
2 **flexibility: Application of specification curve analysis to red meat and all-cause**  
3 **mortality**  
4

5 Yumin Wang

6 *Department of Biomedical Informatics, Harvard Medical School, Boston, MA*

7 Tyler Pitre

8 *Department of Medicine, McMaster University, Hamilton, ON*

9 Joshua D. Wallach

10 *Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA*

11 Russell J. de Souza

12 *Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON*

13 Tanvir Jassal

14 *Department of Anesthesia, McMaster University, Hamilton, ON*

15 Dennis Bier

16 *Department of Pediatrics, Baylor College of Medicine, Houston, TX*

17 Chirag J. Patel

18 *Department of Biomedical Informatics, Harvard Medical School, Boston, MA*

19 Dena Zeraatkar

20 *Department of Anesthesia; Department of Health Research Methods, Evidence, and Impact, McMaster University,*  
21 *Hamilton, ON*

22  
23 **Corresponding author:**

24 Dena Zeraatkar

25 1280 Main St. W

26 Hamilton, Ontario

27 [zeraatd@mcmaster.ca](mailto:zeraatd@mcmaster.ca)

28  
29 **Keywords:** nutrition, red meat, all-cause mortality, multiverse analysis, specification curve analysis,  
30 vibration of effects

31 **Word count:** 4391

32 **Figures/Tables:** 2 tables, 2 figures

33 **Disclaimers:** Dr Wallach reported receiving grant support from the FDA, Arnold Ventures, Johnson &  
34 Johnson through Yale University, and the National Institute on Alcohol Abuse and Alcoholism of the NIH  
35 under award 1K01AA028258; serving as a consultant for Hagens Berman Sobol Shapiro LLP and Dugan  
36 Law Firm APLC; and serving as a medRxiv affiliate.

37 **Funding:** None.

38 **Acknowledgments:** None.

39 **Contributions:** DZ and CJP conceptualized the study. DZ, YW, JDW, RJdS, TP, CJP, and DZ provided input  
40 on the study design and methods. YW, TJ, and TP collected data. YW performed analyses, with input  
41 from DZ and CJP. YW and DZ wrote the first draft of the manuscript and all authors provided critical  
42 comments. The senior author (manuscripts guarantor) affirms that the manuscript is an honest,  
43 accurate, and transparent account of the study being reported and that no important aspects of the  
44 study have been omitted.

The application of specification curve analysis to nutritional epidemiology: red meat and all-cause mortality

45 **Ethics approval:** Not required.

46 **Patient/public engagement:** It was not possible to involve patients or the public in the design, conduct,  
47 reporting, or dissemination plans of our research.

48 **Abstract**

49

50 **Objective:** To present an application of specification curve analysis—a novel analytic method that  
51 involves defining and implementing all plausible and valid analytic approaches for addressing a research  
52 question—to nutritional epidemiology.

53 **Data source:** National Health and Nutrition Examination Survey (NHANES) 2007 to 2014 linked with  
54 National Death Index.

55 **Methods:** We reviewed all observational studies addressing the effect of red meat on all-cause  
56 mortality, sourced from a published systematic review, and documented variations in analytic methods  
57 (e.g., choice of model, covariates, etc.). We enumerated all defensible combinations of analytic choices  
58 to produce a comprehensive list of all the ways in which the data may reasonably be analyzed. We  
59 applied specification curve analysis to NHANES data to investigate the effect of unprocessed red meat  
60 on all-cause mortality, using all reasonable analytic specifications.

61 **Results:** Among 15 publications reporting on 24 cohorts included in the systematic review on red meat  
62 and all-cause mortality, we identified 70 unique analytic methods, each including different analytic  
63 models, covariates, and operationalizations of red meat (e.g., continuous vs. quantiles). We applied  
64 specification curve analysis to NHANES, including 10,661 participants. Our specification curve analysis  
65 included 1,208 unique analytic specifications. Of 1,208 specifications, 435 (36.0%) yielded a hazard ratio  
66 equal to or above 1 for the effect of red meat on all-cause mortality and 773 (64.0%) below 1, with a  
67 median hazard ratio of 0.94 [IQR: 0.83 to 1.05]. Forty-eight specifications (3.97%) were statistically  
68 significant, 40 of which indicated unprocessed red meat to reduce all-cause mortality and 8 of which  
69 indicated red meat to increase mortality.

70 **Conclusion:** We show that the application of specification curve analysis to nutritional epidemiology is  
71 feasible and presents an innovative solution to analytic flexibility.

72 **Limitations:** Alternative analytic specifications may address slightly different questions and investigators  
73 may disagree about justifiable analytic approaches. Further, specification curve analysis is time and  
74 resource-intensive and may not always be feasible.

75

76 **Background**

77 Unlike randomized trials for which investigators typically register protocols and statistical analysis plans  
78 before the collection of any data, when investigators analyze data from observational studies, there are  
79 often hundreds of equally justifiable ways of analyzing the data, each of which may produce results that  
80 vary in direction, magnitude, and statistical significance (1-7). The variability of effect estimates due to  
81 alternative analytical approaches is called ‘vibration of effects’ (2). Empirical evidence shows that results  
82 from observational studies may be highly dependent on analytic choices (1-5).

83 While our empirical and theoretical understanding of the question being investigated should guide our  
84 analytic choices, our knowledge of complex biomedical and environmental systems is limited and even  
85 experienced investigators come to different conclusions about the ideal analytic approach (4, 6, 8-13).

86 While we anticipate that discrepancies in analytic models often result from differences in opinions  
87 regarding the optimal analytic approach among well-intentioned investigators, some investigators may  
88 test many alternative analytic specifications and, intentionally or unintentionally, selectively report  
89 results for the specification that yields the most statistically significant or interesting results or results  
90 that support their preconceived hypotheses. Evidence shows that investigators’ prior beliefs and  
91 expectations influence their results (5). In the presence of very strong opinions, investigators’ beliefs  
92 and expectations may shape the literature to the detriment of empirical evidence (5).

93 ***Nutritional epidemiology***

94 Nutrition is a field particularly amenable to analytic flexibility (14). Trials investigating the health effects  
95 of nutritional exposures are often not feasible and so the evidence is primarily comprised of nutritional  
96 epidemiology studies—observational studies that recruit large groups of people and look for patterns  
97 between diet and health (15, 16).

98 The analysis of nutritional epidemiology data is complex and there is often limited consensus among  
99 experts about the ideal approach (17, 18). Sources of analytic flexibility include the type of analytic  
100 model (e.g., Poisson regression, Cox proportional hazards model), choice of covariates (i.e., investigators  
101 studying the same question will consider different adjusting variables (19)), operationalization of the  
102 exposure variable and covariates in the model (e.g., transformations, categorizations of continuous  
103 variables, functional form), and methods to address missing data, among others (8). Investigators often  
104 present several sensitivity analyses to investigate the effects of these uncertain analytic decisions on the  
105 results, but the choice of sensitivity analyses is also subjective and investigators may be more inclined to  
106 report sensitivity analyses that affirm their primary findings.

107 A large body of evidence shows inconsistency in the results of nutritional studies, some of which may be  
108 explained by analytic flexibility (3, 8, 20, 21). Such inconsistencies have eroded trust in nutritional  
109 epidemiology and subjected the field to criticism (22, 23). Nevertheless, nutritional epidemiology studies  
110 continue to play a crucial role in shaping dietary recommendations and policies, making it imperative to  
111 draw credible inferences from these studies (14, 15, 24).

### 112 ***Specification curve analysis***

113 Specification curve analysis—sometimes called multiverse analysis—is a novel analytic method that  
114 involves defining and implementing all plausible and valid analytic approaches for addressing a research  
115 question (25) (Box 1).

116 Through this approach, investigators define all plausible and justifiable choices for all aspects of the  
117 analysis (e.g., choice of model, covariates, etc.), enumerate all justifiable combinations of these choices  
118 to produce a comprehensive list of all the ways in which the data may be reasonably analyzed (i.e.,  
119 analytic specifications), implement all or a random sample of the ways in which the question may be  
120 analyzed, and draw inferences using the distribution of results from all plausible analyses.

121 Specification curve analysis offers advantages to conventional methods for data analysis. It allows  
122 investigators to draw more credible inferences that are not contingent on arbitrary analytic decisions  
123 and reduces the opportunity for investigators to conduct many analyses and selectively report results  
124 for analyses that yield the most interesting results, though it does not completely eliminate subjectivity  
125 in analytic decisions.

126 While specification curve analysis has been previously applied in psychology and economics, it has  
127 seldom been applied in nutritional and environmental epidemiology (5, 26).

#### **Box 1: Specification curve analysis**

When investigators analyze data from observational studies, they may make numerous potentially justifiable, but still subjective, analytic decisions on which the direction, magnitude, and statistical significance of results may be contingent. Specification curve analysis may mitigate this issue (27).

Specification curve analysis involves defining and implementing all plausible and justifiable analytic methods for investigating a research question. Investigators subsequently interpret the distribution of results across all plausible analyses, instead of focusing on the results of only one analysis.

The implementation of specification curve analysis involves:

- 1) Defining all plausible choices across all aspects of the analysis. This typically includes:
  - Criteria for selecting eligible participants for inclusion in the analysis

- Type of analytic model (e.g., logistic, Poisson, or Cox proportion hazards models)
  - Choice of covariates
  - Operationalizations of the exposure variable and covariates (e.g., transformations, functional form)
- 2) Enumerating all justifiable combinations of these analytic choices to produce a comprehensive list of all the ways in which the data may be reasonably analyzed. For example, three unique choices for five aspects of the analysis yield 243 unique analytic specifications ( $=3^5$ ).
  - 3) Implementing all or a random sample of all reasonable analytic specifications.
  - 4) Ordering the effect estimates from all analyses based on their direction and magnitude and presenting results on a specification curve plot. A specification curve plot reports the results of all analyses at the top and analytic characteristics at the bottom. The specification curve plot visually communicates the distribution of results across all specifications and the aspects of the analysis that are most consequential in influencing the direction and magnitude of findings.

128 **Objectives**

129 We apply specification curve analysis to investigate the effect of unprocessed red meat on all-cause  
130 mortality—a question that has yielded inconsistent results in the literature and produced conflicting  
131 dietary recommendations.

132 A critical limitation of specification curve analysis is the subjectivity involved in selecting justifiable  
133 analytic specifications. Investigators may disagree about justifiable analytic approaches or may present  
134 results of analyses that are only marginally justifiable. To mitigate this issue, our analytic specifications  
135 were informed by the most common analytic methods used in previous published studies addressing the  
136 effects of red meat on all-cause mortality.

137 **Methods**

138 This study was exempt from institutional ethics review because it uses secondary de-identified data. We  
139 report our results according to STROBE reporting guidelines for observational studies (28).

140 **Analytic specifications**

141 We used a published systematic review of observational studies that addressed the effect of red meat  
142 on all-cause mortality to identify justifiable analytic specifications for specification curve analysis (29).

143 We focus only on observational studies because randomized trials typically involve the preparation of  
144 detailed protocols and statistical analysis plans that reduce the analytic decisions available to  
145 investigators. While our objective was to investigate the effects of unprocessed red meat, we did not  
146 anticipate that studies investigating the effects of mixed unprocessed and processed red meat or

147 unspecified types of red meat would use different analytic methods. Hence, we also reviewed studies  
148 that reported on mixed unprocessed and processed red meat and unspecified types of red meat.

149 Two reviewers, working independently and in duplicate, reviewed the primary studies from the  
150 systematic review and collected data on study characteristics and analytic methods, including the type  
151 of analytic model (e.g., Cox proportional hazards model, logistic regression), method of adjustment for  
152 energy (e.g., standard model, nutrient density model), covariates included in the model,  
153 operationalization of covariates (e.g., categorical, linear, quadratic), subgroup analyses (e.g., men vs.  
154 women), and the results of analyses, including secondary and sensitivity analyses, when reported. To  
155 ensure that the primary studies that we used to inform our analytic specifications addressed similar  
156 causal questions and interpreted their findings similarly, we documented the objectives of the primary  
157 studies and the ways in which the authors interpreted their findings.

### 158 ***Study population***

159 The National Health and Nutrition Examination Survey (NHANES) is a repeated cross-sectional  
160 probability survey by the US Centers for Disease Control and Prevention to characterize the health and  
161 nutritional status of the non-institutionalized, civilian US population (30). The survey is based on  
162 household interviews and physical examinations and is representative of the US population by its survey  
163 sampling method. The survey collects demographic, socioeconomic, dietary, and health-related data by  
164 household interview, and medical, dental, physiological measurements, and laboratory tests by physical  
165 examination.

166 For this analysis, we used the continuous 2007-2014 NHANES data linked with the National Death Index  
167 (31) and the Food Patterns Equivalents Database. The National Death Index is a database established by  
168 the National Center for Health Statistics (NCHS) that contains information on all deaths in the US. We  
169 extracted mortality status from the National Death Index up to 31 December 2019. The Food Patterns  
170 Equivalents Database contains information on the composition and nutritional content of individual  
171 foods.

172 We acknowledge that NHANES data is likely suboptimal compared to other nutrition datasets for  
173 investigating the effect of red meat and other nutritional exposures on health outcomes, due to it  
174 including few deaths and only collecting data on diet at a single point in time (30, 32). Our objective,  
175 however, is not to provide conclusive answers about the health effects of red meat but to demonstrate  
176 a proof-of-concept application of specification curve analysis to nutritional epidemiology. We used  
177 NHANES data due its availability to our team and our team's familiarity with its structure.

178 We excluded participants with missing demographic, dietary, or lifestyle information; personal and  
179 family history of disease; pregnant people; and participants with implausible BMI ( $<15$  or  $\geq 60$  kg/m<sup>2</sup>) or  
180 energy intake ( $<500$  kcal/day or  $>4,500$  kcal/day). To minimize missing data, we consolidated related  
181 variables in the database (e.g., when data was missing for the smoking history variable, we classified  
182 participants who endorsed smoking 0 cigarettes in their life as non-smokers).

183 Participants in NHANES completed two 24-hour dietary recalls, each conducted by trained interviewers  
184 and separated by 3-10 days, for which they provided information on intake of foods and beverages on  
185 each recall day (32). For our analysis, we define unprocessed red meat as any mammalian meat (i.e.,  
186 beef, veal, pork, lamb, and game meat) (33).

### 187 **Data analysis**

188 We performed specification curve analysis to investigate the effects of unprocessed red meat on all-  
189 cause mortality, using a Cox proportional hazards regression model with time since 24-hour recalls as  
190 the time variable in the model.

191 For each aspect of the analysis, we used  
192 the most used analytic choices from  
193 previous studies (Box 2) and enumerated  
194 all combinations of these choices to  
195 produce a comprehensive list of all  
196 plausible and reasonable analytic  
197 methods. We reviewed analytic  
198 specifications to confirm that every  
199 combination of analytic choices  
200 throughout the analysis was indeed  
201 justifiable. Although we intended to  
202 exclude specifications comprised of  
203 combinations that were not defensible,  
204 we found no such cases.

#### **Box 2: Aspects of the analysis that varied across analytic specifications**

- 1) Type of nutrition model
  - Standard model
  - Multivariable nutrient density model
- 2) Operationalization of red meat
  - Continuous (per 100 g/day)
  - Quartiles
  - Quintiles
- 3) Subgroups of interest
  - All participants
  - Subgroup based on sex
    - All females
    - All males
  - Subgroups based on age
    - Participants aged 20-39 years old
    - Participants aged 40-59 years old
    - Participants 60-79 years old
- 4) Covariates

205 Aspects of the analysis that varied across  
206 primary studies included the type of nutrition model (i.e., standard model and multivariable nutrient  
207 density model), operationalization of red meat (i.e., continuous, quartiles, quintiles), subgroups of  
208 interest (i.e., only males, only females, 20-39 years old, 40-59 years old, 60-79 years old), and choice of



209 covariates. The standard nutrition model adjusts for total energy in the analytic model while the  
210 multivariable nutrient density model includes total energy as a covariate and divides food intake by total  
211 energy intake (34). We did not consider the residual energy model since it is largely equivalent to the  
212 standard model (34).

213 We constructed two sets of covariates: covariates that we included in all models and covariates that  
214 were adjusted in some models. In all models, we adjusted for a core set of covariates that were  
215 considered in nearly all primary studies: age, sex, smoking, total energy intake, year, menopausal status,  
216 hormone therapy, parity, and oral contraceptives. We also optionally adjusted for a secondary set of  
217 other covariates that were only adjusted in some (but not all) studies: race/ethnicity (Mexican  
218 American/other Hispanic/non-Hispanic white/non-Hispanic black/other race—including multi-racial),  
219 education (less than 9<sup>th</sup> grade/9-11<sup>th</sup> grade/high school graduate/some college or AA degree/college  
220 graduate or above), marital status, alcohol consumption, physical activity, BMI, socioeconomic status,  
221 comorbidities, and dietary variables.

222 We are unable to test for all possible combinations of covariates due to computational feasibility.  
223 Hence, we generated 20 unique combinations of covariates that all adjusted for the core set of variables  
224 and each of which adjusted for a random set of the secondary covariates. We applied specification curve  
225 analysis and computed HRs and 95% confidence intervals corresponding to the effect of red meat intake  
226 on all-cause mortality for each analytic specification.

227 For specifications in which red meat was treated as a continuous variable, we calculated HRs and  
228 associated confidence intervals corresponding to a 100 g/day increase in intake of red meat. For  
229 specifications in which red meat was treated as a categorical variable (e.g., quartiles or quintiles), we  
230 calculated hazard ratios and associated confidence intervals corresponding to the highest versus lowest  
231 quantile of red meat exposure. While these contrasts represent different quantities of red meat intake,  
232 primary observational nutritional epidemiology studies overlook these differences when interpreting  
233 results and systematic reviews and meta-analyses often combine these estimates from studies using  
234 disparate quantities (35). In our supplement, we present results stratified by how red meat is defined in  
235 analytic models (i.e., quartiles, quintiles, or continuous 100 g/day).

236 To test whether models from the specification curve analysis met the proportional hazards assumption,  
237 we selected a sample of all specifications at random and tested the correlation between Schoenfeld  
238 residuals and ranked failure time.

239 We excluded results from models that yielded what we considered to be implausible effect estimates  
240 (i.e., studies that yielded implausibly wide confidence intervals with lower bound HR  $\leq 0.2$  or upper  
241 bound HR  $\geq 5$ ). A review of analytic specifications that yielded results outside of this range suggested  
242 sparse data bias, where there are too few events in certain combinations of explanatory variables  
243 resulting in over- or under-estimation of effect estimates (36). While these thresholds are arbitrary, they  
244 pragmatically excluded specifications that yielded what we considered to be results beyond the range of  
245 effects we would plausibly expect from diet and nutrition on health outcomes.

246 We performed three statistical tests to address (i) whether the median effect estimate across all  
247 specifications is more extreme than would be expected if red meat had no effect on all-cause mortality,  
248 (ii) the proportion of specifications that produced statistically significant effects is more extreme than  
249 would expected if red meat had no effect on all-cause mortality, and (iii) whether Stouffer's averaged Z  
250 value across all specifications is more extreme than would be expected if red meat had no effect on all-  
251 cause mortality (27). To perform these tests, we permuted red meat intake and sampled with  
252 replacement across all participants to yield 500 bootstrapped samples to which we applied specification  
253 curve analysis. Based on the results of the specification curve analysis to the permuted datasets, we  
254 calculate P-values using the percentage of bootstrap sample with results as or more extreme than the  
255 observed results. We used an alpha of 0.05 to indicate statistical significance.

256 We performed all analyses in R (Vienna, Austria; version 4.1.2), using the *specr* package for specification  
257 curve analysis (37). Data from NHANES is publicly accessible and the code to produce the results in this  
258 paper is available on a public repository: [https://github.com/Yumin-Wang/Red-Meat-Consumption---All-  
259 Cause-Mortality](https://github.com/Yumin-Wang/Red-Meat-Consumption---All-Cause-Mortality).

## 260 **Results**

### 261 ***Study characteristics***

262 A systematic review addressing the health effects of red meat identified 15 publications reporting on 24  
263 cohort studies that examined the effect of red meat on all-cause mortality (29) (Supplement Table 1).

264 To ensure that these primary studies addressed similar causal questions and interpreted their findings  
265 similarly, we documented the objectives of the primary studies and the ways in which the authors  
266 interpreted their findings (Supplement Table 2). The primary aim of all except two of these studies was  
267 to investigate the effects of red meat on all-cause mortality. One study investigated the effects of  
268 substituting total and different types of dietary protein for carbohydrates on mortality but also  
269 presented models investigating the effects of isocaloric substitutions of carbohydrates for red meat on

270 mortality (38). The second study investigated the effects of components of a traditional Sami diet,  
271 including red meat, on mortality (39).

272 Studies reported 70 unique methods to investigate the relationship between red meat and all-cause  
273 mortality (Supplement Tables 2 and 3). Studies varied in their choice of analytic model (e.g., Cox  
274 proportional hazards model, Poisson regression), adjustment for energy (e.g., standard model and  
275 nutrient density model), covariates included in the model, operationalizations of variables (e.g.,  
276 functional form in the model), and subgroups. Typical studies performed time-dependent Cox regression  
277 models in which red meat was treated as a categorical variable in quartiles or quintiles and adjusted for  
278 age, sex, smoking, alcohol intake, physical activity, and BMI.

279 Studies reported relative effect estimates of red meat on all-cause mortality ranging between 0.63 to  
280 2.31 (median: 1.14; IQR: 1.02 to 1.23). Supplement Figure 1 presents the results of the analyses reported  
281 in studies.

### 282 ***Participant characteristics***

283 We used data from NHANES 2007 to 2014 and excluded participants without mortality data and missing  
284 or implausible data, leaving 10,661 eligible participants.

285 Table 1 and Supplement Table 5 present participant characteristics. Our study included participants  
286 ranging from young adults to the elderly, with approximately equal representation of men and women.  
287 Most participants were white, non- or light smokers, with a median intake of unprocessed red meat less  
288 than half a serving per day. Figure 1 presents the selection of participants in the analysis.

### 289 ***Specification curve analysis***

290 Using all analytic choices identified in the primary studies, we enumerated all justifiable ways in which  
291 the data may be reasonably analyzed, yielding 1,440 unique analytic specifications. We were able to  
292 accommodate most analytic choices reported in primary studies using data from NHANES (Supplement  
293 Tables 2 and 3). We were unable to implement time-varying variables due to the cross-sectional nature  
294 of the NHANES data.

295 We implemented all 1,440 reasonable specifications and identified 1,208 unique specifications with  
296 plausible results and 232 with implausibly wide confidence intervals (lower bound HR  $\leq 0.2$  or upper  
297 bound HR  $\geq 5$ ). These implausible specifications occurred in analyses of subgroups of the total study  
298 population that included many adjusting covariates, suggesting sparse data bias (36).

299 Figure 2 presents the results of the specification curve analysis. Our specification curve analysis  
300 produced a median hazard ratio of 0.94 [IQR: 0.83 to 1.05] for the effect of red meat on all-cause  
301 mortality. Hazard ratios ranged from 0.51 to 1.75. Of all specifications, 435 (36.0%) yielded hazard ratios  
302 equal to or above 1.0 and 773 (64.0%) below 1.0.

303 Of all specifications, 48 (3.97%) were statistically significant. Of 48 statistically significant results, 40 had  
304 significant point estimates that indicated red meat to reduce all-cause mortality and 8 indicated red  
305 meat to increase all-cause mortality. Among statistically significant effects suggesting benefit, we  
306 observed a median hazard ratio of 0.65 [IQR: 0.58 to 0.69] and, among statistically significant effects  
307 suggesting harm, we observed a median hazard ratio of 1.22 [IQR: 1.19 to 1.27]. We found 45%  
308 (542/1,208) of all specifications to yield point estimates ranging between HR of 0.90 to 1.10.

309 Visual inspection of the specification curve plot suggests subgroup by sex to importantly influence  
310 results, with analyses restricted to women more likely to suggest red meat is beneficial. We observed a  
311 median hazard ratio of 1.05 [IQR: 0.89 to 1.12] for men and 0.85 [IQR: 0.77 to 0.93] for women. We did  
312 not identify other analytic characteristics as consequential.

313 Supplement Figure 2 presents the results of the specification curve analysis stratified by how red meat is  
314 defined in analytic models (i.e., quartiles, quintiles, or continuous 100 g/day).

315 Supplement Tables 6 to 10 and Supplement Figures 3 to 7 show the results of tests for the proportional  
316 hazards assumption and graphical displays of the correlation between Schoenfeld residuals and ranked  
317 failure time. We did not find evidence that the proportional hazards assumption was violated in any  
318 analyses.

319 Finally, we present statistical inferences about the degree to which findings across all specifications are  
320 inconsistent with the null hypothesis (i.e., red meat has no effect on all-cause mortality) (Table 2). We  
321 performed statistical tests addressing whether the median effect estimate across all specifications is  
322 more extreme than expected if red meat had no effect on all-cause mortality, whether the proportion of  
323 specifications that produced statistically significant effects is more extreme than would expected if red  
324 meat had no effect on all-cause mortality, and whether Stouffer's averaged Z value across all  
325 specifications is more extreme than would be expected if red meat had no effect on all-cause mortality.  
326 All three statistical tests yielded P-values > 0.05.

327 **Discussion**

328 ***Main findings***

329 In this study, we applied specification curve analysis—a method that involves defining and implementing  
330 all plausible and valid analytic approaches—to estimate the effect of unprocessed red meat on all-cause  
331 mortality (27). To mitigate the subjectivity involved in selecting analytic specifications, we sourced  
332 analytic approaches from the literature (29). We performed 1,208 unique analyses and found  
333 considerable variability in results, with hazard ratios ranging from 0.51 to 1.75. Our results suggest that  
334 findings in nutritional epidemiology studies may be contingent on analytic methods.

335 In contrast to previous studies addressing red meat, we found few of our analytic specifications to yield  
336 statistically significant effects. This may be because we used more recent data from NHANES, which  
337 includes fewer accumulated deaths (40). The most recent iterations of NHANES, however, are likely  
338 more reflective of the effects of red meat on all-cause mortality in the context of contemporaneous  
339 diets and lifestyles. Nevertheless, our primary objective was not draw inferences about the health  
340 effects of red meat but to provide a proof-of-concept illustration of the application of specification curve  
341 analysis to nutritional epidemiology.

342 Concerns may arise over the impact of various analytical techniques on the interpretation of results. For  
343 example, different methods for energy adjustment may have different implications for how the effect is  
344 interpreted (18, 41). In our study, we show that despite differences in analytic methods authors stated  
345 similar objectives and similarly interpreted their results. This suggests that authors are using disparate  
346 analytic methods to investigate near identical causal questions.

347 In addition to analytic flexibility, researchers criticize observational nutritional epidemiology studies for  
348 producing unreliable results due to biases associated with self-reported dietary data (20, 23).

349 Nevertheless, nutritional epidemiology studies continue to play a critical role in shaping dietary  
350 recommendations and policies (15). While specification curve analysis does not address biases due to  
351 dietary measures, when combined with other tools and methods for producing more reliable dietary  
352 measures, specification curve analysis may have the potential to enhance confidence in the discipline  
353 (42, 43).

354 ***Relation to previous work***

355 Current evidence shows that results from studies may vary due to alternative analytic specifications and  
356 that there is often limited consensus on the optimal approach for data analysis (6, 44). Research to date  
357 has not, however, quantified the magnitude of variation in results for typical epidemiologic questions.

358 Further, while specification curve analysis has been previously applied in psychology and economics, it  
359 has not yet been applied in epidemiology or nutritional epidemiology (26, 45, 46).

### 360 ***Strengths and limitations***

361 The current work offers an innovative solution to analytic flexibility in nutritional epidemiology. To our  
362 knowledge, our work is the first application of specification curve analysis to nutritional epidemiology.

363 Our study also has limitations. There may be disagreements among investigators about what constitutes  
364 a justifiable analytical approach. To mitigate this issue, our choice of analytic specifications was  
365 informed by primary studies and so represents real, published analyses rather than possible,  
366 unpublished analyses that may only be marginally defensible. Further, we confirmed that the research  
367 questions addressed in the primary studies were adequately similar by collecting data on the objectives  
368 of primary studies and the ways in which authors interpreted their results.

369 We emphasize that specification curve analysis does not eliminate the need for subjectivity in selecting  
370 specifications (27). Nonetheless, it does improve on current practice in which investigators can test  
371 many alternative analytic specifications and selectively report results for those that yield interesting or  
372 favorable results. Specification curve analysis can identify findings that are most robust to alternative  
373 analytic specifications and encourage evidence users to interpret the results of nutritional epidemiology  
374 studies considering the typical variation in results expected due to analytic flexibility.

375 Different analytic methods may have implications for how the results are interpreted. For example,  
376 different methods to adjust for energy intake in nutritional epidemiology address different causal  
377 questions (18). Authors of nutritional epidemiology studies, however, seldom acknowledge these issues.  
378 We show that despite differences in analytic methods authors stated similar objectives and similarly  
379 interpreted their results.

380 Specification curve analysis also does not eliminate the need for content knowledge and expertise.  
381 Content knowledge and expertise are essential for selecting justifiable analytic specifications and  
382 interpreting the results of analyses.

383 We only applied specification curve analysis to one question—the effect of red meat on all-cause  
384 mortality. The extent to which results may be contingent on analytic methods may be different for other  
385 questions. We acknowledge that this is a controversial question in the nutrition literature and that the  
386 application of specification curve analytic to less contentious questions in nutritional epidemiology may

387 improve its adoption. Our choice of topic was influenced by our team’s familiarity with red meat and the  
388 related literature (15, 29).

389 Our study likely underestimates the variations in results due to alternative analytic specifications since  
390 the analytic specifications that we could implement were limited by the availability of variables and data  
391 in NHANES. For example, due to the cross-sectional nature of NHANES, we were unable to use time-  
392 varying covariates and explore how alternative ways to account for these variables may influence  
393 results. There are also subjective analytic decisions in translating dietary recalls to nutrient and food  
394 intake, though we could not account for these decisions. For example, nutritional epidemiologists code  
395 dietary recalls according to food classification systems and subsequently use nutrition databases to  
396 estimate individual nutrient components of each item in dietary recalls—all of which involves subjective  
397 decisions.

398 Likewise, the continuous 2007-2014 NHANES data is likely suboptimal for investigating the effect of red  
399 meat and other nutritional exposures on health outcomes, due to it including few deaths and only  
400 collecting data on diet at a single point in time (30, 32). Nevertheless, our primary objective is not to  
401 provide conclusive answers about the health effects of red meat but to demonstrate a proof-of-concept  
402 application of specification curve analysis to nutritional epidemiology.

403 We did not incorporate weights in our analytic models. Sample weights in NHANES are designed to  
404 account for oversampling of specific subgroups and unequal probabilities of selection in the population.  
405 These weights are essential when the objective is to make inferences about population characteristics or  
406 to estimate prevalence rates because they adjust for factors that influence these estimates and ensure  
407 that the results are representative of the target population. However, when focusing on causal  
408 inference, the primary concern is to eliminate or control for confounding factors that may distort the  
409 true relationship between exposure and outcome and sample weights are less important, especially  
410 when variables used to derive sample weights are already included in analytic models (47-49).

411 We excluded results from models that yielded results that we deemed to be implausible based on  
412 pragmatic but arbitrary thresholds (i.e.,  $HR \leq 0.2$  or  $HR \geq 5$ ). We suspect that the observed implausible  
413 specifications were due to sparse data bias—where there are too few events in critical combinations of  
414 explanatory variables (36). It is, however, possible that there were other models that produced results  
415 within this threshold that had too few events to reliably estimate the effect of red meat on all-cause  
416 mortality.

417 Finally, while we attempted to test the proportional hazards assumption using the correlation between  
418 Schoenfeld residuals and ranked failure time, these tests have limited sensitivity (50). We also only  
419 tested a proportion of our models for proportional hazards and it is possible that the proportional  
420 hazards assumption may be violated in models that we did not test.

#### 421 ***Implications***

422 Specification curve analysis allows investigators to test all plausible and justifiable models to explain  
423 conflicting findings or contextualize emerging findings. While this study may provide insights on the  
424 health effects of unprocessed red meat, we believe the most important contribution of this study is to  
425 provide a proof-of-concept demonstrating the feasibility of applying specification curve analysis to  
426 nutritional epidemiology.

427 Nutritional epidemiology has long been criticized for producing sensational and conflicting findings,  
428 which has eroded confidence in the discipline (23). Nevertheless, nutritional epidemiology studies  
429 continue to play a crucial role in shaping dietary recommendations and policies, making it imperative to  
430 draw credible inferences from these studies (14, 15, 24). The broader application of specification curve  
431 analysis to nutritional epidemiology may enhance confidence in nutrition as a field by encouraging  
432 investigators to acknowledge an additional source of uncertainty in studies. When combined with other  
433 tools and methods that also address other limitations of observational nutritional epidemiology studies  
434 (e.g., biases that affect self-reported dietary data) (42), specification curve analysis has the potential to  
435 address a critical issue in epidemiology—analytic flexibility—and identify findings that are most robust  
436 to subjective analytic choices.

437 Findings from our study and future application of specification curve analysis will also be useful to  
438 evidence users who can interpret results in nutritional epidemiology studies in the context of the typical  
439 variation in results expected due to analytic flexibility. When effect estimates exceed the typical  
440 variation due to analytic methods, evidence users can be more certain of the findings, since they are  
441 likely robust to alternative analytic decisions.

442 Our findings may also have implications for precision nutrition that attempts to distinguish between  
443 subgroups of individuals who may differently respond to nutritional interventions or have different  
444 nutritional needs (51-53). Investigators have raised concerns that efforts to identify “responders” and  
445 realize precision nutrition may be highly dependent on the characteristics of analytic models (54).



446 Specification curve analysis may be useful for evaluating the reliability of precision nutrition claims  
447 across a range of defensible models.

448 We acknowledge that the application of specification curve analysis is time consuming and resource  
449 intensive. Sourcing justifiable analytic specifications from primary studies adds to this effort. While the  
450 application of specification curve analysis may not be feasible for all nutritional epidemiology questions,  
451 it can be applied to the most critical, impactful, or contentious questions in the discipline and can serve  
452 as an additional available tool to evaluate the credibility of nutrition claims in the literature.

### 453 **Conclusion**

454 In this study, we apply specification curve analysis—a novel analytic method that involves defining and  
455 implementing all plausible and valid analytic approaches for addressing a research question—to  
456 investigate the effect of red meat on all-cause mortality. We show variability in results across plausible  
457 analytic specifications. Our study demonstrates the feasibility of applying specification curve analysis to  
458 nutritional epidemiology that offers a pragmatic and innovative solution to analytic flexibility.  
459 Specification curve analysis, in combination with other tools and methods, has the potential to improve  
460 the credibility of inferences from such studies.

461 **Tables and Figures**

**Table 1: Participant characteristics**

|  |                   |
|--|-------------------|
| <b>Total participants, N</b>               | 10,661            |
| <b>All-Cause mortality, n (%)</b>          | 1022 (10)         |
| <b>Follow-up (months)</b>                  | 99 (65, 143)      |
| <b>Age (years)</b>                         | 50 (27, 71)       |
| <b>Gender</b>                              |                   |
| Female, n (%)                              | 5150 (48)         |
| Male, n (%)                                | 5511 (52)         |
| <b>Dietary Intakes</b>                     |                   |
| Unprocessed red meat (g/d)                 | 29.5 (0, 120.2)   |
| Total energy intake (kcal/d)               | 1945 (1168, 3099) |
| <b>Years of entering cohort</b>            |                   |
| 2007-2008, n (%)                           | 2311 (22)         |
| 2009-2010, n (%)                           | 2358 (22)         |
| 2011-2012, n (%)                           | 2857 (27)         |
| 2013-2014, n (%)                           | 3135 (29)         |
| <b>Race/Ethnicity</b>                      |                   |
| Mexican American, n (%)                    | 1321 (12)         |
| Other Hispanic, n (%)                      | 988 (9)           |
| Non-Hispanic White, n (%)                  | 5193 (49)         |
| Non-Hispanic Black, n (%)                  | 2235 (21)         |
| Other Race – Including Multi-Racial, n (%) | 924 (9)           |
| <b>Smoking</b>                             |                   |
| Non or light smoker, n (%)                 | 8373 (79)         |
| Moderate smoker, n (%)                     | 437 (4)           |
| Heavy smoker, n (%)                        | 1851 (17)         |
| <b>BMI (kg/m<sup>2</sup>)</b>              | 28.4 (21.9, 38.5) |

*Data presented as numbers and proportions or as medians (10<sup>th</sup> percentile, 90<sup>th</sup> percentile).*

462

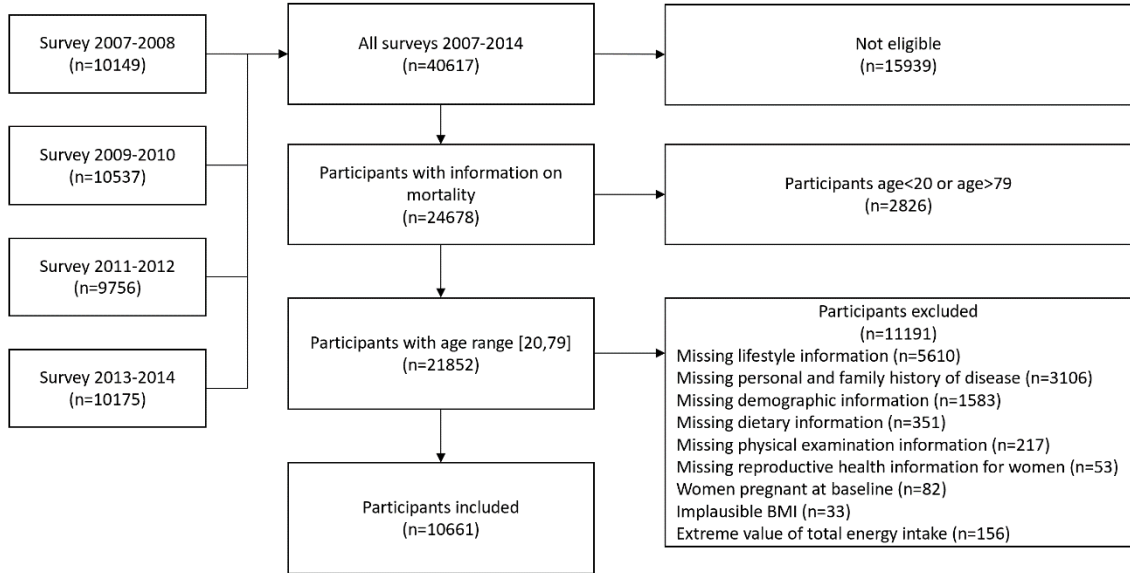
**Table 2: Inferential statistics**

| <b>Test statistics used</b>         | <b>Observed results</b>   | <b>P value</b><br>(% of bootstrap sample with results as or more extreme) |
|-------------------------------------|---------------------------|---|
| <b>Median effect size</b>           | HR=0.94                   | P=0.472   |
| <b>Share of significant results</b> | 48 of 1208 specifications | P=0.998   |
| <b>Aggregate all P values</b>       | Stouffer Z=-11.69         | P=0.732   |

463

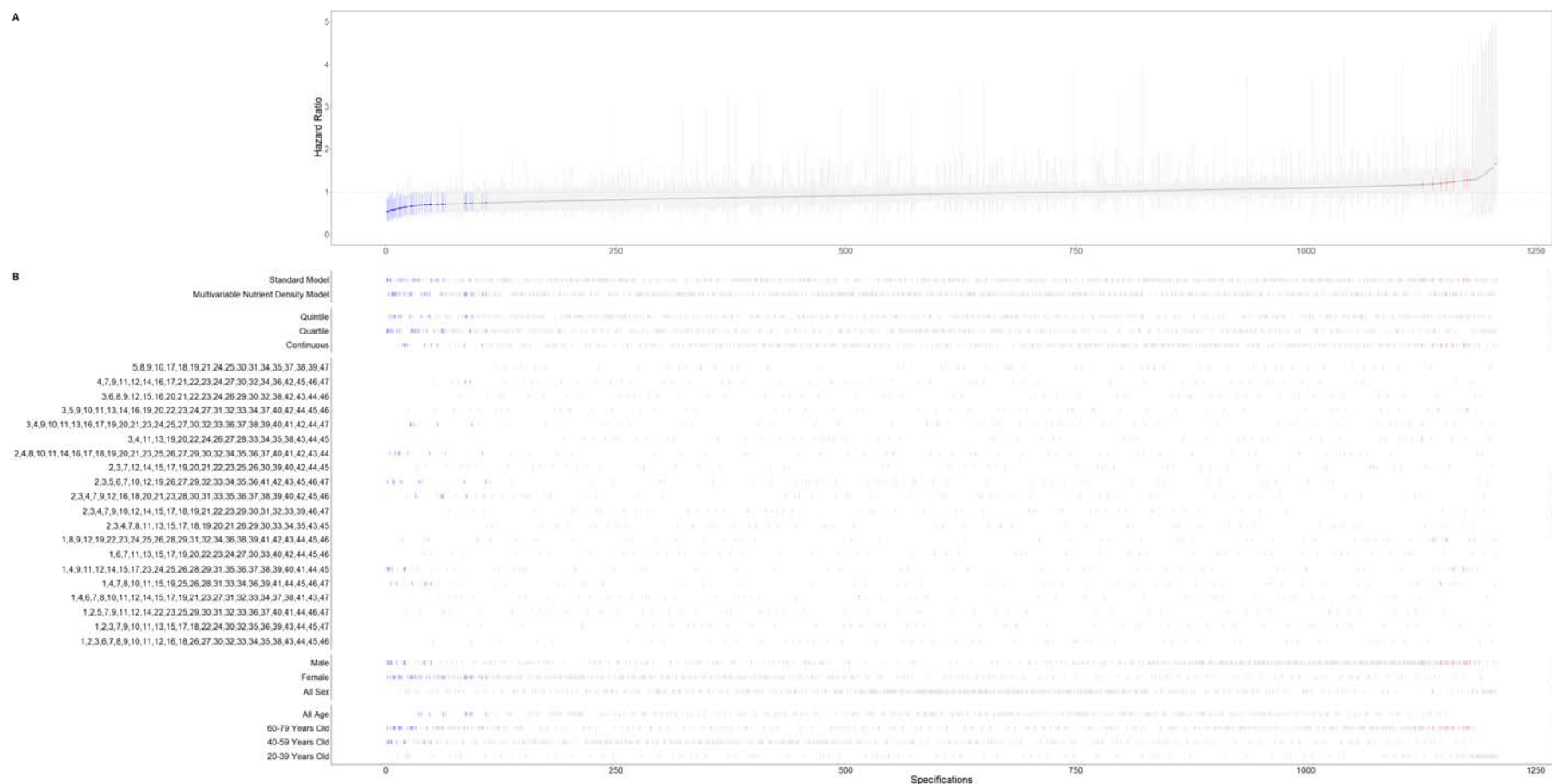
464

465 **Figure 1: Selection of study participants from the National Health and Nutrition Examination Survey**  
466 **(NHANES) for inclusion in the analysis**



467

468 **Figure 2: Results of specification curve analysis.**



469

470 This figure presents the results of the specification curve analysis, including 1,208 unique analytic specifications. The upper portion of the plot  
 471 show HRs representing the effect of red meat on all-cause mortality. On the x-axis are the unique analytic specifications. The y-axis represents  
 472 the magnitude of effect estimates. Each point on the graph represents the results of a unique analytic specification. Point estimates are shown in  
 473 dark grey and 95% confidence intervals as light grey bars. Each point represents the results for the effect of red meat on all-cause mortality for a  
 474 unique model. Points in blue are statistically significant and suggest red meat to prevent all-cause mortality and points in red are statistically  
 475 significant and indicate red meat to increase risk of all-cause mortality.

476 The lower part of the plot show the characteristics of each analysis, including type of analytic model, operationalizations of variables, choice of  
477 covariates, and subgroups of interest. Each vertical line denotes the specific choice applied for each aspect of the analysis. We assigned a unique  
478 number to each covariate (Supplement 4 shows the number corresponding to each variable). Combinations of numbers in the graph represent  
479 combinations of covariates included in the model.

480 **References**

- 481 1. Tierney BT, Anderson E, Tan Y, Claypool K, Tangirala S, Kostic AD, et al. Leveraging vibration of  
482 effects analysis for robust discovery in observational biomedical data science. *PLoS Biol.*  
483 2021;19(9):e3001398.
- 484 2. Patel CJ, Burford B, Ioannidis JP. Assessment of vibration of effects due to model specification  
485 can demonstrate the instability of observational associations. *J Clin Epidemiol.* 2015;68(9):1046-58.
- 486 3. Chu L, Ioannidis JPA, Egilman AC, Vasiliou V, Ross JS, Wallach JD. Vibration of effects in  
487 epidemiologic studies of alcohol consumption and breast cancer risk. *Int J Epidemiol.* 2020;49(2):608-18.
- 488 4. Hoozeveen S, Sarafoglou A, Aczel B, Aditya Y, Alayan AJ, Allen PJ, et al. A many-analysts  
489 approach to the relation between religiosity and well-being. *Religion, Brain & Behavior.* 2022:1-47.
- 490 5. Breznau N, Rinke EM, Wuttke A, Nguyen HHV, Adem M, Adriaans J, et al. Observing many  
491 researchers using the same data and hypothesis reveals a hidden universe of uncertainty. *Proc Natl Acad*  
492 *Sci U S A.* 2022;119(44):e2203150119.
- 493 6. Silberzahn R, Uhlmann EL, Martin DP, Anselmi P, Aust F, Awtrey E, et al. Many analysts, one data  
494 set: Making transparent how variations in analytic choices affect results. *Advances in Methods and*  
495 *Practices in Psychological Science.* 2018;1(3):337-56.
- 496 7. Madigan D, Ryan PB, Schuemie M. Does design matter? Systematic evaluation of the impact of  
497 analytical choices on effect estimates in observational studies. *Therapeutic advances in drug safety.*  
498 2013;4(2):53-62.
- 499 8. Zeraatkar D, Cheung K, Milio K, Zworth M, Gupta A, Bhasin A, et al. Methods for the Selection of  
500 Covariates in Nutritional Epidemiology Studies: A Meta-Epidemiological Review. *Curr Dev Nutr.*  
501 2019;3(10):nzz104.
- 502 9. van Dongen NNN, van Doorn JB, Gronau QF, van Ravenzwaaij D, Hoekstra R, Haucke MN, et al.  
503 Multiple Perspectives on Inference for Two Simple Statistical Scenarios. *The American Statistician.*  
504 2019;73(sup1):328-39.
- 505 10. Landy JF, Jia ML, Ding IL, Viganola D, Tierney W, Dreber A, et al. Crowdsourcing hypothesis tests:  
506 Making transparent how design choices shape research results. *Psychol Bull.* 2020;146(5):451-79.
- 507 11. Schilling KG, Rheault F, Petit L, Hansen CB, Nath V, Yeh F-C, et al. Tractography dissection  
508 variability: What happens when 42 groups dissect 14 white matter bundles on the same dataset?  
509 *NeuroImage.* 2021;243:118502.
- 510 12. Low J, Ross JS, Ritchie JD, Gross CP, Lehman R, Lin H, et al. Comparison of two independent  
511 systematic reviews of trials of recombinant human bone morphogenetic protein-2 (rhBMP-2): the Yale  
512 Open Data Access Medtronic Project. *Systematic Reviews.* 2017;6(1):28.
- 513 13. Scientific Pandemic Influenza Group on Modelling. SPI-M-O: Consensus statement on COVID-19,  
514 8 October 2020. 2020.
- 515 14. Ruxton C. Interpretation of observational studies: the good, the bad and the sensational.  
516 *Proceedings of the Nutrition Society.* 2022;81(4):279-87.
- 517 15. Zeraatkar D, Johnston BC, Guyatt G. Evidence Collection and Evaluation for the Development of  
518 Dietary Guidelines and Public Policy on Nutrition. *Annu Rev Nutr.* 2019;39:227-47.
- 519 16. Willett W. *Nutritional Epidemiology*: Oxford University Press; 2012.
- 520 17. Willett WC, Stampfer M, Tobias DK. Re: Adjustment for energy intake in nutritional research: a  
521 causal inference perspective. *Am J Clin Nutr.* 2022;116(2):608-9.
- 522 18. Tomova GD, Arnold KF, Gilthorpe MS, Tennant PWG. Adjustment for energy intake in nutritional  
523 research: a causal inference perspective. *Am J Clin Nutr.* 2022;115(1):189-98.
- 524 19. Wallach JD, Serghiou S, Chu L, Egilman AC, Vasiliou V, Ross JS, et al. Evaluation of confounding in  
525 epidemiologic studies assessing alcohol consumption on the risk of ischemic heart disease. *BMC Medical*  
526 *Research Methodology.* 2020;20(1):64.

- 527 20. Schoenfeld JD, Ioannidis JP. Is everything we eat associated with cancer? A systematic cookbook  
528 review. *Am J Clin Nutr.* 2013;97(1):127-34.
- 529 21. Gkiouras K, Choleva ME, Verrou A, Goulis DG, Bogdanos DP, Grammatikopoulou MG. A Meta-  
530 Epidemiological Study of Positive Results in Clinical Nutrition Research: The Good, the Bad and the Ugly  
531 of Statistically Significant Findings. *Nutrients.* 2022;14(23).
- 532 22. Hall KD. Challenges of human nutrition research. *Science.* 2020;367(6484):1298-300.
- 533 23. Ioannidis JPA. Unreformed nutritional epidemiology: a lamp post in the dark forest. *European*  
534 *Journal of Epidemiology.* 2019;34(4):327-31.
- 535 24. Ley SH, Ardisson Korat AV, Sun Q, Tobias DK, Zhang C, Qi L, et al. Contribution of the Nurses'  
536 Health Studies to Uncovering Risk Factors for Type 2 Diabetes: Diet, Lifestyle, Biomarkers, and Genetics.  
537 *Am J Public Health.* 2016;106(9):1624-30.
- 538 25. Simonsohn U, Simmons JP, Nelson LD. Specification curve analysis. *Nature Human Behaviour.*  
539 2020;4(11):1208-14.
- 540 26. Rohrer JM, Egloff B, Schmukle SC. Probing birth-order effects on narrow traits using  
541 specification-curve analysis. *Psychological Science.* 2017;28(12):1821-32.
- 542 27. Simonsohn U, Simmons JP, Nelson LD. Specification curve analysis. *Nat Hum Behav.*  
543 2020;4(11):1208-14.
- 544 28. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening  
545 the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting  
546 observational studies. *Lancet.* 2007;370(9596):1453-7.
- 547 29. Zeraatkar D, Guyatt GH, Alonso-Coello P, Bala MM, Rabassa M, Han MA, et al. Red and  
548 Processed Meat Consumption and Risk for All-Cause Mortality and Cardiometabolic Outcomes. *Ann*  
549 *Intern Med.* 2020;172(7):511-2.
- 550 30. Centers for Disease Control and Prevention. National Center for Health Statistics. National  
551 Health and Nutrition Examination Survey Data. 2020.  
552 <https://www.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx> (accessed October 3, 2022).
- 553 31. Centers for Disease Control and Prevention. National Center for Health Statistics. NDI mortality  
554 data. Centers for Disease Control and Prevention. 2020. [https://www.cdc.gov/nchs/data-](https://www.cdc.gov/nchs/data-linkage/mortality.htm)  
555 [linkage/mortality.htm](https://www.cdc.gov/nchs/data-linkage/mortality.htm) (accessed October 3, 2022).
- 556 32. Ahluwalia N, Dwyer J, Terry A, Moshfegh A, Johnson C. Update on NHANES Dietary Data: Focus  
557 on Collection, Release, Analytical Considerations, and Uses to Inform Public Policy. *Advances in*  
558 *Nutrition.* 2016;7(1):121-34.
- 559 33. Wiseman M. The second World Cancer Research Fund/American Institute for Cancer Research  
560 expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. *Proc*  
561 *Nutr Soc.* 2008;67(3):253-6.
- 562 34. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. *Am*  
563 *J Clin Nutr.* 1997;65(4 Suppl):1220S-8S; discussion 9S-31S.
- 564 35. Zeraatkar D, Bhasin A, Morassut RE, Churchill I, Gupta A, Lawson DO, et al. Characteristics and  
565 quality of systematic reviews and meta-analyses of observational nutritional epidemiology: a cross-  
566 sectional study. *Am J Clin Nutr.* 2021;113(6):1578-92.
- 567 36. Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. *Bmj.*  
568 2016;352:i1981.
- 569 37. Masur P, Scharnow M (2020). "speccr: Conducting and Visualizing Specification Curve Analyses  
570 (Version 1.0.0)." <https://CRAN.R-project.org/package=speccr>.
- 571 38. Kelemen LE, Kushi LH, Jacobs DR, Jr., Cerhan JR. Associations of dietary protein with disease and  
572 mortality in a prospective study of postmenopausal women. *Am J Epidemiol.* 2005;161(3):239-49.



- 573 39. Nilsson LM, Winkvist A, Brustad M, Jansson JH, Johansson I, Lenner P, et al. A traditional Sami  
574 diet score as a determinant of mortality in a general northern Swedish population. *Int J Circumpolar*  
575 *Health*. 2012;71(0):1-12.
- 576 40. Kappeler R, Eichholzer M, Rohrmann S. Meat consumption and diet quality and mortality in  
577 NHANES III. *Eur J Clin Nutr*. 2013;67(6):598-606.
- 578 41. Tomova GD, Gilthorpe MS, Tennant PW. Theory and performance of substitution models for  
579 estimating relative causal effects in nutritional epidemiology. *Am J Clin Nutr*. 2022;116(5):1379-88.
- 580 42. Kirkpatrick SI, Baranowski T, Subar AF, Tooze JA, Frongillo EA. Best Practices for Conducting and  
581 Interpreting Studies to Validate Self-Report Dietary Assessment Methods. *J Acad Nutr Diet*.  
582 2019;119(11):1801-16.
- 583 43. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhauser ML, et al. Addressing  
584 Current Criticism Regarding the Value of Self-Report Dietary Data. *J Nutr*. 2015;145(12):2639-45.
- 585 44. Steegen S, Tuerlinckx F, Gelman A, Vanpaemel W. Increasing Transparency Through a Multiverse  
586 Analysis. *Perspect Psychol Sci*. 2016;11(5):702-12.
- 587 45. Orben A, Przybylski AK. The association between adolescent well-being and digital technology  
588 use. *Nature Human Behaviour*. 2019;3(2):173.
- 589 46. Carter EC, Schönbrodt FD, Gervais WM, Hilgard J. Correcting for bias in psychology: A  
590 comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*.  
591 2019;2(2):115-44.
- 592 47. WINSHIP C, RADBILL L. Sampling Weights and Regression Analysis. *Sociological Methods &*  
593 *Research*. 1994;23(2):230-57.
- 594 48. Andrew G. Struggles with Survey Weighting and Regression Modeling. *Statistical Science*.  
595 2007;22(2):153-64.
- 596 49. Solon G, Haider SJ, Wooldridge J. What Are We Weighting For? National Bureau of Economic  
597 Research Working Paper Series. 2013;No. 18859.
- 598 50. Stensrud MJ, Hernán MA. Why Test for Proportional Hazards? *Jama*. 2020;323(14):1401-2.
- 599 51. Kirk D, Catal C, Tekinerdogan B. Precision nutrition: A systematic literature review. *Comput Biol*  
600 *Med*. 2021;133:104365.
- 601 52. Rodgers GP, Collins FS. Precision Nutrition-the Answer to "What to Eat to Stay Healthy". *Jama*.  
602 2020;324(8):735-6.
- 603 53. Bailey RL, Stover PJ. Precision Nutrition: The Hype Is Exceeding the Science and Evidentiary  
604 Standards Needed to Inform Public Health Recommendations for Prevention of Chronic Disease. *Annu*  
605 *Rev Nutr*. 2023;43:385-407.
- 606 54. Fröhlich H, Balling R, Beerenwinkel N, Kohlbacher O, Kumar S, Lengauer T, et al. From hype to  
607 reality: data science enabling personalized medicine. *BMC Med*. 2018;16(1):150.

608

