

1 **Guilt by association: Plant-based foods can be incorporated**
2 **into both healthy and unhealthy plant-based diet indices**
3 **associated with coronary heart disease**

4 Yasaman Jamshidi-Naeini ^{1*+}, Beate Henschel ¹⁺, James M. Shikany ², David B. Allison ¹,
5 Andrew W. Brown ^{3,4*}

6 ¹ Department of Epidemiology and Biostatistics, Indiana University School of Public Health-Bloomington,
7 Bloomington, IN

8 ² Division of Prevention Medicine, Heersink School of Medicine, University of Alabama at Birmingham,
9 Birmingham, AL, USA

10 ³ Department of Biostatistics, University of Arkansas for Medical Sciences, Little Rock, AR, USA

11 ⁴ Arkansas Children's Research Institute, Little Rock, AR, USA

12 ⁺ Equal contributions.

13

14 ***Correspondence:**

15 Andrew W Brown, PhD
16 Arkansas Children's Research Institute
17 13 Children's Way, Slot 842
18 Little Rock, AR 72202
19 awbrown@uams.edu

20

21

22 Abstract

23 **Background:** One approach to test for differential associations between plant foods with health
24 uses a scoring approach: foods categorized into animal or ‘healthy’ plant-based or ‘unhealthy’
25 plant-based groups to construct a plant-based diet index (PDI), healthy PDI (hPDI), and
26 unhealthy PDI (uPDI).

27 **Objective:** To evaluate robustness of associations between diet indices and incident coronary
28 heart disease (CHD) risk when recategorizing food groups in indices.

29 **Methods:** Using REasons for Geographic and Racial Differences in Stroke (REGARDS) data,
30 we replicated a published use of the scoring approach. Using Cox proportional hazards
31 regression, we assessed ramifications of the following on associations between diet indices and
32 CHD risk: 1) reconfiguring foods within and among food groups, using potatoes as an example,
33 2) leave-one-out analysis for each of 12 plant-based food groups, and 3) agnostically redefining
34 each food group as ‘healthy’ or ‘unhealthy’.

35 **Results:** Over 153,286 person-years of follow-up, there were 868 cases of CHD. Replication
36 analyses did not reach statistical significance. General patterns of magnitude of hazard ratios
37 (HRs) in replication and reconfiguration models were PDI HRs < hPDI HRs < uPDI HRs for
38 women, and hPDI < PDI < uPDI for men. Five models reconfiguring potatoes resulted in small,
39 varied differences in PDI, hPDI, and uPDI associations. Leave-one-out analyses resulted in
40 greater variation of associations between indices and CHD. In agnostic models, each plant-based
41 food group was classified in indices as ‘healthy’ and ‘unhealthy’ with statistically significant
42 beneficial or deleterious associations with CHD. Averaged over 4,096 models, HRs’ shifts were
43 small when food groups were moved between ‘healthy’ and ‘unhealthy’.

44 **Conclusion:** Statistically significant associations between hPDI, uPDI, and PDI and incident
45 CHD were not replicated. Small perturbations of the scoring approach had varied impacts on
46 HRs. Agnostically constructing diet indices demonstrated the potential for guilt (or benefit) by
47 association: any of the food groups we studied could be categorized with others in an index
48 showing beneficial or deleterious associations.

49 **Keywords:** Plant-based diet index, Healthy and unhealthy plant-based foods, Agnostic
50 Permutations, Coronary Heart Disease

51 Abbreviations:

52 BMI: Body mass index
53 CHD: Coronary heart disease
54 CI: Confidence interval
55 ECG: Electrocardiogram
56 hPDI: Healthful plant-based diet index
57 HPFS: Health Professionals Follow-Up Study
58 HR: Hazard ratio
59 MI: Myocardial infarction
60 NHS: Nurses’ Health Study
61 NHS2: Nurses’ Health Study 2
62 PDI: Plant-based diet index
63 REGARDS: REasons for Geographic and Racial Differences in Stroke
64 uPDI: Unhealthful plant-based diet index

65 **Introduction**

66 The distinctions between ‘healthy’ and ‘unhealthy’ (sometimes called ‘less-healthy’) plant-based
67 foods are of much interest for diet and health researchers, practitioners, and guideline creators.
68 Research groups have looked at the association of various interpretations of plant-based diets
69 with different health outcomes including overweight/obesity, weight change, cognition,
70 mortality, type 2 diabetes (T2D), overall cardiovascular disease (CVD), coronary heart disease
71 (CHD), breast and overall cancer, biological ageing trajectory, and risk of frailty (1-11).
72 However, the demarcations between ‘healthy’ and ‘unhealthy’ varies depending on factors
73 including glycemic indices and glycemic loads calculated on individual foods rather than meals
74 or diets; median sex-specific splits of self-reported consumption within cohorts; or researcher
75 discretion on choosing what they believe are more or less ‘healthy’. Categories of foods can also
76 vary based on component foods, energy density, numbers and varieties of foods, and other
77 characteristics. The general approach to developing diet indices involves investigators deciding
78 which foods they consider fitting their construct through data-driven approaches, evaluation of
79 prior literature, or ethnographic determinations. Thereafter, researchers will assess adherence to
80 that construct in some way. Such an approach has been used for different food consumption
81 patterns, including Mediterranean diets and ‘plant-based’ diets.

82 One particular extension for plant-based foods advanced by Satija et al. was to categorize food
83 frequency questionnaire (FFQ) data into ‘healthy’ plant-based, ‘less healthy’ plant-based, or
84 animal food groups. The distinctions were driven by their interpretation of “existing knowledge
85 of associations of the foods with T2D, other outcomes (CVD, certain cancers), and intermediate
86 conditions (obesity, hypertension, lipids, inflammation)” (2). Using data from Nurses’ Health
87 Study (NHS), NHS2, and Health Professionals Follow-Up Study (HPFS), the authors presented
88 associations between plant diet indices and T2D (2), cardiometabolic outcomes (3), and weight

89 change (4). Given the adoption of the approach by the original authors and others (e.g., (1-8), the
90 properties of this approach are worth studying to assess how analytical choices may influence
91 conclusions. This is in the spirit of the analogy advanced by Sam Savage of “shaking the ladder”
92 to assess its stability “to minimize the risk that it falls down with you on it” (12).

93 **The scoring approach**

94 The method advanced by Satija et al. (hereafter, the scoring approach) involves using FFQ data
95 to estimate daily servings of consumption of individual foods. The individual foods are then
96 aggregated into food groups (e.g., individual vegetables within the ‘vegetables’ food group). The
97 sum of servings of foods within a group is then divided by quintiles and scored from 1 to 5
98 depending on whether a food is ‘healthy’ or ‘unhealthy’. The scores from each food group are
99 then summed to get an index, and, in the main analysis, the index values are themselves divided
100 by deciles. The ultimate exposure score is therefore highly transformed, and like other index-
101 based approaches, loses much of the identity of the dietary exposure components. The scoring
102 approach has important limitations, some of which are shared by other index-based approaches
103 or nutrition epidemiology in general, including that estimating daily dietary intakes from self-
104 report such as FFQ may not accurately reflect actual intakes (13-16) and removing “implausible”
105 (i.e., extreme) energy intake reporters may not recover unbiased diet-health associations from
106 misreporting (17). Our replication of the scoring approach has a further limitation of
107 extrapolating from FFQ data to grams instead of just servings given the data we used (see
108 Methods).

109 For the purposes of the present investigation, we set aside these limitations and instead focus on
110 how aggregation of foods into categories in the scoring approach may influence the strength and
111 nature of associations between foods and health outcomes. The justification of categorization

112 schemes depends on purpose (e.g., maximizing explained variance; detecting maximal
113 associations; determining associations with food composition; testing associations with
114 sociological dietary constructs). Furthermore, it is improbable that a food is wholly 'unhealthy' or
115 'healthy' for all circumstances, people, and outcomes (18). Thus, there is likely no single
116 'correct' definition of food categories. However, understanding the effect of reclassification of
117 foods on the associations between diet indices and health outcomes is essential to determine the
118 robustness and stability of such associations. Throughout the text, we use quotation marks
119 around 'healthy' and 'unhealthy' to emphasize that these labels are used for categorization, not
120 to indicate a determination that the foods inherently have health promoting or deleterious health
121 effects.

122 **Objectives**

123 Herein, we examine how different ways of categorizing foods in the scoring approach affect the
124 dietary associations with cardiometabolic outcomes in four ways. First, we attempt to replicate
125 the scoring approach from Satija et al. on associations between diet indices and CHD using a
126 different data set. Second, given that potatoes were uniquely assigned as a single-plant category
127 among all plant-based foods, we use potatoes as a case study to investigate the robustness of the
128 scoring approach to recategorization of foods and a food category. Third, we determine the
129 ramifications of each of the 12 plant food groups using a leave-one-out analysis. Fourth, we use
130 an agnostic approach by assigning plant food groups as 'healthy' or 'unhealthy' to determine the
131 distribution of associations across 4,096 models, and assessing the differences between the 2,048
132 models in which a food group was classified as 'healthy' and the 2,048 in which it was classified
133 as 'unhealthy'. We hypothesized that categorizing foods in different ways (e.g., within and/or
134 between 'healthy' and 'unhealthy' categories as modeled in the scoring approach) would have a

135 marked impact on the associations between diet indices and combined incident nonfatal
136 myocardial infarction (MI) and fatal CHD (together referred to as CHD herein).

137 **Methods**

138 We replicated to the best of our ability the analyses from Satija et al. (3) using the REasons for
139 Geographic and Racial Differences in Stroke (REGARDS) data, including cohort creation
140 (eligibility criteria), outcome ascertainment, selection of adjustment covariates, and other
141 methodologic choices. This allowed us to focus on the effects of changes to the scoring approach
142 within a single statistical associational framework.

143 **Study population**

144 As described elsewhere, REGARDS is an ongoing population-based longitudinal study in 30,239
145 randomly sampled adults from all 48 continental US states to examine reasons for regional and
146 racial differences in stroke mortality (19, 20). The REGARDS study was previously approved by
147 all associated institutional review boards (19). Briefly, between January 2003 and October 2007,
148 participants (Black or White race; aged ≥ 45 years) completed a 45-min computer-assisted
149 telephone interview on socio-demographics, medical history, and health status followed by an in-
150 home examination to collect fasting blood and urine samples, electrocardiograms (ECG), and
151 data on blood pressure measures, anthropometric measures, and medications taken in the 2
152 weeks before the visit via pill bottle review. An FFQ was left for self-administration. REGARDS
153 participants are contacted every six months by phone and interviewed about stroke symptoms,
154 hospitalizations, and general health status (19). . By the time of the current analysis, the
155 REGARDS study had conducted two in-home visits for data collection: one at baseline and a

156 follow-up visit approximately 10 years later. The data collection during the second in-home visit
157 was similar in terms of the variables on which data were collected to the first visit.

158 In the current analysis, we excluded individuals with history of self-reported MI or atrial
159 fibrillation, coronary artery bypass graft, angioplasty, or evidence of MI via ECG; cancer¹;
160 stroke or transient ischemic attack; and missing or implausible energy intake (described in the
161 scoring approach as <600 or >3,500 kcal/day for women and <800 or >4,200 kcal/day for men).
162 We additionally excluded participants with missing covariate data at baseline (Supplementary
163 Figure 1). Our final analytic sample included 13,684 individuals, 8,255 of whom had additional
164 data from the follow-up visit.

165 **Adjustment covariates**

166 To emulate Satija et al., we used the 'updated' covariate data from the second (i.e., most recent)
167 REGARDS data collection for participants for whom those data were available, and baseline
168 covariate data for the rest. Adjustment covariates include age at baseline (in years), updated
169 multivitamin use² (yes, no), family history of MI (yes, no), updated margarine intake (by
170 quintiles³), updated energy intake (by quintiles), diabetes⁴ at baseline (yes, no),
171 hypercholesterolemia at baseline⁵ (yes, no), hypertension⁶ at baseline (yes, no), updated smoking
172 status (never, past, current (1 to <15 cigarettes per day), current (15-<25 cigarettes per day),

¹ Satija et al. excluded individuals with cancer but with an exception for nonmelanoma skin cancer, we could not distinguish among cancer types in the REGARDS data we received.

² REGARDS question: "During the past year, have you taken any vitamins or minerals regularly, at least once a month?" (yes, no).

³ Quintiles were used, but reported intakes below q1 and q2 were combined because of imbalance in the number of participants from zero consumption.

⁴ Has diabetes if fasting glucose \geq 126/non-fasting glucose \geq 200 or pills or insulin, self-reported diabetes and taking insulin, self-reported diabetes and taking glucose lowering pills, or self-reported diabetes.

⁵ Has dyslipidemia if TC \geq 240 or LDL \geq 160 or HDL \leq 40 or on medication, self-reported elevated lipids, or self-reported use of lipid lowering medication only in those answering yes to having lipidemia.

⁶ Has hypertension if SBP \geq 140 or DBP \geq 90 or self-reported current medication use to control blood pressure.

173 current (25 or more cigarettes per day)), updated physical activity (none, 1-3 times per week, 4
174 or more times per week), updated alcohol intake⁷, updated aspirin use (yes, no), updated body
175 mass index (BMI) category (BMI<21, 21 to <23, 23 to <25, 25 to <27, 27 to <30, 30 to <33, 33
176 to <35, 35 to <40, and 40 kg/m² or higher), race (Black, White) and updated REGARDS region
177 of residence (stroke belt, stroke buckle, non-belt region⁸). Race and region are specific to
178 REGARDS and not included in the Satija et al. analysis. In the analysis stratified by gender, we
179 also adjusted for hormone use (never, past, current) and past oral contraceptive use (yes, no)
180 among female participants.

181 **Outcome ascertainment**

182 CHD was defined as incident nonfatal MI and fatal CHD. In the REGARDS dataset, fatal CHD
183 was defined as “death within 28 days of an adjudicated definite or probable MI, or CHD or
184 sudden death as the adjudicated cause of death.[...] The main underlying cause of death was
185 determined by 2 trained adjudicators who examined all available information including
186 interviews with next of kin, death certificates, autopsy reports, medical history, and the National
187 Death Index. [...] [In REGARDS, nonfatal MI ascertainment] required a clinical presentation
188 consistent with ischemia, a rising and falling pattern of troponin over at least 6 hours with a peak
189 at least twice the upper limit of normal, or imaging findings consistent with ischemia (20).”
190 Events up to December 31, 2019 were available for the current analysis.

⁷ REGARDS dataset includes a variable for ‘[number of] alcoholic drinks per week’. We divided values by 7 (to obtain per day value) and multiplied the result by 14 to obtain gram alcohol consumed per day (National Institute on Alcohol Abuse and Alcoholism. What is a standard drink? <https://www.niaaa.nih.gov/alcohols-effects-health/overview-alcohol-consumption/what-standard-drink>, Accessed June 21, 2023). We then categorized gram alcohol values per Satija et al.: 0, >0 to <5, 5 to <10, 10 to <15, or ≥15 g/day.

⁸ Stroke belt (North Carolina, South Carolina, Georgia, Tennessee, Alabama, Mississippi, Arkansas, and Louisiana), stroke buckle (the coastal plain of North Carolina, South Carolina, and Georgia), and from elsewhere in the continental US.

191 **Dietary assessment**

192 REGARDS dietary and alcohol data were collected from a self-administered 110-item Block 98
193 FFQ (NutritionQuest, Berkeley, CA) given to participants during in-home visits to complete and
194 return by mail to the REGARDS Operations Center, where they were reviewed for completeness,
195 scanned, and forwarded to NutritionQuest for processing and analysis (21, 22). The dataset we
196 received had frequency and amount of consumption of each FFQ item transformed into
197 consumed grams per day.

198 **o *Replicating the scoring approach food groups using REGARDS FFQ items:***

199 We assigned each REGARDS FFQ item to one of 18 food groups defined by the scoring
200 approach (2, 3). We relied on the details in the Satija et al. publications to determine the items
201 that composed each of the 18 food groups, resulting in three situations: 1) A REGARDS FFQ
202 item was matched with an identical item of the food groups in the scoring approach (e.g.,
203 broccoli, dark bread, strawberries, eggs, hot dog, yogurt, coffee); 2) A REGARDS FFQ item was
204 not identically listed in Satija et al. (2, 3), but based on the available details and 1984 NHS FFQ,
205 the FFQ item clearly fit into one of the food groups. For instance, Raisin Bran cereal, Chinese
206 dish, and different types of chicken from REGARDS did not have an exact match, but could be
207 reasonably fit under the food groups whole grains, miscellaneous animal-based foods, and meat,
208 respectively; and 3) A REGARDS FFQ item could not be matched with a specific item, nor did it
209 clearly fit into one of the food groups, and were thus excluded (e.g., breakfast bars, gravy, non-
210 dairy creamer, water, salsa). The scoring approach used cumulatively averaged dietary data over
211 the follow-up durations of NHS, NHS2 and HPFS. We thus averaged values of the two
212 REGARDS data points (i.e., first and second visits) for dietary information if dietary data from
213 the second visit were available, unless type 2 diabetes was diagnosed before the second data

214 capture because this may have resulted in substantial dietary change. When follow-up data were
215 not available, baseline dietary data were carried forward.

216 o ***Plant-based diet indices in the scoring approach:***

217 To calculate PDI, hPDI, and uPDI of the scoring approach, Satija and colleagues (2, 3)
218 calculated the sum of foods in a food group, created quintiles for these summed intakes of food
219 groups, and assigned a score ranging from 1 to 5 per food group per participant. For positive
220 scoring, participants received higher scores for higher reported intakes and vice versa. For
221 reverse scoring, participants received lower scores for higher reported intakes. Scoring schemes
222 for calculating PDI, hPDI, and uPDI are presented in Table 1. To obtain the PDI, hPDI, and
223 uPDI scores, the scores of all 18 food groups were summed. Thus, all three indices range in
224 theory from 18 to 90.

225 **Statistical analysis**

226 We replicated the analytical procedures as previously reported as closely as possible (2, 3).
227 Briefly, we used Cox proportional hazards regression to estimate hazard ratios (HR) and 95%
228 confidence intervals (CIs) evaluating the associations of continuous index scores and deciles of
229 each index separately with incident CHD. Person-time was calculated from age at questionnaire
230 return date until CHD diagnosis, death, or end of follow-up (December 31, 2019). We used age
231 (in years) as the time scale, but we were not able to stratify by calendar time as done by Satija
232 and colleagues because of shorter follow-up. Indices were averaged over the two dietary data
233 captures except for the cases with the diagnosis of T2D. Models were adjusted for the
234 aforementioned covariates.

235 For completeness in replication, we conducted p-for-trend analyses by assigning median values
236 to each decile bin (e.g., bin 1 from below decile 1; bin 10 from above decile 9), and this variable

237 was used in the analysis. These analyses and results are presented in the supplementary
238 materials.

239 Satija et al.'s analysis was in practice stratified by sex because they used sex-specific cohorts
240 (women: NHS and NHS2; men: HPFS). Thus, we used gender in REGARDS data as a
241 stratification variable. As in Satija et al., we combined the stratified results using a fixed-effects
242 meta-analysis. Variable recoding was done using SAS software version 9.4 (SAS Institute Inc.,
243 Cary, North Carolina) and all analyses were performed in R version 4.2.3 and RStudio version
244 2023.03.0 (R Core Team (2023). R: A language and environment for statistical computing. R
245 Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>). Statistical
246 significance was set at a 2-tailed p-value of <0.05.

247 **Recategorization schemes for potatoes**

248 In the scoring approach 'French fries', 'baked or mashed potatoes', and 'potato or corn chips' are
249 grouped as 'potatoes' as an 'unhealthy' food group. Our replicated 'potatoes' group included
250 white potato⁹, French fries¹⁰, salty snacks¹¹, and low-fat salty snacks¹². Much like the scoring
251 approach, we included FFQ items that were not exclusively potatoes if the item was
252 predominantly potatoes, despite creating a bit of a misnomer in the 'potatoes' group. To examine
253 the influence of food categorization schemes on associations between diet indices and CHD, we
254 used five alternative categorization schemes by moving all or some of these items within or
255 between 'healthy' and 'unhealthy' food groups. We then re-calculated the PDI, hPDI, and uPDI
256 for each recategorization and estimated HRs and 95% CIs for each model. Thus, we have results

⁹ REGARDS FFQ item: "White potatoes not fried, incl. boiled, baked, mashed and potato salad".

¹⁰ REGARDS FFQ item: "French fries, fried potatoes or hash browns".

¹¹ REGARDS FFQ item: "Snack like potato chips, corn chips, popcorn (not pretzels)".

¹² REGARDS FFQ item: "Snack like potato chips, corn chips, popcorn (not pretzels) - low fat".

257 from six models: the scoring approach replication and five alternative models. Table 2 shows the
258 five alternative approaches that we used to categorize potato food items. When the number of
259 food groups was either increased (model 4) or decreased (models 1 and 3) we rescaled the diet
260 indices to match the original scale that ranged from 18 to 90 points. Note that the PDI scores
261 from model 2 are identical to the scoring approach because in calculating PDI all plant food
262 groups are scored positively.

263 All food groups categorized under plant foods (either ‘healthy’ or ‘unhealthy’) are scored
264 positively in the PDI. Thus, it is expected that associations of PDI with CHD should remain
265 largely unchanged in the recategorization approach. “Potatoes” as a food group is categorized
266 under the ‘unhealthy’ category in the scoring approach and scored in reverse in hPDI calculation
267 (i.e., higher consumption of potato food items leads to lower hPDI scores). If one assumes that
268 items in the “potatoes” food group have a meaningful contribution to any detectable association
269 with the risk of CHD, and if all or some of the items in the “potatoes” group are miscategorized
270 as ‘unhealthy’ when in fact they should be ‘healthy’, moving them to ‘healthy’ should result in
271 stronger, inverse associations (lower HRs) between hPDI and CHD risk (that is: as hPDI scores
272 go up, CHD risk goes down). Similarly, for uPDI, if potatoes are incorrectly categorized as
273 ‘unhealthy’ when in fact they should be categorized as ‘healthy’, moving all or some of them to
274 ‘healthy’ should make the associations between uPDI and CHD risk stronger. As such, hPDI is
275 expected to have the strongest beneficial association followed by PDI, while uPDI is expected to
276 have the strongest harmful association with CHD ($hPDI < PDI < uPDI$).

277 **Leave-One-Out Analyses**

278 PDI, hPDI, and uPDI scores were calculated as done in the scoring approach but leaving one of
279 the 12 plant-based food groups out of the calculation at a time. This analysis aims to determine

280 the influence of perturbing an entire food group on the model results, which reveals the
281 importance of each plant-based food group as well as creating a benchmark against which to
282 compare the magnitude of the recategorization results. When potatoes were left out, this
283 approach is equivalent to Model 1 from the recategorization approach above.

284 **Agnostic permutation analyses**

285 Indices were calculated using all possible combinations of each of the 12 plant-based food
286 groups scored either positively or reversely ($2^{12}=4096$ combinations). Because uPDI and hPDI
287 are coded oppositely, and all combinations were tested, uPDI results mirror the results of hPDI.
288 Therefore, we only present the combinations as hPDI. We examined the distribution of HRs from
289 positive and reverse scoring using paired t-tests of the log HRs (separately for each plant-based
290 food group).

291 **Results and Interpretations:**

292 During 153,286 person-years of follow-up, there were 868 cases of incident CHD (388 women
293 with 92,334 person-years and 480 men with 60,952 person-years).

294 **Replicating the Scoring Approach of Satija et al.**

295 We obtained comparable median values within decile bins of PDI, hPDI, and uPDI in the
296 replicated models compared with Satija et al.'s median values (3) (see Supplementary Table 1).
297 For instance, men's hPDI median value was 42.0 in decile bin 1 and 65.0 in bin 10 in our
298 replicated calculation, which compares with 43.0 in bin 1 and 66.0 in bin 10 in (3). We did not
299 observe statistically significant associations of PDI and uPDI scores and the risk of CHD for
300 women, men, or both combined (see Supplementary Tables 2 and 4). Most associations for hPDI
301 scores and risk of CHD were also not statistically significant; only the following HRs were

302 significant at $p < 0.05$: decile bin 7 for women and combined, bin 9 for men, bin 5 for combined, p
303 for trend in combined (see Supplementary Table 3).

304 **Reconfiguring potatoes within the scoring approach**

305 Supplementary Table 1 displays the median values for each decile bin of PDI, hPDI, and uPDI
306 scores by gender obtained in the reconfiguration approaches from Table 2. Because scores are
307 constrained by quantiles (i.e., if one individual's score goes up another must come down),
308 differences in the scores were small (< 2 in any decile category of index and configuration).

309 Comparing the mean diet scores among the five reconfigurations (see Supplementary Figure 2),
310 we similarly see only minimal changes (< 0.2 within or between patterns and configurations).

311 In total, the use of deciles resulted in 9 comparisons (decile bin 2 to bin 1; bin 3 to bin 1, etc...),
312 plus the continuous analysis per 10-unit change in diet index score, and a 'p-for-trend' analysis
313 that uses means of decile bins in a single regression. All decile comparisons are reported for
314 transparency (Supplementary Tables 2-4), but for ease of interpretation, and consistent with the
315 concept of dose-response and monotonic effect, only the extreme decile comparison will be
316 discussed of the decile comparisons, in addition to the continuous analyses. Point estimates of
317 HRs were qualitatively evaluated as moving towards or away from the null (null association
318 shown as $HR = 1.00$) compared with the scoring approach. A lower HR for hPDI and a higher HR
319 for uPDI, as compared with the scoring approach, suggest that the reconfigured models show
320 stronger associations. Stronger associations are consistent with recategorization of potato food
321 items more effectively capturing the concept of 'healthy' and 'unhealthy' food groups.

322 Figure 1 shows the associations between PDI, hPDI, or uPDI (HRs, continuous per 10-unit
323 change in diet indices in upper panel; highest versus lowest decile in lower panel) with CHD risk

324 in the five recategorization schemes for potatoes by the p-value of the HRs. Points moving in the
325 horizontal direction (compared with the replicated scoring approach) indicate a change in the
326 strength of the association, while points moving in the vertical direction indicate a change in the
327 significance level.

328 Overall, the recategorization approaches resulted in minimal changes in continuous HRs (per
329 increments of 10-units in diet scores). For instance, HRs for all five models with PDI in women
330 were 0.82 with almost identical 95% CIs. In the analysis for men, HRs in the recategorization
331 moved slightly towards the null ($HR=1$) with increasing p-values for hPDI and uPDI scores, but
332 not for PDI scores. Thus, we are not able to make strong quantitative distinctions because there
333 were only slight differences in precision among models, with even smaller differences in HRs.
334 Comparing patterns overall, higher PDI and hPDI scores were consistent with lower CHD risk,
335 and a higher uPDI score was consistent with higher CHD risk in women, which is consistent
336 with the results reported by Satija et al., albeit not statistically significant. However, the HRs of
337 PDI for women are lower than those of hPDI, in contrast to Satija et al.

338 Recategorization resulted in greater HR dispersion when using the diet scores as deciles
339 (Figure 1, bottom panels), but the results were not fully consistent between men and women. For
340 example, HRs for uPDI scores moved away from the null for men, while they moved towards the
341 null for women in the recategorization approach compared with the scoring approach, with
342 model 3 (all potato items into vegetable food group) reaching statistical significance. Comparing
343 patterns overall, higher PDI and hPDI scores again were consistent with lower CHD risk, and a
344 higher uPDI score was consistent with higher CHD risk in women, which is again consistent
345 with the results reported by Satija et al., albeit typically not statistically significant. In contrast
346 again, the HRs of PDI for women are lower than those of hPDI.

347 **Leave-one-out analyses**

348 We conducted a leave-one-out analysis to determine the impact of removing other entire food
349 groups on associations of diet indices with CHD.

350 Supplementary Figure 3 shows the baseline mean values of diet indices in leave-one-out models
351 for women and men, and Supplementary Table 5 shows median values of the decile bins of PDI,
352 hPDI and uPDI by gender and model variations in the leave-one-out approach. Similar to the
353 reconfiguration approach, differences in the scores were small across leave-one-out patterns
354 because of constraints of the quantile approach.

355 Figure 2 shows the multivariable adjusted associations of 10-unit changes (top panel) and highest
356 vs lowest decile (bottom panel) in PDI, hPDI, and uPDI scores with incident CHD, leaving each
357 of the 12 plant-based food groups out one at a time. For specific numerical values, see
358 Supplementary Tables 6 (for PDI), 7 (for hPDI), and 8 (for uPDI). The general pattern of the
359 magnitude of HRs (PDI HRs < hPDI HRs < uPDI HRs for women, hPDI < PDI < uPDI for men)
360 was similar to the recategorizing approach and our replication of the scoring approach. Similar to
361 the recategorization models (as shown in Figure 1), changes in magnitude of point estimates
362 compared with the original scoring approach (depicted in black) are relatively small for
363 continuous diet scores and much more spread out for the diet scores by deciles (Figure 2). The
364 directional change in the point estimates (i.e., HRs) when leaving one out compared with our
365 replication of the scoring approach varied. For instance, removing the food groups of vegetables
366 or fruit juices resulted in the strongest beneficial associations between PDI continuous diet
367 scores and CHD among women, with p-values below 0.05 (Figure 2, top panel). Conversely, all
368 leave-one-out for the uPDI diet scores by deciles attenuated associations for women while for
369 men associations were strengthened compared with the scoring approach (bottom panel).

370 **Agnostic permutation of foods**

371 The appropriateness of categorizing certain foods as ‘healthy’ or ‘unhealthy’ raises questions,
372 particularly considering the shifts in associations seen with the leave-one-out approach. We
373 therefore agonistically redefined each food group as ‘healthy’ (positively coded) or ‘unhealthy’
374 (reversely coded) for all food groups, resulting in 4,096 hPDI models (and 4,096 uPDI models
375 that are exact complements; thus, only hPDIs are shown here for simplicity).

376 If the scoring approach is appropriately specified, then changing a food group that is actually
377 healthy and labeled as ‘healthy’ to label it as ‘unhealthy’ (or the converse for unhealthy food
378 groups), would be expected to weaken the associations (that is, move the hPDI toward the null
379 HR of 1). We did not observe a consistent pattern of changes in the expected direction when
380 ‘healthy’ or ‘unhealthy’ foods were recategorized. Instead, we observed a mix of changes in both
381 directions and, in many cases, little difference at all. Tables 3 and 4 show comparison statistics
382 for the distribution of agnostically defined ‘healthy’ and ‘unhealthy’ food groups, respectively.

383 Food groups for which the positive and reverse coding resulted in a shift in HR in the expected
384 direction for the association between hPDI score and CHD are highlighted in gray. For women,
385 HRs were on average below 1 when whole grains, fruits, nuts, and vegetables were positively
386 scored, indicating that higher consumption of foods in those food groups were associated with
387 lower risk of CHD. When those same four food groups were reversely scored, the HRs were on
388 average larger than 1, which means that lower consumption of foods in those food groups were
389 associated with higher risk of CHD. The same trend was observed in the male subsample for
390 fruits, legumes, and vegetable oil. Sweets and desserts for women is an example where the shifts
391 of HRs were not in the expected direction. Scoring sweets and desserts positively (higher
392 consumption, higher hPDI score) resulted in HRs that were smaller than 1 (i.e., higher

393 consumption associated with lower risk of CHD). In the reverse scoring (higher consumption,
394 lower hPDI score), HRs were on average larger than one (i.e., lower consumption associated
395 with higher risk of CHD).

396 Violin plots in Figure 3 display the distribution of HRs for the association of continuous hPDI
397 scores with CHD risk, with the distribution of HRs for the five ‘unhealthy’ food groups from the
398 scoring approach on the left side and the seven ‘healthy’ food groups on the right side. hPDI
399 scores by deciles (Supplementary Figure 4) were qualitatively similar to the results from
400 continuous hPDI scores. Empirical cumulative density functions of the associations from Figure
401 3 are shown in Supplementary Figures 5 and 6 to visualize differences in distributions between
402 positive and reverse coding of each food group.

403 **Discussion**

404 In this methodological exploration of a plant-based dietary scoring approach, 1) we attempted to
405 replicate associations between plant-based diet indices and risk of CHD; 2) we adopted a
406 pragmatic approach in rearranging the foods within and among categories to test the robustness
407 of associations between diet indices and CHD risk, using potatoes as a particular case study; 3)
408 we tested the ramifications of leaving each of the 12 plant-based food groups out of the
409 calculation one at a time; and 4) we calculated hPDI scores using all possible combinations of
410 each of the 12 plant-based food groups scored either positively or reversely (i.e., agnostically
411 assigning plant-based food groups as ‘healthy’ or ‘unhealthy’), and compared the distributions.
412 Our replications of the scoring approach were not statistically significant, but were generally
413 similar in magnitude to the results of Satija et al.; that is, the plant diet index associations were of
414 small magnitude, but generally in the same direction, with the exception of PDI resulting in more
415 extreme HR point estimates than the hPDI for women. We acknowledge, however, that the

416 REGARDS dataset is smaller than the datasets used by Satija et al. (3), and our study is not an
417 exact replication of their study. Nonetheless, the focus of our study was to explore robustness
418 and performance of the scoring approach, rather than to draw conclusions regarding actual
419 associations of healthfulness of foods and food groups given the other limitations of the
420 approach. PDI scores did not differ much across the models that reconfigured potatoes as a case
421 study (i.e., models 1-5). This is not surprising based on the way the indices were constructed
422 using a quintile-decile approach; the approach focuses on relative rankings of individuals by
423 quintiles of consumption rather than the specific foods consumed. Thus, moving foods within
424 and among categories (as done in models 1-5) may not appreciably alter the overall diet scores.
425 Evaluating the healthfulness of individual food groups using results from indices, as opposed to
426 providing inferences in terms of the intact index, can contribute to confusion regarding the
427 relationship between food and disease. For instance, an index from the scoring approach very
428 well may support an association with CHD; however, support for conclusions regarding any one
429 food group within the index may be inappropriate. As an example, ‘sweets and desserts’ showed
430 more beneficial associations with CHD in women, consistent with the discussion about
431 “Nutrition Science’s Most Preposterous Result” (23); yet, the differences were on par with the
432 deleterious associations with CHD of ‘sugar sweetened beverages’ in men. To disregard one
433 while upholding the other at the individual food group level is thus logically and
434 methodologically incongruous, and thus interpretations of intact indices must be considered
435 without being able to make inferences about individual items within the index.
436 Others comparing plant-based dietary indices have also shown variability and inconsistencies.
437 Kim and colleagues (24) contrasted the association of 5 different plant-based dietary indices
438 previously developed by others, namely the 3 indices of the scoring approach (PDI, hPDI, uPDI)

439 (2), provegetarian diet index (1), and overall PDI (i.e., PDI-Rotterdam) (25). They assessed
440 whether these indices are associated with risk of incident hypertension. Participants above the
441 highest vs the lowest quintile PDI, hPDI, and provegetarian diet index were associated with
442 lower risk of hypertension. No significant association was found for the PDI-Rotterdam, and
443 uPDI associations were inconsistent depending on the model. The authors concluded that
444 operational differences could affect indices' ability to detect diet and disease associations. Still, it
445 remained unclear how one should interpret similar associations detected for differently
446 operationalized indices. "Finding no association between the PDI-Rotterdam and hypertension"
447 was described as 'surprising' because PDI-Rotterdam had shown moderate to strong correlations
448 with other indices which had association with hypertension risk, and that in general, there have
449 been strong inverse association between plant-rich dietary patterns that include dairy (e.g.,
450 DASH diet) and hypertension risk. If the results obtained from indices should align with pre-
451 existing evidence, it is important to evaluate the information provided by these indices, and
452 whether the inconsistencies stem from methodological limitations or if they offer valuable new
453 insights. If the indices do not carry any better evidence than a random food classification,
454 emergence of inconsistent evidence only results in confusion, which requires researchers to
455 speculate, selectively report, and deviate from objective interpretation of evidence.
456 One key finding of our agnostic permutations approach is that every scoring method yielded
457 various associations between food groups and CHD, so that each food group can be linked to
458 both beneficial and harmful effects on CHD risk through multiple scores. For instance, there are
459 indices with statistically significant associations with CHD, where one has whole grains
460 positively coded and the other has whole grains reversely coded. Similarly, multiple scores for
461 potatoes show statistically significant associations with CHD in both positive and reverse coding

462 scenarios. The direction of change in associations sometimes aligns with expectations, and
463 sometimes does not.

464 Our methodological evaluation of a plant-based diet scoring approach revealed that alterations of
465 the scoring approach result in varied and unpredictable diverse associations between food groups
466 and CHD risk. Each food group demonstrated the potential for being both beneficially and
467 deleteriously associated with CHD risk. The observed variability in association directions
468 underscores the potential for 'guilt by association,' (or, conversely, 'benefit by association') in
469 which any of the food groups we studied could be categorized with either 'healthy' or
470 'unhealthy' foods and be found to be part of a dietary index beneficially or deleteriously
471 associated with CHD.

472 *"If the same mathematical pattern can yield such disparate interpretations, what claim can*
473 *either have upon reality?" - Stephen Jay Gould, Mismeasure of Man (26)*

474

475 **Acknowledgements**

476 The REGARDS study is supported by cooperative agreement U01 NS041588 co-funded by the
477 National Institute of Neurological Disorders and Stroke (NINDS) and the National Institute on
478 Aging (NIA), National Institutes of Health, Department of Health and Human Service.
479 Additional funding for REGARDS CHD outcomes was provided by R01HL080477. The content
480 is solely the responsibility of the authors and does not necessarily represent the official views of
481 the NINDS, NHLBI, or the NIA. Representatives of the NINDS were involved in the review of
482 the manuscript but were not directly involved in the collection, management, analysis or
483 interpretation of the data. The authors thank the other investigators, the staff, and the participants
484 of the REGARDS study for their valuable contributions. A full list of participating REGARDS
485 investigators and institutions can be found at: <https://www.uab.edu/soph/regardsstudy/>

486 AWB, DBA, and YJ-N designed research; BH analyzed data; YJ-N, AWB, and BH wrote the
487 paper; DBA, JMS, AWB, YJ-N, and BH reviewed and edited the manuscript. AWB and YJ-N
488 had primary responsibility for final content. All authors read and approved the final manuscript.

489 **Data Availability:** Researchers who wish to reproduce our analyses can submit a project
490 proposal to the REGARDS team (<https://www.uab.edu/soph/regardsstudy/>). We did not have any
491 access privilege that others would not have. The analytic code is publicly available at
492 <https://osf.io/c3b4z/>.

493 **Funding:** The present analysis is supported by the Alliance for Potato Research & Education.
494 The funder did not have any role in design and conduct of the analyses; interpretation of the
495 results; preparation of the manuscript; and decision to submit the manuscript for publication.
496

497 **Author Disclosures:**

498 In the 36 months prior to the initial submission, DBA has received personal payments or
499 promises for same from: Amin Talati Wasserman for KSF Acquisition Corp (Glanbia); General
500 Mills; Kaleido Biosciences; Law Offices of Ronald Marron; Medpace/Gelesis; Novo Nordisk
501 Foundation; and Zero Longevity Science (as stock options). Donations to a foundation have been
502 made on his behalf by the Northarvest Bean Growers Association. DBA's institution, Indiana
503 University, and the Indiana University Foundation have received funds or donations to support
504 his research or educational activities from: Alliance for Potato Research and Education;
505 American Egg Board; Arnold Ventures; Eli Lilly and Company; Mars, Inc.; National
506 Cattlemen's Beef Association; National Pork Board; Pfizer, Inc.; USDA; WW (formerly Weight
507 Watchers); and numerous other for-profit and non-profit organizations to support the work of the
508 School of Public Health and the university more broadly. AWB has received travel expenses
509 from Alliance for Potato Research and Education, International Food Information Council,
510 International Food Information Council Foundation, and Soy Nutrition Institute Global; speaking
511 honoraria from Alliance for Potato Research and Education, American Society for Nutrition,

512 Calorie Control Council, Eastern North American Region of the International Biometric Society,
513 International Food Information Council Foundation, Potatoes USA, Purchaser Business Group
514 on Health, The Obesity Society, and University of Arkansas for Medical Sciences; consulting
515 payments from National Cattlemen’s Beef Association, and Soy Nutrition Institute Global; and
516 grants through his institution from Alliance for Potato Research & Education, American Egg
517 Board, National Cattlemen’s Beef Association, NIH/NHLBI, NIH/NIDDK, NIH/NIGMS, and
518 NSF. He has been involved in research for which his institution or colleagues have received
519 grants or contracts from Alliance for Potato Research & Education, Gordon and Betty Moore
520 Foundation, Hass Avocado Board, NIH/NCATS, NIH/NCI, NIH/NIA, NIH/NIGMS, and
521 NIH/NLM. His wife is employed by Reckitt. BH has been involved in research for which her
522 institution, Indiana University, has received grants from NIH; WW International, Inc., the
523 National Pork Board, the California Walnut Commission, the Alliance of Potato Research &
524 Education, and the National Cattlemen’s Beef Association. YJ-N has received honoraria from
525 The Alliance for Potato Research and Education. JMS reports no disclosures.

526

527 **References**

- 528 1. Martínez-González MA, Sánchez-Tainta A, Corella D, Salas-Salvadó J, Ros E, Arós F, et
529 al. A provegetarian food pattern and reduction in total mortality in the Prevención con Dieta
530 Mediterránea (PREDIMED) study. *The American Journal of Clinical Nutrition*.
531 2014;100(suppl_1):320S-8S.
- 532 2. Satija A, Bhupathiraju SN, Rimm EB, Spiegelman D, Chiuve SE, Borgi L, et al. Plant-
533 Based Dietary Patterns and Incidence of Type 2 Diabetes in US Men and Women: Results from
534 Three Prospective Cohort Studies. *PLOS Medicine*. 2016;13(6):e1002039.
- 535 3. Satija A, Bhupathiraju SN, Spiegelman D, Chiuve SE, Manson JE, Willett W, et al.
536 Healthful and Unhealthful Plant-Based Diets and the Risk of Coronary Heart Disease in U.S.
537 Adults. *Journal of the American College of Cardiology*. 2017;70(4):411-22.
- 538 4. Satija A, Malik V, Rimm EB, Sacks F, Willett W, Hu FB. Changes in intake of plant-
539 based diets and weight change: results from 3 prospective cohort studies. *The American Journal*
540 *of Clinical Nutrition*. 2019;110(3):574-82.
- 541 5. Gómez-Donoso C, Martínez-González M, Martínez JA, Gea A, Sanz-Serrano J, Perez-
542 Cueto FJA, Bes-Rastrollo M. A Provegetarian Food Pattern Emphasizing Preference for Healthy
543 Plant-Derived Foods Reduces the Risk of Overweight/Obesity in the SUN Cohort. *Nutrients*.
544 2019;11(7).
- 545 6. Liu X, Dhana K, Barnes LL, Tangney CC, Agarwal P, Aggarwal N, et al. A healthy
546 plant-based diet was associated with slower cognitive decline in African American older adults:
547 a biracial community-based cohort. *The American Journal of Clinical Nutrition*.
548 2022;116(4):875-86.
- 549 7. Shah S, Mahamat-Saleh Y, Ait-Hadad W, Koemel NA, Varrasso R, Boutron-Ruault M-C,
550 Laouali N. Long-term adherence to healthful and unhealthful plant-based diets and breast cancer
551 risk overall and by hormone receptor and histologic subtypes among postmenopausal females.
552 *The American Journal of Clinical Nutrition*. 2023;117(3):467-76.
- 553 8. Thompson AS, Tresserra-Rimbau A, Karavasiloglou N, Jennings A, Cantwell M, Hill C,
554 et al. Association of Healthful Plant-based Diet Adherence With Risk of Mortality and Major
555 Chronic Diseases Among Adults in the UK. *JAMA Network Open*. 2023;6(3):e234714-e.
- 556 9. Shang X, Liu J, Zhu Z, Zhang X, Huang Y, Liu S, et al. Healthy dietary patterns and the
557 risk of individual chronic diseases in community-dwelling adults. *Nature Communications*.
558 2023;14(1):6704.
- 559 10. Wang S, Li W, Li S, Tu H, Jia J, Zhao W, et al. Association between plant-based dietary
560 pattern and biological aging trajectory in a large prospective cohort. *BMC Medicine*.
561 2023;21(1):310.
- 562 11. Maroto-Rodriguez J, Delgado-Velandia M, Ortolá R, Carballo-Casla A, García-Esquinas
563 E, Rodríguez-Artalejo F, Sotos-Prieto M. Plant-based diets and risk of frailty in community-
564 dwelling older adults: the Seniors-ENRICA-1 cohort. *GeroScience*. 2023;45(1):221-32.
- 565 12. Savage SL, Danziger J. *The flaw of averages : why we underestimate risk in the face of*
566 *uncertainty*. 1st edition ed. Hoboken, New Jersey: John Wiley & Sons; 2009.
- 567 13. Schatzkin A, Kipnis V, Carroll RJ, Midthune D, Subar AF, Bingham S, et al. A
568 comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological
569 cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN)
570 study. *International Journal of Epidemiology*. 2003;32(6):1054-62.

- 571 14. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TIA, et
572 al. Energy balance measurement: when something is not better than nothing. *International*
573 *Journal of Obesity*. 2015;39(7):1109-13.
- 574 15. Kristal AR, Peters U, Potter JD. Is It Time to Abandon the Food Frequency
575 Questionnaire? *Cancer Epidemiology, Biomarkers & Prevention*. 2005;14(12):2826-8.
- 576 16. Kristal AR, Potter JD. Not the Time to Abandon the Food Frequency Questionnaire:
577 Counterpoint. *Cancer Epidemiology, Biomarkers & Prevention*. 2006;15(10):1759-60.
- 578 17. Yamamoto N, Ejima K, Zoh RS, Brown AW. Bias in nutrition-health associations is not
579 eliminated by excluding extreme reporters in empirical or simulation studies. *eLife*.
580 2023;12:e83616.
- 581 18. ConscienHealth. The Circular Logic of Healthy Foods 2016 [Available from:
582 <https://conscienhealth.org/2016/03/the-circular-logic-of-healthy-food/>].
- 583 19. Howard VJ, Cushman M, Pulley L, Gomez CR, Go RC, Prineas RJ, et al. The Reasons
584 for Geographic and Racial Differences in Stroke Study: Objectives and Design.
585 *Neuroepidemiology*. 2005;25(3):135-43.
- 586 20. Safford MM, Reshetnyak E, Sterling MR, Richman JS, Muntner PM, Durant RW, et al.
587 Number of Social Determinants of Health and Fatal and Nonfatal Incident Coronary Heart
588 Disease in the REGARDS Study. *Circulation*. 2021;143(3):244-53.
- 589 21. Boucher B, Cotterchio M, Kreiger N, Nadalin V, Block T, Block G. Validity and
590 reliability of the Block98 food-frequency questionnaire in a sample of Canadian women. *Public*
591 *Health Nutrition*. 2006;9(1):84-93.
- 592 22. Troeschel AN, Byrd DA, Judd S, Flanders WD, Bostick RM. Associations of dietary and
593 lifestyle inflammation scores with mortality due to CVD, cancer, and all causes among Black
594 and White American men and women. *British Journal of Nutrition*. 2022:1-12.
- 595 23. Johns DM. Nutrition Science's Most Preposterous Result. Studies show a mysterious
596 health benefit to ice cream. Scientists don't want to talk about it. *The Atlantic*. 2023.
- 597 24. Kim H, Rebholz CM, Garcia-Larsen V, Steffen LM, Coresh J, Caulfield LE. Operational
598 Differences in Plant-Based Diet Indices Affect the Ability to Detect Associations with Incident
599 Hypertension in Middle-Aged US Adults. *The Journal of Nutrition*. 2020;150(4):842-50.
- 600 25. Chen Z, Zuurmond MG, van der Schaft N, Nano J, Wijnhoven HAH, Ikram MA, et al.
601 Plant versus animal based diets and insulin resistance, prediabetes and type 2 diabetes: the
602 Rotterdam Study. *European Journal of Epidemiology*. 2018;33(9):883-93.
- 603 26. Gould SJ. *Mismeasure of man*: WW Norton & company; 1996.

604

605

606 **Table 1. Scoring schemes for creating plant-based diet indices.**

	PDI	hPDI	uPDI
Healthy Plant Food Groups			
Whole grains	Positive	Positive	Reverse
Fruits			
Vegetables			
Nuts			
Legumes			
Vegetable oils			
Tea and coffee			
Less Healthy Plant Food Groups			
Fruit juices	Positive	Reverse	Positive
Refined grains			
Potatoes			
Sugar sweetened beverages			
Sweets and desserts			
Animal Food Groups			
Animal fat	Reverse	Reverse	Reverse
Dairy			
Egg			
Fish or seafood			
Meat			
Miscellaneous animal-based foods			

607

608

609 **Table 2. Recategorization Approaches for ‘Potatoes’ food group¹**

Recategorization approach	Description
Model 1: "exclude potato group"	Exclude all items in “potatoes” group from indices
Model 2: “potato group in healthy”	Move “potatoes” food group as a whole and unchanged group into “Healthy” category
Model 3: "potatoes in vegetable group"	Move all items in “potatoes” group into “vegetables” group
Model 4: "healthy and unhealthy potatoes"	Split up “potatoes” group: - Move white potato as a single item group into “Healthy” category, - Keep French fries, salty snacks, and low-fat salty snacks in “potatoes” under “less-healthy” category
Model 5: "vegetables and unhealthy potatoes"	Split up “potatoes” group: - Move white potato into “Vegetables” group, - Keep French fries, salty snacks, and low-fat salty snacks in “Potatoes” under “less-healthy” category

610 ¹Our replication of the ‘potatoes’ food group using the REGARDS data consists of white potato, French fries, salty
611 snacks, and low-fat salty snacks.

612

Table 3: Summary statistics on agnostically permuted models (hPDI score) – healthy food groups

	Whole grains		Fruits		Vegetables		Nuts		Legumes		Vegetable Oil		Tea & coffee	
	Positive	Reverse	Positive	Reverse	Positive	Reverse	Positive	Reverse	Positive	Reverse	Positive	Reverse	Positive	Reverse
Females														
HR: Mean (empirical 95% spread)	0.95 (0.79, 1.13)	1.07 (0.90, 1.27)	0.99 (0.80, 1.21)	1.02 (0.83, 1.26)	1.03 (0.82, 1.26)	0.98 (0.80, 1.19)	0.96 (0.79, 1.16)	1.05 (0.87, 1.27)	1.00 (0.81, 1.24)	1.01 (0.82, 1.22)	0.95 (0.79, 1.14)	1.06 (0.89, 1.27)	1.00 (0.81, 1.23)	1.01 (0.82, 1.23)
Median (Min, Max)	0.94 (0.73, 1.21)	1.07 (0.85, 1.36)	0.99 (0.73, 1.30)	1.02 (0.78, 1.36)	1.03 (0.77, 1.36)	0.99 (0.73, 1.28)	0.96 (0.73, 1.24)	1.05 (0.81, 1.36)	1.00 (0.73, 1.36)	1.01 (0.76, 1.36)	0.95 (0.73, 1.21)	1.06 (0.84, 1.36)	1.00 (0.73, 1.36)	1.01 (0.75, 1.36)
Difference ^a in means	-0.12		-0.04		0.04		-0.09		-0.00		-0.11		-0.01	
Ratio ^b of means	0.89		0.96		1.04		0.91		1.00		0.90		0.99	
t-test ^c	p<0.001		p<0.001		p<0.001		p<0.001		p<0.001		p<0.001		p<0.001	
Males														
HR: Mean (empirical 95% spread)	1.03 (0.85, 1.24)	0.99 (0.81, 1.17)	0.95 (0.81, 1.10)	1.07 (0.91, 1.24)	1.02 (0.85, 1.24)	0.99 (0.81, 1.17)	1.02 (0.84, 1.23)	1.00 (0.81, 1.20)	0.97 (0.81, 1.17)	1.04 (0.88, 1.24)	0.98 (0.81, 1.17)	1.04 (0.86, 1.23)	1.01 (0.83, 1.22)	1.00 (0.83, 1.21)
Median (Min, Max)	1.03 (0.76, 1.38)	0.99 (0.72, 1.30)	0.95 (0.72, 1.20)	1.07 (0.84, 1.38)	1.02 (0.76, 1.38)	1.00 (0.72, 1.26)	1.02 (0.75, 1.38)	1.00 (0.72, 1.35)	0.97 (0.72, 1.31)	1.04 (0.81, 1.38)	0.98 (0.72, 1.32)	1.04 (0.77, 1.38)	1.01 (0.73, 1.38)	1.01 (0.72, 1.35)
Difference ^a in means	0.04		-0.12		0.03		0.02		-0.07		-0.06		0.00	
Ratio ^b of means	1.04		0.89		1.03		1.02		0.94		0.95		1.00	
t-test ³	p<0.001		p<0.001		p<0.001		p<0.001		p<0.001		p<0.001		p<0.001	

Note: Gray background highlights food groups with expected direction of associations where ‘healthy’ food groups had HR in the positive scoring (higher consumption, higher hPDI score) below 1 and HRs in the reverse scoring (higher consumption, lower hPDI score) of larger than 1; ^adifference = $\exp(\text{mean of } \log(\text{HR}_{\text{positive}})) - \exp(\text{mean of } \log(\text{HR}_{\text{reverse}}))$; ^bratio = $\exp(\text{mean of } \log(\text{HR}_{\text{positive}})) / \exp(\text{mean of } \log(\text{HR}_{\text{reverse}}))$; ^cPaired t –test comparing $\log(\text{HR})$.

Table 4: Summary statistics on agnostically permuted approach (hPDI score) – unhealthy food groups

	Fruit juices		Refined grains		Potatoes		Sugar sweetened beverages		Sweets and desserts	
	Positive	Reverse	Positive	Reverse	Positive	Reverse	Positive	Reverse	Positive	Reverse
Females										
HR: Mean (empirical 95% spread)	1.02 (0.83, 1.25)	0.99 (0.80, 1.21)	1.01 (0.82, 1.25)	1.00 (0.81, 1.22)	1.00 (0.81, 1.24)	1.01 (0.82, 1.22)	1.00 (0.81, 1.22)	1.01 (0.82, 1.24)	0.96 (0.79, 1.16)	1.05 (0.87, 1.27)
Median (Min, Max)	1.02 (0.76, 1.36)	0.99 (0.73, 1.29)	1.01 (0.74, 1.36)	1.00 (0.73, 1.34)	1.01 (0.73, 1.36)	1.01 (0.75, 1.36)	1.00 (0.73, 1.33)	1.01 (0.75, 1.36)	0.97 (0.73, 1.25)	1.05 (0.81, 1.36)
Difference ^a in means	0.04		0.01		-0.00		-0.01		-0.08	
Ratio ^b of means	1.04		1.01		1.00		0.99		0.92	
t test ^c	p<0.001		p<0.001		p<0.001		p<0.001		p<0.001	
Males										
HR: Mean (empirical 95% spread)	1.03 (0.85, 1.23)	0.98 (0.81, 1.18)	0.98 (0.81, 1.18)	1.04 (0.86, 1.23)	1.03 (0.85, 1.24)	0.98 (0.81, 1.16)	1.05 (0.88, 1.24)	0.97 (0.81, 1.16)	1.03 (0.85, 1.23)	0.99 (0.81, 1.19)
Median (Min, Max)	1.03, 0.77, 1.38	0.98, 0.72, 1.31	0.98, 0.72, 1.33	1.03, 0.77, 1.38	1.03, 0.76, 1.38	0.99, 0.72, 1.26	1.05, 0.80, 1.38	0.97, 0.72, 1.28	1.03, 0.76, 1.38	0.99, 0.72, 1.32
Difference ^a in means	0.05		-0.05		0.05		0.08		0.04	
Ratio ^b of means	1.05		0.95		1.05		1.08		1.04	
t test ^c	p<0.001		p<0.001		p<0.001		p<0.001		p<0.001	

Note: gray background highlights food groups with expected direction of associations where ‘unhealthy’ food groups had HR in the positive scoring (higher consumption, higher hPDI score) larger than 1 and HR in the reverse scoring (higher consumption, lower hPDI score) of below 1; ^adifference = exp(mean of log(HR_{positive})) – exp(mean of log(HR_{reverse})); ^bratio = exp(mean of log(HR_{positive})) / exp(mean of log(HR_{reverse})). ^cPaired t test comparing log(HR).

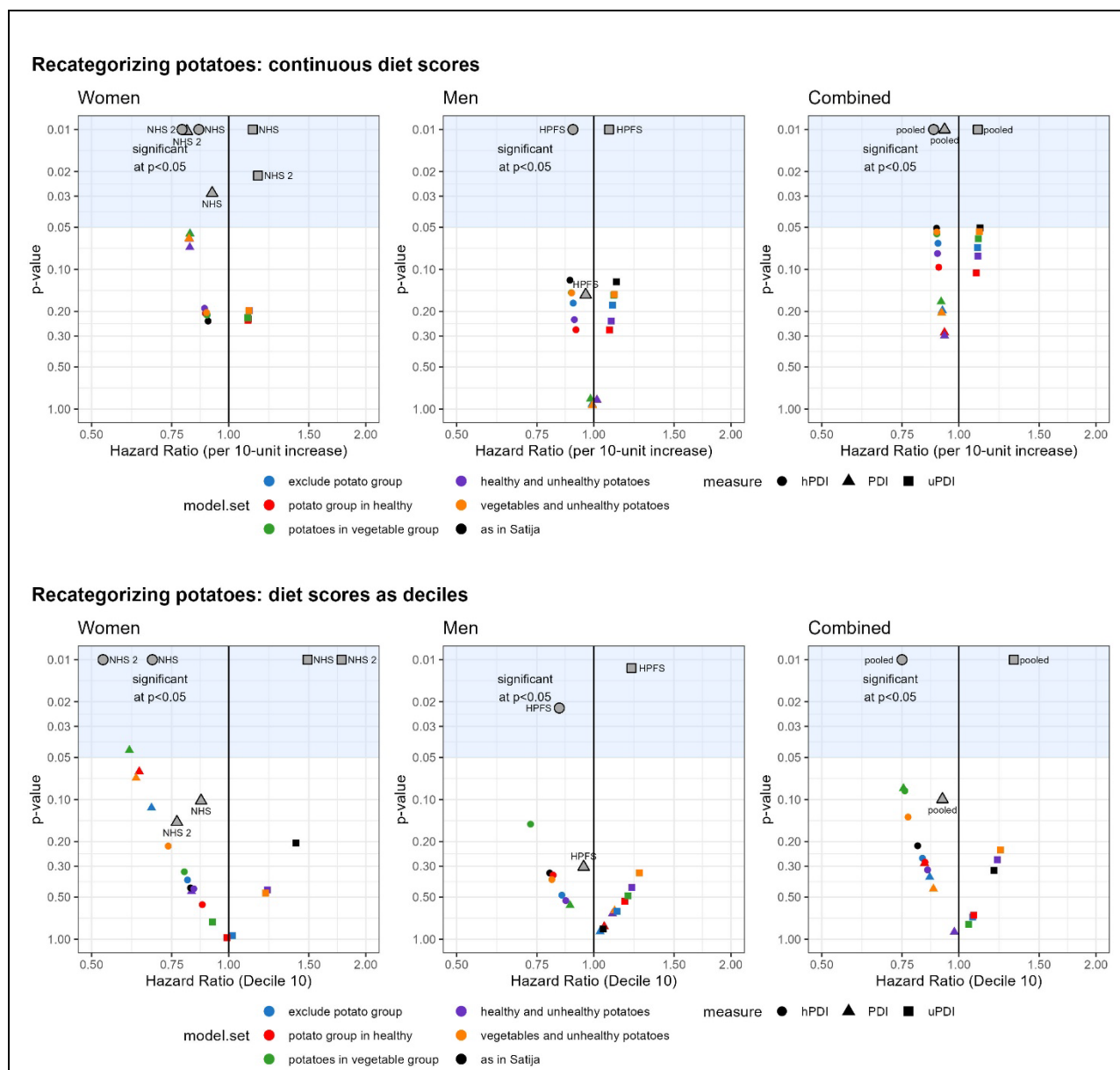


Figure 1: Associations of PDI, hPDI, and uPDI indices and CHD for five different recategorization schemes for potatoes (see Table 2) and our replication of the scoring approach. Top panel: per 10 units using continuous diet scores. Bottom panel: diet scores as deciles, shown are HR for highest decile compared with lowest decile. For specific numerical values, refer to Supplementary Table 2 (for PDI), Supplementary Table 3 (for hPDI), and Supplementary Table 4 (for uPDI). Black – (Scoring approach (as in Satija): Replication of the Scoring Approach; Blue – “exclude potato group” (Model 1): Exclude all items in “Potatoes” group from indices; Red – “potato group in healthy” (Model 2): Move “Potatoes” group as a whole intact group into “Healthy” super category; Green – “potatoes in vegetable group” (Model 3): Move all items in “Potatoes” group into “Vegetables” group; Purple – “healthy and unhealthy potatoes” (Model 4): Split “Potatoes” group by moving white potatoes as a single item group into “Healthy” category and keeping French fries, salty snacks, and low-fat salty snacks in “Potatoes” under “less-healthy” category; Orange – “vegetables and unhealthy potatoes” (Model 5): Split “Potatoes” group by moving white potatoes into “Vegetables” group and keeping French fries, salty snacks, and low-fat salty snacks in “Potatoes” under “less-healthy” category.

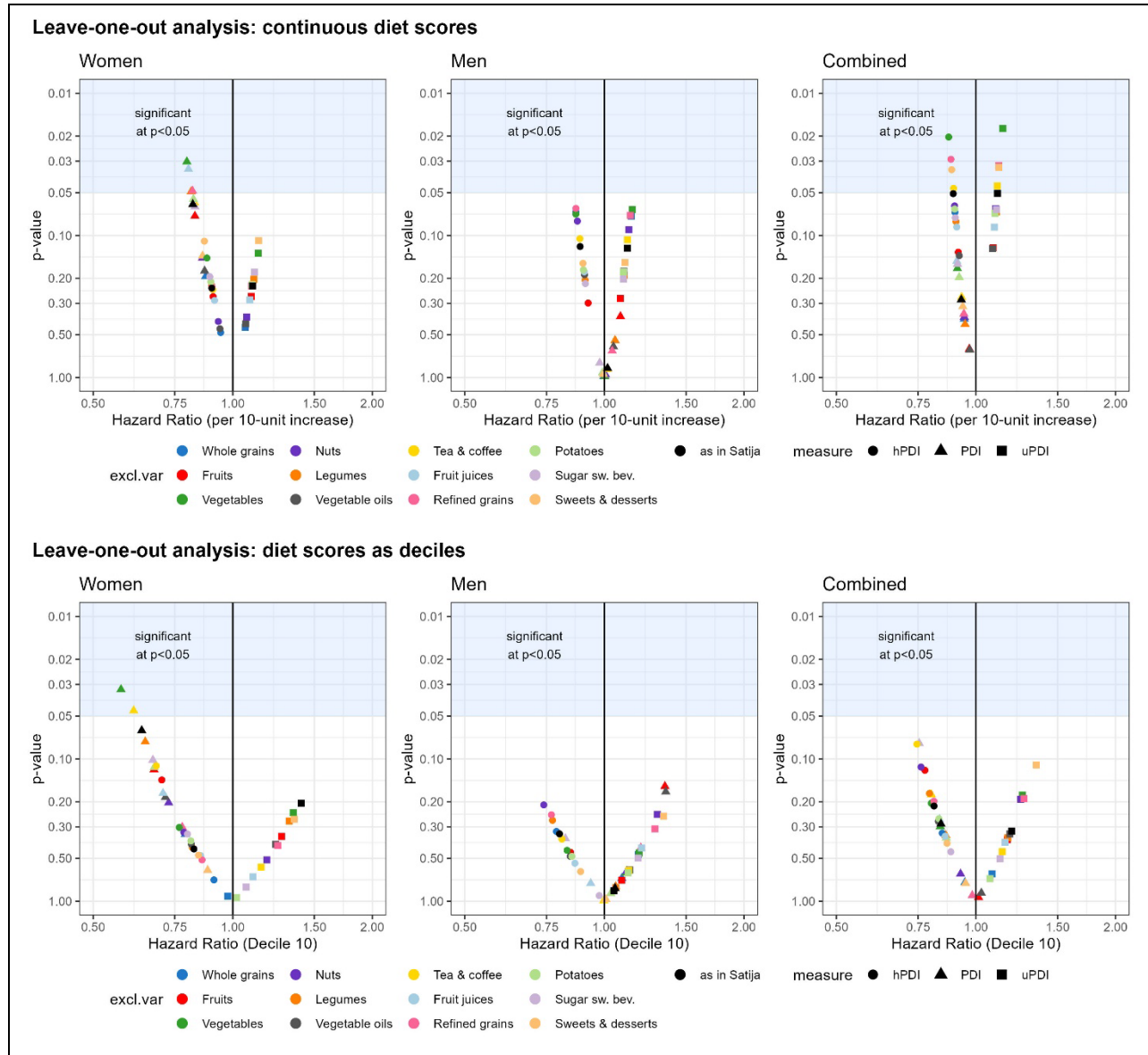


Figure 2: Association of PDI, hPDI, and uPDI indices and CHD for leave-one-out models. Top panel: per 10 units increment using continuous diet scores. Bottom panel: diet scores by decile, shown are HR for highest decile compared with lowest decile.

Agnostic permuted models: continuous diet scores (hPDI)

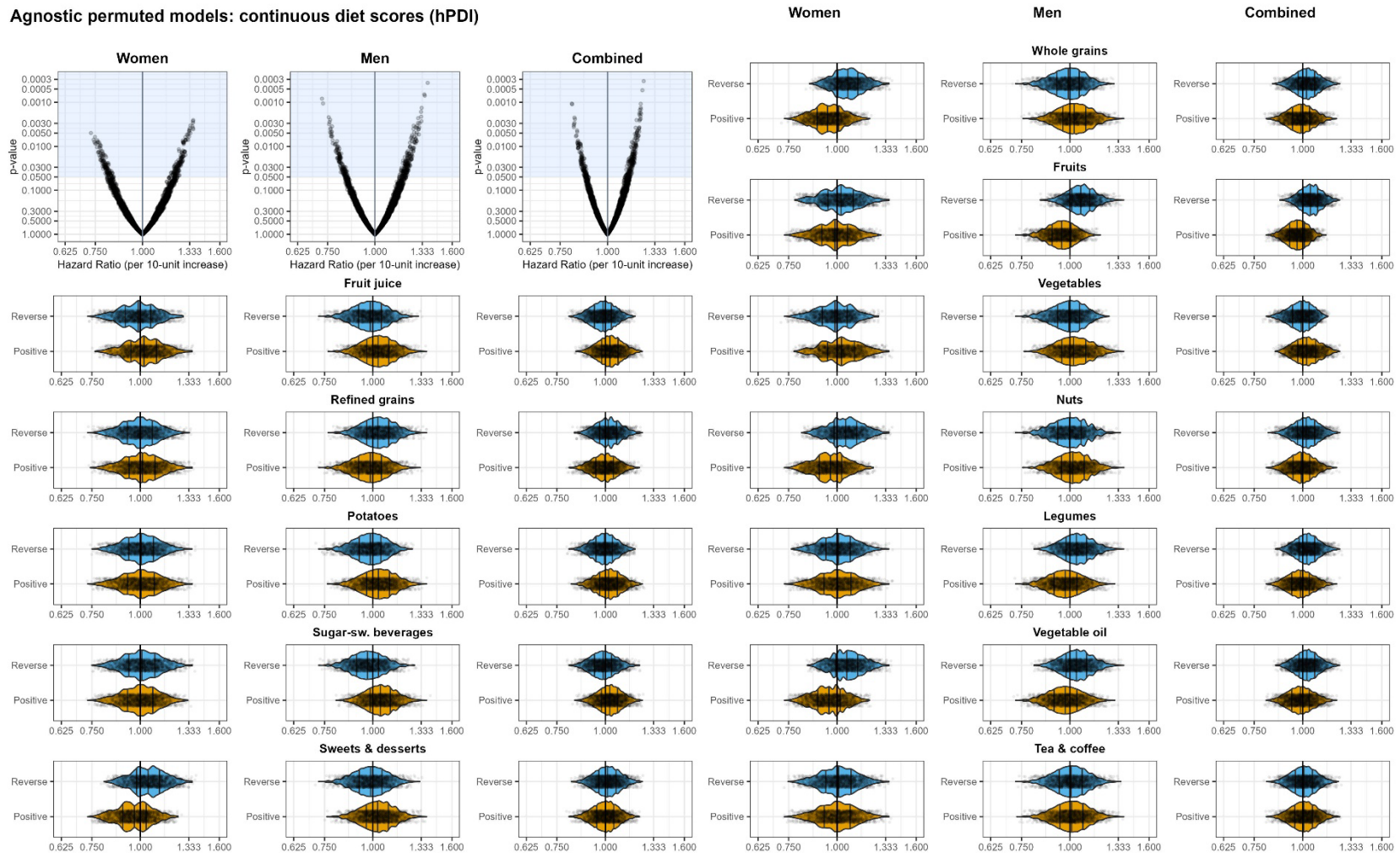


Figure 3: Agnostically permuted models of 12 plant-based food groups. Each of the 12 plant-based food groups are either positively (i.e., more consumption is better; depicted in orange) or reversely (less consumption is better, depicted in blue) scored, using all combinations for a total of 4,096 models. The upper left volcano plots show all 4,096 models. Each panel within sex represents the same set of models but stratified by whether that particular food group is positively or reversely coded. The 5 food groups on the left side were originally coded as ‘unhealthy’ in the scoring approach and the 7 food groups on the right side were coded as ‘healthy’. Comparisons of distributions are shown in Tables 3 and 4.