

Protocol

# Estimating the Effect of Adhering to the Recommendations of the 2019 Canada's Food Guide on Health Outcomes in Older Adults: Protocol for a Target Trial Emulation

Didier Brassard<sup>1</sup>, RD, PhD; Nancy Presse<sup>2</sup>, RD, PhD; Stéphanie Chevalier<sup>1,3</sup>, RD, PhD

<sup>1</sup>School of Human Nutrition, McGill University, Sainte-Anne-de-Bellevue, QC, Canada

<sup>2</sup>Faculty of Medicine and Health Sciences, University of Sherbrooke, Sherbrooke, QC, Canada

<sup>3</sup>Research Institute of the McGill University Health Centre, McGill University, Montreal, QC, Canada

**Corresponding Author:**

Stéphanie Chevalier, RD, PhD

School of Human Nutrition, McGill University

2111 Lakeshore Rd

Sainte-Anne-de-Bellevue, QC, H9X3V9

Canada

Phone: 1 514 398 8603

Email: [stephanie.chevalier@mcgill.ca](mailto:stephanie.chevalier@mcgill.ca)

## Abstract

**Background:** The 2019 Canada's Food Guide provides universal recommendations to individuals aged  $\geq 2$  years. However, the extent to which these recommendations are appropriate for older adults is unknown. Although ideal, conducting a large randomized controlled trial is unrealistic in the short term. An alternative is the target trial emulation framework for causal inference, a novel approach to improve the analysis of observational data.

**Objective:** This study aims to describe the protocol for a target trial emulation in older adults, with an emphasis on key aspects of a hypothetical sustained diet and physical activity intervention.

**Methods:** To emulate the target trial, nonexperimental data from the Quebec Longitudinal Study on Nutrition and Successful Aging (NuAge; N=1753 adults aged  $\geq 67$  years) will be used. NuAge includes 4 yearly measurements of dietary intakes, covariates, and outcomes. The per-protocol causal contrast will be the primary causal contrast of interest to account for nonadherence. The sustained intervention strategy will be modeled using the parametric g-formula. In the hypothetical trial, participants will be instructed to meet sex-specific minimal intakes for vegetables and fruits, whole grains, animal- and plant-based protein foods, milk and plant-based beverages, and unsaturated fats. The eligibility criteria, follow-up, intervention, outcomes, and causal contrast in the emulation will closely align with those of the target trial, with only minor modifications. We will attempt to emulate the randomization of treatment by adjusting for baseline covariates and prebaseline dietary habits.

**Results:** Data collection for NuAge was completed in June 2008. For this study, the main analysis was started in May 2024. Submission of the manuscript is expected by February 2025.

**Conclusions:** Emulating a target trial will provide the first evidence of the adequacy of the 2019 Canada's Food Guide recommendations for older adults in relation to health outcomes.

**International Registered Report Identifier (IRRID):** DERR1-10.2196/65182

(*JMIR Res Protoc* 2025;14:e65182) doi: [10.2196/65182](https://doi.org/10.2196/65182)

**KEYWORDS**

aged; Canada's Food Guide; diet; dietary guidelines; target trial emulation; hypothetical trial; Healthy Eating Food Index-2019; HEFI-2019

## Introduction

### Background

The latest edition of Canada's Food Guide (CFG) was published in 2019 [1]. Compared with the previous edition, key changes include the removal of the prespecified number of servings to consume each day, a shift toward qualitative (eg, "eat plenty of...") instead of quantitative recommendations, and the provision of universal recommendations instead of age- and sex-specific recommendations. In addition, CFG recommendations primarily aim to reduce chronic disease risk. The evidence supporting CFG recommendations focuses on reducing the risk of cardiovascular disease, cancer, and type 2 diabetes risk [2]. However, evidence from a nationally representative survey of adults aged  $\geq 65$  years from Canada suggested that greater adherence to recommendations was insufficient to meet calcium, vitamin D, and folate requirements [3]. In Canada, one-third of community-dwelling older adults are at high nutrition risk [4,5], highlighting the importance of maintaining adequate nutritional status in this stratum of the population. Similarly, in the absence of specific recommendations on the amount of protein foods to eat, older adults may be eating less protein than required to maintain muscle mass [6-8]. CFG also provides brief physical activity recommendations but without explicit acknowledgment of the importance of these recommendations for older adults [1]. Indeed, performing a minimal amount of physical activity is paramount to maintaining muscle mass [6,7]. Thus, the universal recommendations in CFG may not be appropriate for older adults because they face unique challenges in consuming a healthy diet [9] and may require specific nutritional strategies [10].

Ideally, a randomized controlled trial (RCT) should be conducted to investigate the adequacy of CFG recommendations in older adults. However, such an RCT is unlikely to be conducted in the short term. An alternative is the target trial emulation framework for causal effect estimation using observational data [11-14]. Informally, the target trial emulation framework aims to emphasize and resolve design challenges in observational data analysis by explicitly emulating a hypothetical trial to estimate a causal effect [11]. In nutritional epidemiology, common issues with design and analyses can yield results that are largely inconsistent with those of randomized trials [15]. For example, the lack of consideration of the compositional nature of diet can dramatically influence effect estimates, that is, increasing the intake of one food must be compensated by decreasing the intake of another food in substitution modeling [15-17]. Furthermore, diet is a lifelong sustained exposure. In an observational study, the effect of diet assessed at a given time may actually reflect the cumulative exposure to previous dietary habits. In turn, ignoring previous dietary habits may result in a misalignment of "time zero" [18], as dietary habits are not randomly assigned at the beginning of the observational study. In other words, ignoring previous dietary habits makes it impossible to distinguish the effect of prospective or hypothetical dietary modification from the effect of retrospective dietary habits. The target trial framework is a helpful tool to highlight and address common issues in

nutritional epidemiology. Ultimately, a successful emulation of the target trial based on observational data could yield effect estimates that more closely align with those of a hypothetical future RCT. Example applications of the target trial framework include the emulation of interventions on diet [19-22], physical activity [23], or both [24].

### Objectives

To the best of our knowledge, a target trial framework has not been used to assess the effect of adhering to CFG recommendations. Accordingly, this study aims, first, to describe the protocol for the emulation of a target trial using observational data from the Quebec Longitudinal Study on Nutrition and Successful Aging (NuAge [25]), which includes a cohort of 1753 adults aged 67 to 84 years at baseline, and, second, to address key aspects of the target trial emulation in the context of a sustained lifestyle intervention strategy involving diet and physical activity. Key aspects are the description of the sustained lifestyle intervention strategy, the attempt to emulate randomization, as well as assumptions and limitations specific to diet intervention. Notably, more general introductory texts to the target trial framework are available elsewhere [11,12,14].

## Methods

### Research Question and Hypothesis

Explicitly acknowledging the causal nature of a research question is a prerequisite to causal effect estimation using observational data [26-28]. This study aims to examine the adequacy of the universal dietary recommendations provided by CFG for older adults. Expressed as a counterfactual statement, we aim to answer the following question: What would be the difference in a given health outcome at the end of the follow-up if all eligible participants had increased their adherence to CFG recommendations on healthy food choices, compared to if they had maintained their habitual diet?

Specifically, among adults aged 67 to 84 years followed for 3 years, and compared with maintenance of habits, we aim to (1) estimate the causal effect of adhering to CFG dietary recommendations on markers of muscle health (eg, physical function and muscle strength), general health (eg, waist circumference, blood pressure, and glucose), and cognitive health (ie, Modified Mini-Mental State Examination) and (2) estimate the causal effect of adhering to a *reformulation* of CFG dietary recommendations, including more protein foods and a minimal physical activity recommendation, to amplify the positive health effects.

Accordingly, we hypothesize the following:

- *Hypothesis 1: adhering to recommendations positively influences general and cognitive health but does not influence muscle health.*
- *Hypothesis 2: increasing the consumption of protein-rich foods positively influences muscle health, and meeting minimal physical activity recommendations ( $\geq 30$  min per day) further amplifies positive health effects.*

## Study Design and Sample

Data from the NuAge prospective cohort will be used to emulate the target trial [25]. The NuAge cohort comprised 1753 generally healthy community-dwelling adults aged 67 to 84 years at baseline and followed for 3 years. The baseline and each annual follow-up evaluation included a comprehensive assessment of sociodemographic data, diet, physical activity, functional status, as well as physical and mental health status [25].

The NuAge sample is relevant to our research question. The target population of CFG recommendations comprises all individuals aged  $\geq 2$  years, which is compatible with the NuAge target sample of generally healthy older adults from the greater Montreal, Sherbrooke, and Laval areas of the province of Quebec, Canada.

## Target Trial

The target trial framework has been suggested as a potential solution to improve the analysis of nutritional epidemiology

studies aiming at causal inference [13]. Informally, the target trial framework helps to align the observational data analysis with that of a hypothetical trial. This framework is appropriate for the research question in this study, as we aim to estimate the effect of adhering to a hypothetical diet and physical activity intervention using observational data. The first step of a target trial emulation is the description of the target trial, that is, the protocol for the hypothetical RCT we wish to conduct [11,14]. The second step is the emulation, that is, describing how the target trial is emulated and conducting the study described in this protocol.

Table 1 presents the target trial and its emulation using observational data from NuAge. Key differences between the target trial and its emulation are that participants will be required to provide complete dietary assessment and covariate data at baseline (ie, eligibility component) and that we will attempt to emulate the randomized assignment by adjusting for dietary intakes before baseline as well as baseline covariates (ie, assignment component). Each part of the target trial and its emulation are described in the subsequent sections.

**Table 1.** Emulation of a dietary intervention target trial using observational data from the Quebec Longitudinal Study on Nutrition and Successful Aging (NuAge [25]).

Trial component	Target trial specification	Target trial emulation
Eligibility criteria	<ul style="list-style-type: none"> <li>NuAge inclusion criteria: individuals aged 67-84 years; living in Montreal, Laval, or Sherbrooke; not cognitively impaired; and free of disabilities in activities of daily living</li> <li>Exclusion criteria: class II heart failure, chronic obstructive pulmonary disease requiring home oxygen therapy or oral steroids, inflammatory digestive diseases, and cancer treatment in the past 5 years</li> </ul>	<ul style="list-style-type: none"> <li>The inclusion and exclusion criteria are the same as those specified in the target trial.</li> <li>Furthermore, participants will be required to have complete baseline dietary assessment (at least one 24-hour recall with <math>\geq 500</math> kcal and FFQ<sup>a</sup>) and provide baseline covariate data.</li> </ul>
Interventions <sup>b</sup>	<ul style="list-style-type: none"> <li>Each individual would be assigned to 1 of the 4 following strategies: <ul style="list-style-type: none"> <li>Control group (habitual diet, ie, typical North American diet)</li> <li>Adherence to Canada's Food Guide recommendations on healthy food choices</li> <li>Adherence to Canada's Food Guide recommendations on healthy food choices and including a high-protein reformulation</li> <li>Adherence to Canada's Food Guide recommendations on healthy food choices, including a high-protein reformulation and a minimal physical activity component</li> </ul> </li> <li>Each strategy is followed until the end of follow-up. Participants assigned to a lifestyle strategy are expected to maintain their dietary intake or amount of physical activity at or above the prespecified threshold by the corresponding intervention strategy.</li> </ul>	<ul style="list-style-type: none"> <li>The intervention component will be the same as that specified in the target trial.</li> <li>Furthermore, we will assume that each dietary assessment period (ie, within 2 months beginning at each time point) accurately reflects the average diet in the interval between follow-ups.</li> </ul>
Assignment	<ul style="list-style-type: none"> <li>Participants are randomly assigned to a dietary strategy but are not blinded to their assignment.</li> </ul>	<ul style="list-style-type: none"> <li>We will attempt to emulate randomized assignment by adjusting for dietary intakes before baseline and baseline covariates.</li> </ul>
Outcomes	<ul style="list-style-type: none"> <li>The outcomes are physical function and muscle strength, general health indicators, and cognitive health.</li> </ul>	<ul style="list-style-type: none"> <li>The outcomes will be physical function and muscle strength, general health indicators, and cognitive health.</li> </ul>
Time zero and follow-up	<ul style="list-style-type: none"> <li>The study starts at baseline and ends at incomplete follow-up or 3 years after baseline, whichever occurs first.</li> </ul>	<ul style="list-style-type: none"> <li>The study will start at baseline and end at incomplete follow-up or 3 years after baseline, whichever occurs first.</li> <li>An incomplete follow-up is defined as missing data for questionnaires (nonresponse or loss to follow-up) or missing outcome data at the end of follow-up.</li> </ul>
Causal contrast <sup>c</sup>	<ul style="list-style-type: none"> <li>Intention-to-treat effect</li> <li>Per-protocol effect</li> </ul>	<ul style="list-style-type: none"> <li>Observational analog of both contrasts: <ul style="list-style-type: none"> <li>Secondary: intention-to-treat effect</li> <li>Primary: per-protocol effect</li> </ul> </li> </ul>
Statistical analysis	<ul style="list-style-type: none"> <li>Intention-to-treat analysis: apply inverse probability weighting with adjustment for prebaseline and baseline factors associated with incomplete follow-up to account for study dropouts</li> <li>Per-protocol analysis: apply the parametric g-formula algorithm to compare postintervention outcomes between groups receiving each treatment strategy, adjusting for pre- and postbaseline factors associated with adherence to intervention strategies and incomplete follow-up.</li> </ul>	<ul style="list-style-type: none"> <li>Analysis will be the same as that specified in the target trial for both contrasts.</li> <li>However, the observational analog will require additional adjustments for confounding at baseline and before baseline due to previous dietary pattern.</li> </ul>

<sup>a</sup>FFQ: food-frequency questionnaire.<sup>b</sup>Refer to the *Hypothetical Interventions* section for detailed intervention.<sup>c</sup>The observational analog of the intention-to-treat contrast corresponds to the baseline values of the intervention, which are assigned and initiated at the same time.

## Eligibility Criteria

The inclusion criteria for the target trial are the same as those for NuAge [25]. In the emulation, participants will be required to have at least one 24-hour dietary recall completed at baseline with  $\geq 500$  calories, as well as complete covariate data at baseline, as identified in the *Assignment* section.

## Hypothetical Interventions

### Overview

The hypothetical intervention strategies evaluated will be as follows:

1. No change in dietary habits or physical activity (similar to a control intervention)
2. Adherence to CFG recommendations
3. Adherence to CFG recommendations, including reformulation (ie, higher intake of protein foods)
4. Adherence to CFG recommendations, including reformulation (ie, higher intake of protein foods) and performing at least 30 minutes of aerobic physical activity

Physical activity recommendations are not traditionally at the forefront of CFG recommendations. However, CFG does mention that “at least 150 minutes of moderate-to-vigorous-intensity aerobic physical activity per week ... is recommended to achieve health benefits” [1]. Thus, recognizing the key role of exercise in maintaining health and muscle for older adults, the fourth hypothetical intervention includes a formal physical activity recommendation. In the target trial, the physical activity corresponds to performing aerobic exercise of light to vigorous intensity for at least 30 minutes per day [6].

### The Challenge of a Well-Defined Nutritional Intervention

Emulating a well-defined dietary intervention for CFG recommendations is challenging. First, recommendations in the latest edition of CFG are qualitative and flexible (eg, “eat plenty of vegetables and fruits” [1,29]). Thus, various suitable yet distinct dietary patterns can align with the recommendations. Second, CFG recommendations target both food intakes (eg, vegetables and fruits) and nutrients (eg, saturated fats). The nutrient-based recommendations can be met by modifying the consumption of various foods. For example, to achieve the hypothetical intervention of “decreasing consumption of calories from saturated fats,” one could decrease saturated fats from dairy, nuts, and low nutritive value foods altogether. Arguably, the relationship between these food categories and a given health outcome may vary greatly.

To estimate a causal effect using observational data, the hypothetical interventions must be clearly defined to the point where “no meaningful variation” in the intervention remains [30,31]. In other words, the hypothetical diet interventions should be elaborated until no additional dietary characteristics are deemed impactful regarding the outcome of interest. Another consideration is that the modeling of hypothetical interventions

should ideally be conducted with dietary intakes expressed using the same units. For example, mixing food intakes expressed in servings and grams in a statistical model may cause poor estimation of causal effects [17]. Finally, the statistical approach used to account for total energy or total food intake also affects the causal effect of interest and should be consistent with the research question [17,31-33].

### Diet Simulations

For this study, adherence to CFG recommendations was defined based on simulated diets generated by Health Canada [34] and summarized in [Multimedia Appendix 1](#) [34]. The simulated diets were designed to meet CFG recommendations on healthy food choices and nutrient requirements (Dietary Reference Intake).

These diets achieve near-perfect Healthy Eating Food Index (HEFI)-2019 scores ( $>78/80$ ) through relatively high intake of recommended foods (ie, vegetables and fruits, whole grain foods, protein foods, and unsweetened milk and plant-based beverages with protein) and null intakes of foods not recommended (ie, non-whole grain foods; other low nutritive value foods; juice, sugary drinks and alcohol; and fatty foods rich in saturated fats). The HEFI-2019 score indicates the extent to which dietary intakes are consistent with CFG recommendations on healthy food choices [29,35].

Notably, the HEFI-2019 could have been used as a main exposure to measure adherence to CFG. However, the use of a composite score metric would not completely satisfy the criterion of a well-defined intervention to estimate a causal effect. First, high HEFI-2019 scores and high adherence to CFG recommendations can be achieved through many different strategies or dietary patterns. In the context of observational data, the specific strategies through which individuals achieve a high HEFI-2019 score would be based on dietary habits and patterns self-selected by the participants. This approach is similar to asking hypothetical trial participants to modify their intakes without clearly specifying how, which would obscure the estimated causal effect. Second, the HEFI-2019 score includes recommendations on foods and nutrients. As described earlier, mixing servings and grams in statistical models may cause poor estimation of causal effects [17].

We stress that the diets simulated by Health Canada were not actually consumed by older adults. Therefore, the simulated values for vegetables and fruits, whole grain foods, and plant-based protein foods *exceed* the 99th percentile of the distribution of usual intakes of these food categories, as estimated in adults aged  $\geq 65$  years from the Canadian Community Health Survey 2015–Nutrition [3]. In [Table 2](#), the target intakes for vegetables and fruits, whole grains, and plant-based protein foods in the adherence to the 2019 CFG recommendations intervention were revised to correspond, at most, to the 90th percentile of the distribution of usual intakes among Canadians aged  $\geq 65$  years in 2015 [3].

**Table 2.** Emulation<sup>a</sup> of hypothetical diet and exercise interventions by sex in the Quebec Longitudinal Study on Nutrition and Successful Aging (NuAge) cohort [25].

Sex	Recommended foods (RA <sup>b</sup> /d)						Dietary supplement <sup>c</sup>	Foods and beverages not recommended (RA/d) <sup>d</sup>			Physical activity (min/d) <sup>e</sup>	
	Vegetables and fruits	Whole grains	Plant-based protein foods	Animal-based protein foods	Milk and plant-based beverages with protein	Unsaturated oils and fats		Other foods	Sugary drinks and alcohol	Non-whole grains		
<b>Control (no change)<sup>f</sup></b>												
Male individuals	— <sup>g</sup>	—	—	—	—	—	—	—	—	—	—	—
Female individuals	—	—	—	—	—	—	—	—	—	—	—	—
<b>Adhering to the 2019 Canada's Food Guide recommendations on healthy food choices<sup>h</sup></b>												
Male individuals	6	1.5	1.0	2.0	1.0	1	No change	Minimum	Minimum	Minimum	No change	
Female individuals	5	1.5	0.8	1.5	1.0	1	No change	Minimum	Minimum	Minimum	No change	
<b>Adhering to the 2019 Canada's Food Guide recommendations on healthy food choices, including extra protein<sup>i</sup></b>												
Male individuals	6	1.5	1.5	3.5	1.5	1	No change	Minimum	Minimum	Minimum	No change	
Female individuals	5	1.5	1.3	3.0	1.5	1	No change	Minimum	Minimum	Minimum	No change	
<b>Adhering to the 2019 Canada's Food Guide recommendations on healthy food choices, including extra protein and physical activity</b>												
Male individuals	6	1.5	1.5	3.5	1.5	1	No change	Minimum	Minimum	Minimum	30 or more	
Female individuals	5	1.5	1.3	3.0	1.5	1	No change	Minimum	Minimum	Minimum	30 or more	

<sup>a</sup>The emulation of all hypothetical interventions will be implemented using a substitution approach in statistical models. In all models, 1 variable for total food intake and 1 variable for total beverage intake will be included, and foods not recommended will be left out from the models (ie, non-whole grain foods; other low nutritive value foods; juice, sugary drinks and alcohol; and fatty foods rich in saturated fats).

<sup>b</sup>RA: reference amount.

<sup>c</sup>Dietary supplements were not intervened on but were, nonetheless, excluded from foods and beverages not recommended to avoid being considered in the substitution. In other words, participants would not be instructed to modify their dietary supplements in the hypothetical trial.

<sup>d</sup>Minimum indicates that consumption would be set at the smallest amount, permitting a concomitant increase in recommended foods to meet Canada's Food Guide targets. Portions for foods not recommended may vary on an individual basis.

<sup>e</sup>Physical activity corresponds to aerobic exercise of moderate or higher intensity [6].

<sup>f</sup>Values will be the averages observed at baseline in the NuAge cohort. In other words, values will be the observed intakes for the food categories or amount of physical activity when no change is applied.

<sup>g</sup>Not applicable.

<sup>h</sup>Values are derived from Health Canada's simulated composite diets of adults aged  $\geq 71$  years. Participants would be expected to meet these targets for each food category. The specific food choices within these categories would be at the participants' discretion. Values for vegetables and fruits, whole grain foods, and plant-based protein foods were truncated to correspond, at most, to the 90th percentile of the distribution of usual intakes among Canadians aged 65 years in 2015.

<sup>i</sup>Extra protein foods were added as follows: +0.5 RA of plant-based protein foods (eg, 25 g of nuts), +1.5 RA of animal-based protein foods (eg, 150 g of cooked unprocessed red meat, fish, or poultry or 3 small eggs), +0.5 RA of milk or plant-based beverage with protein (eg, 125 mL of milk or plant-based beverages with sufficient protein).

Because the 2019 CFG does not have a portion size system, reference amounts (RAs) were used as a proxy for servings. RAs are regulated quantities of foods that reflect the portion size typically consumed at 1 sitting in Canada. RAs were used by Health Canada to simulate a diet consistent with the 2019

CFG recommendations and dietary reference intake ([Multimedia Appendix 1 \[34\]](#)); therefore, RAs are adequate for this study.

## Implementation

In the target trial, the sustained intervention strategy could be implemented as follows:

- Step 1: the participants' usual dietary intake and physical activity would be assessed by research dietitians at each study visit.
- Step 2: if reported food intakes and duration of physical activity were equal to or above the prespecified thresholds (Table 2), no change would be suggested to the participants' diet or physical activity. If food intakes or duration of physical activity were below the prespecified thresholds, participants would be instructed to increase food consumption to exactly the prespecified portions or increase physical activity duration to 30 minutes per day (when applicable).
- Step 3: if changes are required, participants would be instructed to decrease consumption of foods not recommended by the same amount as the increase in step 2. For example, if a 2-serving increase in vegetables and fruits is required to meet the prespecified intervention thresholds, participants would be instructed to substitute 2 servings of vegetables and fruits for non-whole grain foods; other low nutritive value foods; juice, sugary drinks and alcohol; and fatty foods rich in saturated fats.

In the emulation, for all hypothetical interventions, substitution will be implemented by including total intakes as a covariate and excluding foods not recommended from the models. More precisely, a variable reflecting total food intake (in RA/d) and a variable reflecting total beverage intake (in RA/d) will be included in all models. Hence, total food and beverage intakes will be constant across hypothetical diet interventions. In this approach to account for total energy, all model coefficients reflect the action of increasing intakes of recommended foods and a concomitant decrease in *any* of the foods not recommended [32]. On one hand, this approach can be potentially confusing [17,33], as the default interpretation of model coefficients assumes increasing the intake of each *food included in the model* while simultaneously decreasing the intake of foods *excluded from the model* [32]. On the other hand, the standard model is generally consistent with the implementation of dietary intervention in feeding trials [16,36,37]. The standard model also reduces the number of variables to be considered as intervention variables. Otherwise, 4 additional dietary components would have to be modeled for foods not recommended (ie, non-whole grain foods; other low nutritive value foods; juice, sugary drinks and alcohol; and fatty foods rich in saturated fats). Finally, the explicit description of the intervention strategies in the target trial protocol clarifies the estimand of interest, as done in an earlier study [19,31].

Notably, nutrient-based recommendations in CFG (ie, saturated fats, free sugars, and sodium intake) are not explicitly modeled to avoid the problems associated with mixed-unit models [17]. In the target trial, we assume that nutrient-based targets would be met by reducing consumption of foods not recommended (ie, non-whole grain foods; other low nutritive value foods;

juice, sugary drinks, and alcohol; and fatty foods rich in saturated fats). In that regard, food-level substitution analyses in Canadians support this assumption for saturated fats [38,39].

The extra protein intervention consists of increasing the intake of both animal-based and plant-based protein foods by 1.5 and 0.5 RA per day, respectively, as well as milk and plant-based beverages with sufficient protein by 0.5 RA per day. Regarding the amount of food, this corresponds to adding 150 g of cooked unprocessed red meat, fish, or poultry or 3 small eggs; 25 g of nuts and seeds; and 125 mL of milk or plant-based beverages with sufficient protein while proportionally decreasing the intake of foods not recommended. In a previous RCT [40], older women aged 60 to 90 years were able to consume an additional 160 g of cooked lean red meat without substitution, thereby supporting the feasibility of the protein intervention in this hypothetical study.

## Assignment

We will attempt to emulate random allocation or randomization by adjusting for dietary intakes in the year before the intervention, as well as adjusting for covariates at the start of the study. Covariates were identified using the causal diagrams depicted in Figure 1, based on background knowledge of the relationship between the hypothetical lifestyle intervention and outcomes.

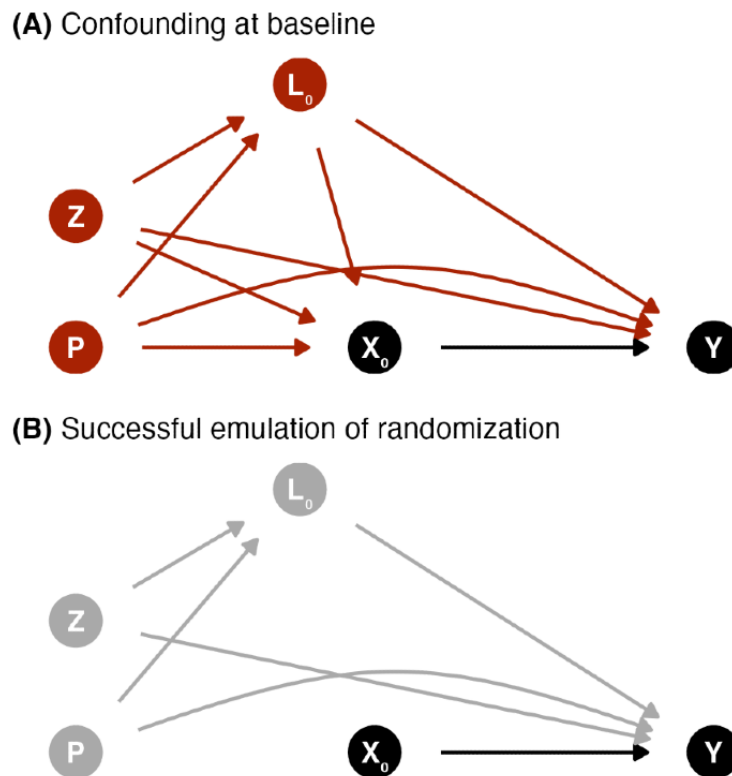
Dietary components that are the foundation for healthy eating in CFG include intakes of vegetables and fruits, whole grains, protein foods (plant- and animal-based protein foods, milk, and plant-based beverages with protein), and unsaturated oils and fats [1].

Covariates, including age at baseline, biological sex, region, education, living alone, smoking and drinking (alcohol) habits, major chronic diseases (ie, hypertension, diabetes, cancer, and heart disease), the number of medications, supplement use (eg, vitamins and minerals), and height and weight, are as follows:

1. Z, baseline covariates: age, sex, region, education, history of smoking, height, and former cancer history
2. P, previous exposure (ie, exposure of time-varying intervention before baseline): dietary habits before baseline
3. L, (time-varying) covariates: weight, number of medications, supplement use, living alone, major chronic diseases (eg, hypertension, diabetes, cancer, and heart disease), and smoking and alcohol habits
4. X, (time-varying) treatment: diet and physical activity habits
5. Y, end of follow-up outcome: muscle health, general health, and cognitive health

Contrary to dietary habits, data on physical activity habits before baseline were not collected in NuAge. In this case, the potential effect of previous physical activity habits will not be accounted for in the models that aim to emulate the sustained physical activity intervention strategy. For the models emulating the sustained diet intervention strategy only, physical activity habits during the study will be used as a covariate, hence mitigating the confounding of previous physical activity, at least to some extent.

**Figure 1.** Causal directed acyclic graph (DAG) depicting (A) confounding and (B) successful emulation of randomization using g-methods at baseline between the intervention strategy (X) and outcome (Y). Baseline covariates (both time-invariant [Z] and time-varying [L]) and previous diet and physical activity habits (P) must be considered to emulate randomization. Time-varying treatment and covariates are not shown in this DAG to focus on randomization emulation. Subscripts indicate the time points, where 0 represents baseline.



A successful emulation of randomization requires that there is no unmeasured confounding. However, this is never guaranteed with observational data. Thus, we emphasize our assumptions that (1) the causal graph accurately depicts the relationship under study and (2) the covariates included are a sufficient set of covariates to address confounding.

### Outcomes

The primary outcomes will be the mean end of follow-up values for muscle strength (ie, handgrip using vigorimeter, elbow flexor, and knee extensor) and physical function (ie, normal and fast walking and “timed up-and-go”). Outcome values were measured according to a standardized protocol in NuAge [25].

For secondary outcomes, mean end of follow-up values for a set of relevant variables will be considered by domains:

- *General health:* waist circumference, blood pressure (systolic and diastolic), blood glucose, and estimated glomerular filtration rate
- *Cognition:* the modified Mini-Mental State Examination score

### Time Zero and Follow-Up

In the target trial of a sustained lifestyle intervention, participants would be met at baseline and then regularly to ensure that diet and physical activity habits are consistent with the intervention assigned by the random allocation. The hypothetical diet and physical activity intervention would be assigned and initiated at baseline. In the emulation, annual follow-ups with comprehensive diet, physical activity, and covariate data collection are available to emulate the

hypothetical intervention. Hence, participants would be followed from the study baseline (time point 0; ie, the time at which the intervention strategy would also be assigned and would begin), at each year (ie, time points 1 and 2), and until the end of the study (time point 3). We also assume that the diet and physical activity habits measured at each follow-up time adequately reflect the habits during the entire year.

The end of follow-up outcome measurements will be used to estimate the effect of the sustained lifestyle intervention strategy. Measures of dietary intake and physical activity throughout the study (ie, time point 0 to time point 3) will be used to emulate the sustained lifestyle intervention. Dietary intakes in the year preceding the intervention will be estimated using the frequency questionnaire completed at baseline (ie, time point 0). Missing covariate data at a given follow-up will be carried forward once, after which participants will be considered as having incomplete follow-up.

### Causal Contrast

The estimand of interest in this study, the target causal effect of a sustained lifestyle intervention strategy, is

$$E(Y^{1,1,1,1}|C=0) - E(Y^{0,0,0,0}|C=0) \quad (1)$$

that is, the expected value of a given health outcome  $Y$  at the end of follow-up if all participants had increased their adherence to CFG recommendations on healthy food choices and physical activity, when applicable, at all 4 time points ( $X_k=1$ , *always intervene*) versus if all participants had maintained their habitual diet and physical activity ( $X_k=0$ , *never intervene*). The estimand



(equation 1) also indicates that all participants completed the intervention ( $C=0$ ), that is, in the absence of incomplete follow-up.

The causal contrasts of interest are the observational analogs of intention-to-treat and per-protocol analyses [11]. Given the observational design, participants are not expected to have followed a treatment strategy unknown to them at the time of data collection. Therefore, the primary analysis will be the per-protocol contrast of a sustained lifestyle intervention strategy. In the per-protocol analysis, nonadherence to the hypothetical interventions can be accounted for. In the target trial, participants with a condition after baseline that would have prevented or limited participation in a hypothetical lifestyle intervention would be allowed to discontinue the intervention (eg, lengthy hospitalization, prolonged bed rest, and incident cancer). In the emulation, if such conditions occur in a sufficiently large number of participants, these participants will be “excused” from following the hypothetical intervention [23]. In other words, participants who would have been unable to pursue the study due to major events will not be considered as having incomplete follow-up if they attended the annual assessment. Allowing participants to discontinue adhering to the hypothetical intervention strategy mitigates confounding by the disease burden [23].

The intention-to-treat analysis will be a secondary analysis of a hypothetical point intervention, for example, dietary counseling at baseline only. Notably, it will not be possible to conduct an intention-to-treat analysis identical to that of a controlled study where the interest is to estimate the effect of being assigned to an intervention [11]. However, it is possible to conduct an observational analog of the intention-to-treat analysis. In the observational analog, the intention-to-treat analysis aims to estimate the impact of a hypothetical intervention in which adherence is measured at baseline only.

## Statistical Analysis

### Modeling of Hypothetical Interventions

Stratification and multivariable regression (ie, covariate adjustment) are conventional statistical approaches to address confounding in nutritional epidemiology. However, the conventional approaches are not adequate to estimate cumulative treatment effects (eg, diet over time) in the presence of time-varying confounding (eg, weight status over time) and treatment (eg, previous diet) [41,42]. In this study, nonadherence to the hypothetical interventions and incomplete follow-up will be considered using general g-methods for the per-protocol analysis [11,42]. Among g-methods, the parametric g-formula provides the most flexibility for analyses involving hypothetical dietary interventions, as used in previous studies [19,22,43]. Briefly, in the context of an observational study, the parametric

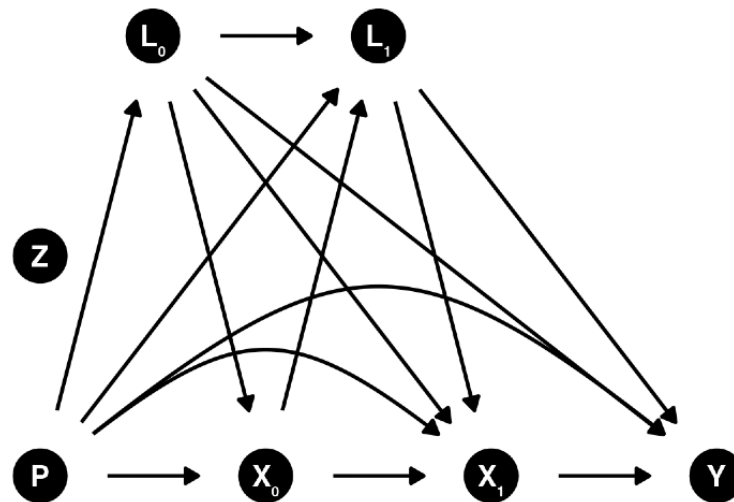
g-formula and its implementation into an R package (R Foundation for Statistical Computing) [44] use parametric models to predict the joint history of previous diet and physical activity habits (ie, the hypothetical sustained intervention strategy) and confounding variables. For example, linear regression models are used to predict continuous covariates (eg, body weight), while logistic regressions are used to predict binary or categorical variables (eg, indicator variable for dietary supplement use). The per-protocol causal contrast of the hypothetical intervention presented in Table 2 is then emulated based on Monte Carlo simulated data generated using the g-formula algorithm [44]. The parametric g-formula correctly accounts for time-varying confounding in the presence of feedback between the intervention and the confounding variables, as confounding is addressed using standardization [41,42]. Furthermore, standardization allows for estimating an average causal effect (ie, marginal effect) consistent with the estimand of interest (equation 1) rather than a conditional effect. In summary, “threshold interventions” that depend on the reported dietary intakes or amount of physical activity [43,45] and the parametric g-formula [44,46] will be used to emulate the intervention of *consuming at least x servings of food and doing at least x minutes of light to vigorous physical activity*.

Figure 2 presents the causal directed acyclic graph (DAG) of the hypothesized relationship between a sustained lifestyle intervention strategy ( $X_0, X_1$ ) and an end of follow-up outcome  $Y$  for 1 follow-up after baseline (year 1). The model is limited to year 1 for clarity, but the hypothesized causal structure extends to additional follow-ups. The exposure of interest  $X_k$  is the joint and cumulative effect of a sustained diet and physical activity intervention strategy measured at baseline and follow-ups.

In the context of this target trial emulation (Figure 2),  $P$  includes dietary habits before the baseline assessment.  $P$  can have an effect on baseline dietary habits (eg, previous healthy habits increase the likelihood of baseline healthy habits) and dietary habits throughout the target trial emulation (eg, previous healthy habits increase the likelihood of adhering to healthy habits). Furthermore,  $P$  influences baseline and time-varying confounding. Finally, given the long-term effect of chronic exposure,  $P$  potentially affects  $Y$  directly.

The intention-to-treat analysis is similar to the per-protocol analysis. However, only baseline diet and physical activity habits and covariates are considered, as well as prebaseline diet and physical activity. In both the per-protocol and intention-to-treat analyses, loss to follow-up (eg, nonresponse or missing follow-up and health outcomes not measured) will be accounted for using g-methods such as inverse probability weighting.

**Figure 2.** Directed acyclic graph depicting the hypothesized relationship among previous dietary exposure, time-varying interventions (diet and physical activity), and covariates. Arrows from the Z node are not shown for clarity but would point toward all baseline and time-varying nodes, as well as the outcome. Only 1 follow-up is shown for visualization purposes, but the hypothesized causal structure extends to additional follow-ups. Subscripts indicate the time points, where 0 represents baseline and 1 represents time point 1.



### Dietary Assessment

In NuAge, diet during the year before baseline was assessed using 1 semiquantitative food-frequency questionnaire [47]. Dietary intakes at baseline (ie, time point 0) and each annual follow-up (ie, time points 1, 2, and 3) were assessed using 3 repeated face-to-face interviewer-administered 24-hour dietary recalls.

Dietary intakes measured using 24-hour dietary recall are more accurate (ie, have less systematic error or bias) than food-frequency questionnaire [48-50] but are particularly affected by random measurement error (ie, within-individual random error) [50]. When several variables measured with errors are considered simultaneously in a regression model, the regression coefficients may be biased in any direction [51]. To account for random measurement error, the National Cancer Institute Markov Chain Monte Carlo (NCI MCMC) multivariate method could be applied [52]. However, the combination of the parametric g-formula and multivariate measurement error correction using the NCI MCMC method is not feasible. The NCI MCMC method estimates time-invariant measurement error-corrected intakes, while the g-formula algorithm is designed for time-varying exposures.

Recognizing the importance of accounting for measurement error, the correction for measurement error will be reserved for the secondary intention-to-treat analysis. For the intention-to-treat contrast, the time-varying values of exposure and the time-varying values of confounding are not considered. Thus, the NCI MCMC method will be used to obtain measurement error-corrected estimates of the relationship between dietary intakes measured at baseline and outcome at the end of follow-up. Notably, the three 24-hour dietary recalls collected at each time point contribute to reducing random errors, at least to some extent, even in the absence of measurement error correction.

### Sensitivity Analysis to Assess the Impact of Measurement Error

For the secondary intention-to-treat analysis, results based on the measurement error-corrected and -uncorrected dietary intakes will be compared. The difference between the estimated relationships will allow to extrapolate the impact of not accounting for random errors in the primary *per-protocol* analysis.

### Physical Activity Assessment

Physical activity throughout the study will be assessed using the Physical Activity Scale for the Elderly questionnaire [53,54]. Rather than the total Physical Activity Scale for the Elderly score, specific questions estimating the total time of physical activities will be used to be consistent with the intervention strategy.

### Covariates and Subgroups

To the extent permitted by the number of observations for each outcome, continuous covariates will be modeled using restricted cubic splines with 3 to 5 knots (at percentiles 10-50-90, 5-35-65-95, or 5-27.5-50-77.5-95) [55]. Categorical covariates will be modeled to ensure a sufficient sample size at each level.

The effect of dietary changes will be estimated for the entire sample. The sample will also be stratified by biological sex to reflect both potential biological differences and, to some extent, gender differences (although not reported).

### Variance Estimation

The variance will be estimated using a minimum of 200 bootstrap sample replicates to consider uncertainty at each step of the estimation [56].

### Software and Code

The main statistical analyses will be conducted using R software (version 4.3.1 or greater) and the *gfoRmula* package [44,57]. The manuscript results will be generated using Quarto markdown (Posit PBC). Codes for main analyses and generation

of manuscript results will be shared in a publicly available code repository.

## Ethical Considerations

The original NuAge protocol was approved by the research ethics boards of the *Institut universitaire de gériatrie de Montréal* and the *Institut universitaire de gériatrie de Sherbrooke* (Quebec, Canada). The NuAge Database and Biobank [58] has received approval by the *Centre intégré universitaire de santé et de services sociaux de l'Estrie—Centre hospitalier universitaire de Sherbrooke* Research Ethics Board. Secondary analyses of data from the NuAge Database and Biobank for the study described in this protocol are approved by the McGill University Research Ethics Board Office (#22-11-041).

All participants of NuAge provided informed consent. From the initial cohort of 1793 participants, 1753 (97.77%) agreed to the integration of their data and biological samples into the NuAge Database and Biobank for future studies.

Secondary analyses based on the NuAge Database and Biobank use deidentified data, which do not allow participants to be identified by the investigators.

NuAge participants voluntarily consented to participate and were not provided with monetary compensation.

## Results

Data collection for NuAge was completed in June 2008. For this study, the main analysis based on the final curated data started in May 2024. The manuscript will be written according to the Strengthening the Reporting of Observational Studies in Epidemiology statement. We anticipate the submission of the manuscript to a peer-reviewed academic journal by February 2025.

## Discussion

### Principal Findings

In this study protocol, we have described a target trial to assess the effect of adhering to CFG recommendations on healthy food choices. The emulation will be performed using data from the NuAge Database and Biobank [58]. Benefiting from the flexibility of observational data, we also aim to compare adherence to multiple reformulations of CFG recommendations, including the effect of increasing the intake of protein-rich foods and the amount of aerobic physical activity on selected health outcomes. Furthermore, we have outlined the rationale for using simulated diets to emulate adherence to CFG recommendations, the process of selecting covariates to attempt to emulate randomization with causal diagrams, and the challenges of addressing random measurement error.

We emphasize that the purpose of emulating a target trial using observational data is to improve the quality of observational analysis [11,14]. In other words, the target trial framework aims to support the coherence between the causal research question and the observational data analysis [26]. However, estimating causal effects with nonexperimental observational data depends

on strong assumptions. The key assumptions are that there are no unmeasured confounders, no measurement errors, and no model misspecifications (eg, functional form of covariates and model outcome distribution) [19]. We first recognize that the absence of residual or unmeasured confounding cannot be guaranteed. The extent to which this assumption is sufficiently satisfied depends on the appraisal of covariates considered. In that regard, we have used graphical tools, DAG, to explicitly describe our analytical assumptions and to identify confounding variables [59-61]. Second, the absence of measurement error assumption will not be satisfied considering the use of dietary intake data measured with 24-hour dietary recalls. On one hand, 24-hour dietary recalls have the least systematic error or bias compared with other common instruments, such as food-frequency questionnaires [48,49]. On the other hand, 24-hour dietary recalls are largely affected by within-individual random errors [50], which can cause bias in any direction in multivariable models [51], as in this study. This issue is mitigated, at least to some extent, by using average data from 3 repeated 24-hour dietary recalls at each follow-up. Sensitivity analyses comparing estimates based on measurement error-corrected and -uncorrected baseline dietary intakes will be used to assess the impact of random measurement errors. Third, the absence of model misspecification will be assessed by examining differences between the observed value of time-varying covariates and the predicted value of time-varying covariates as modeled with the g-formula.

### Limitations

The strengths of this study and protocol include the explicit emulation of a hypothetical trial, the thorough description of the emulation of the sustained dietary intervention, and the use of background knowledge and DAG to derive a sufficient set of confounders. Limitations must be addressed. First, the NuAge sample size is relatively limited (N=1753), although comprehensive nutrition and covariate data were collected. Second, the target food intakes based on diet simulations from Health Canada exceeded the 99th percentile of the usual intake distribution of Canadians aged  $\geq 65$  years from Canada in 2015 [3]. A revision of the dietary intervention targets may be needed if observed dietary intakes in NuAge deviate significantly from targets (Table 2), as was done in a previous nutrition target trial emulation [22]. Third, the presence of random measurement error associated with 24-hour dietary recalls may bias estimates. Finally, the target trial emulation cannot replace an actual RCT. Evidence from an RCT will be required to confirm the value of either CFG recommendations or the enhanced CFG recommendations.

### Conclusions

In conclusion, the target trial framework is useful for estimating the causal effect of adhering to CFG recommendations using nonexperimental data when an RCT is impractical [11,19]. Coupled with key assumptions, including the absence of unmeasured confounding, the absence of measurement error, and no model misspecification, we believe that the emulation will provide timely evidence regarding the effect of adhering to CFG recommendations in older adults and inform on the added value of a reformulation.

## Acknowledgments

This work was supported by a Canadian Institutes of Health Research Fellowship award (MFE-181852) granted to DB.

## Data Availability

The datasets analyzed during this study are available in the GitHub repository [62].

## Authors' Contributions

The study was conceptualized by all authors. DB and SC were involved in the methodology. DB was involved in data curation, formal analysis, visualization, and writing of the original draft, and SC supervised the study. All authors reviewed and edited the paper.

## Conflicts of Interest

DB was a casual employee of Health Canada (2019 to 2020) and held a doctoral training award from the Fonds de recherche du Québec - Santé (2019 to 2021). DB has no conflicts of interest. SC receives research funding from the Canadian Institutes of Health Research, Fonds de recherche du Québec, Canadian Foundation for Dietetics Research, Canadian Foundation for Innovation, and Canadian Cancer Society. None of these agencies has funded or been involved in this study. NP is the NuAge database administrator; NP and SC serve as NuAge steering committee members. NuAge was supported by a research grant from the Canadian Institutes of Health Research (MOP-62842). The NuAge Database and Biobank is supported by the Fonds de recherche du Québec (2020-VICO-279753); the Quebec Network for Research on Aging, a thematic network funded by the Fonds de recherche du Québec - Santé; and the Merck-Frosst Chair funded by La Fondation de l'Université de Sherbrooke.

## Multimedia Appendix 1

Healthy Eating Food Index-2019 dietary constituents and score among simulated diets by Health Canada, by age and sex group. [[DOCX File , 20 KB-Multimedia Appendix 1](#)]

## References

1. Canada's dietary guidelines - for health professionals and policy makers. Health Canada. 2019. URL: <https://food-guide.canada.ca/en/guidelines/> [accessed 2024-04-29]
2. Food, nutrients and health: Interim evidence update 2018 for health professionals and policy makers. Health Canada. 2019. URL: <https://tinyurl.com/3wt7t7at> [accessed 2024-04-29]
3. Brassard D, Chevalier S. Relationship between adherence to the 2019 Canada's food guide recommendations on healthy food choices and nutrient intakes in older adults. *J Nutr.* Sep 2023;153(9):2699-2708. [[FREE Full text](#)] [doi: [10.1016/j.tjnut.2023.07.005](https://doi.org/10.1016/j.tjnut.2023.07.005)] [Medline: [37460086](https://pubmed.ncbi.nlm.nih.gov/37460086/)]
4. Ramage-Morin PL, Garriguet D. Nutritional risk among older Canadians. *Health Rep.* Mar 2013;24(3):3-13. [[FREE Full text](#)] [Medline: [24257971](https://pubmed.ncbi.nlm.nih.gov/24257971/)]
5. Mills CM, Keller HH, DePaul VG, Donnelly C. Factors associated with the development of high nutrition risk: data from the Canadian Longitudinal Study on Aging. *Can J Aging.* Mar 2024;43(1):153-166. [doi: [10.1017/S0714980823000545](https://doi.org/10.1017/S0714980823000545)] [Medline: [37749058](https://pubmed.ncbi.nlm.nih.gov/37749058/)]
6. Bauer J, Biolo G, Cederholm T, Cesari M, Cruz-Jentoft AJ, Morley JE, et al. Evidence-based recommendations for optimal dietary protein intake in older people: a position paper from the PROT-AGE Study group. *J Am Med Dir Assoc.* Aug 2013;14(8):542-559. [[FREE Full text](#)] [doi: [10.1016/j.jamda.2013.05.021](https://doi.org/10.1016/j.jamda.2013.05.021)] [Medline: [23867520](https://pubmed.ncbi.nlm.nih.gov/23867520/)]
7. Deutz NE, Bauer JM, Barazzoni R, Biolo G, Boirie Y, Bosy-Westphal A, et al. Protein intake and exercise for optimal muscle function with aging: recommendations from the ESPEN Expert Group. *Clin Nutr.* Dec 2014;33(6):929-936. [[FREE Full text](#)] [doi: [10.1016/j.clnu.2014.04.007](https://doi.org/10.1016/j.clnu.2014.04.007)] [Medline: [24814383](https://pubmed.ncbi.nlm.nih.gov/24814383/)]
8. Paddon-Jones D, Campbell WW, Jacques PF, Kritchevsky SB, Moore LL, Rodriguez NR, et al. Protein and healthy aging. *Am J Clin Nutr.* Jun 2015;101(6):1339S-1345S. [[FREE Full text](#)] [doi: [10.3945/ajcn.114.084061](https://doi.org/10.3945/ajcn.114.084061)] [Medline: [25926511](https://pubmed.ncbi.nlm.nih.gov/25926511/)]
9. Rémond D, Shahar DR, Gille D, Pinto P, Kachal J, Peyron MA, et al. Understanding the gastrointestinal tract of the elderly to develop dietary solutions that prevent malnutrition. *Oncotarget.* Jun 10, 2015;6(16):13858-13898. [[FREE Full text](#)] [doi: [10.18632/oncotarget.4030](https://doi.org/10.18632/oncotarget.4030)] [Medline: [26091351](https://pubmed.ncbi.nlm.nih.gov/26091351/)]
10. Shlisky J, Bloom DE, Beaudreault AR, Tucker KL, Keller HH, Freund-Levi Y, et al. Nutritional considerations for healthy aging and reduction in age-related chronic disease. *Adv Nutr.* Jan 2017;8(1):17-26. [[FREE Full text](#)] [doi: [10.3945/an.116.013474](https://doi.org/10.3945/an.116.013474)] [Medline: [28096124](https://pubmed.ncbi.nlm.nih.gov/28096124/)]
11. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol.* Apr 15, 2016;183(8):758-764. [[FREE Full text](#)] [doi: [10.1093/aje/kwv254](https://doi.org/10.1093/aje/kwv254)] [Medline: [26994063](https://pubmed.ncbi.nlm.nih.gov/26994063/)]

12. Kutcher SA, Brophy JM, Banack HR, Kaufman JS, Samuel M. Emulating a randomised controlled trial with observational data: an introduction to the target trial framework. *Can J Cardiol*. Sep 2021;37(9):1365-1377. [doi: [10.1016/j.cjca.2021.05.012](https://doi.org/10.1016/j.cjca.2021.05.012)] [Medline: [34090982](https://pubmed.ncbi.nlm.nih.gov/34090982/)]
13. Tobias DK, Lajous M. What would the trial be? Emulating randomized dietary intervention trials to estimate causal effects with observational data. *Am J Clin Nutr*. Aug 02, 2021;114(2):416-417. [FREE Full text] [doi: [10.1093/ajcn/nqab169](https://doi.org/10.1093/ajcn/nqab169)] [Medline: [34041528](https://pubmed.ncbi.nlm.nih.gov/34041528/)]
14. Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. *JAMA*. Dec 27, 2022;328(24):2446-2447. [doi: [10.1001/jama.2022.21383](https://doi.org/10.1001/jama.2022.21383)] [Medline: [36508210](https://pubmed.ncbi.nlm.nih.gov/36508210/)]
15. Stern D, Ibsen DB, MacDonald CJ, Chiu YH, Lajous M, Tobias DK. Improving nutrition science begins with asking better questions. *Am J Epidemiol*. Nov 04, 2024;193(11):1507-1510. [doi: [10.1093/aje/kwae110](https://doi.org/10.1093/aje/kwae110)] [Medline: [38992167](https://pubmed.ncbi.nlm.nih.gov/38992167/)]
16. Tobias DK. What eggsactly are we asking here? Unscrambling the epidemiology of eggs, cholesterol, and mortality. *Circulation*. May 17, 2022;145(20):1521-1523. [doi: [10.1161/CIRCULATIONAHA.122.059393](https://doi.org/10.1161/CIRCULATIONAHA.122.059393)] [Medline: [35576317](https://pubmed.ncbi.nlm.nih.gov/35576317/)]
17. Tomova GD, Arnold KF, Gilthorpe MS, Tennant PW. Adjustment for energy intake in nutritional research: a causal inference perspective. *Am J Clin Nutr*. Jan 11, 2022;115(1):189-198. [FREE Full text] [doi: [10.1093/ajcn/nqab266](https://doi.org/10.1093/ajcn/nqab266)] [Medline: [34313676](https://pubmed.ncbi.nlm.nih.gov/34313676/)]
18. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. Nov 2016;79:70-75. [FREE Full text] [doi: [10.1016/j.jclinepi.2016.04.014](https://doi.org/10.1016/j.jclinepi.2016.04.014)] [Medline: [27237061](https://pubmed.ncbi.nlm.nih.gov/27237061/)]
19. Chiu YH, Chavarro JE, Dickerman BA, Manson JE, Mukamal KJ, Rexrode KM, et al. Estimating the effect of nutritional interventions using observational data: the American Heart Association's 2020 dietary goals and mortality. *Am J Clin Nutr*. Aug 02, 2021;114(2):690-703. [FREE Full text] [doi: [10.1093/ajcn/nqab100](https://doi.org/10.1093/ajcn/nqab100)] [Medline: [34041538](https://pubmed.ncbi.nlm.nih.gov/34041538/)]
20. Berkowitz SA, Basu S, Hanmer J. Eliminating food insecurity in the USA: a target trial emulation using observational data to estimate effects on health-related quality of life. *J Gen Intern Med*. Aug 2023;38(10):2308-2317. [doi: [10.1007/s11606-023-08095-6](https://doi.org/10.1007/s11606-023-08095-6)] [Medline: [36814050](https://pubmed.ncbi.nlm.nih.gov/36814050/)]
21. Bonekamp NE, Crujnsen E, Visseren FL, van der Schouw YT, Geleijnse JM, Koopal C. Compliance with the DASH diet and risk of all-cause and cardiovascular mortality in patients with myocardial infarction. *Clin Nutr*. Aug 2023;42(8):1418-1426. [doi: [10.1016/j.clnu.2023.06.033](https://doi.org/10.1016/j.clnu.2023.06.033)] [Medline: [37433229](https://pubmed.ncbi.nlm.nih.gov/37433229/)]
22. Ibsen DB, Chiu YH, Gémes K, Wolk A. Hypothetical 22-year intervention with the dietary approaches to stop hypertension and risk of heart failure in a general population. *Am J Epidemiol*. Jan 08, 2024;193(1):96-106. [FREE Full text] [doi: [10.1093/aje/kwad181](https://doi.org/10.1093/aje/kwad181)] [Medline: [37656615](https://pubmed.ncbi.nlm.nih.gov/37656615/)]
23. Dickerman BA, Giovannucci E, Pernar CH, Mucci LA, Hernán MA. Guideline-based physical activity and survival among US men with nonmetastatic prostate cancer. *Am J Epidemiol*. Mar 01, 2019;188(3):579-586. [FREE Full text] [doi: [10.1093/aje/kwy261](https://doi.org/10.1093/aje/kwy261)] [Medline: [30496346](https://pubmed.ncbi.nlm.nih.gov/30496346/)]
24. Guo F, McGee EE, Chiu YH, Giovannucci E, Mucci LA, Dickerman BA. Evaluating recommendation-based dietary and physical activity strategies for prostate cancer prevention: a target trial emulation in the Health Professionals Follow-up Study. *Am J Epidemiol*. Jul 05, 2024;kwae184. [doi: [10.1093/aje/kwae184](https://doi.org/10.1093/aje/kwae184)] [Medline: [38973750](https://pubmed.ncbi.nlm.nih.gov/38973750/)]
25. Gaudreau P, Morais JA, Shatenstein B, Gray-Donald K, Khalil A, Dionne I, et al. Nutrition as a determinant of successful aging: description of the Quebec longitudinal study Nuage and results from cross-sectional pilot studies. *Rejuvenation Res*. Sep 2007;10(3):377-386. [doi: [10.1089/rej.2007.0596](https://doi.org/10.1089/rej.2007.0596)] [Medline: [17708689](https://pubmed.ncbi.nlm.nih.gov/17708689/)]
26. Hernán MA. The C-word: scientific euphemisms do not improve causal inference from observational data. *Am J Public Health*. May 2018;108(5):616-619. [doi: [10.2105/AJPH.2018.304337](https://doi.org/10.2105/AJPH.2018.304337)] [Medline: [29565659](https://pubmed.ncbi.nlm.nih.gov/29565659/)]
27. Goetghebeur E, le Cessie S, de Stavola B, Moodie EE, Waernbaum I, “on behalf of” the topic group Causal Inference (TG7) of the STRATOS initiative. Formulating causal questions and principled statistical answers. *Stat Med*. Dec 30, 2020;39(30):4922-4948. [FREE Full text] [doi: [10.1002/sim.8741](https://doi.org/10.1002/sim.8741)] [Medline: [32964526](https://pubmed.ncbi.nlm.nih.gov/32964526/)]
28. Haber NA, Wieten SE, Rohrer JM, Arah OA, Tennant PW, Stuart EA, et al. Causal and associational language in observational health research: a systematic evaluation. *Am J Epidemiol*. Nov 19, 2022;191(12):2084-2097. [FREE Full text] [doi: [10.1093/aje/kwac137](https://doi.org/10.1093/aje/kwac137)] [Medline: [35925053](https://pubmed.ncbi.nlm.nih.gov/35925053/)]
29. Brassard D, Elvidge Munene LA, St-Pierre S, Guenther PM, Kirkpatrick SI, Slater J, et al. Development of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. *Appl Physiol Nutr Metab*. May 2022;47(5):595-610. [FREE Full text] [doi: [10.1139/apnm-2021-0415](https://doi.org/10.1139/apnm-2021-0415)] [Medline: [35030038](https://pubmed.ncbi.nlm.nih.gov/35030038/)]
30. Hernán MA. Does water kill? A call for less casual causal inferences. *Ann Epidemiol*. Oct 2016;26(10):674-680. [FREE Full text] [doi: [10.1016/j.annepidem.2016.08.016](https://doi.org/10.1016/j.annepidem.2016.08.016)] [Medline: [27641316](https://pubmed.ncbi.nlm.nih.gov/27641316/)]
31. Chiu YH. Well-defined interventions for nutritional studies: from target trials to nutritional modeling. *Am J Clin Nutr*. Jan 11, 2022;115(1):3-5. [FREE Full text] [doi: [10.1093/ajcn/nqab343](https://doi.org/10.1093/ajcn/nqab343)] [Medline: [34849525](https://pubmed.ncbi.nlm.nih.gov/34849525/)]
32. Kipnis V, Freedman LS, Brown CC, Hartman A, Schatzkin A, Wacholder S. Interpretation of energy adjustment models for nutritional epidemiology. *Am J Epidemiol*. Jun 15, 1993;137(12):1376-1380. [doi: [10.1093/oxfordjournals.aje.a116647](https://doi.org/10.1093/oxfordjournals.aje.a116647)] [Medline: [8333419](https://pubmed.ncbi.nlm.nih.gov/8333419/)]

33. Tomova GD, Gilthorpe MS, Tennant PW. Theory and performance of substitution models for estimating relative causal effects in nutritional epidemiology. *Am J Clin Nutr*. Nov 2022;116(5):1379-1388. [FREE Full text] [doi: [10.1093/ajcn/nqac188](https://doi.org/10.1093/ajcn/nqac188)] [Medline: [36223891](https://pubmed.ncbi.nlm.nih.gov/36223891/)]
34. Simulated composite diets. Health Canada. 2022. URL: <https://tinyurl.com/yc4xjdua> [accessed 2024-04-29]
35. Brassard D, Elvidge Munene LA, St-Pierre S, Gonzalez A, Guenther PM, Jessri M, et al. Evaluation of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. *Appl Physiol Nutr Metab*. May 2022;47(5):582-594. [FREE Full text] [doi: [10.1139/apnm-2021-0416](https://doi.org/10.1139/apnm-2021-0416)] [Medline: [35030069](https://pubmed.ncbi.nlm.nih.gov/35030069/)]
36. Willett W. *Nutritional Epidemiology*. 3rd edition. Oxford, UK. Oxford University Press; 2013.
37. Willett WC, Stampfer M, Tobias DK. Re: Adjustment for energy intake in nutritional research: a causal inference perspective. *Am J Clin Nutr*. Aug 04, 2022;116(2):608-609. [FREE Full text] [doi: [10.1093/ajcn/nqac114](https://doi.org/10.1093/ajcn/nqac114)] [Medline: [35482596](https://pubmed.ncbi.nlm.nih.gov/35482596/)]
38. Harrison S, Brassard D, Garriguet D, Lemieux S, Lamarche B. A food-level substitution analysis assessing the impact of replacing regular-fat dairy with lower fat dairy on saturated fat intake at a population level in Canada. *Am J Clin Nutr*. Nov 08, 2021;114(5):1830-1836. [FREE Full text] [doi: [10.1093/ajcn/nqab251](https://doi.org/10.1093/ajcn/nqab251)] [Medline: [34320157](https://pubmed.ncbi.nlm.nih.gov/34320157/)]
39. Harrison S, Lemieux S, Lamarche B. Assessing the impact of replacing foods high in saturated fats with foods high in unsaturated fats on dietary fat intake among Canadians. *Am J Clin Nutr*. Mar 04, 2022;115(3):877-885. [FREE Full text] [doi: [10.1093/ajcn/nqab420](https://doi.org/10.1093/ajcn/nqab420)] [Medline: [34958344](https://pubmed.ncbi.nlm.nih.gov/34958344/)]
40. Daly RM, O'Connell SL, Mundell NL, Grimes CA, Dunstan DW, Nowson CA. Protein-enriched diet, with the use of lean red meat, combined with progressive resistance training enhances lean tissue mass and muscle strength and reduces circulating IL-6 concentrations in elderly women: a cluster randomized controlled trial. *Am J Clin Nutr*. Apr 2014;99(4):899-910. [FREE Full text] [doi: [10.3945/ajcn.113.064154](https://doi.org/10.3945/ajcn.113.064154)] [Medline: [24477043](https://pubmed.ncbi.nlm.nih.gov/24477043/)]
41. Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. *Int J Epidemiol*. Apr 01, 2017;46(2):756-762. [FREE Full text] [doi: [10.1093/ije/dyw323](https://doi.org/10.1093/ije/dyw323)] [Medline: [28039382](https://pubmed.ncbi.nlm.nih.gov/28039382/)]
42. Igelström E, Craig P, Lewsey J, Lynch J, Pearce A, Katikireddi SV. Causal inference and effect estimation using observational data. *J Epidemiol Community Health*. Sep 06, 2022;76(11):960-966. [doi: [10.1136/jech-2022-219267](https://doi.org/10.1136/jech-2022-219267)]
43. Taubman SL, Robins JM, Mittleman MA, Hernán MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol*. Dec 2009;38(6):1599-1611. [FREE Full text] [doi: [10.1093/ije/dyp192](https://doi.org/10.1093/ije/dyp192)] [Medline: [19389875](https://pubmed.ncbi.nlm.nih.gov/19389875/)]
44. McGrath S, Lin V, Zhang Z, Petito LC, Logan RW, Hernán MA, et al. gfoRmula: an R package for estimating the effects of sustained treatment strategies via the parametric g-formula. *Patterns (N Y)*. Jun 12, 2020;1(3):100008. [FREE Full text] [doi: [10.1016/j.patter.2020.100008](https://doi.org/10.1016/j.patter.2020.100008)] [Medline: [32656541](https://pubmed.ncbi.nlm.nih.gov/32656541/)]
45. Young JG, Hernán MA, Robins JM. Identification, estimation and approximation of risk under interventions that depend on the natural value of treatment using observational data. *Epidemiol Methods*. Dec 2014;3(1):1-19. [FREE Full text] [doi: [10.1515/em-2012-0001](https://doi.org/10.1515/em-2012-0001)] [Medline: [25866704](https://pubmed.ncbi.nlm.nih.gov/25866704/)]
46. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Model*. 1986;7(9-12):1393-1512. [FREE Full text] [doi: [10.1016/0270-0255\(86\)90088-6](https://doi.org/10.1016/0270-0255(86)90088-6)]
47. Shatenstein B, Nadon S, Godin C, Ferland G. Development and validation of a food frequency questionnaire. *Can J Diet Pract Res*. 2005;66(2):67-75. [doi: [10.3148/66.2.2005.67](https://doi.org/10.3148/66.2.2005.67)] [Medline: [15975195](https://pubmed.ncbi.nlm.nih.gov/15975195/)]
48. Freedman LS, Commins JM, Moler JE, Arab L, Baer DJ, Kipnis V, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. *Am J Epidemiol*. Jul 15, 2014;180(2):172-188. [FREE Full text] [doi: [10.1093/aje/kwu116](https://doi.org/10.1093/aje/kwu116)] [Medline: [24918187](https://pubmed.ncbi.nlm.nih.gov/24918187/)]
49. Freedman LS, Commins JM, Moler JE, Willett W, Tinker LF, Subar AF, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake. *Am J Epidemiol*. Apr 01, 2015;181(7):473-487. [FREE Full text] [doi: [10.1093/aje/kwu325](https://doi.org/10.1093/aje/kwu325)] [Medline: [25787264](https://pubmed.ncbi.nlm.nih.gov/25787264/)]
50. Thompson FE, Kirkpatrick SI, Subar AF, Reedy J, Schap TE, Wilson MM, et al. The National Cancer Institute's dietary assessment primer: a resource for diet research. *J Acad Nutr Diet*. Dec 2015;115(12):1986-1995. [FREE Full text] [doi: [10.1016/j.jand.2015.08.016](https://doi.org/10.1016/j.jand.2015.08.016)] [Medline: [26422452](https://pubmed.ncbi.nlm.nih.gov/26422452/)]
51. Keogh RH, Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: part 1-basic theory and simple methods of adjustment. *Stat Med*. Jul 20, 2020;39(16):2197-2231. [FREE Full text] [doi: [10.1002/sim.8532](https://doi.org/10.1002/sim.8532)] [Medline: [32246539](https://pubmed.ncbi.nlm.nih.gov/32246539/)]
52. Zhang S, Midthune D, Guenther PM, Krebs-Smith SM, Kipnis V, Dodd KW, et al. A new multivariate measurement error model with zero-inflated dietary data, and its application to dietary assessment. *Ann Appl Stat*. Jun 01, 2011;5(2B):1456-1487. [FREE Full text] [doi: [10.1214/10-AOAS446](https://doi.org/10.1214/10-AOAS446)] [Medline: [21804910](https://pubmed.ncbi.nlm.nih.gov/21804910/)]
53. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol*. Feb 1993;46(2):153-162. [doi: [10.1016/0895-4356\(93\)90053-4](https://doi.org/10.1016/0895-4356(93)90053-4)] [Medline: [8437031](https://pubmed.ncbi.nlm.nih.gov/8437031/)]
54. Washburn RA, McAuley E, Katula J, Mihalko SL, Boileau RA. The physical activity scale for the elderly (PASE): evidence for validity. *J Clin Epidemiol*. Jul 1999;52(7):643-651. [doi: [10.1016/s0895-4356\(99\)00049-9](https://doi.org/10.1016/s0895-4356(99)00049-9)] [Medline: [10391658](https://pubmed.ncbi.nlm.nih.gov/10391658/)]

55. Kyle RP, Moodie EE, Klein MB, Abrahamowicz M. Evaluating flexible modeling of continuous covariates in inverse-weighted estimators. *Am J Epidemiol*. Jun 01, 2019;188(6):1181-1191. [FREE Full text] [doi: [10.1093/aje/kwz004](https://doi.org/10.1093/aje/kwz004)] [Medline: [30649165](https://pubmed.ncbi.nlm.nih.gov/30649165/)]
56. Bland JM, Altman DG. Statistics notes: bootstrap resampling methods. *BMJ*. Jun 02, 2015;350:h2622. [doi: [10.1136/bmj.h2622](https://doi.org/10.1136/bmj.h2622)] [Medline: [26037412](https://pubmed.ncbi.nlm.nih.gov/26037412/)]
57. Lin V, McGrath S, Zhang Z, Logan RW, Petito LC, Li J, et al. gfoRmula: parametric g-formula. *Cran R. URL: <https://CRAN.R-project.org/package=gfoRmula>* [accessed 2024-04-29]
58. Database and biobank of the Quebec longitudinal study on nutrition and successful aging. *Banques NuAge. URL: <https://nuage.recherche.usherbrooke.ca>* [accessed 2024-04-29]
59. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. Jan 1999;10(1):37-48. [Medline: [9888278](https://pubmed.ncbi.nlm.nih.gov/9888278/)]
60. Tennant PW, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *Int J Epidemiol*. May 17, 2021;50(2):620-632. [FREE Full text] [doi: [10.1093/ije/dyaa213](https://doi.org/10.1093/ije/dyaa213)] [Medline: [33330936](https://pubmed.ncbi.nlm.nih.gov/33330936/)]
61. Lipsky AM, Greenland S. Causal directed acyclic graphs. *JAMA*. Mar 15, 2022;327(11):1083-1084. [doi: [10.1001/jama.2022.1816](https://doi.org/10.1001/jama.2022.1816)] [Medline: [35226050](https://pubmed.ncbi.nlm.nih.gov/35226050/)]
62. Brassard D. didierbrassard / NuAge\_protocol. *GitHub*. 2024. URL: [https://github.com/didierbrassard/NuAge\\_protocol](https://github.com/didierbrassard/NuAge_protocol) [accessed 2024-04-29]

## Abbreviations

**CFG:** Canada's Food Guide

**DAG:** directed acyclic graph

**HEFI:** Healthy Eating Food Index

**NCI MCMC:** National Cancer Institute Markov Chain Monte Carlo

**NuAge:** Quebec Longitudinal Study on Nutrition and Successful Aging

**RA:** reference amount

**RCT:** randomized controlled trial

*Edited by A Schwartz; submitted 07.08.24; peer-reviewed by JM Hutchinson; comments to author 12.11.24; revised version received 27.11.24; accepted 29.11.24; published 23.01.25*

*Please cite as:*

*Brassard D, Presse N, Chevalier S*

*Estimating the Effect of Adhering to the Recommendations of the 2019 Canada's Food Guide on Health Outcomes in Older Adults: Protocol for a Target Trial Emulation*

*JMIR Res Protoc 2025;14:e65182*

*URL: <https://www.researchprotocols.org/2025/1/e65182>*

*doi: [10.2196/65182](https://doi.org/10.2196/65182)*

*PMID:*

©Didier Brassard, Nancy Presse, Stéphanie Chevalier. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 23.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.