Dietary Behaviors and Incident COVID-19 in the UK Biobank

Thanh-Huyen T. Vu, Kelsey J. Rydland, Chad J. Achenbach, Linda Van Horn and Marilyn C. Cornelis

Abstract: Background: Nutritional status influences immunity but its specific association with susceptibility to COVID-19 remains unclear. We examined the association of specific dietary data and incident COVID-19 in the UK Biobank (UKB). Methods: We considered UKB participants in England with self-reported baseline (2006–2010) data and linked them to Public Health England COVID-19 test results—performed on samples from combined nose/throat swabs, using real time polymerase chain reaction (RT-PCR)—between March and November 2020. Baseline diet factors included breastfed as baby and specific consumption of coffee, tea, oily fish, processed meat, red meat, fruit, and vegetables. Individual COVID-19 exposure was estimated using the UK’s average monthly positive case rate per specific geo-populations. Logistic regression estimated the odds of COVID-19 positivity by diet status adjusting for baseline socio-demographic factors, medical history, and other lifestyle factors. Another model was further adjusted for COVID-19 exposure. Results: Eligible UKB participants (n = 37,988) were 40 to 70 years of age at baseline; 17% tested positive for COVID-19 by SAR-CoV-2 PCR. After multivariable adjustment, the odds (95% CI) of COVID-19 positivity was 0.90 (0.83, 0.96) when consuming 2–3 cups of coffee/day (vs. <1 cup/day), 0.88 (0.80, 0.98) when consuming vegetables in the third quartile of servings/day (vs. lowest quartile), 1.14 (1.01, 1.29) when consuming fourth quartile servings of processed meats (vs. lowest quartile), and 0.91 (0.85, 0.98) when having been breastfed (vs. not breastfed). Associations were attenuated when further adjusted for COVID-19 exposure, but patterns of associations remained. Conclusions: In the UK Biobank, consumption of coffee, vegetables, and being breastfed as a baby were favorably associated with incident COVID-19; intake of processed meat was adversely associated. Although these findings warrant independent confirmation, adherence to certain dietary behaviors may be an additional tool to existing COVID-19 protection guidelines to limit the spread of this virus.

Keywords: nutrition; COVID-19; immunity; coffee; breastfeeding; epidemiology

1. Introduction

Coronavirus (COVID-19) is an infectious disease caused by SARS-CoV-2 that predominantly attacks the respiratory system and has spread rapidly in the last year [1]. This unprecedented global public health crisis necessitates the expansion of our understanding of COVID-19 for purposes of prevention and mitigation of its severity and mortality risk in infected individuals [2].

Cross-sectional analyses identify patient characteristics and clinical data subsequent to a positive test, thereby informing the progression and severity of COVID-19. However, identifying factors underlying potential susceptibility a priori to the virus offers public health benefits. The immune system plays a key role in an individual’s susceptibility and response to infectious diseases, including COVID-19 [3,4]. A major modifiable factor affecting immune function is dietary behavior that influences nutritional status [5–7].
Ecological studies of COVID-19 report favorable correlations with specific vegetables and dietary patterns, such as the Mediterranean diet [8–10]. Some dietary supplements were found to have an association with SARS-CoV-2 infection [11]. To our knowledge, no population data have been examined regarding the role of specific dietary intakes in prevention of COVID-19.

Using data from the UK Biobank (UKB), we examined the associations between dietary behaviors measured in 2006–2010 and incident COVID-19 infection in 2020. We additionally linked UKB geo-data to UK COVID-19 surveillance data to account for COVID-19 exposure in our analyses, which to our knowledge has never been done in COVID-19 studies of the UKB.

2. Materials and Methods

2.1. UK Biobank

The UKB is an international health resource of over 500,000 participants aged 37–73 years at 22 centers across England, Wales, and Scotland. Details of the study methods have been described previously [12]. Briefly, participants underwent physical measurements, assessments about health and risk factors (including lifestyle and dietary behaviors), and blood sampling at baseline (2006–2010) and agreed to follow-up on their health status. Details of data collected are provided on the Showcase tab of the UKB website [13]. UKB ethical approval was from the National Research Ethics Service Committee North West–Haydock (approval letter dated 17 June 2011, Ref 11/NW/0382), and all study procedures were performed in accordance with the World Medical Association Declaration of Helsinki ethical principles for medical research. The current analysis was approved under the UKB application #21394 (PI, M.C.C).

2.2. COVID-19 Diagnosis

COVID-19 test results from Public Health England have been dynamically linked to the UKB since 16 March 2020 [14]. The regularly updated COVID-19 data table provided to UKB researchers includes participant ID, record date, test location (mouth, nose, throat, trachea, etc.), testing laboratory (71 labs listed), and test results (negative or positive). The vast majority of samples tested are from combined nose/throat swabs that are transported in a medium suitable for viruses and subject to polymerase chain reaction testing.

2.3. COVID-19 Exposure

UKB provides grid co-ordinate data at the 1 km resolution as a measure of residential location [15]. The data are provided to researchers in the projection of the Ordnance Survey National Grid (OSGB1936) geographic reference system; measures refer to easting and northing with a reference point near the Isles of Sicily. These data were used in conjunction with country-wide surveillance data to identify UKB participants exposed to COVID-19 [16]. The UKB geo-data and UK COVID-19 surveillance data was imported and projected in ArcGIS (Environmental Systems Research Institute, Redlands, CA, USA) for visual inspection. To make the UKB geo-data (North/East) compatible with the latitude/longitude surveillance data, we converted the former to the latter using standard equations already built into ArcGIS [17]. Hotspot analysis was performed on the UK surveillance data to assign values (“Gi” statistic) back to each UKB participant representing a COVID-19-exposure risk factor or score. We calculated monthly positive case rates per specific geo-populations, and the average of these monthly rates was used in the analysis.

2.4. Baseline Dietary Data

The UKB touchscreen at baseline included a semi-quantitative food frequency questionnaire (FFQ) that assessed self-reported usual intake of 17 foods and beverages [18]. Participants were asked to report the number of pieces/teaspoons/cups of each item they consumed or to choose one of several pre-specified frequency categories. Among these choices, baseline servings of vegetables (cooked, raw), fruit (fresh, dried), oily fish, pro-
cessed meat, red meat (beef, lamb/mutton, or pork), tea, and coffee have specifically been identified as contributing nutritional factors implicated in immunity [19–24] and were thus targeted for consideration in this study. The details of the questions and possible answers for these diet factors are provided in Supplemental Table S1. Based on the participants’ responses, we computed portion sizes as cups or servings per day. Participants were also asked if they were breastfed as a baby and responded yes, no, or don’t know.

2.5. Other Covariates

Age, sex, race, education, employment status, type of accommodation lived in and number of co-habitants, smoking behaviors, and current health status were self-reported at baseline using the touchscreen. Townsend Deprivation Index, a measure of SES, was derived using census data and postal codes of participants, where higher scores represent higher deprivation. Body mass index (BMI) was calculated using height and weight measured at the assessment center. Physical activity was assessed using questions adapted from the validated short International Physical Activity Questionnaire [25]. Participants were asked how many minutes they engaged in each of the activities on a typical day. Duration of moderate and vigorous activities (minutes/day) was used in this analysis. History of diabetes, history of any heart diseases, and hypercholesterolemia or hypertensive medication use were obtained via self-reporting (and verified during a verbal interview) and hospital databases.

Because individuals with lung diseases may have an increased risk of COVID-19 infection [26], we further classified participants by their pre-existing compromised pulmonary function (yes/no respiratory diseases excluding common colds); data extracted from health records during follow-up. We also used health records to calculate incident diabetes and heart disease.

2.6. Analysis Sample

Because no COVID-19 test data were available for UKB assessment centers in Scotland and Wales, only participants in England were included. For this analysis, we included participants with test results between 16 March and 30 November 2020, which was before vaccines were rolled out in the UK (8 December 2020) [27]. There were 74,878 tests performed on 42,699 UKB participants. We excluded participants who had missing baseline data on nutritional factors and covariates (n = 4711). The final study sample included 37,988 UKB participants (Supplemental Figure S1).

2.7. Statistical Analysis

All analyses were performed using SAS (SAS Institute, Inc., Cary, NC, USA). Mapping was performed using ArcGIS. F-tests or χ² tests were used to compare differences in baseline characteristics across participants’ age and race/ethnicity.

Our main outcome of interest was whether a person had any confirmed COVID-19 infection (defined as having any positive test result). Our exposures of interest included consumption of coffee, tea, oily fish, processed meat, red meat, fruit, and vegetables; and were analyzed in quartiles of servings/d [each modeled: lowest quartile/Q1(referent), Q2, Q3, and Q4] as well as tea and coffee [each modeled: none or <1 (referent), 1, 2–3, and ≥4 cups/d], and breastfed as a baby [yes (referent), no, and don’t know]. Because the cut-points to define Q2 and Q3 for oily fish were similar, Q3 and Q4 were combined.

We examined the association between each dietary behavior and incident COVID-19 using logistic regression, first adjusting for age, sex, and race (White/Asian/Black/Mixed-Others) (Crude model), then further adjusting for Townsend-deprivation index (quartiles), education (6 qualification classes), employment status (employed/retired/other), type of living accommodation (house/apartment/other), number of co-habitants (1,2,3 or ≥4), income (4 levels), physical activity (quartiles of moderate or vigorous activities, min/day), smoking (never/past/current), BMI (<25, 25–<30, and ≥30 kg/m²), self-rated health (excellent/good/fair/poor), hypercholesterolemia medication use (yes/no), hypertensive
medication use (yes/no), diabetes (yes/no), and any heart diseases (yes/no) (Model 1). Model 2 was similar to Model 1 but with diet behaviors assessed mutually (i.e., including all 7 other dietary factors). Model 3 included variables from Model 2 and further adjusted for COVID-19 exposure score. Finally, Model 4 further adjusted for pre-existing compromised pulmonary function (yes/no). Statistical significance was defined as \( p < 0.05 \). No adjustments were made for multiple testing as all tests were a priori.

In sensitivity analyses, we excluded individuals self-reporting diabetes, heart diseases, or using cholesterol-lowering or antihypertension medication because they may recently have made dietary changes consequential to the disease. We also adjusted for incident diabetes and heart disease during follow-up. Moreover, as shown in Supplemental Figure S2, there were two waves of new cases between March and November 2020 (i.e., from March–August 2020, or Wave 1, and September–November 2020, or Wave 2) in the UK, with greater numbers in Wave 2 than Wave 1. The differences may reflect a change in testing-rate, a true change in risk of exposure to the virus or some other factor. Regardless, analyses were also conducted for Wave 1 and Wave 2 separately.

We screened for effect modification (interaction) by sex (male or female), age (<55 or \( \geq 55 \) years of age) and race (White, Asian, Black, or mixed/others) by including the cross-product term of each nutritional factor (e.g., coffee consumption, 4 categories) and the interacting variable in multivariable regression models. Statistically significant interactions were defined as \( p < 0.002 \); accounting for 24 tests (8 nutrition factors \( \times \) 3 effect modifiers) performed.

3. Results

3.1. Participant Characteristics

Of 37,988 eligible participants who were tested for COVID-19 from March to November 2020, 6482 (17%) tested positive. In general, participants in the analysis sample had baseline demographic and diet characteristics similar to those of the full UKB cohort (\( N = 502,633 \), Table 1). However, those in the analysis sample tended to report poorer health and more comorbidities than those in the full UKB cohort.

Characteristics of the analysis sample, stratified by baseline age and race, are presented in Supplemental Tables S2 and S3. Briefly, compared to the older age group, the younger age group tended to be female, employed, and better educated, with better income and rated better health; they also tended to consume less tea, fruit, vegetables, fish, and red meat, with fewer having been breastfed. Compared to non-white participants, white participants tended to consume more coffee, tea, processed meat, less fruit or vegetables, with fewer having been breastfed. Among non-whites, Black people tended to live in higher deprivation areas, be employed, have higher BMI, and consume more red meat, while Asians tended to report poorer health and consume more fruit and vegetables. The Asian group also had a higher incidence of COVID-19 positivity than other racial groups.

The exposure risk for the UKB sample in the context of the nation overall is illustrated in Figure 1 and Supplemental Figure S3, in which geo-positive tests are shown by Waves for UKB and UK data separately. Areas with a high number of positive tests in UKB were generally similar to those in the UK overall.

3.2. Nutritional Factors and COVID-19 Positivity

With adjustment for age, race, and sex (Crude model), consumption of coffee, moderate tea, oily fish, and vegetables; and being breastfed as a baby were significantly associated with lower odds of COVID-19 positivity, while consuming processed meat was associated with higher odds of COVID-19 positivity (Supplemental Table S4). After further adjusting for other socio-demographics, medical and lifestyle factors, associations were attenuated; consumption of tea and fish oil were no longer significantly associated with COVID-19 infection (Model 1–Table 2). The odds (95% CI) of COVID-19 positivity were 0.90 (0.83, 0.98), 0.90 (0.83, 0.96), and 0.92 (0.84, 1.00) when consuming 1 cup, 2–3 cups, and 4+ cups of coffee/day (vs. <1 cup/day), respectively (Model 2). The odds (95%CI) of
COVID-19 infection was 0.88 (0.80, 0.98) for individuals in the 3rd quartile of vegetable intake (vs. lowest quartile), 1.14 (1.01, 1.29) for individuals in the 4th quartile of processed meat intake (vs. lowest quartile), and 0.91 (0.85, 0.98) for those who had been breastfed as a baby (vs. not breastfed). We observed similar associations between these dietary factors assessed individually (Model 1) or mutually (Model 2) with COVID-19 infection, suggesting independent associations. Associations were attenuated when further adjusted for COVID-19 exposure, but patterns of associations remained (Model 3). Further adjusting for pre-existing compromised pulmonary function, or incident diabetes and heart disease during follow-up did not alter results (Model 4, data not shown).

Table 1. Baseline characteristics of the analysis sample and the full UK Biobank cohort.

| Baseline Characteristics * | Analysis Sample | Full UKB Cohort |
|----------------------------|-----------------|-----------------|
| Number of Persons          | 37,988          | 502,633         |
| Age, yr, mean(sd)          | 57.36 (8.23)    | 56.53 (8.10)    |
| Female                     | 20,026 (52.72)  | 273,453 (54.41) |
| Townsend-deprivation index, mean(sd) | −1.21 (3.12)    | −1.29 (3.10)    |
| White/British              | 35,793 (94.22)  | 472,801 (94.06) |
| Household Income, £ < 18,000 | 8264 (21.75)    | 97,220 (19.34)  |
| College or university degree | 11,000 (28.96)  | 161,198 (32.07) |
| Currently employed         | 20,354 (53.58)  | 287,210 (57.14) |
| Homeowner                  | 34,295 (90.28)  | 447,795 (89.36) |
| Number of co-habitants ≥ 4 | 6821 (17.96)    | 93,987 (18.87)  |
| Current smoker             | 4205 (11.07)    | 55,954 (11.13)  |
| BMI (kg/m²), mean(sd)      | 27.93 (4.95)    | 27.43 (4.80)    |
| Physical activity, minutes/day, mean(sd) | 75.50 (97.52)   | 74.23 (96.19)   |
| Poor overall health rating | 2412 (6.35)     | 22,780 (4.56)   |
| Using cholesterol medication | 8195 (21.57)    | 86,907 (17.29)  |
| Using blood pressure medication | 9610 (25.30)    | 104,024 (20.70) |
| Diabetes                   | 2680 (7.05)     | 26,552 (5.28)   |
| Heart diseases             | 3003 (7.91)     | 29,166 (5.80)   |
| Breastfed as baby          | 21,051 (55.41)  | 277,596 (55.39) |
| Coffee consumption, cups/day, mean(sd) | 2.02 (2.06)     | 2.01 (2.02)     |
| Tea consumption, cups/day, mean(sd) | 3.43 (2.73)     | 3.40 (2.69)     |
| Oily fish consumption, servings/day, mean(sd) | 0.16 (0.15)     | 0.16 (0.15)     |
| Processed meat, servings/day, mean(sd) | 0.22 (0.20)     | 0.21 (0.20)     |
| Red meat, servings/day, mean(sd) | 0.30 (0.21)     | 0.30 (0.21)     |
| Fruit (fresh/dried), servings/day, mean(sd) | 3.04 (2.60)     | 3.05 (2.62)     |
| Vegetables (cooked/raw), servings/day, mean(sd) | 0.82 (0.55)     | 0.81 (0.56)     |

* Data drawn from baseline (2006–2010). Values are numbers (%) unless stated otherwise.

Patterns of associations were similar when excluding individuals reporting diabetes, CVD, or the use of cholesterol-lowering or antihypertension medication from the sample (data not shown), or when analyzing data for Wave 1 and Wave 2 separately (Supplemental Table S5). Associations between nutrition factors and COVID-19 positivity were not significantly modified by age, sex, and race (p > 0.002 for interactions, Supplemental Tables S6–S8).
Figure 1. Total COVID-19 Positive Tests in Wave 1 compared to Wave 2 by UKB (A) and National (UK) (B) Data. The UKB geo-data and National (UK) COVID-19 surveillance data was imported, projected, and converted to be compatible with each other in ArcGIS for visual inspection (see COVID-19 exposure in Materials and Methods section for more details). Wave 1: Total positive tests for March-August 2020; Wave 2: Total positive tests for September–November 2020; Source: UKB and https://coronavirus.data.gov.uk/ accessed on 4 April 2021.

Table 2. Adjusted * OR (95% CI) of having a positive COVID-19 test by nutritional factors.

| Nutritional Factor | Model 1 | Model 2 | Model 3 |
|--------------------|---------|---------|---------|
| **Coffee, Cups/Day** |         |         |         |
| None or <1 cup     | Reference | Reference | Reference |
| 1 cup              | 0.90 (0.83, 0.97) | 0.90 (0.83, 0.98) | 0.93 (0.86, 1.01) |
| 2–3 cups           | 0.89 (0.83, 0.95) | 0.90 (0.83, 0.96) | 0.92 (0.85, 0.99) |
| ≥4 cups            | 0.92 (0.85, 0.996) | 0.92 (0.84, 0.999) | 0.91 (0.83, 0.99) |
| **Tea, cups/day**  |         |         |         |
| None or <1 cup     | Reference | Reference | Reference |
| 1 cup              | 0.90 (0.83, 0.97) | 0.90 (0.83, 0.98) | 0.93 (0.86, 1.01) |
| 2–3 cups           | 0.89 (0.83, 0.95) | 0.90 (0.83, 0.96) | 0.92 (0.85, 0.99) |
| ≥4 cups            | 1.00 (0.92, 1.08) | 0.98 (0.90, 1.06) | 0.96 (0.88, 1.05) |
| **Oily Fish, Servings/Day** |         |         |         |
| Quartile 1 (0–<0.07) | Reference | Reference | Reference |
| Quartile 2 (0.07–<0.14) | 0.93 (0.85, 1.02) | 0.94 (0.86, 1.03) | 0.97 (0.88, 1.06) |
| Quartiles 3 and 4 (≥0.14) | 0.95 (0.87, 1.04) | 0.98 (0.90, 1.07) | 1.00 (0.91, 1.09) |
Table 2. Cont.

| Nutritional Factor | Model 1 OR (95%CI) | p ** | Model 2 OR (95%CI) | p ** | Model 3 OR (95%CI) | p ** |
|--------------------|-------------------|------|-------------------|------|-------------------|------|
| Processed Meat, Servings/Day | | | | | | |
| Quartile 1 (0–<0.07) Reference | | | Reference | | | Reference |
| Quartile 2 (0.07–<0.14) 1.02 (0.91, 1.14) 0.699 | 1.05 (0.93, 1.19) 0.410 | 1.04 (0.92, 1.17) 0.547 |
| Quartile 3 (0.14–<0.43) 1.07 (0.96, 1.20) 0.248 | 1.09 (0.97, 1.24) 0.155 | 1.07 (0.94, 1.21) 0.314 |
| Quartile 4 (≥0.43) 1.12 (1.00, 1.25) 0.053 | 1.14 (1.01, 1.29) 0.036 | 1.12 (0.98, 1.26) 0.091 |
| Red Meat, Servings/Day | | | | | | |
| Quartile 1 (0–<0.21) Reference | | | Reference | | | Reference |
| Quartile 2 (0.21–<0.28) 0.96 (0.89, 1.04) 0.344 | 0.95 (0.87, 1.04) 0.236 | 0.99 (0.90, 1.08) 0.811 |
| Quartile 3 (0.28–<0.35) 1.01 (0.93, 1.11) 0.771 | 1.00 (0.90, 1.10) 0.948 | 1.01 (0.91, 1.11) 0.903 |
| Quartile 4 (≥0.35) 0.99 (0.92, 1.08) 0.878 | 0.98 (0.89, 1.07) 0.600 | 0.98 (0.90, 1.08) 0.697 |
| Fruit (Fresh/Dried), Servings/Day | | | | | | |
| Quartile 1 (0–<1.00) Reference | | | Reference | | | Reference |
| Quartile 2 (1.00–<2.25) 1.02 (0.92, 1.12) 0.778 | 1.05 (0.95, 1.16) 0.376 | 1.06 (0.95, 1.17) 0.300 |
| Quartile 3 (2.25–<4.00) 0.97 (0.87, 1.09) 0.622 | 1.02 (0.91, 1.14) 0.762 | 1.03 (0.91, 1.15) 0.679 |
| Quartile 4 (≥4.00) 0.97 (0.88, 1.09) 0.635 | 1.03 (0.92, 1.15) 0.660 | 1.03 (0.92, 1.16) 0.596 |
| Vegetables (Cooked/Raw), Servings/Day | | | | | | |
| Quartile 1 (0–<0.50) Reference | | | Reference | | | Reference |
| Quartile 2 (0.50–<0.67) 0.92 (0.85, 0.99) 0.034 | 0.93 (0.85, 1.00) 0.060 | 0.96 (0.88, 1.04) 0.271 |
| Quartile 3 (0.67–<1.00) 0.87 (0.79, 0.96) 0.006 | 0.88 (0.80, 0.98) 0.015 | 0.93 (0.84, 1.03) 0.186 |
| Quartile 4 (≥1.00) 0.90 (0.83, 0.98) 0.015 | 0.92 (0.84, 0.998) 0.046 | 0.96 (0.88, 1.05) 0.337 |
| Breastfed as a Baby | | | | | | |
| No Reference | | | Reference | | | Reference |
| Yes 0.91 (0.85, 0.98) 0.010 | 0.91 (0.85, 0.98) 0.013 | 0.95 (0.88, 1.02) 0.125 |
| Don’t know 0.97 (0.91, 1.07) 0.750 | 0.98 (0.90, 1.07) 0.696 | 0.99 (0.91, 1.08) 0.789 |

* Model 1: Adjusted for Townsend deprivation index, baseline age, sex, race, education, income, employment status, home ownership, number of co-habitants, BMI level, smoking status, physical activity, self-rated health, cholesterol-lowering medication use, antihypertension medication use, history of diabetes, and history of cardiovascular disease. Individual diet factors assessed in separate models. Model 2: Adjusted for all covariates listed in Model 1, with all diet factors included in the model (i.e., mutual adjustment). Model 3: Adjusted for all covariates listed in Model 2 and COVID-19 exposure score. ** An α of 0.05 was used as the cutoff for significance.

4. Discussion

In this study, consuming more coffee, vegetables, and being breast fed as well as consuming less processed meat intake were independently associated with lower odds of COVID-19 positivity. These associations were attenuated when accounting for the UK’s COVID-19 case rate (i.e., exposure).

Much research has focused on characterizing individuals who test positive for COVID-19 and infection progression and outcomes. Individuals with suppressed immune systems, such as the elderly and those with existing comorbidities including cardiovascular diseases, hypertension, diabetes, and obesity, are more likely to suffer with progression toward severe outcomes of COVID-19 [3,4,28,29]. Less attention has focused on modifiable risk factors preceding COVID-19 infection. While nutrition may theoretically impact COVID-19 susceptibility [3,6,30–32], few investigations have specifically tested the hypothesis a priori. Low vitamin D status is associated with COVID-19 infection, severity, and mortality [33,34]. Del Ser et al. [35] recently explored a variety of risk factors for the incidence, severity, and mortality of COVID-19 in a Spanish cohort of 913 volunteers aged 75–90 years. Sixty-two cases reported symptoms compatible with COVID-19; 6 of them died. High alcohol and lower coffee and tea consumption were associated with disease severity; other dietary behaviors were not considered.

In the UKB, habitual consumption of 1 or more cups of coffee per day was associated with about a 10% decrease in risk of COVID-19 compared to less than 1 cup/day. Coffee is
not only a key source of caffeine, but contributes dozens of other constituents; including many implicated in immunity [21]. Among many populations, coffee is the major contributor to total polyphenol intake, phenolic acids in particular [36,37]. Coffee, caffeine, and polyphenols have antioxidant and anti-inflammatory properties [38–44]. Coffee consumption favorably correlates with inflammatory biomarkers such as CRP, interleukin-6 (IL-6), and tumor necrosis factor α (TNF-α) [38,45–51], which are also associated with COVID-19 severity and mortality [52,53]. Coffee consumption has also been associated with lower risk of pneumonia in elderly [54]. Taken together, an immunoprotective effect of coffee against COVID-19 is plausible and merits further investigation.

Fruits and vegetables are rich dietary sources of vitamins, folate, fiber, and several phytochemicals such as carotenoids and flavonoids [55,56]. These substances have anti-inflammatory, antibacterial, and antiviral properties and are thus immune-protective [19,57,58]. In the current study, consumption of at least 0.67 servings/d of vegetables (cooked or raw, excluding potatoes) was associated with a lower risk of COVID-19 infection. Recent ecological studies of COVID-19 report that countries with high consumption of foods with potent antioxidant or anti angiotensin-converting enzyme (ACE) activity such as raw or fermented cabbage have a lower COVID-19 death rate compared to other countries [59,60]. Studies of vegetables and fermented-foods with COVID-19 mortality in Europe reported that each g/day increase in the average national consumption of head cabbage, cucumber or fermented vegetables decreased the mortality risk for COVID-19 by 11–35%; consumption of broccoli surprisingly increased COVID-19 mortality [9,10]. COVID-19 mortality was not our outcome of interest, and while we did not analyze vegetables individually, the UK’s national consumption of broccoli is above Europe’s average and thus our findings conflict with the previous inverse relationship between broccoli consumption and COVID-19 mortality [9,10]. Fruit (fresh and dried) consumption was not associated with COVID-19 risk in the UKB. While fruits and vegetables share several health benefits, the specific bioactive compounds in fruits and in vegetables can vary [61]. Fruits are also relatively higher in sugar (fructose) while vegetables contain more starch. Follow-up studies to our current findings might therefore focus on vegetable constituents distinct from those of fruit.

Processed meat consumption of as little as 0.43 servings/d was associated with a higher risk of COVID-19 in UKB. However red meat consumption presented no risk, suggesting meat per se does not underlie the association we observed with processed meats. Processed meat refers to any meat that has been transformed through salting, curing, fermenting, smoking, or other process to enhance flavor or improve preservation [62]. In the UK, sausages, bacon, and ham are the major contributors to processed meat intake [63] and these often contain salt enriched with nitrates/nitrites. Preservatives and other additives are increasingly being used yet difficult to measure in observational studies [64]. These include milk-, soy- and wheat-based ingredients, spices, ascorbic acid, phosphates, antioxidants, monosodium glutamate, food colorings, blood plasma, gelatin, and transglutaminase [62]. Total and saturated fat concentrations are also relatively higher in processed meats [65]. Processed meats are also characteristic of a western-style diet, which may adversely impact immunity [23] and thus other dietary behaviors that correlate with processed meat intake may underlie the association we observed with COVID-19 susceptibility.

Consumption of human milk has also emerged as an early-life factor impacting immunity in both infancy and adulthood, with links to lower rates of allergy, influenza, asthma, and other respiratory infections [66,67]. Breastfeeding has also been linked to epigenetic changes of toll-like-receptors, which play significant roles in innate immunity [68]. We found a long-term favorable association between being breastfed as a baby and COVID-19 infection in UKB and thus contribute to the growing evidence in support of nutrition early in life for optimal immunity for life.

The strengths of this study include access to the largest cohort established years before the COVID-19 pandemic, with detailed health, lifestyle, and nutrition data and ongoing follow-up. Although much COVID-19 research [69] has incorporated this database, we are the first to leverage independent COVID-19 surveillance database of the UK to estimate
COVID-19 exposure [70]. COVID-19 data from UKB are consistent with public government data, which adds confidence to our approach and results.

The study has limitations. First, only a portion of overall UKB participants were tested for COVID-19 (about 10%) during the study timeframe, and these participants were slightly older, less educated, and less employed, while reporting poorer health than the original UKB cohort. Those factors were associated with higher odds of COVID-19 infection in our analysis sample. We note, however, that the full UKB is not representative of the sampling population, with evidence of a ‘healthy volunteer’ selection bias [71]. However, participants in these two samples had many baseline characteristics in common, including dietary behaviors. Any effects of selection bias would likely lead to an underestimation of the true associations between nutrition and COVID-19. Second, diet assessment tools are generally prone to measurement error: limitations inherent to closed-ended options, inaccurate portion size estimates, and social desirability bias are among the sources contributing to error [12,72]. Nevertheless, we were uniquely positioned to pursue this investigation with novel aims and analyses; and although the effect sizes may be imprecise, establishing a relationship between diet and COVID-19 is informative in itself and can be inferred from the direction and statistical significance of the association [12,72]. Finally, the current study is observational and cannot infer causality. While we accounted for a comprehensive set of potential confounders, we cannot discount the possibility of residual confounding; whether it be positive or negative. Moreover, we had no concurrent pandemic data on other established risk factors for COVID-19 infection, such as participant social distancing behavior, work environment, and face mask-use; some of these factors may correlate with diet behaviors.

5. Conclusions

Our results support the hypothesis that nutritional factors may influence distinct aspects of the immune system, hence susceptibility to COVID-19. Encouraging adherence to certain nutritional behaviors (e.g., increasing vegetable intake and reducing processed meat intake) may be an additional tool to existing COVID-19 protection guidelines to limit the spread of this virus. Nevertheless, our findings warrant confirmation in other populations.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/nu13062114/s1: Supplemental Table S1: Description of the dietary touchscreen questionnaire and possible responses, Supplemental Table S2: Characteristics of the analysis sample by age groups, Supplemental Table S3: Characteristics of the analysis sample by racial groups, Supplemental Table S4: Age-race-sex adjusted OR (95% CI) of having positive COVID-19 test by nutritional factors, Supplemental Table S5: Adjusted OR (95% CI) of having positive COVID-19 test by nutritional factors, stratified by Wave, Supplemental Table S6: Adjusted OR (95% CI) of having positive COVID-19 test by nutritional factors, stratified by baseline age group, Supplemental Table S7: Adjusted OR (95% CI) of having positive COVID-19 test by nutritional factors, stratified by sex, Supplemental Table S8: Adjusted OR (95% CI) of having positive COVID-19 test by nutritional factors, stratified by racial group, Supplemental Figure S1. Flow chart depicting the study design and the analysis sample, Supplemental Figure S2: National (UK) COVID-19 Cases Reported Between March and December 1 2020, Supplemental Figure S3. Total COVID-19 Positive Tests in UKB (A) Compared to National Data by Wave (B and C)*.

Author Contributions: Conceptualization, M.C.C.; methodology, T.-H.T.V., K.J.R., and M.C.C.; analysis, T.-H.T.V. and K.J.R.; writing, T.-H.T.V. and M.C.C.; funding acquisition, M.C.C. All authors (T.-H.T.V., K.J.R., L.V.H., C.J.A., and M.C.C.) discussed the results and helped critically revise the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Institute on Aging (K01AG053477 to M.C.C.).

Institutional Review Board Statement: UKB ethical approval was from the National Research Ethics Service Committee North West–Haydock (approval letter dated 17 June 2011, Ref 11/NW/0382), and all study procedures were performed in accordance with the World Medical Association Declaration
of Helsinki ethical principles for medical research. The current analysis was approved under the UK Biobank application #21394.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data that support the findings of this study are available from the UK Biobank upon approved request.

Acknowledgments: This research was conducted using the UK Biobank Resource (Application 21394). Computations in this manuscript were run on the Quest cluster supported in part through the computational resources and staff contributions provided for the Quest high performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Johns Hopkins University Coronavirus Resource Center. Available online: https://coronavirus.jhu.edu/ (accessed on 21 December 2020).
2. NIA. Notice of Special Interest. Available online: https://grants.nih.gov/grants/guide/notice-files/NOT-AG-20-022.html (accessed on 16 March 2020).
3. Calder, P.C. Nutrition, immunity and COVID-19. BMJ Nutr. Prev. Health 2020, 3, 74–92. [CrossRef] [PubMed]
4. Shi, Y.; Wang, Y.; Shao, C.; Huang, J.; Gan, J.; Huang, X.; Bucli, E.; Piacentini, M.; Ippolito, G.; Melino, G. COVID-19 infection: The perspectives on immune responses. Cell Death Differ. 2020. [CrossRef] [PubMed]
5. De Frel, D.L.; Atsma, D.E.; Pijl, H.; Seidell, J.C.; Leenen, P.J.M.; Dik, W.A.; van Rossum, E.F.C. The Impact of Obesity and Lifestyle on the Immune System and Susceptibility to Infections Such as COVID-19. Front. Nutr. 2020, 7, 597600. [CrossRef] [PubMed]
6. Zabetakis, I.; Lordan, R.; Norton, C.; Tsoupras, A. COVID-19: The Inflammation Link and the Role of Nutrition in Potential Mitigation. Nutrients 2020, 12, 1466. [CrossRef] [PubMed]
7. Childs, C.E.; Calder, P.C.; Miles, E.A. Diet and Immune Function. Nutrients 2019, 11, 1933. [CrossRef]
8. Greene, M.W.; Roberts, A.P.; Frue, A.D. Negative Association between Mediterranean Diet Adherence and COVID-19 Cases and Related Deaths in Spain and 23 OECD Countries: An Ecological Study. Front. Nutr. 2021, 8, 591964. [CrossRef]
9. Fonseca, S.C.; Rivas, I.; Romaguera, D.; Quijal-Zamorano, M.; Czarlewski, W.; Vidal, A.; Fonseca, J.; Ballester, J.; Anto, J.M.; Basagana, X.; et al. Association between consumption of vegetables and COVID-19 mortality at a country level in Europe. medRxiv 2020. [CrossRef]
10. Fonseca, S.C.; Rivas, I.; Romaguera, D.; Quijal-Zamorano, M.; Czarlewski, W.; Vidal, A.; Fonseca, J.; Ballester, J.; Anto, J.M.; Basagana, X.; et al. Association between consumption of fermented vegetables and COVID-19 mortality at a country level in Europe. medRxiv 2020. [CrossRef]
11. Louca, P.; Murray, B.; Klaser, K.; Graham, M.S.; Mazidi, M.; Leeming, E.R.; Thompson, E.; Bowyer, R.; Drew, D.A.; Nguyen, L.H.; et al. Dietary supplements during the COVID-19 pandemic: Insights from 1.4M users of the COVID Symptom Study app—A longitudinal app-based community survey. medRxiv 2020. [CrossRef]
12. Boushey, C.J.; Coulston, A.M.; Rock, C.L.; Monsen, E. Nutrition in the Prevention and Treatment of Disease; Elsevier: Amsterdam, The Netherlands, 2001.
13. UK Biobank. Data Showcase. Available online: https://biobank.ctsu.ox.ac.uk/crystal/browse.cgi?id=100080&cd=category (accessed on 30 December 2020).
14. Armstrong, J.; Rudkin, J.K.; Allen, N.; Crook, D.W.; Wilson, D.J.; Wyllie, D.H.; O’Connell, A.M. Dynamic linkage of COVID-19 test results between Public Health England’s Second Generation Surveillance System and UK Biobank. Microb. Genom. 2020, 6. [CrossRef]
15. UK Biobank. Deriving the Grid Coordinates. Available online: http://biobank.ndph.ox.ac.uk/showcase/showcase/docs/UKgrid.pdf (accessed on 30 December 2020).
16. UK Summary. People Tested Positive. Available online: https://coronavirus.data.gov.uk/ (accessed on 4 April 2021).
17. Maling, D.H. Coordinate Systems and Map Projections; Elsevier: Amsterdam, The Netherlands, 2013.
18. Bradbury, K.E.; Young, H.J.; Guo, W.; Key, T.J. Dietary assessment in UK Biobank: An evaluation of the performance of the touchscreen dietary questionnaire. J. Nutr. Sci. 2018, 7. [CrossRef]
19. Hosseini, B.; Berthon, B.S.; Saedisomeolia, A.; Starkey, M.R.; Collison, A.; Wark, P.A.B.; Wood, L.G. Effects of fruit and vegetable consumption on inflammatory biomarkers and immune cell populations: A systematic literature review and meta-analysis. Am. J. Clin. Nutr. 2018, 108, 136–155. [CrossRef] [PubMed]
20. Kamath, A.B.; Wang, L.; Das, H.; Li, L.; Reinhold, V.N.; Bukowski, J.F. Antigens in tea-beverage prime human Vgamma 2Vdelta 2 T cells in vitro and in vivo for memory and nonmemory antibacterial cytokine responses. Proc. Natl. Acad. Sci. USA 2003, 100, 6009–6014. [CrossRef] [PubMed]
21. Melamed, I.; Kark, J.D.; Spire, Z. Coffee and the immune system. Int. J. Immunopharmacol. 1990, 12, 129–134. [CrossRef]
