Introduction

There is extensive epidemiological evidence that the amount of fruit and vegetable intake has an impact on cancer prevention particularly with cancers of the digestive system. However, the differences in behavior between those that have cancer and those that just have a family history of cancer is not known. In this paper, we will investigate the differences in servings per day of fruit and vegetables and minutes of physical activity between these two groups.

These data are from the Health Information National Trends Survey (HINTS)data from the National Cancer Institute. I chose HINTS 4 Cycle 2 as a starting point for these analyses because it had questions on BMI, cancer history, questions pertaining to screening history and information on diet and exercise. I initially was interested in exploring relationships between healthy lifestyle behaviors and incidence of prostate cancer in people with and without a family history. However, there are few in the dataset that have or have had prostate cancer, therefore I am expanding the question to differences between family history of any cancer to those that have or have had any cancer. In the future, these may be better limited to those with digestive tract cancers since this is where the most evidence lies about fruit and vegetable intake.

Methods

Data

The HINTS data is a nationwide and population-based cross-sectional computer-assisted telephone survey that collects various cancer-related information, focusing on health technology and communication. It is not conducted every year and not all questions are continued each year. HINTS 4 Cycle 2 was conducted in 2012 and contained questions about cancer history, body mass index, fruit and vegetable intake and physical activity habits in adults aged 18 or older.

Study population

I included all participants that answered yes to either having a family history of cancer or having had cancer currently or in the past. I realize there could be a cohort of cancer patients with a family history, but this will be left for future analyses. XX were excluded from the sample because of missing. The American Cancer Society recommends meeting the Physical Activity Guidelines of 150 minutes of moderate intensity or 75 minutes of vigorous intensity aerobic activity per week (no mention of strength training) and consuming at least 2.5 cups of fruits and vegetables per day, which is less than the USDA recommendations. I further restricted the data to exclude those missing physical activity and fruit and vegetable serving information (fruit missing n= 135, vegetable missing n= 134). The physical activity variable is calculated from two variables in the dataset. The first question asks "In a typical week, how many days do you do any physical activity of at least moderate intensity?" (missing = 16)and the next question is "On the days that you do any physical activity or exercise of at least moderate intensity, how long ar you typically doing these activities?"(missing n=70)This question is in hours and minutes, the hours variable is converted to minutes and added to the minute variable, then mulitplied by the number of days per week. It is known that education level can mediate fruit and vegetable intake and physical activity level, so this was used as a covariate. Age and smoking status were also used as covariates. Smoking behavior questions are limited in the HINTS surveys, but there is a recoded variable available called smokeStat. This combines two variables were used to determine current smoking status. The question "Have you smoked at least 100 cigarettes in your entire life?" was used to define never smokers and "How often do you now smoke cigarettes?" was used to define this smoke Stat variable. Education is categorized by less than high school, high school graduate, some college and college graduate or higher (missing n= 89). The variables Fruit and Vegetables were combined to compare with ACS recommendations and are reported in cups. This loaded dataset was simplified in SPSS before importing to R. Missing data are coded in a variety of ways in this dataset. Most ways are with a negative number of some kind. I recoded them as NA, so they would be ignored when used in physical activity calculations.

Statistical Analysis

All results are displayed as unweighted population estimates and were conducted using R version 3.6.1 (2019). Univariate analysis included tests to determine the differences between demographics and health behavior variables differed between those with a family history of cancer and those that have had or currently have cancer. Then to determine association between cancer status and fruit and vegetable intake and then physical activity I used a multivariate logistic regression model that controlled for statistically significant or clinically meaningful variables. The two models were compared using ROC curves and confusion matrices.

Table 1. Characteristics overall and by current cancer status

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Characteristic | Overall (n=3630) | Current Cancer (n=464) | No Current Cancer (n=3135) | p-value |
| Age [mean(SD)] | 51.8 (19.7) | 64.7 (15.7) | 50.4 (18.9) | <0.001 |
| Gender |  |  |  | 0.090 |
| (%male) | 1390 (38%) | 3(10%) | 200(43%) |  |
| (%female) | 2172 (60%) | 18(58%) | 258(56%) |  |
| Missing | 68 (2%) | 10(32%) | 6(1%) |  |
| Race/Ethnicity |  |  |  | <0.001 |
| Hispanic | 511(14%) |  |  |  |
| Non-Hispanic White | 2043(56%) | 4 (12.9) | 320 (69.0) |  |
| Non-Hispanic Black or African American | 496(14%) | 1 (3.2) | 39 (8.4) |  |
| Non-Hispanic American Indian or Alaska Native | 14(<1%) | 0 (0.0) | 2 (0.4) |  |
| Non-Hispanic Asian | 98(3%) | 1 (3.2) | 6 (1.3) |  |
| Non-Hispanic Native Hawaiian or other Pacific Islander | 6(<1%) | 0 (0.0) | 0 (0.0) |  |
| Non-Hispanic Multiple Races Mentioned | 90(3%) | 0 (0.0) | 10 (2.2) |  |
| Missing | 372(10%) | 24 (77%) | 45(10%) |  |
| Marital Status (%) |  |  |  | <0.001 |
| Married or Living as Married | 1857(51%) | 4 (13%) | 246(53%) |  |
| Divorced or Separated | 666(18%) | 2(6%) | 90(19%) |  |
| Widowed | 377(10%) | 2(7%) | 79(17%) |  |
| Single, never married | 628(17%) | 1(3%) | 40(9%) |  |
| Missing | 88(2%) | 22(71%) | 9(2%) |  |
| Education (%) |  |  |  | <0.001 |
| Less than high school | 329(9%) | 2(6%) | 42(9%) |  |
| 12 years or high school graduate | 775(21%) | 2(7%) | 106(23%) |  |
| Vocational or technical school | 267(7%) | 0(0%) | 32(7%) |  |
| Some College | 790(22%) | 3(10%) | 119(26%) |  |
| College graduate | 845(23%) | 0(0%) | 67(14%) |  |
| Postgraduate | 535(15%) | 0(0%) | 93(20%) |  |
| Missing | 89(3%) | 24(77%) | 5(1%) |  |
| Household Income |  |  |  | 0.027 |
| Less than $20,000 | 740(20%) | 6(19%) | 84(18%) |  |
| $20,000 to < $35,000 | 501(14%) | 1(3%) | 64(14%) |  |
| $35,000 to < $50,000 | 459(13%) | 1(3%) | 59(13%) |  |
| $50,000 to < $75,000 | 524(14%) | 0(0%) | 64(14%) |  |
| $75,000 or More | 926(26%) | 0(0%) | 111(24%) |  |
| Missing | 480(13%) | 23(74%) | 82(18%) |  |
| Smoking Status (%) |  |  |  | <0.001 |
| Current | 586(16%) | 3(10%) | 63(14%) |  |
| Former | 939(26%) | 6(19%) | 165(36%) |  |
| Never | 2052(57%) | 14(45%) | 229(49%) |  |
| Missing | 53(2%) | 8(26%) | 7(2%) |  |
| Fruit (cups) [mean(SD)] | 2.0 (2.4) | 2.0 (2.5) | 2.0 (2.4) | 0.938 |
| Vegetables (cups) [mean(SD)] | 2.3 (2.5) | 2.3 (2.6) | 2.3 (2.4) | 0.939 |

Table 2. Characteristic differences between those with Family History of Cancer and those without

|  |  |  |  |
| --- | --- | --- | --- |
| Characteristic | Family History of Cancer (n=2412) | No Family History of Cancer (n=870) | P-value |
| Age [mean(SD)] | 53.3 (17.5) | 52.4 (18.6) | 0.167 |
| Gender |  |  | <0.001 |
| (%male) | 888 (37%) | 377 (43%) |  |
| (%female) | 1498 (62%) | 477 (55%) |  |
| Missing | 26 (1%) | 16 (2%) |  |
| Race/Ethnicity |  |  | <0.001 |
| Hispanic | 297 (12%) | 155 (18%) |  |
| Non-Hispanic White | 1487 (62%) | 442 (51%) |  |
| Non-Hispanic Black or African American | 309 (13%) | 126 (15%) |  |
| Non-Hispanic American Indian or Alaska Native | 7 (0.3%) | 4 (0.5%) |  |
| Non-Hispanic Asian | 41 (2%) | 51 (6%) |  |
| Non-Hispanic Native Hawaiian or other Pacific Islander | 4 (0.2%) | 2 (0.2%) |  |
| Non-Hispanic Multiple Races Mentioned | 65 (3%) | 20 (2%) |  |
| Missing | 202 (8%) | 70 (8%) |  |
| Marital Status (%) |  |  | 0.432 |
| Married or Living as Married | 1246 (52%) | 469 (54%) |  |
| Divorced or Separated | 440 (18%) | 157 (18%) |  |
| Widowed | 260 (11%) | 79 (9%) |  |
| Single, never married | 431 (18%) | 148 (17%) |  |
| Missing | 29 (1%) | 17 (2%) |  |
| Education (%) |  |  | 0.674 |
| Less than high school | 200 (9%) | 76 (9%) |  |
| 12 years or high school graduate | 505 (21%) | 187 (22%) |  |
| Vocational or technical school | 166 (7%) | 73 (8%) |  |
| Some College | 547 (23%) | 173 (20%) |  |
| College graduate | 582 (24%) | 216 (25%) |  |
| Postgraduate | 381 (16%) | 136 (16%) |  |
| Missing | 21 (1%) | 9 (1%) |  |
| Household Income |  |  | 0.322 |
| Less than $20,000 | 469 (19%) | 164 (19%) |  |
| $20,000 to < $35,000 | 339 (14%) | 120 (14%) |  |
| $35,000 to < $50,000 | 298 (12%) | 124 (14%) |  |
| $50,000 to < $75,000 | 368 (15%) | 121 (14%) |  |
| $75,000 or More | 631 (26%) | 247 (28%) |  |
| Missing | 307 (13%) | 94 (11%) |  |
| Smoking Status (%) |  |  | <0.001 |
| Current | 413 (17%) | 113 (13%) |  |
| Former | 652 (27%) | 200 (23%) |  |
| Never | 1322 (55%) | 543 (62%) |  |
| Missing | 25 (1%) | 14 (2%) |  |
| Fruit (cups) [mean(SD)] | 2.0 (2.3) | 2.1 (2.4) | 0.380 |
| Vegetables (cups) [mean(SD)] | 2.4 (2.4) | 2.3 (2.5) | 0.465 |

Results

There are 3630 total survey participants in these data.