THE EFFECT OF DIETARY INTAKE ON COLON CANCER AND WOMEN'S BREAST CANCER MORTALITY RATE

by

Emman Dawood

A Thesis Submitted in Partial Fulfillment

of the Requirements for the degree

Master of Science

West Texas A&M University

Canyon, Texas

August 2019

Abstract

Diet has been known to play a key role as a risk factor for chronic diseases. Female breast, and colorectal cancers are the top three cancer types in terms of incidence and are ranked within the top five in terms of mortality. The goal of this analysis is to identify the dietary types that effect colon cancer and breast cancer in women mortality rates for the developed countries. Results: Positive association between animal fat intake and colon cancer mortality. The socioeconomic factor that include infant mortality rate, crude birth rate, fertility rate was negatively associated with colon cancer mortality rate. Animal fat and starchy roots were positively associated with breast cancer mortality rate.

ACKNOWLEDGMENTS

I would like to pay special thankfulness, warmth and appreciation to D. Lockwood who made my research successful and assisted me at every point to reach my goal. I would like to thank the committee members for their help and guidance.

Approved:

Member, Thesis Committee **Date**

Member, Thesis Committee **Date**

Member, Thesis Committee **Date**

Dean, Academic College **Date**

Dean, Graduate School **Date**

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Introduction

Maintaining healthy dietary habits throughout one's life helps to prevent malnutrition in all its forms as well as many conditions and non-communicable diseases (NCDs). However, rising incomes, increased production of processed foods, and rapid urbanization have led to a global shift in lifestyle habits, in which traditional diets are replaced by diets higher in refined sugars, refined fats, oils, and meats [1]. As a consequence, individuals consume less fruit, vegetables, and other dietary fibers such as whole grains [2].

Diet has been known to play a key role as a risk factor for chronic diseases. Existing epidemiological evidence clearly shows that many individuals in the United States have suboptimal diets, and that there is great potential for disease prevention by improved nutrition [4]. The diet that based on consuming foods such as fruits, vegetables, nuts, beans, and fish like the Mediterranean-style diet is well documented to protect against chronic diseases [3]. The traditional Mediterranean-style diet has a higher fat content (about 40 percent of calories) than that of the typical American diet (34 percent of calories), but most of the fat from this diet comes from olive oil and other plant sources.

Cancer is the second leading cause of mortality in the world currently, second only to cardiovascular disease, as it caused 9.6 million deaths in 2018. Female breast, and colorectal cancers are the top three cancer types in terms of incidence and are ranked within the top five in terms of mortality [5]. Cancers of the female breast is the leading type worldwide in terms of the number of new cases. For breast cancer, approximately 2.1 million diagnoses have been estimated for 2018, contributing about 11.6% of the total cancer incidence burden. Colorectal cancer (1.8 million cases, 10.2% of the total burden) is the third most commonly diagnosed cancer [6].

The research conducted at the National Cancer Institute (NCI) in the United States indicated that the food component of daily diet may change (increase or reduce) the risk of cancer, but that they are not solely responsible for causing cancer in human beings. The role of fluoride and vitamins, particularly Vitamin D, has been helpful in decreasing the risk of cancer, as some research studies of the NCI have revealed [7]. The NCI presented probable effects of food and dietary habits in the form of a depiction; the favorable diets are presented in blue while the likely harmful food items are indicated in green (Figure 1). From the NCI's conclusions in this study, the trends for garlic, fish, and calcium were inconclusive.

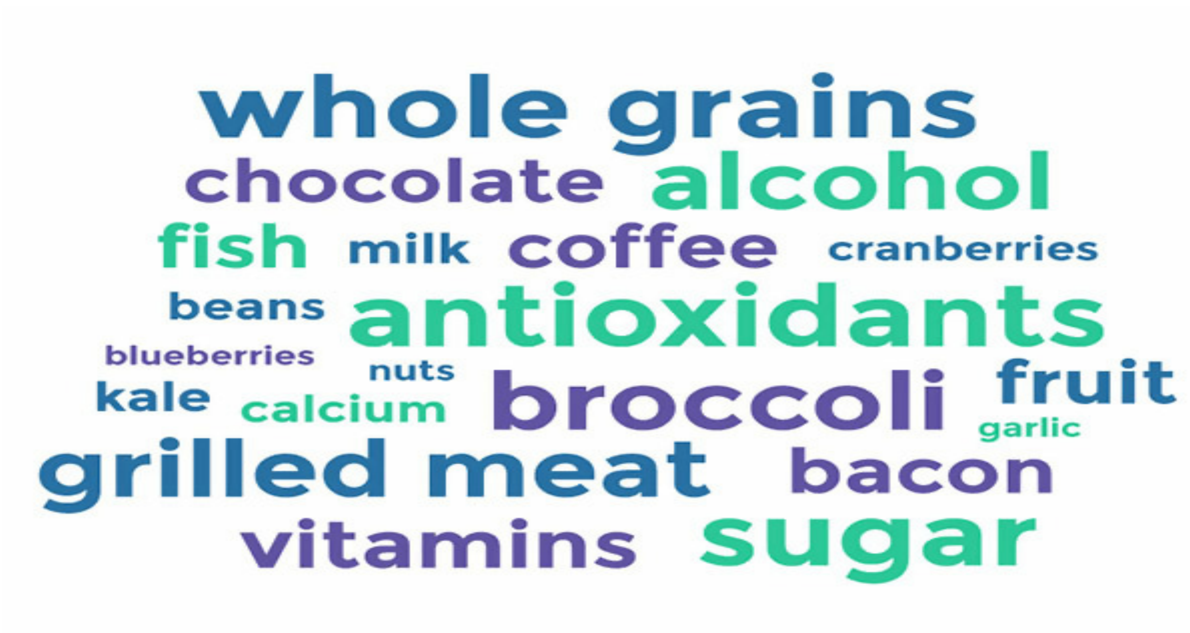

Figure 1: Daily diet increasing (green) or decreasing (blue) chances of various cancers (National Cancer Institute).

The major objective of this investigation was to identify the significant food variables, such as dietary habits, which increase or decrease the risk of colon cancer incidence in males and breast cancer incidence in females. The goal is to provide guidance to allow individuals to learn about how to maintain a healthy diet and prevent the dire consequences associated with the failure to do so. As a result, promoting the effects of maintaining a healthy lifestyle may allow for a reduction of the annual death toll from these diseases.

Literature review

In industrialized nations throughout the twentieth century, changes have been observed in the quality of nutrition programs, policy, and research. The goals of these activities throughout the century have ranged from promoting the identification and prevention of nutrient deficiency diseases to applying knowledge of nutritional requirements to everyday life to focusing on the increase in mortality due to infectious and chronic diseases. Since then, society has focused its attention on investigating the role of diet in maintaining a healthy lifestyle and reducing the risk of chronic diseases such as cancer and heart disease. As a result, epidemiological, clinical, and laboratory research have since shown that diet plays a crucial role in the etiology of these chronic conditions [4].

The relative importance of cancer as a cause of death is increasing, mostly because of the increasing number of older individuals in the population, and partly because of reduction in mortality due to other causes, such as infectious diseases. The rapid urbanization of developing countries allows for the observation of stronger patterns of cancer in more economically developed countries. Between 2000 and 2020, the total number of cases of cancer in the developing world is predicted to increase by 73% and, in the developed world, to increase by 29%, largely as a result of an increase in the number of individuals surviving to an older age [2]. The incidence of cancers of the lung, colon and rectum, breast and prostate is directly proportional to economic development [2]. Poverty factors like infant mortality rates, lack of sanitation, and lack of safe drinking water is inversely associated with prostate cancer [27]. About two-thirds of colorectal cancer cases and about 60 per cent of colorectal cancer deaths occur in countries characterized by high or very high indices of development and/ or income Over the next 15 years, the global burden of colorectal of cancer is expected to increase by 60 per cent to more than 2.2 million new cases and 1.1 million deaths. Colorectal cancer is considered one of the clearest markers of epidemiological and nutritional transition, with incidence rates of this cancer [29].

Breast cancer is the most common cancer in women worldwide, with nearly 1.7 million new cases diagnosed in 2012, representing about 25 per cent of all cancers in women. It is also the most frequent cause of cancer death in women from regions characterized by lower indices of development and/or income (14.3 per cent of deaths), and the second most frequent from regions characterized by higher indices of development and/or income (15.4 per cent of deaths), after lung cancer. Studies of women who migrate from areas of low risk to areas of high risk show that they assume the rate in the host country within one or two generations. This shows that environmental factors are important in the development of the disease [30].

Dietary habits strongly influence different lifestyle outcomes, such as the consumption of red meat and incidence of colon cancer [12,13,8] and with the red and the processed meat [16,17,18,21,22,25]. The red meat intake was linked with breast cancer risk [10,14] and with the red and processed meat [16,17]. Estimates show that around 30% of cancers in the industrialized nations are influenced by dietary habits, and up to 80% of large bowel, breast, and prostate cancers are accounted for by poor dietary choices [2,17]. Western style diet that is high in red and processed meat, sweet, desserts, french fries, and refined grains has shown a positive significant association with the risk of colon cancer [23]. There was association between western (fat- rich) pattern and the risk of breast cancer while a salad vegetable dietary pattern is associated with reduction in breast cancer risk [11]. As a result, the lifestyle factor of dietary habits is second only to tobacco consumption as a preventive measure for cancer [2].

Different analyses show strong correlations between per capita consumption of fat and colon cancer risk [4,20,24,26]. Intake of animal fat mainly from red meat and high fat dairy food during premenopausal years is associated with an increased risk of breast cancer [10]. Other studies linked dietary fat intake with the breast cancer [8,14,19]. The study of P. Correa demonstrated strong and consistent correlations between death rates of cancers of colon and breast and the per capita of total fat and nutrients derived from animal sources [20]. In contrast, a review of studies on the relationship between dietary fat intake and breast cancer among women was inconsistent [10]. The study of Hebert et.al. didn't support the role of fat intake and breast cancer, it supported the role of red meat intake and colon cancer but not due to its fat content [27]. Kushi, L. & Giovannucci, E. concluded that consumption of whole grains, vegetables, fruits and decreasing red meat intake are likely be more effective in decreasing risk of breast and colorectal cancers than decreasing total fat intake [13]. Likewise, a review of large cohort study in the U.S. on men found that animal fat, including dairy products, poultry, and fish as well as vegetable fat, were slightly inversely related to risk of colon cancer and no clear association between fiber or vegetables intake and risk of colon cancer [25]. A systematic review and meta- analysis studies that Glycemic load and carbohydrate intake were associated with increased risk of breast cancer among women with hormone receptor negative tumor [15], contradict another study of negative association between carbohydrate intake and breast cancer [19].

According to the American Institute for Cancer Research that consuming wholegrains, dietary fiber will decrease the risk of colon cancer and consuming red and processed meat and two or more alcoholic drink per day increases the risk of colon cancer [29]. In contrast on a meta-analysis investigation concluded that vegetarian in low income countries have not shown low death rates of colorectal cancer [12], the same investigation found that consuming meat, dairy products, fruits, and vegetable, fiber are inconsistent with the risk of breast cancer.

Furthermore, the incidence of colorectal cancer is approximately ten-fold higher in developed countries than in developing countries, and studies suggest that diet-related factors may account for up to 80% of the differences in rates between countries [2]. Likewise, physical activity has consistently been associated with a reduced risk of colon cancer (but not of rectal cancer). There is a strong evidence that consuming alcohol is associated to breast cancer, while limited evidence that non-starchy vegetables, and diet high in calcium, food contains carotenoids decreases the risk of breast cancer according to the American Institute for Cancer Research [30].

Some components of a “westernized” diet have frequently been associated with increasing or decreasing risks of cancer such as the increased risk associated with greater consumption of meat or fats and decreased risk associated with greater consumption of fruits, vegetables, fiber, folate, and calcium. None of these hypotheses, however, have been firmly established [2].

Data Sources

Data for cancer mortality were obtained from the World Health Organization (WHO) [31]. The WHO provides an enormous amount of information regarding public health. Its databases allow for the sorting of data by cause of death, gender, and age group ranges by country. The respective response variables for this study are breast cancer mortality for females aged 35-74 and colon cancer mortality for males and females aged 50-74 represented in table 1.

Table 1*:* Colon cancer deaths per 100,000population 50-74 years and Breast cancer deaths per 100,000 female population 35-74 years.

Country	colon cancer mortality rate	breast cancer mortality rate
Australia	13.8318	32.9369
Austria	22.6922	20.2323
Belgium	26.9449	45.1783
Canada	24.8463	33.1866
Chile	18.0777	24.1219
Cyprus	14.4723	26.7275
Denmark	35.4616	45.0583
Estonia	27.167	41.0175
Finland	17.9441	37.1153
France	23.6704	38.5857
Germany	25.8284	43.4651
Hungary	54.154	47.6323
Iceland	21.6966	29.234
Ireland	20.621	39.8476
Israel	25.5275	38.5694
Italy	29.2726	37.7052
Japan	26.8234	24.8633
Luxembourg	27.0495	33.7444
Malta	32.3107	48.3183
Mexico	10.9037	20.1232
Netherlands	35.0563	44.4458
Newzealands	30.6514	39.598
Norway	32.8296	29.5761
Poland	32.2002	37.3362
Portugal	34.3367	32.9636
Slovakia	35.8791	39.2526
Slovenia	40.5085	36.9345
Spain	33.1519	27.2098
Sweden	24.9936	33.5874
Switzerland	20.3982	36.0736
UK	21.0402	40.0419
USA	22.9113	34.5436

The age range for Colon cancer and Breast cancer were chosen where cancer mortality rates and reported data were highest. Cancer mortality rates were averaged for a five-year period from 2009 through 2013. The true response variables of death specific age per 100,000 individuals in the population was calculated as the number of deaths was divided by the size of the population at risk multiplied by 100,000.

National nutritional data was obtained from Food Balance Sheets for the period from 1999 through 2003. This data sheet was compiled and published by the Food and Agriculture Organization (FAO) of the United Nations [32]. The FAO provides a profound amount of information that allows for the selection of a wide variety of foods with defining characteristics such as daily kilocalories (kcal) per caput, the daily grams of fat per caput, and more. The average food intake between 1999-2003 was calculated for each country. Depicts all of the food variables utilized in this study. The Food Balance Sheet presents a comprehensive picture of a country's food supply pattern during a specific reference period. The Food Balance Sheet shows for each food item, i.e. each primary commodity and a number of processed commodities potentially available for human consumption, the source of supply and its utilization. The total quality of food items produced in a country, added to the total quantity imported and adjusted to any change in stocks that may have occurred since the beginning of the reference period gives available during that period. On the utilization side, a distinction is made between the quantities exported, fed to livestock, used for seed, put to manufacture for food use and non-food uses, losses during storage and transportation, and food supplies available for human consumption. The per caput supply of each food item available for human consumption is then obtained by dividing the respective quantity by the related data for the population actually partaking of it. Data on per caput food supplies are expressed in terms of quantity and by applying appropriated food consumption factors for all primary and processed products, also in terms of caloric value and protein and fat content. Table 2.A and 2.B illustrate food variables for 32 countries.

Table 2.A: Dietary Variables for the 32 countries.

Country	Total energy *	Total fat **	Animal fat energy *	Animal energy *	Animal protein **	Meat energ y *	Milk energy *	Egg energy *	Fish energy *	Vegetable energy *
Australia	3046.8	131.382	124.6	981.8	68.414	471.8	301.6	21.6	31	67.2
Austria	3677.6	156.494	326	1147.4	63.774	402.8	342.4	49.2	23.4	60.8
Belgium	3730.5	163.585	405	1185.25	60.8425	296	380.75	45.75	47	117.5
Canada	3504	146.074	232.8	936.6	59.834	385.2	232.6	43	38.6	90.4
Chile	2858.4	83.368	57	627.4	38.51	355.6	158.2	21.4	27.2	78
Cyprus	2656.4	107.872	50	775.8	51.556	357.6	280.6	39.8	36.2	90
Denmark	3339.2	132.384	412	1159.2	65.912	289.6	292.4	63.6	50.8	67.4
Estonia	3064.6	95.754	104.8	862.2	52.952	272.8	396.6	43.8	33.6	55.4
Finland	3160.2	126.194	151.4	1173	63.542	485.4	432.4	34.2	64	48
France	3618.4	169.822	289	1370.4	78.172	540.8	385.4	60.8	66.8	79.2
Germany	3359	139.774	307.2	1037.4	57.676	351.2	283.6	47.4	38.2	64.6
Hungary	3080.2	130.598	322.6	1000.2	51.166	379.8	219	62.4	8.4	84.8
Iceland	3182	132.562	162.6	1386.2	90.81	518.8	496	28.6	155.8	35.65
Ireland	3665.6	136.536	183.8	1198.4	74.822	436.2	472.6	27.8	43.4	66.6
Israel	3582.4	133.952	31.8	734.4	67.19	361.6	251	38.4	30.4	131.8
Italy	3663.2	155.186	165.6	950.6	61.73	396.6	285.2	46.4	45.4	103.6
Japan	2876	89.108	38.4	597.8	54.636	169.2	125.4	77.2	176.8	78.6
Luxembourg	3484	150.122	44.25	1206.5	72.33	620.7 5	437.5	39.75	45	67
Malta	3351.8	106.682	200.8	906.8	57.966	284.8	303.2	49.2	59.6	125.8
Mexico	3063.2	84.772	61.4	568	37.456	258.6	160	53.8	19.4	44.2
Netherlands	3238.4	142.438	149.2	1144.4	74.1025	432.8	441.4	67.2	47	77.6
Newzealands	3143.6	90.792	291.2	963.2	53.082	444.6	129.2	37.4	45.8	98.2
Norway	3422	142.73	256.2	1139.8	64.082	362.2	349	37.8	128	49.6
Poland	3414.8	114.53	229	915.8	50.414	356.2	250.2	43	28.6	78
Portugal	3540	133.84	238.2	1032.8	67.614	384.4	270.4	37.4	84.2	111
Slovakia	2807.2	103.828	282	736.6	35.028	238	149	45.8	12.8	54.8
Slovenia	3073.4	113.044	156.8	895	59.132	347.6	315.4	40	14.4	54
Spain	3335.8	152.036	69	955.2	73.158	471.2	250.6	54.4	90.8	106.8
Sweden	3119	124.14	214	1083.2	70.156	319.2	426	42.8	75	53.8
Switzerland	3402.4	151.252	223	1149	58.596	466	379.4	38.6	27	65.4

UK	3395.6	138.776	144.8	1015.4	56.108	448.4	342.2	37.6	34	67
USA	3739	154.848	104.8	1028.4	73.382	447	387.4	55.6	31.2	80.4

Table 2.A 1:Dietary variables for 32 countries.

***:** Energy in Kcal/Day/Caput
**: Gram of fat/day/caput

Table 2.B: Dietary variables for 32 countries.

Country	Vegetable oil energy *	Starchy roots energy *	Sugar energy *	Cereal energy *	Pulses energy *	Tree nut energy *	Oil crops energy *	Fruit energy *	Alcohol energy *
Australia	451	93.4	409	695.6	11	29	34.4	115.8	140.6
Austria	458.8	113.8	441.6	926.2	6.8	37.6	37.4	156.8	262.2
Belgium	534.5	203.75	515.25	797.75	21.25	44	20	82.25	192.75
Canada	554	133	486.6	863.4	70.2	22.6	67	118.6	136.4
Chile	249.4	124.2	439.2	1154.6	37.6	4.6	8	66	62
Cyprus	297.6	66.8	346.2	665	34.6	31	76	117.8	96.6
Denmark	243.8	139	469.6	808.2	11.2	33.8	9	118.4	220.2
Estonia	189.8	231.2	483	859.4	10.8	8.6	3.6	89	185.8
Finland	249.2	135.6	347.4	867	11.8	7.6	15.2	87	187.4
France	428.8	119.6	382.6	878.6	18.8	24.2	21.2	91	168
Germany	400.6	133.2	433	810	9.6	36.4	29.8	114.4	248.6
Hungary	368.6	120.6	378.8	754.4	32.8	2	13.6	88.4	208.4
Iceland	183	91.2	490	638.2	8.6	6.4	23.2	98.8	108.6
Ireland	362.6	207	370.8	936.2	24	6.2	21.4	90	347.4
Israel	604.2	86	446.8	1115.6	72.2	37.6	89.6	194.6	35.8
Italy	641.6	71	296	1167.6	51.8	39.6	13.4	170.2	140.4
Japan	363.6	69.8	280.8	1108.6	17.8	10.2	121.2	52.8	140.8
Luxembourg	370.75	82	278	745	13.5	6.75	10.75	171	399.5
Malta	170.2	125.4	502	1111	50.6	27.4	40.2	114	92.4
Mexico	227.6	30	475.6	1382	117.8	14.6	27	106.8	54.6
Netherlands	445.6	171.4	449.6	542	17.4	30	23.4	140.2	173.6
Newzealands	216.2	115.6	564.4	799.4	45	12	46	150.4	122
Norway	369.6	131.4	432.8	969.8	9.8	23.2	14.8	112.2	114.8
Poland	297.8	241.6	439.2	1189.8	19.4	6	12.8	64.2	139.4
Portugal	407.2	147.6	303	999	36.2	29.6	11.8	151.2	261
Slovakia	273.8	133.6	323	944.8	18.6	9.2	21.8	69.6	181.8
Slovenia	306.4	114.2	186.8	1069.6	11.2	25.2	26.4	174.4	150.6
Spain	654.8	135.2	288.2	731.2	51	41.8	35.8	135.2	183.2

Sweden	354.8	96	434.8	775.6	14.8	16.6	21.2	101.5	136.6
Switzerland	402.2	87.4	523.8	768.6	12.4	64.5	17.2	111	182
UK	442.4	217.2	381.2	848.4	52.4	10.6	36.6	103.2	186.2
USA	627	106.4	642.2	825.8	35.6	22.8	65	120.8	159.8

***:** Energy in Kcal/Day/Caput
**: Gram of fat/day/caput

Dietary food data from 32 countries of the United Nations was collected from two categories: developed countries and countries with characteristics of both developed and developing countries. Some countries, like Qatar, Kuwait, United Arab Emirates, and Hong Kong, were excluded because they either did not report data on the Food Balance Sheet or the data on socioeconomic variables were inconsistent. Turkey was also excluded due to its geographic location, as it shares a border with Iraq and Syria, at war during the timing of data acquisition. The incidence of war negatively impacts the infrastructure of a country, the socioeconomic status of its inhabitants, and therefore disease prevalence. Turkey accepted many immigrants from these two countries at war, and the period after this time saw an increase in the rates of cancer mortality. As a result, the cancer mortality data of Turkey could have been affected by the influx of Iraqi and Syrian immigrants. Saudi Arabia was also removed from this list since they reported cancer mortality rates for only two years (2009 and 2012). Similarly, Italy did not report data of drinking water availability and level of sanitation for the years of 2000-2004; the data was instead collected from the year of 2006 for which data was available. These outliers could potentially affect the analysis outcome for many food variables. Outliers were observed with some of the food variables (palm oil, sesame oil, spices, wine, beer, pig meat) under exploratory analysis for colon cancer mortality. Outliers were also observed with other variables (rice kcal, spices kcal, pig meat kcal, pig meat gram protein) under exploratory analysis for breast cancer mortality when calculating the breast cancer single regression.

Infant mortality rate for the year 2004 is the rate of number of infants dying before reaching one year of age, per 1,000 live births in 2004.

Life expectancy at birth for the year 2000 indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life. Annual population average growth for the years (1994-2004) is the average of population counts for all residents regardless of legal status or citizenship. Fertility rate for the years (1994-1998) represents the number of children that would be born to a woman if she were to live to the end of her childbearing years and bear children in accordance with age-specific fertility rates of the specified year. GDP per capita for the year (2004), PPP GDP, is the gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United States. GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant 2011 international dollars.

Crude birth rate for the years (1994-1998) indicates the number of live births occurring during the year, per 1,000 population estimated at midyear. Subtracting the crude death rate from the crude birth rate provides the rate of natural increase, which is equal to the rate of population change in the absence of migration. Data for these variables were collected from the World Bank database [33]. The World Bank database allows for the selection of a large range of years for each country. The other socioeconomic variables were collected from WHO [34] included the number of hospital beds per 10,000 individuals in the population for the years (2000-2004), hospital beds include inpatient beds available in public, private, general, and specialized hospitals and rehabilitation centers. Average was calculated using the data from each year for each variable of the following variables: Number of medical doctors (physicians) for the years (2000-2004), including generalist and specialist medical practitioners, per 1000 population, the percentage of the population using safety managed water for the years (2000-2004), the percentage of the population using safety sanitation for the years (2000-2004). Both breast and colon cancer mortality data were collected ten years following that of the food and socioeconomic variables. The following table summarizes each of the variables described above.

Table 3: The socioeconomic variables for 32 countries.

Country	average growth *	life expectancy	infant mortality rate **	fertility rate***	crude birth rate #	GDP §	Physician density ¶	hospital beds £	percentage with managed drinking water service	percentage using safely managed sanitation services
Australia	1.18531	79.2341	4.96	1.8	13.9	39627.7	10	38	100	65
Austria	0.301218	78.1268	4.4	1.42	10.94	36583.9	38	77.5	98	97
Belgium	0.298537	77.722	4.54	1.582	11.4	37919.9	42	71.9	100	70
Canada	0.945862	79.2366	5.22	1.5961	12.14	38410.9	19	33	100	73.4
Chile	1.27552	76.793	8.44	2.27	18.8124	15083.1	11	21	92.8	34.4
Cyprus	1.9121	78.01	4.98	2.0346	15.7344	31455.1	23	40.6	96	76.2
Denmark	0.370608	76.5927	4.4	1.766	12.98	42756.6	32	39.4	94	93
Estonia	-0.838044	70.4171	7.46	1.354	9.26	18113.5	33	58.6	93.6	88
Finland	0.285654	77.4659	3.34	1.774	11.9	36370.5	33	70.3	93.4	89.6
France	0.537219	79.0561	4.16	1.754	12.72	35461.1	37	75.8	93	88.2
Germany	0.15107	77.9268	4.2	1.3	9.62	37149.5	35	85.7	99	96
Hungary	-0.222458	71.2463	7.68	1.472	10.3	19617.1	28	77.1	52.6	52.6
Iceland	0.928185	79.6537	2.88	2.086	15.86	35402.7	38	44.67	90.8	69
Ireland	1.17627	76.5366	5.38	1.888	13.94	41887.8	31	53.4	92.8	40.4
Israel	2.34476	78.9537	5.16	2.924	21.38	26074.3	36	38.1	99	81
Italy	0.135513	79.778	4.3	1.204	9.34	37084.4	37	40.9	87.6	96
Japan	0.232423	81.0761	3.02	1.4238	9.648	34219.9	21	137.5	97	98.6
Luxembourg	1.2904	77.8732	3.5	1.716	13.26	84414.7	29	58	98	93

Malta	0.705433	78.2	6.46	1.916	12.64	25099.1	34	73	100	93
Mexico	1.51208	74.364	20.36	2.9524	25.8984	15443.3	29	16	93.4	22.4
Netherlands	0.571123	77.9878	4.88	1.564	12.44	42148.2	39	46.1	100	97
Newzealands	1.22502	78.6366	5.8	1.96	15.36	29280.1	21	62	77	76
Norway	0.571785	78.6341	3.68	1.86	13.68	59567.6	39	44.7	95	76
Poland	-0.06623	73.7488	7.34	1.96	11.14	15455.3	20	65.8571	94	71.6
Portugal	0.461734	76.3146	4.74	1.448	11	26233.1	34	35.3	93.8	60
Slovakia	0.07984	73.0512	7.62	1.492	9.36	17043.8	31	70.6	93	82
Slovenia	0.024004	75.4122	4.14	1.274	11.36	24232.4	24	49.1	78.2	73.8
Spain	0.787272	78.9659	4.16	1.154	9.18	31278	38	34.1	98	94
Sweden	0.282285	79.6439	3.28	1.646	11.12	38233.4	36	30.4	98	91
Switzerland	0.572957	79.6805	4.56	1.484	11.56	50934.7	40	56.5	93	98
UK	0.350575	77.7415	5.36	1.722	12.6	34773.1	21	37.4	96	97
USA	1.08307	76.6366	6.94	1.9851	14.5	46837.5	27	31	100	89

Percentage of annual growth. **: Rate per 1000 lives at birth. ***: Total birth per women. #: Birth rate per 1000 population. §: per capita in international dollars. ¶: Physicians per 1000 people. £: Hospital beds per 1000 population

Statistical Analysis

Two Multi- variable regression models were developed for each response variable: 1) colon cancer age truncated mortality rate and 2) women's breast cancer age truncated mortality rate. Initially, all variables were inspected for adherence to assumptions for linear regression, investigating distribution, presence of outliers, independence and normality.

Factor Analysis was used to reduce the large collection of socioeconomic predictor variables to a smaller set of orthogonal predictors which retained the variance of the full set. Similarly, factor analysis was used to examine interrelationships between the large set of food variables. Factor analysis identified food variables that were not linear combinations of the others by selecting one variable from each factor with the highest load or weight.

Six different regression model development strategies were used to reach a final model for colon cancer mortality rate and breast cancer mortality rate, including a manual elimination regression method, stepwise regression methods (Forward AIC model, Backward AIC model, Forward BIC, Backward BIC model), and the best subset regression method. Each method uses different criteria to calculate and select variables in the final model [35]. For each outcome variable of breast and colon cancer mortality, a final model was selected for discussion and comparison with prior knowledge from the literature.

Table 4 displays summery statistics for the dependent variable age-truncated colon cancer mortality rate, summary statistics for the dependent variable sex-specific age-truncated breast cancer rate, and food and socioeconomic independent variables to be used in the statistical analysis. The residuals of multivariable regression models met the assumption of normality.

Results

Table 4: Summary statistics for colon cancer specific age mortality rate, Female Breast cancer specific age mortality rate, food, and socioeconomic factors.

Variable	Mean	Standard deviation	Minimum	Maximum
Colon Cancer Mortality	26.97665	8.598842	10.90374	54.15395
Breast Cancer Mortality	35.60082	7.472302	20.12321	48.31826
Total Energy	3299.834	287.2945	2656.4	3739
Total Fat	129.2024	23.90899	83.368	169.822
Animal energy	995.7547	203.6347	568	1386.2
Animal Fat	188.4141	106.2241	31.8	412
Animal protein	61.37953	12.04258	35.028	90.81
Meat Energy	386.0234	95.09193	169.2	620.75
Milk Energy	310.2078	102.0572	125.4	496
Egg Energy	44.74062	12.61685	21.4	77.2
Fish Energy	51.86875	38.94392	8.4	176.8
Vegetable Energy	76.6609	24.16609	35.65	131.8
Vegetable oil Energy	379.6078	137.5355	170.2	654.8
Starchy root Energy	127.3172	49.61712	30	241.6
Sugar Energy	413.7891	95.94951	186.8	642.2
cereal Energy	898.3859	187.0473	542	1382

pulses Energy	29.89219	24.44111	6.8	117.8
Tree nuts oil Energy	22.55156	14.81311	2	64.5
Oil crops Energy	31.71094	26.2022	3.6	121.2
Fruit Energy	114.9234	34.77906	52.8	194.6
Alcohol Energy	169.3578	76.76202	35.8	399.5
Annual Average Growth	0.6397217	0.6454695	-0.8380437	2.344765
Life expectancy	77.33488	2.475199	70.41708	81.0761
Infant Mortality Rate	5.541875	3.076632	2.88	20.36
Fertility Rate	1.743187	0.4155047	1.154	2.9524
Crude Birth Rate	12.96791	3.620781	9.18	25.8984
GDP	34506.01	14018.11	15083	84414.73
Physicians Density	30.1875	8.372141	10	42
Hospital Beds	53.54772	23.70316	16	137.5
Population Drink Safe Water	93.34375	9.236706	52.6	100
Population Use Safe Sanitation	78.81875	19.5916	22.4	98.6

Results of preliminary univariate analysis for all considered predictor variables against the response variable of age specific and truncated colon cancer mortality rate are displayed in Table 5. Displayed in Table 6 are results from univariate analysis for breast cancer age specific and truncated mortality rate versus all considered food and economic predictor variables under consideration. The tables display the associated simple linear regression coefficient (b), its 95% confidence interval, along with the correlation coefficient (r) between mortality rate and the identified predictor variable.

Displayed in Table 6 are the univariable regression with female cancer mortality rate of individuals aged 35-74.

Table 5: Results of a single linear regression coefficient (b) along with the 95% confidence interval and correlation coefficient for candidate independent variables with age-specific and truncated colon cancer mortality rates.

Variable	B	lower	upper	R	p-value
Total fat energy	0.5723058	-2.626803	3.771414	0.0666	0.7174
Total energy	0.3019852	-2.902255	3.506225	0.0351	0.8487
Animal energy	0.6506148	-2.546412	3.847642	0.0757	0.6806
animal protein energy	-0.2599001	-3.464653	2.944853	-0.0302	0.8696
Animal fat energy	3.621778	0.7138329	6.529724	0.4212	0.0164
Meat energy	-1.487927	-4.64578	1.669925	-0.173	0.3436
Milk energy	-1.438237	-4.599289	1.722815	-0.1673	0.3602
Egg energy	3.615144	0.7060511	6.524237	0.4204	0.0166
Fish energy	-0.197499	-3.402871	3.007873	-0.023	0.9007
Vegetable energy	1.956531	-1.165588	5.07865	0.2275	0.2104
Vegetable oil energy	0.1694737	-3.036121	3.375069	0.0197	0.9147
Starchy roots energy	2.090058	-1.020008	5.200123	0.2431	0.1801
Sugar energy	-2.092278	-5.202135	1.01758	-0.2433	0.1796
Cereal energy	-0.585373	-3.784153	2.613407	-0.0681	0.7112
Pulses energy	-1.709041	-4.851294	1.433212	-0.1988	0.2755
Treenut energy	-0.5547955	-3.754333	2.644742	-0.0645	0.7257
Oilcrops energy	-1.764262	-4.902269	1.373745	-0.2052	0.26
Alcohol energy	1.619763	-1.529058	4.768583	0.1884	0.3019
Fruits energy	1.091612	-2.088665	4.27189	0.1269	0.4887
socioeconomic factor1	-4.595261	-7.346602	-1.843919	-0.5286	0.0019
socioeconomic factor2	-1.900717	-5.285403	1.483969	-0.2049	0.2605

Table 6: Result of a single linear regression coefficient (b) with 95% confidence interval and correlation coefficient for candidate independent variables with age female age-specific and truncated breast cancer mortality rates.

Variable	B	lower	upper	R	p-value
Total fat energy	1.471033	-1.260613	4.202679	0.1969	0.2802
Total energy	1.685309	-1.029071	4.399689	0.2255	0.2145
Animal energy	2.312552	-0.3368299	4.961934	0.3095	0.0848
animal protein energy	0.6226537	-2.153826	3.399133	0.0833	0.6503
Animal fat energy	3.52369	1.066761	0.980618	0.4716	0.0064
Meat energy	-0.0981928	-2.884121	2.687736	-0.0131	0.9431
Milk energy	1.594868	-1.127098	4.316835	0.2134	0.2408
Egg energy	1.155269	-1.597399	3.907938	0.1546	0.3982

Fish energy	-2.125084	-4.796204	0.5460368	-0.2844	0.1147
Vegetable energy	2.163799	-0.5029964	4.830595	0.2896	0.1079
Vegetable oil energy	-0.2618997	-3.046357	2.522558	-0.035	0.849
Starchy roots energy	3.638289	1.204694	6.071885	0.4869	0.0047
Sugar energy	1.122776	-1.631761	3.877313	0.1503	0.4117
Cereal energy	-1.872324	-4.569611	0.8249637	-0.2506	0.1666
Pulses energy	-1.473325	-4.204799	1.258149	-0.1972	0.2794
Tree nut energy	0.3648457	-2.418	3.147692	0.0488	0.7907
Oil crops energy	-2.01671	-4.699486	0.6660669	-0.2699	0.1352
Alcohol energy	1.607041	-1.11393	4.328012	0.2151	0.2372
Fruits energy	-0.0379109	-2.824044	2.748222	-0.0051	0.978
socioeconomic factor1	-2.608921	-5.252212	0.0343698	-0.3454	0.0529
socioeconomic factor2	-0.7196023	-3.712641	2.273437	-0.0893	0.627

Factor analysis

Factor analysis is a reduction technique. The basic assumption of factor analysis is that for a collection of observed variables there are a set of *underlying* variables called **factors** (smaller than the observed variables), that can explain the interrelationships among those variables.

The factor is a set of linear combinations of predictor variables. These factors are not correlated with each other. Each factor is associated with an eigenvalue, a factor with an eigenvalue equal to or greater than 1 is retained.

In the analysis of the socioeconomic variables, factor analysis was applied and each factor is associated with an eigenvalue. The results are shown in table 7. Two factors were retained, each one had a different load or weight of the socioeconomic variables. The variable with highest load or highest weight was retained. Factor 1 is associated with fertility rate, crude birth rate, and infant mortality; is related to poverty. Factor 2 is associated with life expectancy, the percentage of annual average growth, and GDP per capita. The Varimax rotation option creates factors that are not correlated with each other. To compute factors, a Factor command was used with a Varimax rotation option in STATA (Stata software, version 15.1). Table 8 shows the weighting of each individual socioeconomic variable in the factor analysis.

Table 7: Results of factor analysis for the socioeconomic variables and the eigenvalues associated with each factor.

Factor	Eigen Value	Difference	Proportion
Factor 1	3.66929	1.5452	0.5929
Factor 2	2.12409	1.75695	0.3432
Factor 3	0.36714	0.12336	0.0593
Factor 4	0.24378	0.10034	0.0394
Factor 5	0.14343	0.11801	0.0232
Factor 6	0.02543	0.06295	0.0041
Factor 7	-0.03752	0.0442	-0.0061
Factor 8	-0.08172	0.0207	-0.0132
Factor 9	-0.10242	0.05992	-0.0165
Factor 10	-0.16234	.	-0.0262

Table 8: Factors retained from principle component analysis

Variable	Factor 1	Factor 2
Annual Average Growth	0.6548	0.6166
Life Expectancy	-0.1868	0.8086
Infant Mortality Rate	0.7305	-0.4208
Fertility Rate	0.8948	0.24
Crude Birth Rate	0.9492	0.2538
GDP	-0.3041	0.5731
Physician Density	-0.3149	0.1406
Hospital Beds	-0.5034	-0.1476
Population Drink Safe Water	-0.0287	0.5572
Population Use Safe Sanitation	-0.7241	0.3324

Factor analysis was used (Stata version 15.1) without a rotation option, which maintains the factor-loading weights unchanged. The factor analysis was used to identify the food variables that are not linear combinations of the others by selecting one variable of each factor with the highest load or weight. Only six variables were retained from the analysis: animal energy, animal fat, fish energy, vegetable oil energy, starchy roots energy, and sugar energy. Table 9 shows the results of factor analysis and the eigenvalues associated with each variable. Table10 shows the weight of each individual food variable in factor analysis.

Table 9: Results of factor analysis for the food variables and the eigenvalues associated with each factor.

Factor	Eigen Value	Difference	Proportion
Factor 1	5.48262	2.33735	0.303
Factor 2	3.14528	1.15526	0.1738
Factor 3	1.99002	0.40991	0.11
Factor 4	1.58011	0.27549	0.0873
Factor 5	1.30462	0.26744	0.0721
Factor 6	1.03718	0.18955	0.0573
Factor 7	0.84764	0.20658	0.0468
Factor 8	0.64106	0.08931	0.0354
Factor 9	0.55174	0.05417	0.0305
Factor 10	0.49757	0.14214	0.0275
Factor 11	0.35543	0.12002	0.0196
Factor12	0.23541	0.05117	0.013
Factor 13	0.18424	0.03603	0.0102
Factor 14	0.14822	0.06889	0.0082
Factor 15	0.07933	0.05857	0.0044
Factor 16	0.02076	0.01941	0.0011
Factor 17	0.00134	0.00513	0.0001
Factor 18	-0.00378	0.00126	-0.0002
Factor 19	-0.00504	.	-0.0003

Table 10: Factors retained from principle component analysis

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Total energy	0.6783	0.4499	0.35	-0.0317	0.0424	0.3681
Total Fat energy	0.8754	0.353	0.1091	0.0616	-0.0844	-0.0143
Animal energy	0.9334	-0.2371	0.0071	0.0996	0.1032	-0.0258
Animal protein	0.8088	0.1128	-0.4035	0.1827	0.1127	0.1564
Animal Fat	0.3489	-0.194	0.7036	0.2966	-0.052	-0.2944
Meat energy	0.7033	0.0173	-0.3489	-0.4116	0.1529	0.0124
Milk energy	0.7962	-0.2728	-0.1925	-0.0502	0.1834	0.1462
Egg energy	-0.0785	0.2236	0.1201	0.5894	-0.3233	-0.1127
Fish energy	0.0863	-0.0941	-0.5321	0.6365	-0.0301	0.3086
Vegetable energy	-0.0392	0.707	0.2411	0.0806	-0.0256	0.1364
Vegetable oil energy	0.3879	0.764	0.0394	0.0338	-0.1655	0.0905
Starchy roots energy	0.2464	-0.3673	0.5519	0.0085	-0.0634	0.4547
Sugar energy	0.0737	0.0471	0.3592	0.2167	0.8395	-0.0467
Cereal energy	-0.6077	0.2082	0.2125	-0.1955	-0.1183	0.4377
Pulses energy	-0.4737	0.5582	0.0622	-0.2378	0.214	0.1902
Tree nut energy	0.3083	0.5639	0.1641	0.1082	0.0224	-0.3591
Oil crops energy	-0.3416	0.5153	-0.351	0.4044	0.0254	0.0424
Fruit energy	0.2918	0.5977	-0.1875	-0.4155	-0.0777	-0.2549
Alcohol energy	0.5985	-0.2484	0.1166	-0.149	-0.5508	0.0332

Regression Model Development

Two separate linear regression models were developed to address the objectives of this study. The first regression model described the association between colon cancer mortality per 100,000 individuals in the population (both male and female) aged 50-74 and predictor food variables controlling for socioeconomic factors. The second regression model describes the association between breast cancer mortality per 100,000 individuals in the population (female) aged 35-74 and predictor food variables. The response variables were weighted for the specific population across countries by age groups of 50-74 for colon cancer and 35-74 for breast cancer.

A log transformation was applied to the colon cancer mortality age-specific rate, for female breast cancer mortality age-specific rate, and food for the appropriate variables. The results of this transformation were not significantly different from those obtained using untransformed data. The log transformation was not retained since model assumptions were met with non-transformed data and ease in interpretation of results. All food variables and the socioeconomic variables were standardized prior to analysis.

Variable Selection in Linear Regression

Multivariate regression models were formulated in STATA using the `vselect` command. Multiple models were developed for each response variable using manual elimination, stepwise methods, and best subset regression so as to determine the most consistently appearing predictor variables across methodologies. Stepwise methods included both forward selection and backward elimination. Five relevant criteria were implemented for evaluation of linear models: Mallows's information criterion (C), Akaike's information criterion (AIC), Akaike's corrected (AICC), Bayesian information criterion (BIC), and an adjusted R- squared procedure [35].

Results

Colon Cancer

A manual elimination method results in the retention of the socioeconomic control Factor 1, animal fat energy, and sugar energy presented in Table 11. None of these variables is significant, but the animal fat energy variable was close to significance at (p= 0.066).

Table 11: Manual elimination regression results.

Variable	Coefficient	P-Value	95% Confidence interval

Factor 1	-2.688755	0.113	-6.051124	0.6736131
Animal fat energy	3.025013	0.066	-0.2159069	6.265933
Sugar energy	-1.965891	0.209	-5.100588	1.168805
Model summery				
Number of observations	32	Adjusted R-squared	0.2995	
Probability > F	0.0046	Root MSE	7.1968	
R-squared	0.3673			

A forward stepwise method using AIC criteria showed a negative association between colon cancer mortality and the socioeconomic Factor 1 (p=0.016) (forward AIC). Table 12 shows the regression results of this forward AIC method.

Table 12: Forward AIC regression results.

Variable	Coefficient	P-Value	95% Confidence interval	
Factor 1	-3.720465	0.016	-6.680992	-0.7599381
Animal fat energy	2.118526	0.15	-0.8100061	5.047059
Model summery				
Number of observations	32	Adjusted R-squared	0.2838	
Probability > F	0.003	Root MSE	7.277	
R-squared	0.33			

None of the food variables or the socioeconomic factors were significant with backward stepwise regression using the AIC criteria. Table 13 shows the results of backward AIC regression results.

Table 13: Backward AIC regression results.

Variable	Coefficient	P-Value	95% Confidence interval	
Factor 1	-2.764059	0.099	-6.084764	0.5566469
Factor 2	-1.876872	0.187	-4.718184	0.9644391
Animal fat energy	2.851119	0.079	-0.3585702	6.060808
Sugar energy	-2.063711	0.183	-5.161283	1.03386
Model summery				
Number of observations	32	Adjusted R-squared	0.3199	
Probability > F	0.0055	Root MSE	7.0915	
R-squared	0.4076			

Using forward and backward stepwise regression and BIC criteria the effects of socioeconomic Factor 1 were significant but no food variables remained in the model. Table 14 displays the results of these models.

Table 14: Forward BIC & Backward BIC regression results.

Variable	Coefficient	P-Value	95% Confidence interval	
Factor 1	-4.595261	0.002	-7.346602	-1.843919
Model summery				
Number of observations	32	Adjusted R-squared	0.2554	
Probability > F	0.0019	Root MSE	7.4198	
R-squared	0.2794			

Results of the best subset regression methodology for the colon cancer mortality model are displayed in Table 15. Best subset methods agree with stepwise methods suggesting a model with only the socioeconomic Factor 1 for inclusion or at most Factor 1 and Animal Fat Energy. Sugar energy also appears with a three variables predictor model where the socioeconomic Factor 2 appears in a four predictor variables model. All methodologies seem to support a single food variable of Animal Fat energy as positively impacting colon cancer mortality.

Table15: The best model results.

Number of predictors	Adjusted R-squared	C	AIC	AICC	BIC
1	0.2554305	0.8272445	221.0126	221.8698	223.9441
2	0.283816	0.803978	220.684	222.1654	225.0812
3	0.2995233	1.312108	220.8514	223.1591	226.7144
4	0.3198554	1.699635	220.7451	224.1051	228.0737
5	0.300092	3.485018	222.454	227.1206	231.2484
6	0.2805586	5.211973	224.0797	230.3406	234.3399
7	0.2574072	7.000571	225.787	233.9688	237.5129
8	0.2251397	9	227.7862	238.2624	240.9778
Predictors for each model					
1	Factor1				
2	Animal Fat energy Fator1				

3	Animal Fat energy Fator1 Sugar energy
4	Animal Fat energy Fator1 Sugar energy Factor2
5	Animal Fat energy Fator1 Sugar energy Factor2 Animal energy
6	Animal Fat energy Fator1 Sugar energy Factor2 Fish energy Vegetable oil energy
7	Animal Fat energy Fator1 Sugar energy Factor2 Fish energy Vegetable oil energy Animal energy
8	Animal Fat energy Fator1 Sugar energy Factor2 Fish energy Vegetable oil energy Animal energy Starchy roots energy

Breast Cancer

The manual elimination method was utilized initially and resulted in no significant coefficients. Table 16 shows the results of the manual elimination regression model

Table 16: Results of manual elimination regression model.

Variable	Coefficient	P-Value	95% Confidence interval	
Factor 1	-1.384167	0.419	-4.865576	2.097242
Factor 2	1.427928	0.457	-2.478276	5.334132
Animal energy	0.5630824	0.737	-2.862527	3.988692
Animal fat	1.547813	0.359	-1.874234	4.96986
Fish energy	-2.355853	0.11	-5.289519	0.5778126
Vegetable oil energy	-0.9642339	0.495	-3.842787	1.914319
Starchy roots energy	2.329338	0.121	-0.6590842	5.317761
Sugar energy	0.6375973	0.662	-2.340145	3.61534
Model summery				
Number of observations	32	Adjusted R-squared	0.228	
Probability > F	0.0729	Root MSE	6.5653	
R-squared	0.4273			

All of the stepwise regression methodologies and criteria resulted in the same model illustrated in Table 17. These results indicate a statistically significant positive relationship between Animal Fat Energy and Starchy Roots Energy with breast cancer mortality. Table 18 and 19 show the results of the best subset method. These results would agree with the step-wise methods above, with criteria recommending the two variables model for breast cancer mortality as predicted by Animal Fat Energy and Starchy Roots Energy. The best subset method based on the adjusted R squared criteria recommends the addition of a third variable in Fish Energy; however, Fish Energy is not statistically significant. The model of these three variable predictors is displayed in Table 18 below.

Table 17: Results of Forward (AIC, BIC), Backward (AIC, BIC).

Variable	Coefficient	P-Value	95% Confidence interval	
Animal fat energy	2.618327	0.035	0.1934949	5.043159
Starchy roots energy	2.788036	0.026	0.3632045	5.212868
Model summery				
Number of observations	32	Adjusted R-squared	0.3019	
Probability > F	0.0021	Root MSE	6.2434	
R-squared	0.3469			

Table 18: Results of the best model.

Number of predictors	Adjusted R-squared	C	AIC	AICC	BIC
1	0.2116441	2.637253	213.8541	214.7112	216.7855
2	0.3018699	0.2265159	210.8798	212.3613	215.277
3	0.3153648	0.8326765	211.1323	213.4399	216.9952
4	0.3114251	2.083588	212.1521	215.5121	219.4808
5	2937353	3.787405	213.7561	218.4228	222.5505
6	0.2794626	5.334731	215.1413	221.4022	225.4014
7	0.2564878	7.115623	216.8394	225.0212	228.5653
8	0.2280419	9	218.6789	229.1551	231.8706
Predictors for each model					
1	Starchy roots energy				
2	Starchy roots energy Animal Fat energy				
3	Fish energy Starchy roots energy Animal Fat energy				
4	Fish energy Starchy roots energy Animal Fat energy Factor2				
5	Fish energy Starchy roots energy Animal Fat energy Factor1 Factor2				
6	Fish energy Starchy roots energy Animal Fat energy Factor1 Factor2 Vegetable oil energy				

7	Fish energy Starchy roots energy Animal Fat energy Factor1 Factor2 Vegetable oil energy Sugar energy
8	Fish energy Starchy roots energy Animal Fat energy Factor1 Factor2 Vegetable oil energy Sugar energy Animal energy

Table 19: Results of best model method.

Variable	Coefficient	P-Value	95% Confidence interval	
Animal fat energy	2.498411	0.043	0.0854301	4.911392
Starchy roots energy	2.611364	0.036	0.1890885	5.033639
Fish energy	-1.413277	0.22	-3.722522	0.8959677
Model summery				
Number of observations	32	Adjusted R-squared	0.3154	
Probability > F	0.0034	Root MSE	6.1828	
R-squared	0.3816			

DISCUSSION

Colon Cancer

In this study, our results indicated a negative association with the socioeconomic control Factor1 using Forward AIC (p=0.016), therefore, at α=0.05 level of significance there is sufficient evidence to infer that for each standard deviation of the socioeconomic Factor1 away from the mean, there are expected to be 3.72 fewer deaths due to colon cancer per 100,000 individuals in the population aged 50-74. Forward and Backward BIC results also supported a Factor 1 significance with a p-value of p=0.002. The significance of Factor1 is consistent with the findings of previous research [2] which states that Colorectal cancer incidence rates are approximately ten-fold higher in developed than in developing countries. Almost 55% of colorectal cancer occur in more developed region [23]. Factor1 did not display significant in the implementation of the manual elimination method. The study showed a close but not significant association of colon cancer with

Animal fat using the manual elimination method with a p-value of p=0.066. Thus, there is sufficient evidence at the α=0.05 level to infer that for each standard deviation of animal fat energy intake (106.2241 kcal) away from the mean (88.4141), there are expected to be 3.025013 more deaths due to colon cancer per 100,000 individuals in the population aged 50-74. Implementation of the backward AIC method resulted in a p-value of p=0.079 indicating that.

Factor 1 and Animal fat association was close to significance at the α = 0.05 level. This result could likely be due to a small sample size (n=32 countries). Some previous researches indicated that Animal fat had a significant impact on colon cancer mortality rate [2,13,17], and a strong correlation between animal fat and colon cancer rate (p=0.01) [16]. There are limited studies with no evidence of the effect of animal fat energy on colon cancer [24.] This association may occur because of lower rates of animal consumption. Countries with higher levels of poverty likely experience lower rates of colon cancer due to the relative high cost of animal products.

Breast Cancer

In this study, our results showed a positive association with the consumption of animal fat (p-value =0.035) and starchy roots (potato, sweet potato, yams, and other kinds of roots) (p-value =0.026) in Forward AIC and BIC, as well as Backward AIC and BIC methods, Thus, there is sufficient evidence, at the α =0.05 level of significance ,to infer that each standard deviation of starchy roots (49.61712 kcal) away from the mean (127.3172), 2.788 more deaths are expected due to breast cancer of females aged 35-74 per 100,000 individuals in the population. This result is consistent with previous meta- analysis study that find weak positive association between carbohydrate intake and breast cancer only among women with hormone receptor negative tumors [16], in contrast another study finding was negative association between carbohydrate intake and breast cancer (r= -0.71) [20].

Additionally, at the α=0.05 level, there is sufficient evidence to infer that for each standard deviation of animal fat (106.2241 Kcal) away from the mean (188.4141), 2.62 more deaths are expected due to breast cancer of females aged 35-74 per 100,000 individuals in the population. This result support previous findings [8,12], there were correlation between total fat consumption per capita and the incidence of breast cancer (r=0.9) [15], (r= 0.94) [20], (r=0.74) [21]. Total fat consumption correlate with breast cancer mortality (r= 0.89)[22]. On the other hand, a review of studies on the relationship between dietary fat intake and breast cancer among women was inconsistent [13], and between total fat and subtypes of fat intake with breast cancer intake [10]. No supported data found between fat content and breast cancer [19]. There is limited- no conclusion studies on potatoes, total fat and saturated fat, starch, carbohydrate intake and breast cancer according to American Institute for Cancer Research [30].

High fat diet consumption cause to gain more weight than low- fat diet consumption, even when the caloric intake is similar, high fat consumption will cause a higher caloric intake. There is some evidence that cancer develops more rapidly in obese animals [4]. A higher caloric intake can lead to an increase in body fat, which puts women at a higher risk of developing breast cancer. The results of the manual elimination method did not provide sufficient evidence at the α=0.05 significance level of association found between the socioeconomic control factors and the breast cancer mortality rates. This lack of association may be due to the increasing awareness of breast cancer. Overall survival rates for breast cancer vary worldwide, but in general, they have improved. This likely occurs because access to medical care is improving in many nations, and the majority of breast cancer cases are diagnosed at an earlier and localized stage [32].

In both colon cancer and breast cancer mortality rates risks and their association with animal fat intake could be attributable to other components in food containing animal fat (e.g. red meat and high fat dairy foods) [11]. Cooking meats at high temperatures, prolonged exposure to heat and cooking by various types of grilling results in the formation of heterocyclic amines and polycyclic aromatic hydrocarbons both of which have been linked to colorectal cancer development in experimental studies. Polycyclic aromatic hydrocarbons, which are formed when organic substances like meat are burnt incompletely, may also have carcinogenic potential. Grilling (broiling) and barbecuing (charbroiling) meat, fish, or other foods with intense heat over a direct flame results in fat dropping on the hot fire, causing flames; these flames contain polycyclic aromatic hydrocarbons that stick to the surface of food [36].

High- fat dairy foods contain fat- soluble hormones growth factors, which may be related to breast cancer risk [37].

Limitations of the Study

The databases from WHO and FAO provide a wealth of information, but some countries do not report their mortality statistics accurately. Food balance sheets present a comprehensive picture of the pattern of a country's food supply during a specified reference period. A food balance sheet shows for each food item, i.e. each primary commodity and a number of processed commodities potentially available for human consumption, the sources of supply and its utilization. The total quantity of foodstuffs produced in a country added to the total quantity imported and adjusted for any change in stocks that may have occurred since the beginning of the reference period gives the supply available during that period. On the utilization side, a distinction is made between the quantities exported, fed to livestock, used for seed, put to manufacture for food use and non-food uses, lost during storage and transportation, and available as food for human consumption at the retail level. The per capita supply of each such food item available for human consumption is then obtained by dividing its respective quantity by the related data on the population actually partaking of it. Data on per capita food supplies are expressed in terms of quantity and also in terms of caloric value, protein and fat content. The accuracy of food balance sheets, which are in essence derived statistics, is of course dependent on the reliability of the underlying statistics of supply and utilization of food and of population.

Food available for human consumption = Total food supply - Feed - Seed - Industrial uses - Waste.

Further studies will be needed to determine if significant association exist for organic animal fat and meat increased incidence of increased mortality rates of cancer.

In this study, mortality rate was considered instead of cancer incidence rate due to the availability of the data. These measurements are not the same because mortality is a combination of incidence, stage and age at diagnosis, and survival.

Despite extensive research, few specific dietary determinates of cancer risk have been evaluated for colon cancer and breast cancer. However, most researchers agree that diet probably has an important impact on the occurrence of these chronic conditions. The main factors that have held back progress are the inaccuracy of methods for estimating food and nutrient intake and the biases in case-control studies. The results of existing large studies and controlled trials should advance our understanding of the role of diet in cancer incidence in the future. The best advice for now is to maintain a healthy balanced diet. Although these healthy choices are made by individuals, they may be facilitated or impeded by the social, physical, economic, and regulatory environment in which people live. Community efforts are therefore essential to create an environment that facilitates healthy food choices and physical activity.

References

1- Tilman, D., & Clark, M. (2014). Global diets link environmental sustainability and human health. *Nature,515*(7528), 518-522. doi:10.1038/nature13959.

2- Population nutrient intake goals for preventing diet-related chronic diseases. (2007, June 22). Retrieved from https://www.who.int/nutrition/topics/5_population_nutrient/en/index15.html.

3- Food and Diet. (n.d.). Retrieved from https://www.hsph.harvard.edu/

4- Motulsky, A. G. (1989). Diet and health: Implications for reducing chronic disease risk. Washington, D.C.: National Academy Press.

5- Cancer. (n.d.). Retrieved from https://www.who.int/news-room/fact-sheets/detail/cancer

6- Latest global cancer data: Cancer burden rises to 18.1 million new cases and 9.6 million cancer deaths in 2018. (n.d.). Retrieved from https://www.iarc.fr/featured-news/latest-global-cancer-data-cancer-burden-rises-to-18-1-million-new-cases-and-9-6-million-cancer-deaths-in-2018/

7- Diet. (n.d.). Retrieved from https://www.cancer.gov/about-cancer/causes-prevention/risk/diet

8- Armstrong, B., & Doll, R. (2006, July 18). Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices - Armstrong 1975 - International Journal of Cancer - Wiley Online Library. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/ijc.2910150411?sid=nlm:pubmed

9- Khodarahmi, M., & Azadbakht, L. (2014). The association between different kinds of fat intake and breast cancer risk in women. *International journal of preventive medicine*, *5*(1), 6–15.

10- Eunyoung Cho, Donna Spiegelman, David J. Hunter, Wendy Y. Chen, Meir J. Stampfer, Graham A. Colditz, Walter C. Willett, Premenopausal Fat Intake and Risk of Breast Cancer, *JNCI: Journal of the National Cancer Institute*, Volume 95, Issue 14, 16 July 2003, Pages 1079–1085, https://doi.org/10.1093/jnci/95.14.1079

11- Sieri, S., Krogh, V., Pala, V., Muti, P., Micheli, A., Evangelista, A., . . . Berrino, F. (2004, April 01). Dietary Patterns and Risk of Breast Cancer in the ORDET Cohort. Retrieved from http://cebp.aacrjournals.org/content/13/4/567.long

12- Key, T. J., Allen, N. E., Spencer, E. A., & Travis, R. C. (2002). The effect of diet on risk of cancer. The Lancet*,360*(9336), 861-868. doi:10.1016/s0140-6736(02)09958-0

13- Kushi, L., & Giovannucci, E. (2002). Dietary fat and cancer. *The American Journal of Medicine,113*(9), 63-70. doi:10.1016/s0002-9343(01)00994-9

14- Gray, G. E., Pike, M. C., & Henderson, B. E. (1979). Breast-cancer incidence and mortality rates in different countries in relation to known risk factors and dietary practices. *British Journal of Cancer,39*(1), 1-7. doi:10.1038/bjc.1979.1

15- Schlesinger, S., Chan, D. S., Vingeliene, S., Vieira, A. R., Abar, L., Polemiti, E., . . . Norat, T. (2017). Carbohydrates, glycemic index, glycemic load, and breast cancer risk: A systematic review and dose–response meta-analysis of prospective studies. *Nutrition Reviews,75*(6), 420-441. doi:10.1093/nutrit/nux010

16- Cross, A. J., Leitzmann, M. F., Gail, M. H., Hollenbeck, A. R., Schatzkin, A., & Sinha, R. (2007, December). A prospective study of red and processed meat intake in relation to cancer risk. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/18076279

17- Cummings, J. H., & Bingham, S. A. (1998). Diet and the prevention of cancer. *BMJ (Clinical research ed.)*, *317*(7173), 1636–1640. doi:10.1136/bmj.317.7173.1636

18- Willett W. C. (1995). Diet, nutrition, and avoidable cancer. *Environmental health perspectives*, *103 Suppl 8*(Suppl 8), 165–170. doi:10.1289/ehp.95103s8165

19- Kolonel, L. N., Hankin, J. H., Lee, J., Chu, S. Y., Nomura, A. M., & Hinds, M. W. (1981). Nutrient intakes in relation to cancer incidence in Hawaii. *British Journal of Cancer,44*(3), 332-339. doi:10.1038/bjc.1981.189

20- Correa, P. (1981, September 01). Epidemiological Correlations between Diet and Cancer Frequency. Retrieved from http://cancerres.aacrjournals.org/content/41/9_Part_2/3685.long

21- Vieira, A. R., Abar, L., Chan, D. S., Vingeliene, S., Polemiti, E., Stevens, C., . . . Norat, T. (2017). Foods and beverages and colorectal cancer risk: A systematic review and meta-analysis of cohort studies, an update of the evidence of the WCRF-AICR Continuous Update Project. *Annals of Oncology,28*(8), 1788-1802. doi:10.1093/annonc/mdx171

22- Perera, P. S., Thompson, R. L., & Wiseman, M. J. (2012). Recent Evidence for Colorectal Cancer Prevention Through Healthy Food, Nutrition, and Physical Activity: Implications for Recommendations. *Current Nutrition Reports,1*(1), 44-54. doi:10.1007/s13668-011-0006-7

23- Fung, T. (2003, February 10). Major Dietary Patterns and the Risk of Colorectal Cancer in Women. Retrieved from https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/215072

24- Boyle, P., & Langman, J. S. (2000, September 30). ABC of colorectal cancer: Epidemiology. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1118620

25- Giovannucci, E., Rimm, E. B., Stampfer, M. J., Colditz, G. A., Ascherio, A., & Willett, W. C. (1994, May 01). Intake of Fat, Meat, and Fiber in Relation to Risk of Colon Cancer in Men. Retrieved from http://cancerres.aacrjournals.org/content/54/9/2390.long

26- Willett, W. C., Stampfer, M. J., Colditz, G. A., Rosner, B. A., & Speizer, F. E. (1990, December 13). Relation of meat, fat, and fiber intake to the risk of colon cancer in a prospective study among women. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2172820

27- Hebert, J. R., Hurley, T. G., Olendzki, B. C., Teas, J., & Ma, Y. (1999). RESPONSE: Re: Nutritional and Socioeconomic Factors in Relation to Prostate Cancer Mortality: A Cross-national Study. JNCI Journal of the National Cancer Institute*, 91*(8), 726-726. doi:10.1093/jnci/91.8.726

28- Doll R, Peto R. Epidemiology of cancer. In: Weatherall DJ, Ledingham JGG, Warrell DA, eds. Oxford textbook of medicine. Oxford, Oxford University Press, 1996:197-221.

29- Diet, nutrition, physical activity and colorectal cancer. (n.d.). Retrieved from https://www.aicr.org/continuous-update-project/reports/colorectal-cancer-2017-report.pdf

30- Diet, nutrition, physical activity and breast cancer. (n.d.). Retrieved from https://www.aicr.org/continuous-update-project/reports/breast-cancer-report-2017.pdf

31- Cancer Mortality Statistics. (n.d.). Retrieved from http://apps.who.int/healthinfo/statistics/mortality/causeofdeath_query/start.php

32- Food Balance Sheet. (n.d.). Retrieved from http://www.fao.org/faostat/en/#data/FBS

33- World Development Indicators. (n.d.). Retrieved from https://databank.worldbank.org/data/source/world-development-indicators

34- Socioeconomic indicators. (n.d.). Retrieved from http://apps.who.int/gho/data/node.imr

35- Variable selection in linear regression. (n.d.). Retrieved from https://www.stata-journal.com/article.html?article=st0213

36- Meat, fish & dairy. (2018, May 24). Retrieved from https://www.wcrf.org/dietandcancer/exposures/meat-fish-dairy

37- Outwater, J., Nicholson, A., & Barnard, N. (1997). Dairy products and breast cancer: The IGF-I, estrogen, and bGH hypothesis. *Medical Hypotheses, 48*(6), 453-461. doi:10.1016/s0306-9877(97)90110-9.