Luke Andreesen
Dr. Schultz
5/17/2022

The Relationship Between Wealth and Cancer Incidence Rate

Introduction

While extensive research has established a widely recognized positive correlation between wealth and cancer *survival* rates, there is a large gap in research in regards to the relationship between cancer *incidence* rates and wealth. The causal mechanism behind the relationship between cancer survival rates and wealth is quite straightforward; essentially, wealthier individuals tend to have better access to healthcare, which in turn allows them to receive better cancer care and thus increase their chances of survival. The relationship between wealth and incidence rates of cancer, however, remains much more blurred. Incidence rates, put simply, "refers to the occurrence of new cases of disease or injury in a population over a specified period of time" (CDC 2012). Cancer is characterized by a strong inherent randomness; according to the Johns Hopkins Kimmel Cancer center, "random, unpredictable DNA copying 'mistakes' account for nearly two-thirds of the mutations that cause cancer" (Wasta 2017). However, that is not to say cancer is entirely random — many risk factors also play a role in causing cancer. These factors can range from quite obvious and well-known events such as exposure to radiation, to less obvious causes including HPV-related events. To summarize, cancer incidence, while generally random, is strongly impacted by certain risk factors and behaviors. This study seeks to use wealth to pool together some of these specific risk factors and behaviors.

Theoretical Framework

While the scientific and health communities have yet to establish a consensus in regards to how wealth might impact the incidence rates of certain forms of cancer, there are multiple prior studies that have provided valuable research on this relationship. Two particular studies with very similar goals and methods, Clegg et al (Clegg et al 2009) and Boscoe et al. (Boscoe et al 2014), have provided particularly interesting and important insight on this area of research. Clegg et al, the first of the two studies, used individual level data to determine the impact of socioeconomic status on cancer incidence. The socioeconomic data utilized in this study consisted mainly of self-reported data including income, poverty status, and educational attainment. The authors found elevated incidence rates of lung cancer among individuals who were less educated — those who have less than a highschool education. Additionally, Clegg et al determined that individuals with low family incomes ($12,500 per year or less) have lung cancer incidence rates of nearly double rates for individuals with high family incomes ($50,000). Additionally, this study found a negative correlation between income level and risk of distant-stage breast cancer in women, and distant-stage prostate cancer in men. Data in this study was sourced from a cancer registry known as Surveillance, Epidemiology and End Results (SEER). This data was then linked to the U.S. representative National Longitudinal Mortality Study (NLMS), which provides self-reported socioeconomic data; the effort to link the two was carried out by the National Cancer Institute.

Boscoe et al, the latter of the two studies, also sought to examine the relationship between socioeconomic status and cancer incidence. Using cancer data from the North American Association of Cancer Registries (NAACCR), the authors examined the link between the incidence rates of cancer at 39 different sites — site refers to the area of the body where cancer develops — and poverty rate. The registry utilized by Boscoe et al consisted of cases for 3

million tumors over the course of four years (2005-2009), from 17 U.S. states. The registry contained an assigned poverty level for each case; four poverty levels were determined by the percentage of the population living below the poverty line for each tract that a case occurred in. The study found a monotonic increasing or decreasing relationship between cancer incidence and poverty at 19 different sites. Sites with the strongest association with higher poverty included the larynx, cervix, penis, and liver, while sites most strongly associated with lower poverty included melanoma, thyroid, and testis.

This study seeks to validate the findings of these two earlier studies by producing similar findings with greater external validity. While both previous studies had strong internal validity through extensive research efforts that were based on individual cases, they lack external validity due to the fact that all data was collected only within the United States. It is the goal of this study to provide information that can be used worldwide; relative to the rest of the world, the United States has a relatively high cancer survival rate. While healthcare throughout the world - especially within third-world countries - remains an important issue, it is one that is extremely difficult and costly to solve. By determining certain forms of cancer that have strong negative or positive associations with national wealth, we seek to open the door for more research by scientists and doctors to determine risk factors that might cause these relationships. Policy change on a worldwide scale is quite difficult, but if our data can provide scientists with identifiable risk factors that were previously unnoticed, international health organizations such as the World Health Organization (WHO) may be able to spread this information and construct initiatives that decrease cancer incidence rates. Using a greater scale than those used by Clegg et al and Boscoe et al will not only provide greater external validity, but will ideally also provide the benefits to a broader population.

The research question that guides this study is identical to that of Clegg et al and Boscoe et al: is there a relationship between wealth and the incidence rates of certain forms of cancer? This line of research is guided by the previously mentioned studies, Clegg et al and Boscoe et al. In our study, we are assigning a more general independent variable, wealth, as compared to the more specific and individualized socioeconomic status variable used in prior studies. The dependent variable in our study remains the same as the prior studies: cancer incidence rates. Using measures more general of wealth rather than specific socioeconomic metrics, over a worldwide scale as opposed to a country-wide, will allow us to ideally validate the findings of Clegg et al and Boscoe et al by repeating a similar analysis with greater external validity.

It is not the immediate goal of this study to identify specific risk factors or causes of cancer, but rather to identify and fortify a positive or negative relationship between wealth and certain forms of cancer. This information will be valuable to scientists, doctors, and cancer experts who specialize in identifying these factors by allowing them to understand a critical relationship. While certain risk factors may be more obvious than others and become evident through this study - for example those associated with tobacco use and HPV-causing activities - the effects of certain environmental, social, habitual, or other factors that may create the relationships we observe are beyond the scope of this study.

Based on the results of Clegg et al and Boscoe et al, in our analyses expect a negative relationship for tobacco-related cancers due to higher prevalence of tobacco usage in less wealthy countries due to lower health standards. We expect a negative relationship for HPV-related cancers due to inferior health standards in poorer countries. We expect positive relationships between wealth and incidence rates for cancer of the prostate, testis, thyroid, and melanoma of skin due to higher testing rates due to greater healthcare access among individuals

in wealthier counties. In our study, HPV-related cancers consist of all cancers in which HPV DNA was detected in over 60% cases. There are six sites at which this holds true: anus, cervix, vagina and vulva, penis, and oropharynx (CDC,2021). Tobacco-related cancers consist of lip and oral cavity cancer and lung cancer.

Data and Descriptive Statistics

Cancer incidence data for this study is sourced from The World Health Organization (WHO). WHO is an international health agency that makes a global effort to offer health services, research, and response to health emergencies (World Health Organization). Founded in 1948, WHO operates as a nonprofit branch of the United Nations (U.N). WHO's lack of political agenda and goal to promote global health ensures that there is very low risk of bias in the data that they collect. WHO receives global funding from U.N. nations allowing them to conduct accurate data collection, which in turn creates strong internal validity among the data they provide. Our study utilizes data collected by the International Agency for Research (IAR), a branch of WHO. IAR's "Cancer Today" database provides incidence rates of cancer at each of the 39 sites in the human body for all 193 U.N. member nations. The dataset we work with provides incidence rates measured as the number of cases per 100,000 individuals; this rate is provided on a global scale and for each individual country (Global Cancer Observatory, 2021). Additionally, for each country we have access to the total number of cases for cancer at each site, an uncertainty interval for the incidence rate estimate, and a cumulative risk calculation based on incidence rate. As previously noted, due to the relatively large number of cancer sites, we are limiting our study to a few specific sites that will be described more thoroughly later in this

paper. The most WHO IAR cancer data available is from 2020, which we will be using in our study.

Wealth data in our study is sourced from The World Bank, an international organization consisting of five separate institutions. The World Bank operates with the goals of poverty alleviation and the promotion of "shared prosperity," particularly among developing nations, and consists of 189 member countries (World Bank). As with the WHO, The World Bank has no agenda that would create any particular risk in regards to bias in their collected data. As this organization seeks to promote global well-being and worlds to support developing countries, it is in the best interest of the organization to provide accurate data that allows them to pursue this goal. The wide-scale global support and funding for the organization, along with the fact that the organization has operated since 1944, creates a strong dataset that fortifies the internal validity of this study. Our study utilizes two World Bank datasets. The first is GDP per capita data, which can be used as a very broad indication of economic well-being and living standards (The World Bank, 2022). The second metric we use is GNI per capita. GNI, or gross national income, is another measure of economic output but differs from GDP in that it is representative of the income of a country and its residents that is not restricted to business within the country itself. Both methods prove quite similar information in regards to the general economic state of a country, and can both be used as very general representations of the wealth of a country. We utilize two metrics for the independent variable to allow us to create two sets of regressions to provide greater internal validity to our findings, as we should expect to see a relationship in one metric should be supported by the other metric.

Our study will combine WHO and World Bank datasets to create a new dataset that contains cancer incidence rates and each of the aforementioned national wealth metrics. The

units of analysis for the study will be countries. Combining the datasets requires solutions to a few notable challenges as we seek to create a new set of data that can readily be analyzed. First, we must narrow down the 39 available cancer sites; analyzing the relationship between wealth and incidence rates for each site would create a multiple testing issue. Instead, we will use a total of seven incidence rates. We will create one single HPV-related cancer incidence rate for each country by combining incidence rate data for cancer at the anus, cervix, vagina and vulva, penis, and oropharynx. We will do the same for tobacco-related cancers, which consist of incidence rates for lip/oral cavity and lungs. In addition to these two combined cancer categories, each of which will have one single averaged incidence rate, we will examine incidence rates for prostate cancer, testicular cancer, thyroid cancer, and melanoma of skin. In total, this gives us six incidence rates to analyze for each country.

We will utilize the 189 countries for which the World Bank provides data; the majority of these countries are U.N. member countries and will therefore be included in WHO data, but any country that is not included will be removed. While the World Bank releases data every year, with the most recent being 2021, since the most recent available WHO data is from 2020, both sets of data will come from the year 2020. Essentially, from this combination of data, we can create a dataset for each of the six incidence rates that we are examining. For each of the six incidence rates, we will have the incidence rates for a specific form of cancer for roughly 170 countries — though this varies for each incidence rate — alongside the GDP per capita and GNI per capita of each country. As mentioned earlier, incidence rates are measured as the number of new cases for the year 2020 per 100,000 citizens of a given country, and both of the economic metrics used are measured in U.S. dollars.

Five Number Summary and Observations for all Variables

Table 1

Variable	*Minimum*	*Quartile 1*	*Median*	*Quartile 3*	*Maximum*	*Observations (Countries)*
GDP per Capita (USD)	238.9907	1927.7078	5606.5381	15742.4537	173688.1894	232
GNI per Capita (USD)	230.00	1955.00	5263.85	14460.00	112240.00	232
ASR - HPV	0.285	1.130	1.645	2.335	6.705	181
ASR - Tobacco	0.470	1.700	4.065	6.830	11.440	182
ASR - Testis	0.00	0.06	0.14	0.42	0.95	171
ASR- Prostate	0.21	7.80	12.67	16.78	35.50	181
ASR- Melanoma	0.020	0.120	0.230	0.935	7.790	175

ASR - Thyroid	0.03	0.29	0.58	0.98	2.65	182

For each measure of the independent variable, we have provided a five number summary of the data, along with the number of observations for each of these variables. The number of observations refers to the number of countries examined for each measure of wealth or for each incidence rate.

Variable Distribution Figures

Figure 1

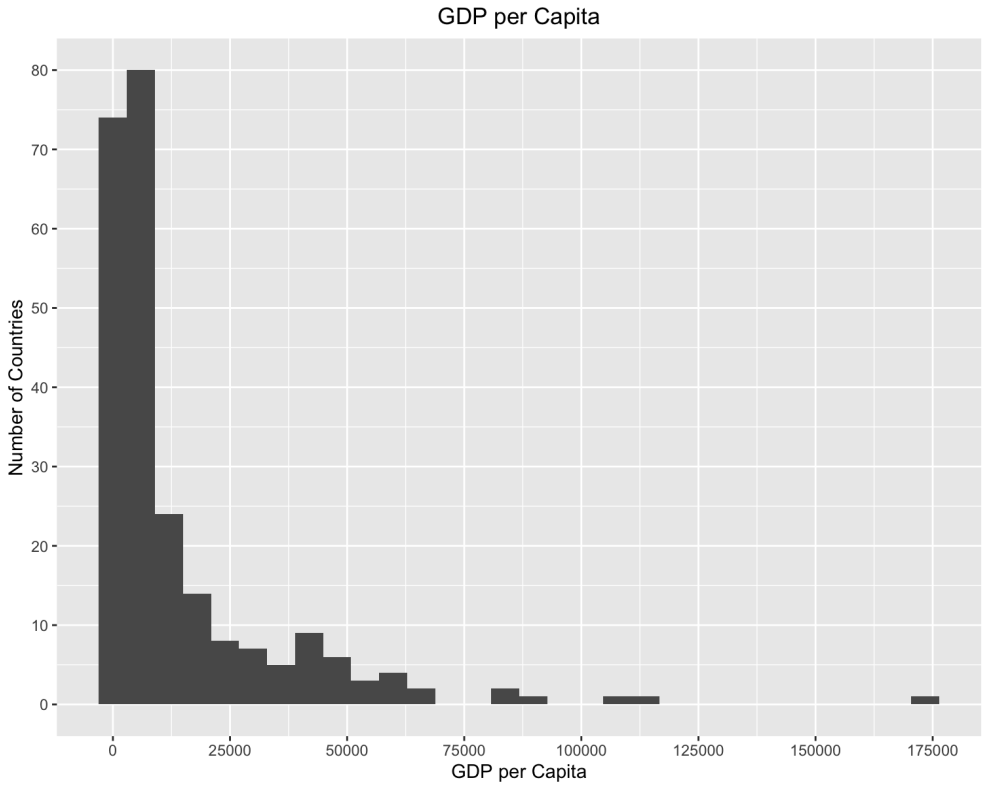


Figure 2

Figure 3

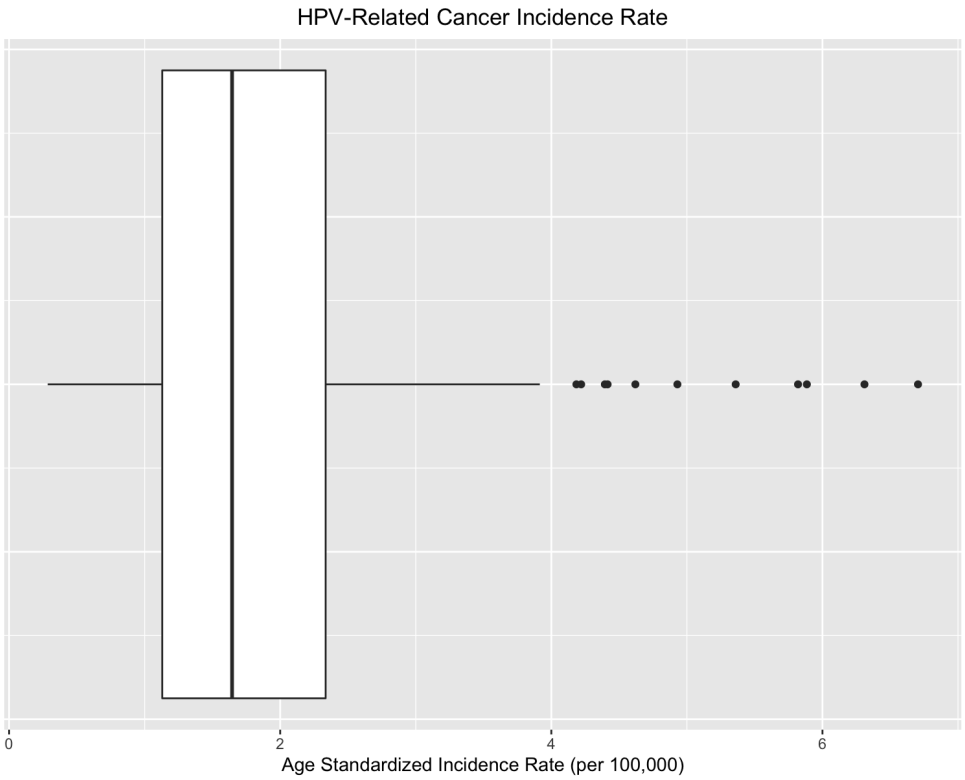


Figure 4

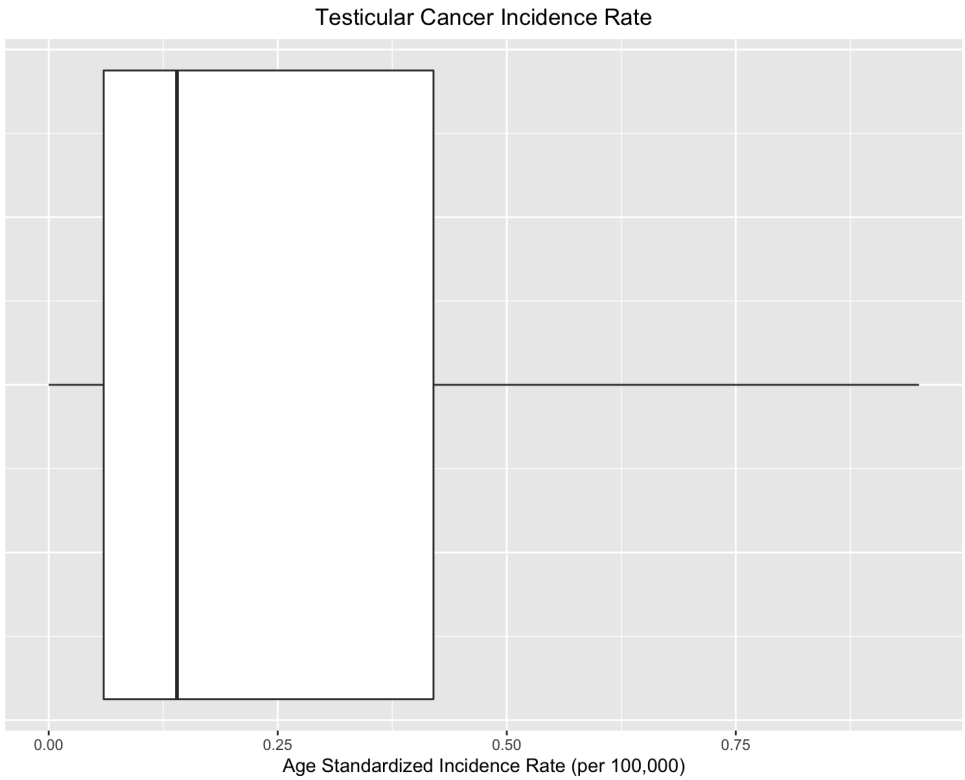


Figure 5

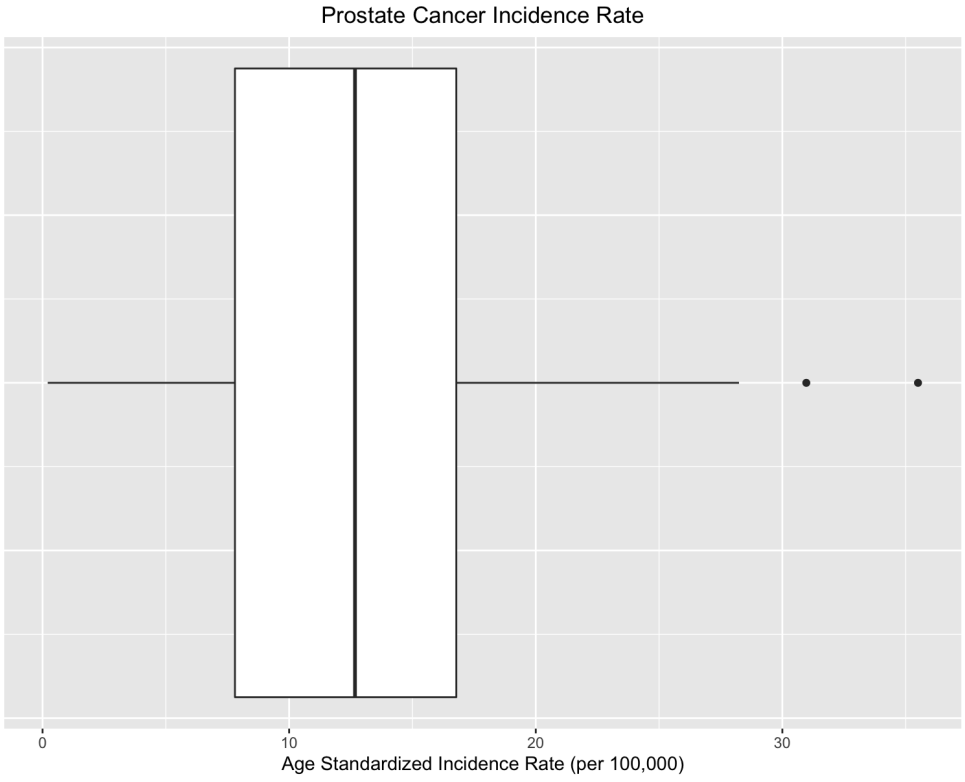


Figure 6

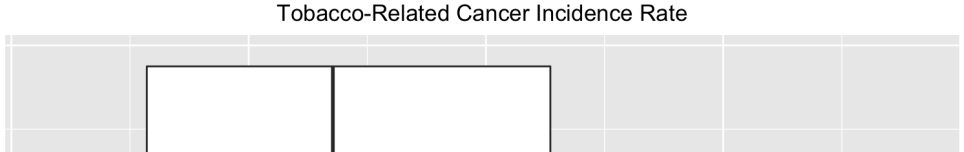


Figure 7

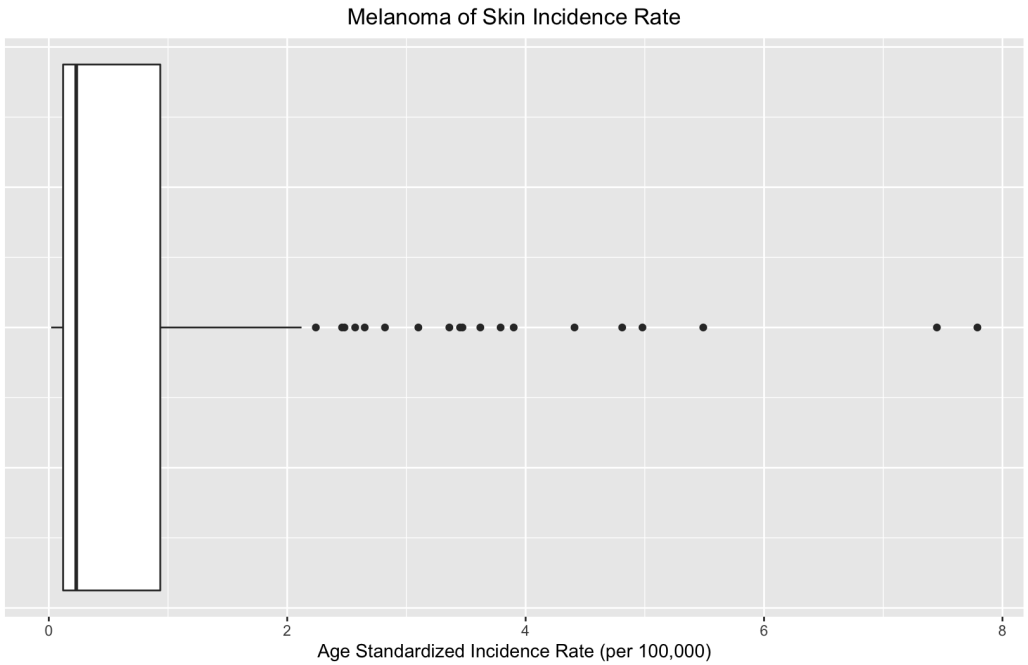
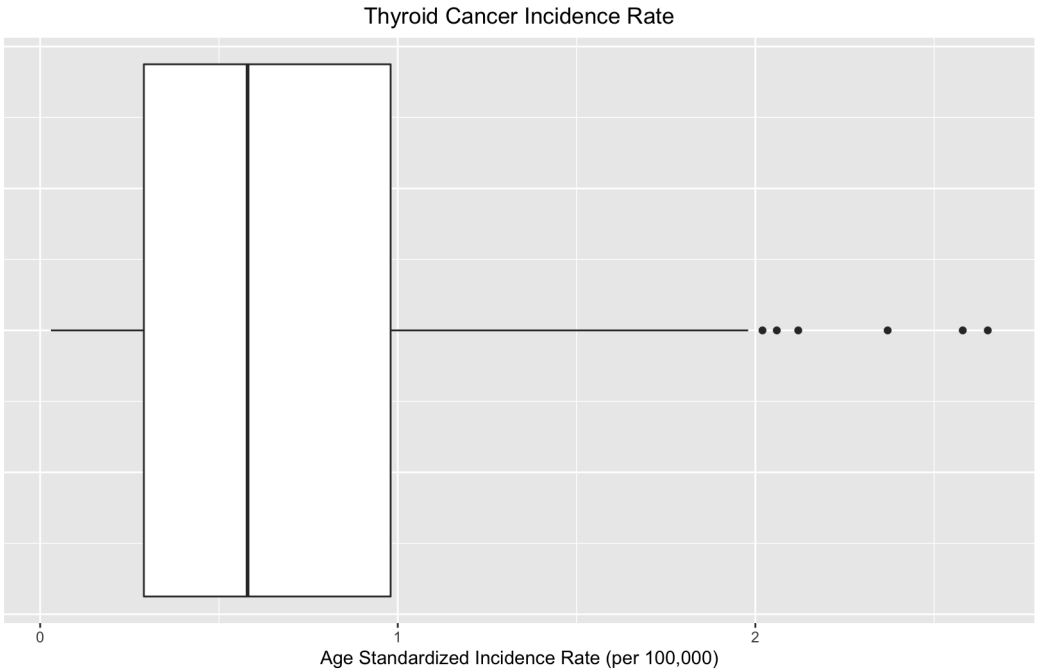


Figure 8



Empirical Framework

Linear regressions will be used as the estimation strategy to evaluate our causal theory. In total, we will construct 12 linear analyses. The first six regressions will use GDP per capita as the independent variable, and incidence rates for one of the six cancer sites as the dependent variable. The second set of six regressions will be used to verify the results of the first six by replacing the GDP per capita variable with the GNI per capita variable. Analysis of the regressions should prove relatively easily; if we see a strong positive or negative relationship that exists among both of the sets of graphs, this will reasonably imply that there is a correlation between the incidence rates of a certain type of cancer and a country's wealth. As mentioned earlier, we simply seek to identify correlations; since each of the sites we are testing were found to be positively or negatively associated with socioeconomic status in the earlier studies, we hope to find the same relationships between each of these indices rates and national wealth. Our regressions do not feature any controls, as the cancer incidence rates provided by WHO are age-standardized, which essentially accounts for the discrepancy in cancer rates among different age groups.

This study does face a few limitations in attempting to establish strong external validity. By examining wealth on a country-wide scale we make sacrifices to internal validity; the very broad scale as opposed to using individual cases as in Clegg et al and Boscoe et al limits our internal validity. Examining rates on worldwide scales inherits many unknown risk factors, cultural effects, political policies, and beyond that cannot all be accounted for, thus we sacrifice some internal validity for this greater external validity. Using multiple different regressions also creates somewhat of a multiple testing issue, though this is largely addressed by repeating each

of the six regressions with two metrics of our independent variable. An additional threat to internal validity is created by an inherent measurement issue. Wealthier countries tend to have greater access to healthcare and thus have increased cancer screening rates as compared to poorer countries. Poorer countries are likely to have lower incidence rates for certain forms of cancer as they simply are not screened for as often. One final limitation faced by our study is the threat of reverse causation; it is possible that high cancer incidence rates may decrease human capital and thus decrease economic output, which would be reflected by GDP per capita and GNI per capita.

Despite the limitations of the study, the data compiled by this research effort will likely be beneficial regardless of whether or not the causal theory is verified. It is important that we continue to look at the issue of cancer incidence on many different levels, and since there is limited past research utilizing a country-wide level to examine this relationship between wealth and cancer incidence rates, our study will be valuable. Though we hope to verify the relationships established in Clegg et al and Boscoe et al with greater external validity, understanding how incidence rates might interact with wealth will be valuable regardless.

Results

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 2

HPV-related					Testis				
Predictors	*Estimates*	Std. Error	*CI*	*p*	*Predictors*	*Estimates*	Std. Error	*CI*	*p*
(Intercept)	5.14767	0.50345	4.15 - 6.14	< 2e-16 ***	(Intercept)	-0.62976	0.09827	-0.82 - -0.44	1.65e-09 ***

Logged GDP	-0.36538	0.05731	-0.48 - -0.25	1.76e-09 ***	Logged GDP	0.10314	0.01127	0.08 - 0.13	2.92e-16 ***
Observations	167				Observations	158			
R^2/R^2 Adjusted	0.1976/0.1928				R^2/R^2 Adjusted	0.3494/0.3453			

Table 3

Melanoma					Prostate				
Predictors	*Estimates*	Std. Error	*CI*	*p*	*Predictors*	*Estimates*	Std. Error	*CI*	*p*
(Intercept)	-3.71008	0.45984	-4.62 - -2.80	1.59e-13 ***	(Intercept)	-2.8686	2.9445	-8.68 - -2.94	0.331
Logged GDP	0.52552	0.05278	0.42 - 0.63	< 2e-16 ***	Logged GDP	1.7883	0.3385	1.12 - 2.46	3.93e-07 *** -
Observations	162				Observations	168			
R^2/R^2 Adjusted	0.3826/0.3787				R^2/R^2 Adjusted	0.144/0.1388			

Table 4

Tobacco-related					Thyroid				
Predictors	*Estimates*	Std. Error	*CI*	*p*	*Predictors*	*Estimates*	Std. Error	*CI*	*p*

(Intercept)	-8.5133	0.9907	-10.47 - -6.56	5.99e-15 ***	(Intercept)	-1.04537	0.21408	-1.47 - -0.62	2.45e-06 ***
Logged GDP	1.5056	0.1139	1.28 - 1.73	< 2e-16 *** -	Logged GDP	0.20152	0.02459	0.15 - 0.25	6.64e-14 *** -
Observations	168				Observations	167			
R^2/R^2 Adjusted	0.5129/0.5099				R^2/R^2 Adjusted	0.2892/0.2849			

Table 5

HPV-related					Testis				
Predictors	*Estimates*	Std. Error	*CI*	*p*	*Predictors*	*Estimates*	Std. Error	*CI*	*p*
(Intercept)	5.18461	0.50399	4.15 - 6.14	< 2e-16 ***	(Intercept)	-0.63141	0.09880	-0.82 - -0.44	1.82e-09 ***
Logged GNI	-0.36962	0.05738	-0.48 - -0.25	1.24e-09 ***	Logged GNI	0.10330	0.01133	0.08 - 0.13	3.57e-16 ***
Observations	167				Observations	158			
R^2/R^2 Adjusted	0.201/0.1961				R^2/R^2 Adjusted	0.3478/0.3436			

Table 6

Melanoma					Prostate				
Predictors	*Estimates*	Std. Error	*CI*	*p*	*Predictors*	*Estimates*	Std. Error	*CI*	*p*

(Intercept)	-3.72871	0.46205	-4.62 - -2.80	1.58e-13 ***	(Intercept)	-2.6920	2.9644	-8.68 - -2.94	0.365
Logged GNI	0.52748	0.05302	0.42 - 0.63	< 2e-16 *** -	Logged GNI	1.7674	0.3407	1.12 - 2.46	6.15e-07 ***
Observations	162				Observations	168			
R^2/R^2 Adjusted	0.3822/0.3783				R^2/R^2 Adjusted	0.1395/0.1343			

Table 7

Tobacco-related					Thyroid				
Predictors	*Estimates*	Std. Error	*CI*	*p*	*Predictors*	*Estimates*	Std. Error	*CI*	*p*
(Intercept)	-8.6040	0.9919	-10.47 - -6.56	3.67e-15 ***	(Intercept)	-1.05806	0.21475	-1.47 - -0.62	2.02e-06 ***
Logged GNI	1.5158	0.1140	1.28 - 1.73	< 2e-16 ***	Logged GNI	0.20294	0.02466	0.15 - 0.25	5.45e-14 ***
Observations	168				Observations	167			
R^2/R^2 Adjusted	0.5157/ 0.5128				R^2/R^2 Adjusted	0.2909/ 0.2866			

Regression Results

Figure 9

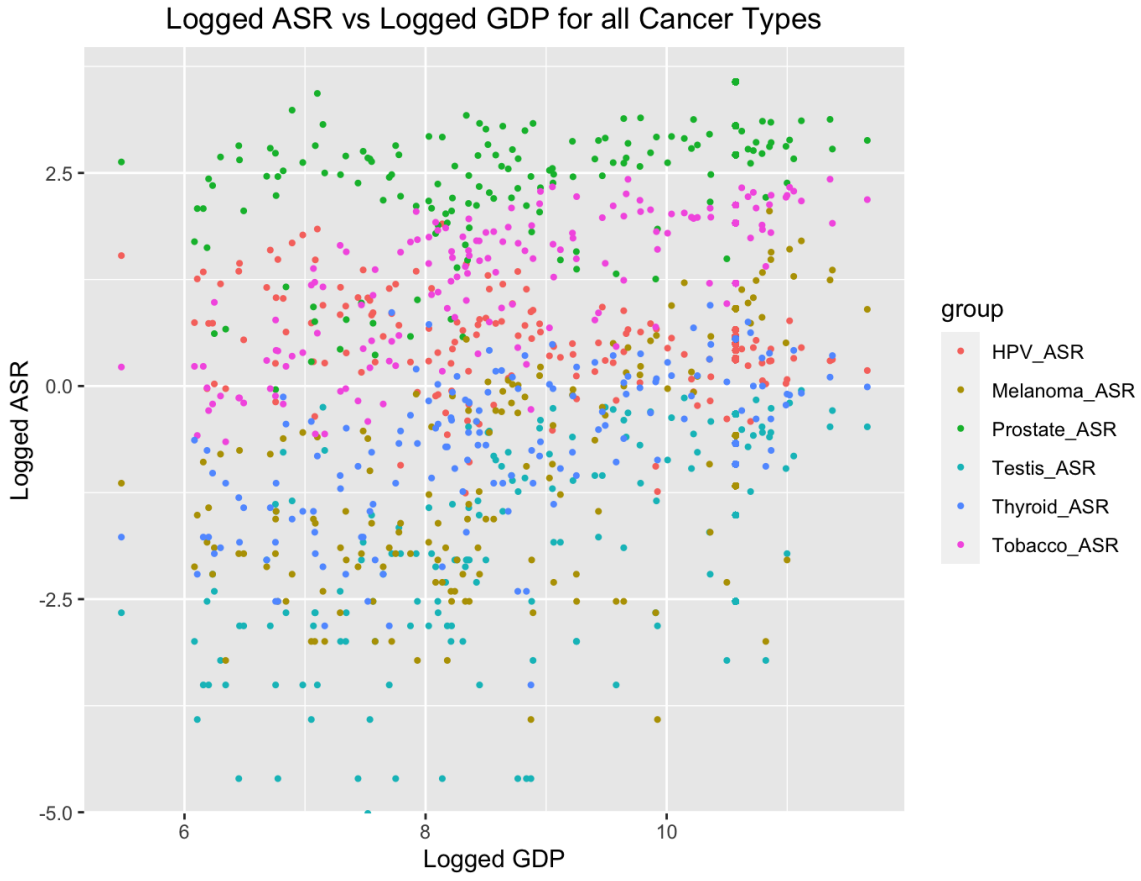


Figure 10

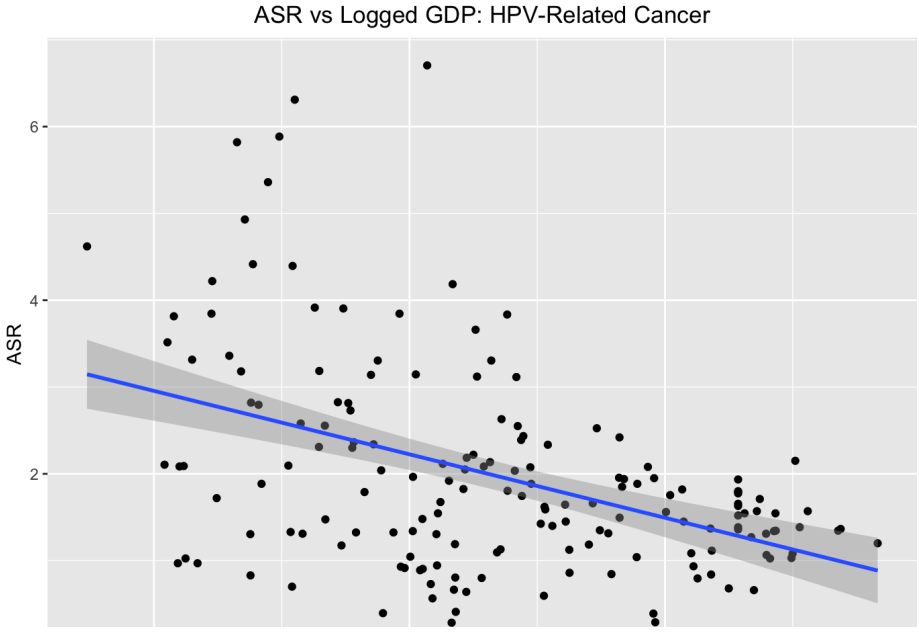
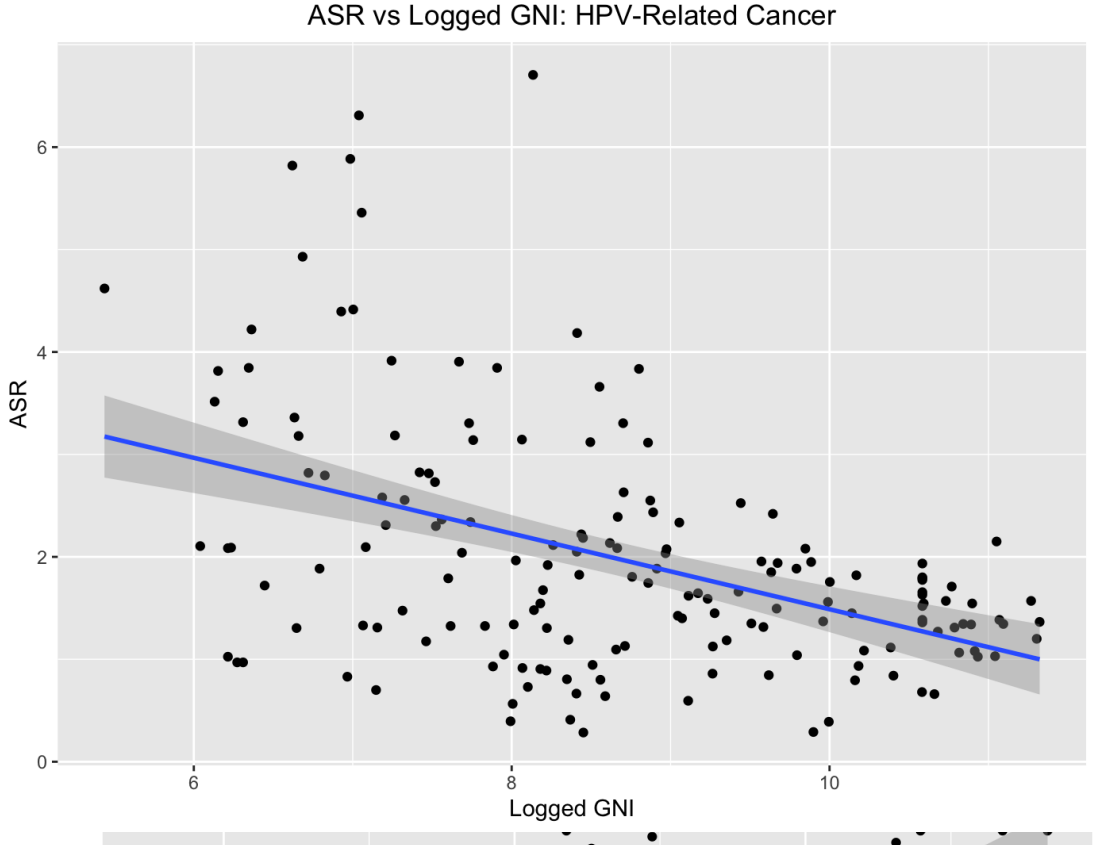


Figure 11



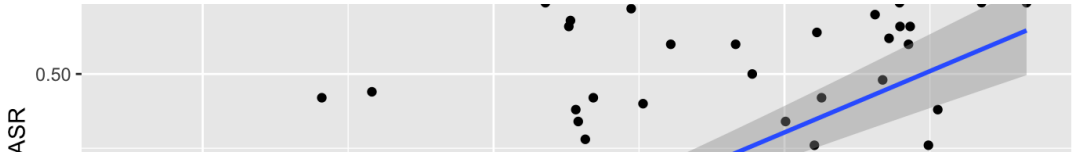


Figure 12

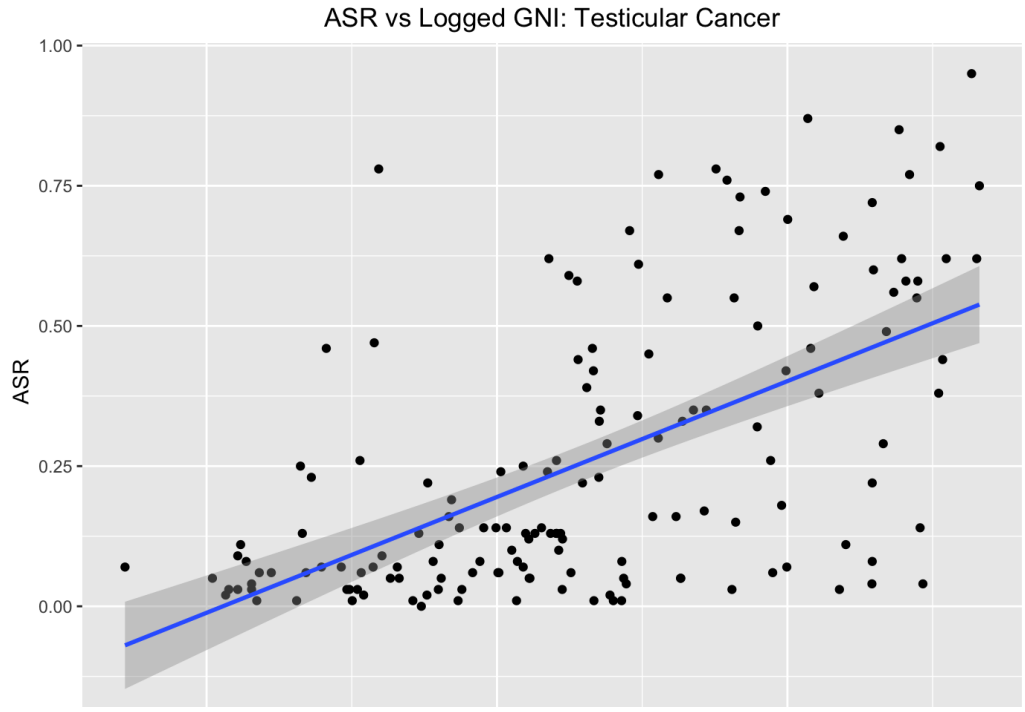


Figure 13

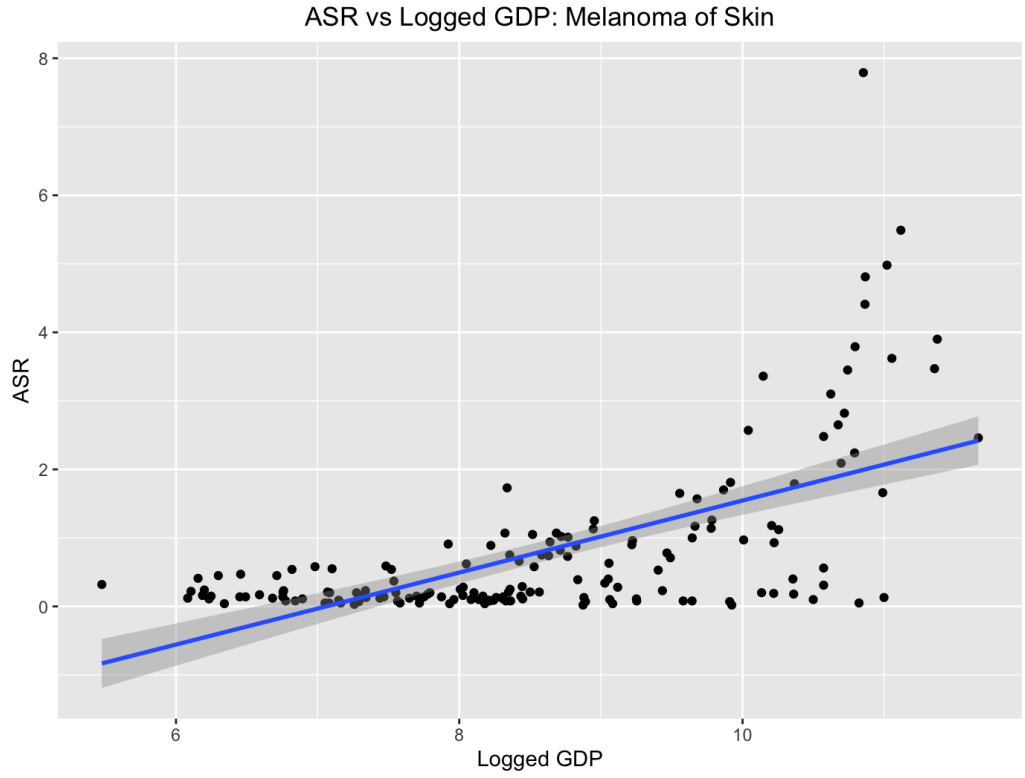


Figure 14

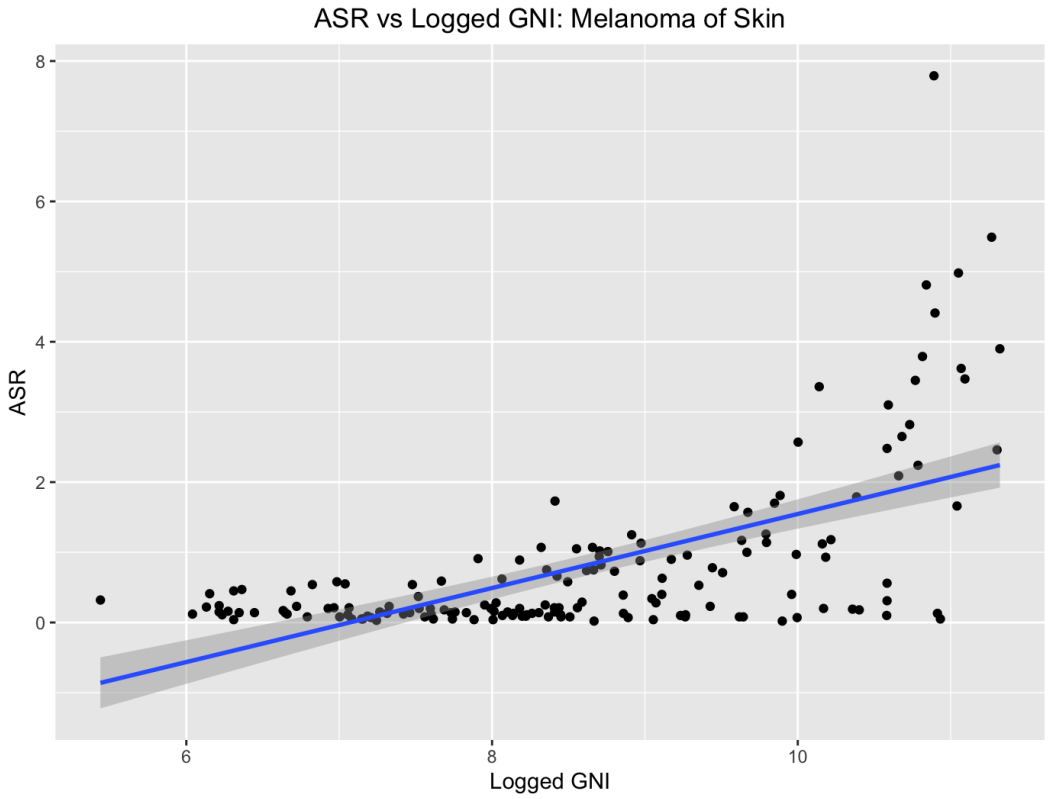


Figure 15

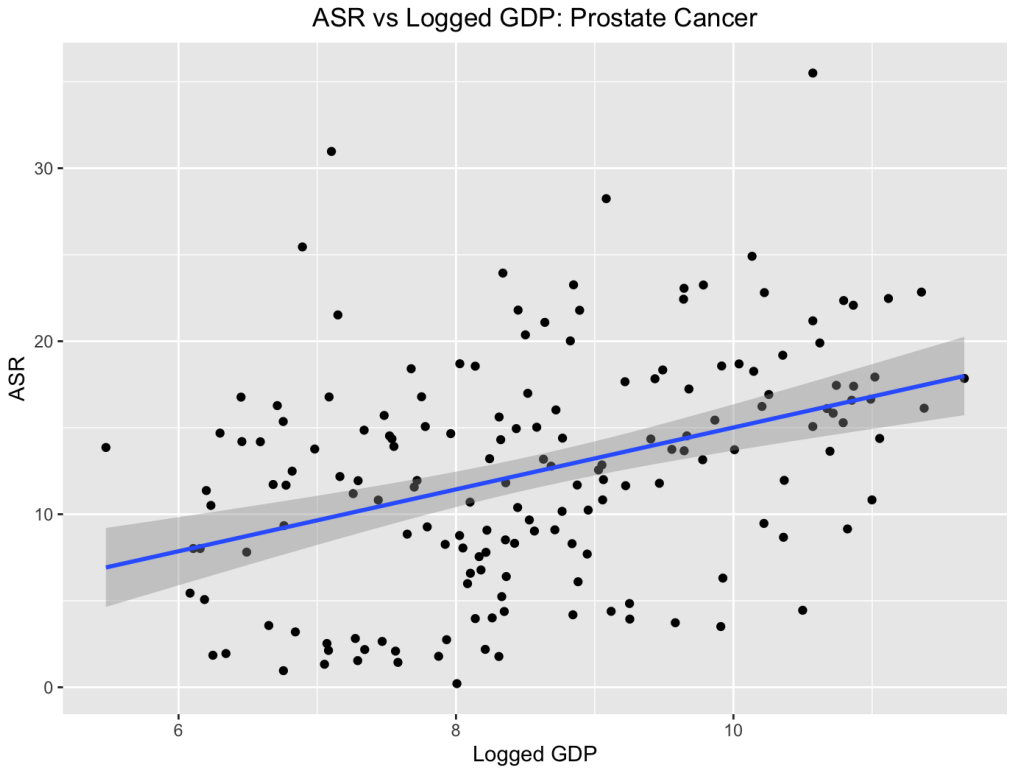


Figure 16

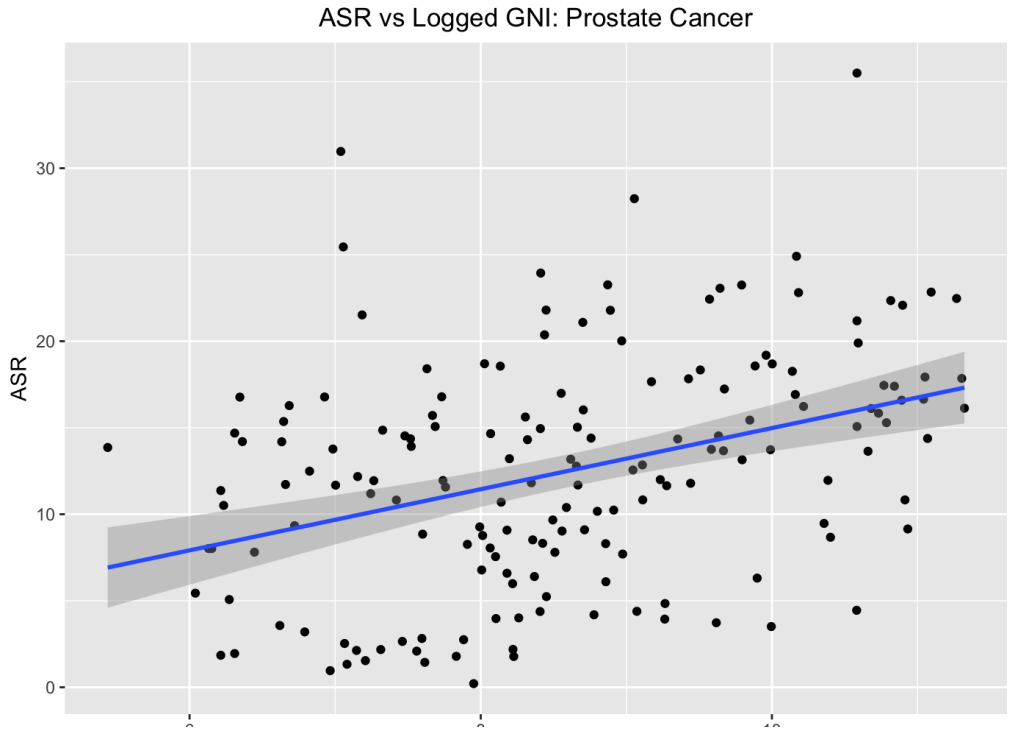


Figure 17

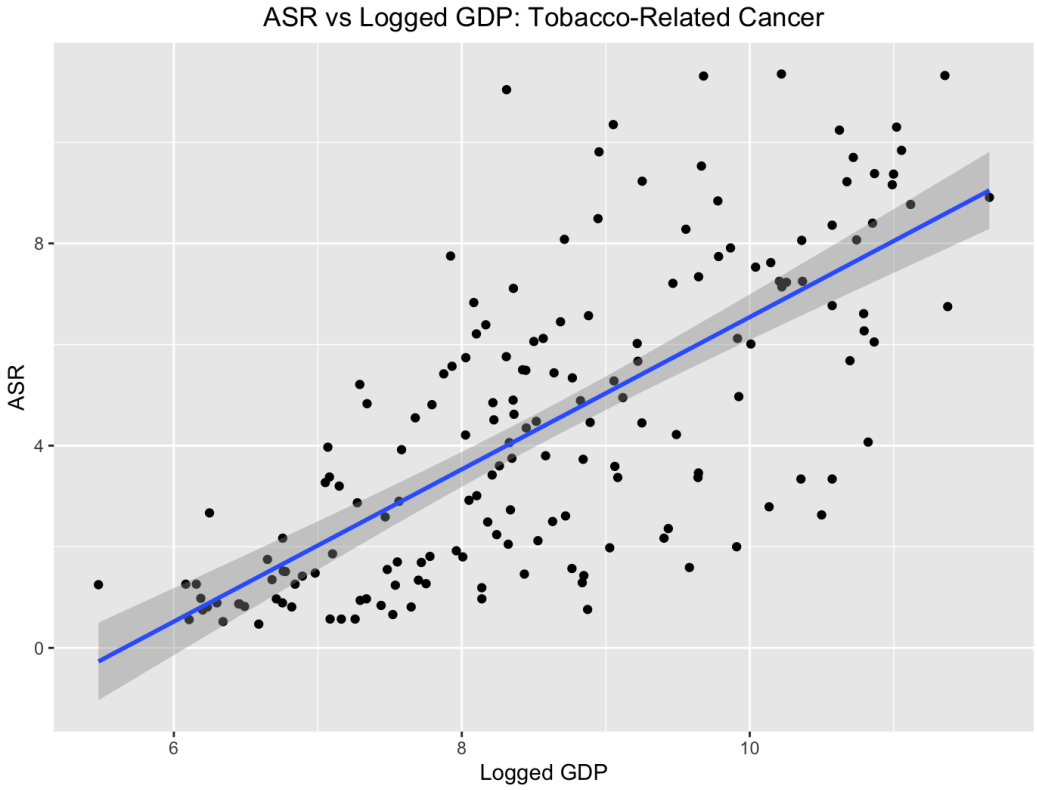


Figure 18

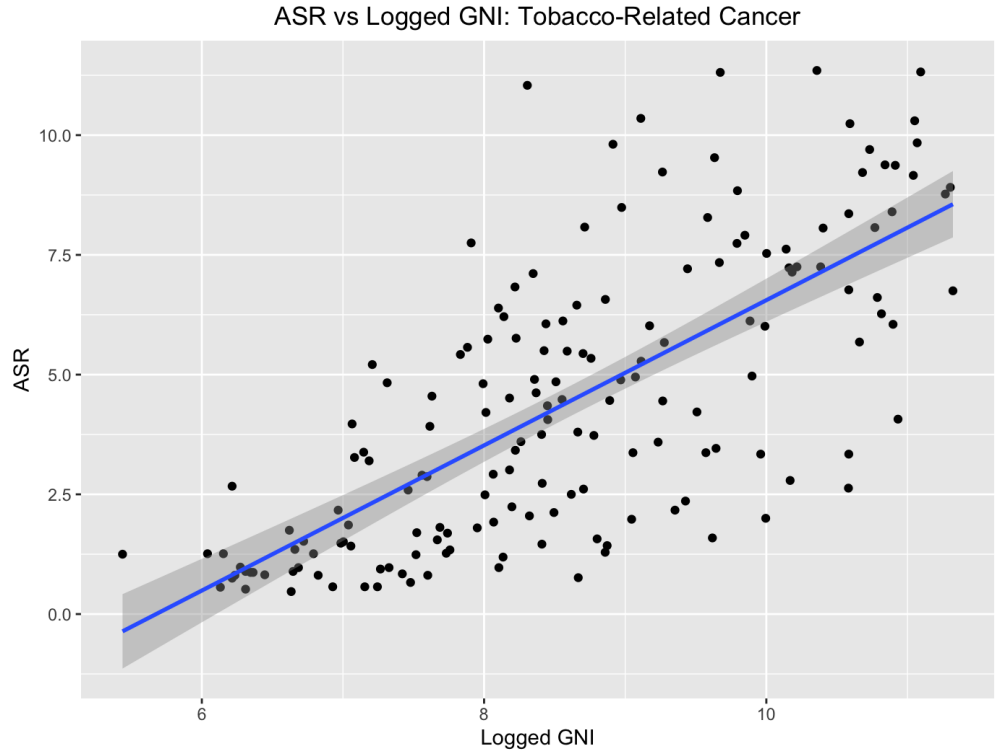


Figure 19

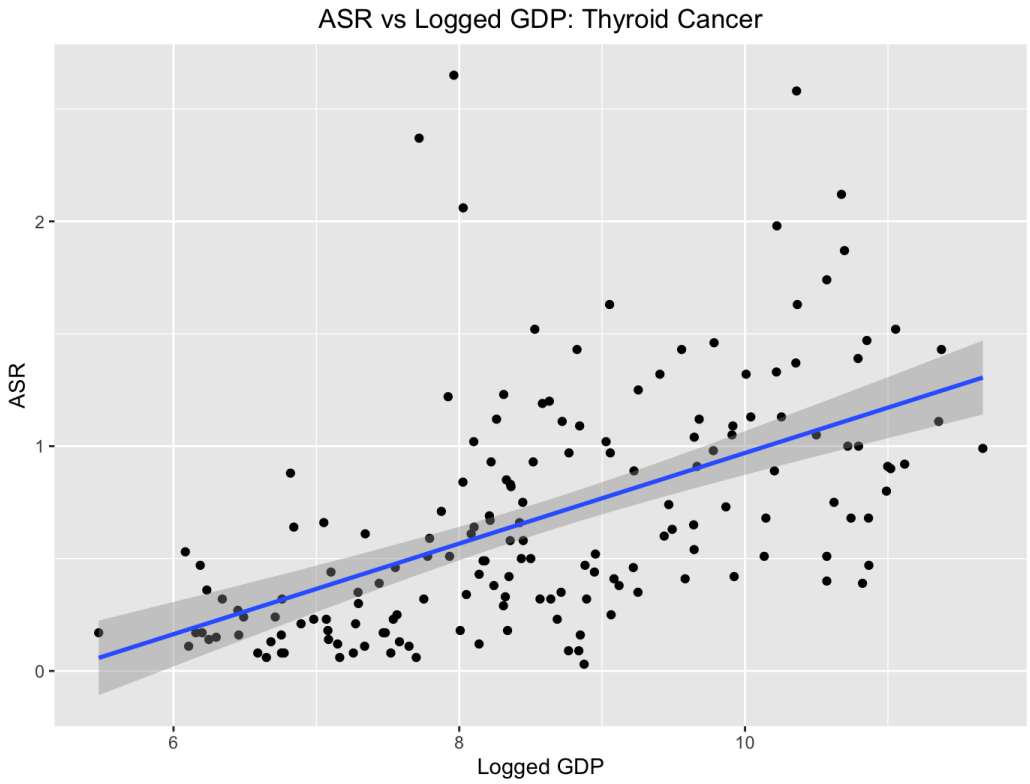
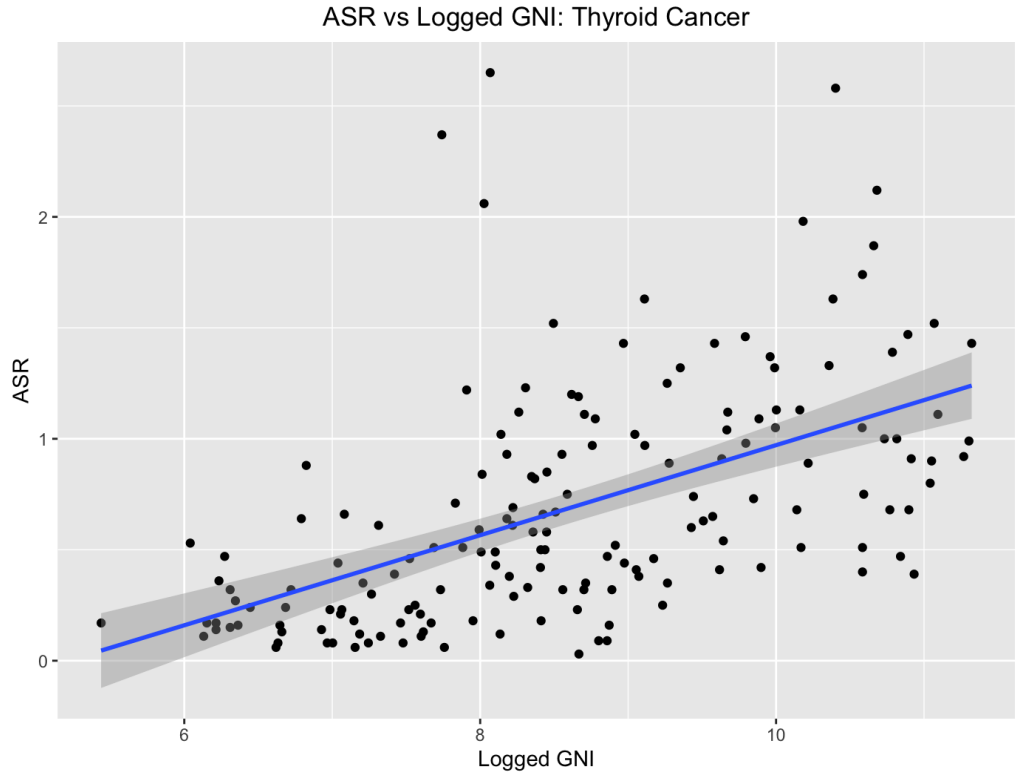


Figure 20



All of the findings displayed in Tables 1-7 and Figures 9-20 display statistically significant results, as the p-value for each of the 12 estimates is below 0.05. In regards to substantive significance, four out of our six metrics demonstrate strong correlations. Though a significant amount of heteroscedasticity exists in each of these metrics, we find substantive significance based on the coefficients of our regressions. In regards to HPV-related cancer, the relatively strong negative coefficient of - 0.3654 for logged GDP and -0.3696 logged GNI is substantial enough to support our hypothesis that the correlation between wealth and HPV-related cancer incidence is negative. For testicular cancer, our regressions yielded positive but relatively small coefficients of 0.1031 and 0.1033 for logged GDP and logged GNI, respectively. The coefficient itself is quite small,

and substantial heteroscedasticity accompanied by a high R^2 value provides us with inconclusive results for the relationship between wealth and testicular cancer incidence. The regression for melanoma of skin yielded substantial heteroscedastic for both logged GDP and GNP, however, the regressions had a very high positive coefficient of 0.5255 and 0.5275 for logged GDP and logged GNI, respectively. Alongside a relatively low standard error of 0.0530, we can determine a positive relationship between wealth and melanoma of skin incidence, thus confirming our hypothesis for this relationship. Our regressions for prostate cancer yielded the most positive correlation for both metrics, with coefficients of 1.7883 and 1.7674 for logged GDP and logged GNI, respectively. These high coefficients are coupled with low R^2 values and relatively minor heteroskedasticity, therefore these results are substantively significant. Thus, we support the hypothesis that wealth is positively correlated with prostate cancer incidence. To our surprise, the regressions for tobacco-related cancers also yielded a high and positive correlation, contrary to our expectation for a negative relationship. The coefficients of 1.5056 and 1.5158 for logged GDP and logged GNI, respectively, along with relatively low standard errors, contradicts our theory that the relationship between wealth and tobacco-related cancer incidence would be negative. Finally, in regard to our thyroid cancer regressions, both metrics yielded relatively low coefficients of 0.20152 and 0.20294 for logged GDP and logged GNI, respectively. As evident in Figures 19 and 20, the thyroid regressions aren't particularly strong and coupled with the low coefficient estimates, our data is inconclusive in regards to the relationship between wealth and thyroid cancer incidence.

Discussion and Conclusion

Using six pairs of linear regression models examining the relationship between six types of cancer among two metrics of wealth, our data analysis has resulted in statistically and substantively significant findings for relationships of four of the six cancer types. Initially, based on the prior research of Clegg et al and Boscoe et al, we expect a negative relationship for tobacco-related cancers due to higher prevalence of tobacco usage in less wealthy countries due to lower health standards. We expected a negative relationship for HPV-related cancers due to inferior health standards in poorer countries. Additionally, we expected positive relationships between wealth and incidence rates for cancer of the prostate, testis, thyroid, and melanoma of skin due to higher testing rates due to greater healthcare access among individuals in wealthier counties. Three of the six expected relationships held true, while two relationships were inconclusive, and one relationship resulted in unexpected findings. We concluded a positive relationship between wealth and incidence of both melanoma of skin and prostate cancer. Additionally, we concluded an unexpected positive relationship between wealth and tobacco-related cancer incidence. Finally, our analyses supported the expected negative relationship between wealth and HPV-related cancer incidence.

As mentioned earlier in this paper, there were many limitations that became evident in our regressions. Many of the regressions had high standard errors, high R^2 values, and substantial heteroskedasticity. Perhaps the most significant limitations in our regressions proved to be using relatively few data points; at just around 165 points per regression, our analyses simply did not yield particularly strong regressions. Increasing data points by perhaps looking at provinces within a set number of countries, or taking set samples of individuals within each country, would yield stronger regressions and better results. Additionally, the introduction of controls would

likely account for the aforementioned limitations. Future studies should consider controlling for income inequality, healthcare infrastructure, as well as testing rates.

Despite the limitations of our study and possibilities for improvement, this study has proved useful in supporting the existing literature on this topic, particularly Clegg et al and Boscoe et al. By confirming three important relationships between wealth and cancer incidence, we have helped to reinforce the findings of these studies while using a more expansive scale that provides greater external validity. Additionally, our unexpected negative relationship between tobacco-related cancer and wealth can serve as a point of contention to the findings of previous studies, and will hopefully encourage further research into perhaps this specific relationship due to the economic and cultural nuance of tobacco purchase and usage. Generally, this study supports previous literature on the relationship between cancer incidence and wealth and has served to provide valuable additional data on a more vast scale than previous studies.

The negative relationship between HPV-related cancer and wealth serves as the strongest candidate for real-world policy implementations. HPV-related cancer incidence can be decreased through very definitive measures such as sexual education and access to sexual health services. This negative relationship indicates that with proper resources allocated to the issue, we can hope to see a decline in HPV-related cancer incidence rates. The three positive relationships don't necessarily lend themselves to any particular policy changes; however, the existence of these relationships should be heavily noted and should encourage further research by doctors and scientists in order to identify potential risk factors.

Overall, this study has proved relatively successful in accomplishing its goal of establishing correlations between wealth and cancer incidence rates. Though there were limitations, unexpected results, and inconclusive results, valuable data has been collected and

shared through this study, and it should be used to guide and encourage future research and focus on this important issue.

References

"About the World Bank." World Bank, https://www.worldbank.org/en/about.

"About WHO" World Health Organization, World Health Organization, https://www.who.int/about.

Boscoe, F.P., Johnson, C.J., Sherman, R.L., Stinchcomb, D.G., Lin, G. and Henry, K.A. (2014), The relationship between area poverty rate and site-specific cancer incidence in the United States. Cancer, 120: 2191-2198. https://doi.org/10.1002/cncr.28632

"Cancers Associated with Human Papillomavirus (HPV)." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 13 Dec. 2021, https://www.cdc.gov/cancer/hpv/basic_info/cancers.htm.

"Cancer Today." Global Cancer Observatory, 2021, https://gco.iarc.fr/today/online-analysis-table?v=2020&mode=cancer&mode_population=continents&population=900&populations=900&key=asr&sex=0&cancer=39&type=0&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=17&group_cancer=1&include_nmsc=0&include_nmsc_other=1.

Clegg, Limin & Reichman, Marsha & Miller, Barry & Hankey, Ben & Singh, Gopal & Lin, Yi & Goodman, Marc & Lynch, Charles & Schwartz, Stephen & Chen, Vivien & Bernstein, Leslie & Gomez, Scarlett & Graff, John & Lin, Charles & Johnson, Norman & Edwards, Brenda. (2009). Impact of Socioeconomic Status on Cancer Incidence and Stage at Diagnosis: Selected Findings from the Surveillance, Epidemiology, and End Results: National Longitudinal Mortality Study. Cancer Causes and Control. 20. 417-435. 10.1007/s10552-008-9256-0

"GDP per Capita (Current US$)." Data, The World Bank, 2022, https://data.worldbank.org/indicator/NY.GDP.PCAP.CD.

"Principles of Epidemiology." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 18 May 2012, https://www.cdc.gov/csels/dsepd/ss1978/lesson3/section2.html.

Wasta, Vanessa. "New Study Finds That Most Cancer Mutations Are Due to Random DNA ..." Hopkinsmedicine.org, Johns Hopkins Medicine, 23 Mar. 2017, https://www.hopkinsmedicine.org/news/media/releases/new_study_finds_that_most_cancer_mutations_are_due_to_random_dna_copying_mistakes.