ELSEVIER

Cancer Letters 99 (1996) 225–231

CANCER LETTERS

Relationship between latitude and melanoma incidence: international evidence

Tapen Sinha[a,*], Rebecca Benedict[b]

[a]*School of Business, Bond University, Gold Coast, Queensland 4229, Australia*
[b]*Griffith University, Nathan, Queensland 4111, Australia*

Received 7 September 1995; revision received 13 November 1995; accepted 14 November 1995

Abstract

We investigate the relationship between latitude and incidence of melanoma among Whites in different age groups using data from different continents. Relevant data for Whites were obtained for 59 regions around the world. A statistical analysis was carried out using regional dummy variables to eliminate spurious statistical correlation due to clustering. Simple correlation between latitude and incidence of melanoma is strongly negative for almost all age groups. However, once the regional dummies were introduced in the analysis, the relation between latitude and incidence rates disappeared for all age groups but the explanatory power of the regression equation increased substantially.

Keywords: Melanoma; Incidence; Latitude; Correlation; Whites; International

1. Introduction

Skin cancer is the fastest rising and most common form of cancer among the White population in the World. In the United States, there are 600 000 new cases reported every year, or about one-third of all cancer incidence. Of these cases, about 75% are basal cell and 20% are squamous cell carcinomas, both of which are highly treatable and rarely metastasise. The remaining 5% of skin cancer cases are malignant, lethal melanomas. They account for 6700 deaths per year [8].

Between 1973 and 1989, the incidence rate for melanoma increased by 80.6%, more than any other cancer site, and far greater than the 16.1% increase for all sites combined. The mortality rate during the same period for all races and both sexes was 32.1% for melanoma, compared to the 6.1% cancer mortality rate for all sites combined [7].

Solar radiation appears to be the primary risk factor for more than 90% of non-melanoma skin cancer cases, and it has also been linked to melanoma. Evidence for the effect of ultraviolet (UV) light exposure, especially the shorter wavelength UVB rays, on skin cancer shows that a 1% relative increase in UVB radiation may result in a 2% increase in skin cancer incidence. UVB is the radiation that produces tanning and burning in human skin [6].

* Corresponding author. Tel.: +61 75 952263; fax: +61 75 951160.

0304-3835/96/$12.00
SSDI 0304-3835(95)04063-A

Concern about the harmful effects of longer wavelength UVA rays, which are more common in sunlight although less mutagenic than UVB rays, is growing among researchers. Expert opinion represented by NIH Consensus Statement proclaimed that "Overall, data appear to indicate that non-melanoma skin cancer is related to annual cumulative exposure, and that melanoma *may be* related to high intensity, intermittent UV radiation exposure (i.e., sunburns) particularly at a young age" [6]. Incidence of skin cancer is also influenced by degree of skin pigmentation and sex [4].

It is hypothesized that there may be a growing risk of increased exposure to UV radiation due to depletion of the Earth's atmospheric ozone. The effects of a significant loss of the ozone layer on human health are not precisely known, but scientists speculate that skin cancer rates could increase as a result.

2. Objective

It is well known that large doses of UVB (and also, to a lesser extent, UVA) cause skin cancer (and melanoma) in rodents [1]. However, the effects of UVR from (natural) sunlight on humans is less understood. Strongest evidence comes from the study of the association of latitude of a place and mortality of Whites for the United States and Canada [3,4]. Their sample consists of 48 states of the US (they excluded Hawaii and Alaska), and 10 provinces of Canada. Their analysis shows that there is a strong negative correlation between the latitude and age standardized mortality rates from melanoma and other skin cancer.

There are several problems with their study. (1) The correlation between mortality and latitude should be studied *after* the correlation between incidence rate and latitude is established, because death is by no means a sure thing after a person gets skin cancer (or even melanoma). (2) The study does not take into account age effect at all as mortality rates are age standardized for each state. However, we know that mortality rates from skin cancer and melanoma vary tremendously with age. (3) Population mobility between states will distort the picture on mortality. For example, it is possible that more 'hardy' people stay on in northern states whereas 'weaker' people move to a less harsh climate, making mortality rates from all causes in southern states higher than in northern states!

3. Methods

To address the problems of previous studies, we study the incidence of melanoma (instead of mortality rates), we take into account different age groups by studying incidence rate for each age group, and finally, our study has 59 regions across the world reducing the mobility induced bias indicated above (the regions are given in the Appendix).

Age specific incidence rates were obtained from the *Cancer Incidence in Five Continents* [2]. In the past studies, researchers ignored the difference in melanoma incidence rates among different age groups within a given population. We explicitly build this difference into our regression model. Our model gives a better estimate (i.e. our estimated coefficients are both efficient and consistent) of the relationship between latitude and incidence. Thus, we have better control for migration of people between regions in a given country.

For each age level 5,10,...,80,85 and for each of the 59 regions across the world for the White population, we have melanoma incidence statistics. In addition, we collected data of the latitude of the population centre for each region to the nearest degree latitude.

We regressed incidence of melanoma on latitude and on one and two dummy variables (dummy variables were used to separate out regional effects). All regressions were done for each age group separately as well as for the age adjusted incidence rate.

The statistical analysis was conducted using SHAZAM (K. White, Department of Economics, University of British Columbia, Canada, version 7) to generate the dummy variables, checking residuals and to create weights. MICROFIT (H. Pesaran, Cambridge University, England, version 2.0) was used for diagnostic testing of the models used. ET-Econometric Tool Kit (W. Greene, University of California, USA, version 3.0) was used for testing heteroskedasticity.

4. Results

4.1. Aggregate data

For all age groups combined, we first estimated the simplest possible equation to discover the relationship between incidence of melanoma and latitude:

$$\text{Incidence} = \underset{(6.21)}{17.1821} - \underset{(-4.11)}{0.2378}\ \text{latitude}$$

The simple regression seems to bear out the strong negative relation between incidence and latitude. Adjusted R^2 for the equation was 21.5%. That is, 21.5% of the variation in incidence rate across different regions was explained by the regression equation. The numbers in parentheses are the t-ratios for the intercept and the slope coefficients. Both are significant at $P < 0.0001$. However, when we looked at the residuals, several observations had residuals that were outside of ±2 standard errors from the mean. Moreover, the residuals showed signs of heteroskedasticity. A chi-square (with 2 degrees of freedom) Lagrange multiplier test statistic (for testing for heteroskedasticity) produced a value of 7.25 ($P < 0.0071$).

For inferences drawn from this statistical analysis to be meaningful the range of variability of the dependent variable should be the same for every single level of the independent variable. In most instances the regression residuals showed signs of heteroskedasticity: variation among different observations was variable. This can be seen in Fig. 1 where the scatter plot spreads out like a fan.

Heteroskedasticity creates the problem where the regression estimates do not have the usual statistical properties (such as the regression coefficient will no longer have a t distribution). Thus the conclusions drawn from regression results would be incorrect; the presence of heteroskedasticity makes the regression estimates unreliable.

To correct for heteroskedasticity a standard procedure that is commonly used was applied. The correction factor is the construction of a standard deviation of incidence of melanoma across different age groups. This standard deviation becomes the weight in the weighted least squares procedure of regression

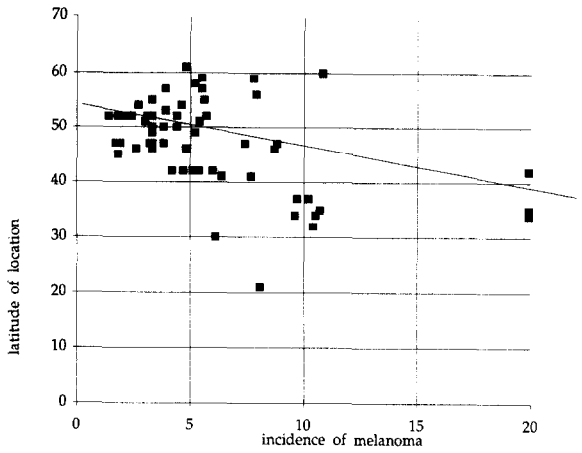

Fig. 1. Scatterplot of incidence and latitude.

analysis [5]. For each of the 59 regions, we had incidence for each age group. Hence, we could calculate the standard deviation of the incidence. The computation required adjustment for population distribution. Each age group for each region produced a weight. These weights were used to calculate the standard deviation of incidence rate for each region.

A weighted least square produced the following estimates:

$$\text{Incidence} = \underset{(6.92)}{21.0590} - \underset{(-4.40)}{0.2895}\ \text{latitude}$$

Adjusted R^2 for the equation was 24.0%. The fit did not improve significantly. The coefficients for the intercept and the slope are still highly significant. However, the residuals still showed signs of heteroskedasticity. A chi-square (with 2 degrees of freedom) Lagrange multiplier test statistic produced a value of 6.28 ($P < 0.0122$).

Inspection of residuals showed that the observations from Australia/New Zealand were still producing errors outside the 2 standard errors from the mean. Thus, it seems that a simple correction for heteroskedasticity did not solve the problem. A deeper problem manifested itself. There seemed to be two sets of observations from the USA and from Australia/New Zealand that were creating an additional heteroskedasticity. This was exhibited by the corresponding residuals from regression being more than 2 standard errors from the mean. To reduce the vari-

ability due to regional effects, the use of dummy variables is commonly warranted. Therefore we introduced two separate dummy variables: one dummy variable for each region Australia/New Zealand and the United States. When the same dummy variable is used for both of the regions the qualitative results (of statistical insignificance of latitude) stay the same (see results below).

We introduced a dummy variable that took a value equal to 1 when the observation was from Australia or New Zealand (a southern Hemisphere dummy variable). Inclusion of such a dummy variable produced a regression equation as follows:

$$\text{Incidence} = \underset{}{11.8524} \underset{(-3.57)}{-0.1389\ \text{latitude}} + \underset{(9.03)}{13.1873\ \text{DANZ}}$$

where DANZ stands for the dummy variable for Australia/New Zealand. The fit of the regression improved dramatically. Adjusted R^2 was 67.5%. The chi square value for heteroskedasticity was 1.91 with 3 degrees of freedom giving us a P value of 0.3811.

Inspection of residuals showed that the observations from the USA were also still producing errors outside the 2 standard errors from the mean. A separate dummy variable was introduced for the observations from the USA. It produced the following regression equation:

$$\text{Incidence} = \underset{}{6.0240} \underset{(-0.60)}{-0.0308\ \text{latitude}} + \underset{(10.01)}{15.0172\ \text{DANZ}} + \underset{(2.97)}{2.9450\ \text{DUSA}}$$

where DANZ stands for the dummy variable for Australia/New Zealand and DUSA is a dummy variable for the USA. The fit of the regression improved again. Adjusted R^2 was 71.5%. The chi square value for heteroskedasticity was 1.99 with 4 degrees of freedom giving us a P value of 0.5736.

Using the same dummy variable (D) for both regions (Australia/New Zealand and the USA) instead of two separate dummy variables produces the following regression results:

$$\text{Incidence} = \underset{}{6.1480} \underset{(-0.4291)}{-0.0332\ \text{latitude}} + \underset{(3.5914)}{5.1753\ \text{D}}$$
$$(1.5482)$$

The adjusted $R^2 = 35.1\%$. The value of the latitude

Table 1

Results from fitted equation: incidence = $a + b_{age}$ (latitude) + error (first set with all the observations, the second set without three observations from Australia and New Zealand)

	Coeff.	t	Adj. R^2 (%)	Coeff.	t	Adj. R^2 (%)
b_5	0.0022	1.04	−0.1	0.0020	0.90	−0.3
b_{10}	−0.0090	−1.89	4.3	−0.0057	1.39	1.7
b_{15}	−0.0898	−4.73	27.0	−0.0612	−4.82	28.8
b_{20}	−0.1905	−4.29	23.1	−0.1161	−4.22	23.4
b_{25}	−0.3412	−4.81	27.6	−0.2296	−4.38	24.8
b_{30}	−0.3491	−4.15	21.9	−0.2355	−3.76	19.3
b_{35}	−0.4488	−4.31	23.3	−0.2710	−3.85	20.1
b_{40}	−0.3329	−3.39	15.3	−0.1815	−2.50	8.7
b_{45}	−0.3838	−3.49	16.2	−0.1957	−2.56	9.2
b_{50}	−0.3967	−3.29	14.5	−0.2112	−2.44	8.2
b_{55}	−0.4098	−3.44	16.2	−0.2020	−2.62	9.6
b_{60}	−0.4352	−4.09	21.3	−0.2749	−3.57	17.6
b_{65}	−0.4896	−4.66	26.3	−0.3737	−4.09	22.2
b_{70}	−0.2977	−2.06	5.3	−0.0542	−0.56	−1.3
b_{75}	−0.2675	−1.92	4.4	−0.0916	−0.76	−0.8
b_{80}	−0.1318	−0.82	−0.7	−0.2288	−0.18	−2.0
b_{85}	−0.2290	−1.19	−0.8	−0.0848	−0.47	−1.6

effect for both regressions is 0.03 and the t statistic is insignificant for both regressions (−0.6 and −0.4, respectively). Thus the results are qualitatively the same for the latitude effect whether the same dummy or two separate dummy variables are used.

The use of two separate dummy variables is the preferred method. The adjusted R^2 is 71.5% for two separate dummy variables and the adjusted R^2 is 35.1% for using one dummy variable for both regions. Note that we have used adjusted R^2 instead of R^2 in the regression analysis. If we add independent variables, R^2 keeps increasing. However, adding independent variables with no explanatory power cannot increase adjusted R^2. Whether adjusted R^2 increases or decreases depends on whether the contribution of the new variable to fit the regression equation more than offsets the correction for the loss of an additional degree of freedom. In fact, adjusted R^2 can decrease or even become negative (in extreme cases). Therefore, an 'over adjustment' problem does not arise with adjusted R^2.

4.2. Analysis of data of different age groups

So far, we have dealt with age adjusted overall incidence of melanoma in the 59 different regions.

Table 2

Results from fitted equation: incidence = a + $b1$ (latitude) + $b2$ (dummy ANZ) + $b3$ (dummy USA) + error

Age	$b1$	$b2$	$b3$	Adj. R^2 (%)
5	0.0019	−0.0268	−0.0405	−0.3
	(0.59)	(−0.28)	(−0.06)	
10	0.0070	0.7658	0.3211	26.7
	(0.25)	(4.28)	(2.72)	
15	−0.0186	4.7621	1.1000	70.0
	(−1.04)	(9.13)	(3.21)	
20	−0.0369	11.7196	2.0483	74.4
	(−0.98)	(10.66)	(2.82)	
25	−0.0475	18.2262	4.8975	72.0
	(−0.73)	(9.61)	(3.91)	
30	−0.0158	20.3824	5.6310	64.0
	(−0.18)	(8.34)	(3.49)	
35	−0.0016	28.2124	7.3511	75.9
	(−0.01)	(11.27)	(4.44)	
40	0.0605	24.6654	6.5048	64.8
	(0.65)	(9.07)	(3.62)	
45	0.0342	28.7176	6.3200	69.9
	(0.35)	(10.14)	(3.38)	
50	−0.0650	30.4546	7.2779	65.4
	(0.57)	(9.25)	(3.35)	
55	0.0225	31.2430	6.1780	72.4
	(0.22)	(10.87)	(3.27)	
60	−0.0338	26.7763	6.2584	67.8
	(−0.33)	(9.14)	(3.24)	
65	−0.2135	19.6574	4.0133	49.8
	(−1.66)	(5.24)	(1.66)	
70	−0.0338	32.2499	0.6808	56.2
	(−0.23)	(7.62)	(0.24)	
75	0.0263	24.4259	3.4500	31.2
	(0.15)	(4.81)	(1.03)	
80	0.4324	23.7830	11.6103	18.9
	(1.89)	(3.47)	(2.72)	
85	0.1091	29.6099	4.5155	19.5
	(0.40)	(3.59)	(0.87)	

But, we know that skin cancer (and melanoma) is a disease affected by sun exposure over a lifetime. Thus, it would be fruitful to see if there is a clear connection between melanoma incidence and latitude for each age group. The estimated equations are similar to the ones described above. But, this time, instead of regressing age adjusted incidence rate for each region, we regressed incidence rate for each age group separately.

The first set with and without Australia/New Zealand is reported in Table 1. There are two striking features. (1) The fit of the equation does not change much when we exclude Australia/New Zealand. (2) The significance of latitude is absent for age groups under 15 and above 70. The overall fit of the equation is good between the age groups above 15 and below 70.

In Table 2, we introduce two dummy variables one for Australia/New Zealand and the other for the US. The striking feature is the t-statistics for the coefficient for latitude is not statistically significant for every age group. Moreover for some age groups the coefficients are of the wrong sign (they are positive as opposed to the hypothesized negative).

Thus, the results of aggregate data for all age groups is repeated for every single age group in the study. In addition, we find the fit of the equations are better for age groups above 10 and below 75.

5. Discussion

Usually the effect of latitude on skin cancer and melanoma is taken for granted. The past evidence is problematic because the studies only concentrated on mortality rather than incidence. Moreover, age specific rates across different regions were not considered. In addition, statistical problems with heteroskedasticity were ignored. With international data, we find that melanoma incidence and latitude show some unusual patterns. Two results are particularly surprising. (1) Without controlling for regional effects by using dummy variables, incidence of melanoma and latitude does not have a monotonic relationship with age, that is, higher age groups do not necessarily show more significant effect of latitude. In fact, the fit of the regression equation worsens as does the significance of the coefficient of latitude. (2) With USA and Australia/New Zealand dummy variables (or even with one dummy variable for both regions), the contribution of latitude to explain melanoma incidence completely disappears. The effects seem to be a regional one.

The obvious question that arises is: How can we explain this result? We do not have any definite answer to this question. However, we have several guesses about it. (1) The data for the USA and Australia/New Zealand are special. If we look at the ethnic background of the dominant groups in these two continents, we find many if not most are of Celtic origin. Celtic ancestry has been linked to skin cancer and melanoma in the literature [8, Table 1]. (2) The Whites who move out of Europe to North America and Australia/New Zealand dramatically alter their lifestyle to indulge in more outdoor activities, absorb more UVR and become more susceptible to skin cancer and melanoma. (3) The possibility cannot be discounted that there is no connection between incidence of melanoma and exposure to UVR from the sun for humans [9].

Acknowledgements

We are indebted to participants at the seminars at various universities. We particularly thank Dr. Tony Hall for his incisive comments on an earlier draft. We also thank the participants of the Western Risk and Insurance Association conference in Honolulu, Hawaii for their input into the preliminary version of the paper in January 1994. Comments from the anonymous referee have improved the exposition of the paper enormously. We, therefore, express our gratitude to the referee. However, we alone are responsible for the contents of this paper.

Appendix. Names of regions for which the melanoma incidence rates were used

CANADA, Alberta, British Columbia, Manitoba, Newfoundland, Ontario, Quebec, Saskatchewan; USA, California Alameda Co Whites, California San Francisco Bay Whites, California Los Angeles Co Other White, Connecticut, Georgia Atlanta Whites, Hawaii Caucasian, Iowa, Louisiana New Orleans Whites, Michigan Detroit Whites, New Mexico Other White, New York State New York City White, Utah, Washington Seattle; ISRAEL, Jews Born Europe and

America; CZECHOSLOVAKIA, W. Slovakia; DENMARK (1973–1976); GERMANY, Hamburg, Saarland, Democratic Republic; FINLAND; FRANCE, BasBhin, Doubs; HUNGARY, Szabolcs Szatmar Co, Vas Co; ITALY, Varese; NORWAY; POLAND, Cieszyn Area, Cracow, Katowice District, Nowy Sacz, Warsaw City, Warsaw Rural; ROMANIA, Cluj Co; SWEDEN; SWITZERLAND, Geneva, Neuchatel, Vaud; ENGLAND, Birmingham W. Mid, N.W. Region, Oxford Region, Trent Region; ENGLAND and WALES, Mersey Reg; South Thames Region; SCOTLAND, East, North, South East, North East, West; YUGOSLAVIA, Slovenia; AUSTRALIA, New South Wales, South Australia; NEW ZEALAND, North Non-Maori.

References

[1] Blum, H. (1969) Quantitative aspects of cancer induction by ultraviolet light. In: The Biological Effects of Ultraviolet Radiation. Editor: F. Urbach. Pergamon Press, London.

[2] IARC (1976) Cancer Incidence in Five Continents, IARC Scientific Publication No. 15, Vol. III. International Agency for Research on Cancer, Lyons.

[3] Elwood, J. et al. (1974) Relationship of melanoma and other skin cancer mortality to latitude and ultraviolet radiation in the United States and Canada. Int. J. Epidemiol., 325–332.

[4] Elwood, J. (1989) The epidemiology of melanoma. Trans. Menzies Foundation, 95–107.

[5] Griffiths, W., Hill, R. and Judge, G. (1993) Learning and Practicing Econometrics. Wiley, New York.

[6] NIH (1989) NIH Consensus Statement. Sunlight, ultraviolet radiation, and the skin. Online May 8–10; 7(8), 1–29.

[7] NIH (1993) NIH Guide, Research in Public and Professional Education for the Prevention and Control of Skin Cancer, Vol. 22, Number 6, February 12.

[8] Preston, D. and Stern, R. (1992) Nonmelanoma cancers of the skin. N. Engl. J. Med., 327, 1649–1662.

[9] Sinha, T. and Benedict, R. (1994) Skin cancer: myths and realities, Paper presented at the Western Risk and Insurance Association Annual Meeting, Honolulu, HI.