Human papilloma virus - A cause of malignant melanoma?

Ilija Barukčić

Internist: Horandstrasse, DE-26441 Jever, Germany.
Email: Barukcic@t-online.de

How to cite this paper: Ilija Barukčić (2017) Human papilloma virus - A cause of malignant melanoma? *viXra-org*, pp. 1-11.
http://vixra.org/author/ilija_barukcic

Received: 28.10.2017
Accepted: 28.10.2017
Published: 28.10.2017

Abstract

Background: The aim of this study is to work out a possible relationship between human papillomavirus (HPV) and malignant melanoma.
Objectives: This systematic review and re-analysis of Roussaki-Schulze et al. available retrospective study of twenty-eight melanoma biopsy specimens and of the control group of 6 patients is performed so that some new inference can be drawn.
Materials and methods: Roussaki-Schulze et al. obtained data from twenty-eight human melanoma biopsy specimens and from six healthy individuals. The presence and types of HPV DNA within biopsy specimens was determined by polymerase chain reaction (PCR). **Statistical Analysis**: In contrast to Roussaki-Schulze et al., the method of the conditio per quam relationship was used to proof the hypothesis that the presence of human papillomavirus (HPV) guarantees the presence of malignant melanoma. In other words, *if* human papillomavirus (HPV) is present, *then* malignant melanoma must also be present. The *mathematical formula of the causal relationship k* was used to proof the hypothesis, whether there is a cause effect relationship between human papillomavirus (HPV) and malignant melanoma. Significance was indicated by a p-value of less than 0.05.
Results: Based on the data published by Roussaki-Schulze et al. we were able to make evidence that the presence of human papillomavirus (HPV) guarantees the presence of malignant melanoma. In other words, human papillomavirus (HPV) is a conditio per quam of malignant melanoma. Contrary to expectation, the data of Roussaki-Schulze et al. based on a very small sample size failed to provide significant evidence that human papillomavirus (HPV) is a cause or the cause of malignant melanoma.
Conclusions: Human papillomavirus (HPV) is a **conditio per quam** of malignant melanoma.

Keywords

Human papilloma virus, Malignant melanoma, Causality

1. Introduction

Melanoma is a rare form of skin cancer and most common in adults. Still, sometimes melanoma is found in adolescents and children too. Melanoma can occur anywhere on the skin and even at locations which are not exposed to ultraviolet (UV) radiation extensively. The incidence and mortality rates of melanoma have risen for many decades while the etiology of melanoma remains unclear. In recent years there has been a dramatic increase [1] in incidence especially in people over the age of 60. Meanwhile, some risk factors for melanoma including exposure to ultraviolet (UV) radiation, high numbers of common naevi, large congenital naevi, multiple and/or atypical (Bauer) naevi (dysplastic naevi) [2] et cetera have been discussed. Thus far, an increased exposure to ultraviolet (UV) radiation [3] is generally considered to be the major environments cause of melanoma. Still, melanoma appears to be an immunogenic tumor too since especially immunosuppressed [4] patients seem to have a higher risk of developing this neoplasm. The primary treatment of melanoma is a surgical excision, sometimes radiotherapy is indicated. Under some specific circumstances a chemotherapy (Decarbazine, Temozolomide, Paclitaxel, Cisplatin, Carboplatin), or an immunotherapy including Interferon (IFN)-Alpha or anti-CTLA-4 antibody ipilimumab or the use of of BRAF/MEK inhibitors (vemurafenib, dabrafenib, trametinib) is offered to patients. The PD-1 antibodies pembrolizumab and nivolumab are approved for therapy of unresectable metastatic melanoma. Melanoma is able to metastasise to human brain with the consequence that patients with brain metastases have a life expectancy of only 3 to 5 months. Even if new melanoma therapies are being developed rapidly, a cure of this many times deadly disease is still not in sight.

2. Material and methods

2.1. Study design of Roussaki-Schulze et al.

Roussaki-Schulze et al. [5] designed a retrospective study to evaluate the presence of human papillomavirus (HPV) DNA in 28 paraffin wax-embedded and formalin-fixed melanoma biopsy specimens and within 6 controls. Roussaki-Schulze et al. determined that five of 28 biopsy melanoma specimens were positive for HPV DNA. In contrast to this, Roussaki-Schulze et al. were not able to detected HPV DNA in any of the biopsy specimens of the control group (0/6). The following 2x2 table (**Table 1**) may illustrate the data as obtained by Roussaki-Schulze et al.

Table 1. The relationshiop between of human papillomavirus (HPV) and malignat melanoma.

		Malignant melanoma		
		YES	NO	SUM
HPV DNA	YES	5	**0**	5
	NO	23	6	29
	SUM	28	6	34

2.2. Methods

All statistical analyses were performed with Microsoft Excel version 14.0.7166.5000 (32-Bit) software (Microsoft GmbH, Munich, Germany).

2.2.1. Conditio per quam

The formula of the *conditio per quam* relationship

$$\mathrm{p}\left(\mathrm{HPV} \rightarrow \mathrm{Malignant\ melanoma}\right) \tag{1}$$

was used to proof the hypothesis: *if* HPV infection *then* malignant melanoma.

Scholium.

Historically, the notion sufficient condition is known since thousands of years. Many authors testified original contributions of the notion material implication only for *Diodorus Cronus.* Still, *Philo the Logician* (~ 300 BC), a member of a group of early Hellenistic philosophers (the Dialectical school), is the main forerunner of the notion *material implication* and has made some groundbreaking contributions [6] to the basics of this relationship. As it turns out, it is very hard to think of the "*conditio per quam*" relationship without considering the historical background of this concept. Remarkable as it is, Philo's concept of the material implications came very close [7] to today's modern concept material implication. In propositional logic, a conditional is generally symbolized as "$p \rightarrow q$" or in spoken language "if p then q". Both q and p are statements, with q the consequent and p the antecedent. Many times, the logical relation between the consequent and the antecedent is called a material implication. In general, a conditional "if p then q" is false only if p is true and q is false otherwise, in the three other possible combinations, the conditional is always true. In other words, to say that p is *a sufficient condition for* q is to say that the presence of p guarantees the presence of q. In other words, it is impossible to have p without q. If p is present, then q must also be present. To show that p is not sufficient for q, we come up with cases where p is present but q is not. In a well-known that the notion of necessary condition can be used in defining what a sufficient condition is (and vice versa). In general, p is a necessary condition for q if it is impossible to have q without p. In fact, the absence of p guarantees the absence of q. A necessary condition is sometimes also called "an essential condition" or a condition sine qua non. To show that p is not a necessary condition for q, it is necessary to find an event or circumstances where q is present but p is not. Especially, necessary and sufficient conditions are converses of each other. Thus far, there is a straightforward way to give a precise and comprehensive account of the meaning of the term necessary (or sufficient) condition itself. On any view, logic has as one of its goals to characterize the most basic, the most simple and the most general laws of objective reality. Especially, in logic, these notions are defined and meanwhile transferred into Biostatistics [8] too. What, then, is a sufficient (or a necessary) condition from the standpoint of (Bio-) statistics? (Bio-) statistics generalizes the notions of a sufficient or a necessary condition *from one single Bernoulli trial to N Bernoulli trials* [9].

2.2.2. Rule of three

In general, describing properties of data (*descriptive statistics*) or drawing conclusions about a population of interest based on a sample drawn from that population (*inferential statistics*) is of key importance in empirical scientific research. Many times, the relation between empirical data and hypotheses is based on a set of measurements of individuals (a sample, a subset of a population) from a certain population (a set of objects which are of interest in a statistical study). The distinction between a sample together with its statistics and a population together with its parameters is of fundamental importance, since every scientific research rests on it. A sample either selected at random or at least representative is used to make inferences about a population from which the same sample was drawn. Generally, the quality of the data is only as good as the sample that produced it. From the sample data various statistics can be calculated. And yet, it is worth noting that despite a long history of progress in statistics, an estimate can be distorted or biased and depends not only on the size of a sample. One such statistics is the key idea of the construction of the i. e. 95% confidence interval. These confidence intervals itself are constructed entirely from the sample data. Confidence intervals for proportions or a population mean of random variables which are not normally distributed in the population can be constructed while relying on the central limit theorem as long as the sample sizes and counts are big enough (i. e. a sample size of n=30 and more). A formula, justified by the central limit theorem, is

$$p_{Crit} = p_{Calc} \pm \left(z_{Alpha/2} \times \left(\sqrt[2]{\frac{1}{N} \times p_{Calc} \times (1 - p_{Calc})} \right) \right) \tag{2}$$

where p_{Calc} is the proportion of successes in a Bernoulli trial process with N trials yielding X successes and N-X failures and z is the 1 – (Alpha/2) quantile of a standard normal distribution corresponding to the significance level alpha. For example, for a 95% confidence level alpha = 0.05 and z is z = 1.96. The Agresti-Coull [10] interval is also another method for calculating binomial confidence intervals. But it is worth noting that another very common technique for calculating binomial confidence intervals was published by Clopper-Pearson [11] too. A faster and an alternative way to determine the lower and upper "exact" confidence interval for p_{Calc} is justified by the F distribution [12]. In this study, we will use *the rule of three* [13] to calculate the confidence interval for p_{Calc}. Briefly sketched, the rule of three can be derived [14] from the binomial model. The rule of three defines that 3/N is an upper 95% confidence bound for a binomial probability p_{Calc} when in N independent trials no events occur [15]. Under conditions where a certain event did not occur [16] in a sample with N subjects (i. e. $p_{Calc} = 0$) the interval from 0 to 3/n is called a 95% classical confidence interval for the binomial parameter for the rate of occurrences in the population. According to the rule of the three the same interval is calculated for a sample sizes of 30-50 or more as

$$p_{Crit} = \left(\frac{3}{N}\right) \tag{3}$$

By symmetry, the one-sided 95 percent confidence interval for only successes (i.e. $p_{Calc}=1$) is

$$p_{Crit} = 1 - \left(\frac{3}{N}\right) \tag{4}$$

The rule of three applies to any Bernoulli trial done n times. The 95% confidence interval for a certain event in the population is the interval from 0 to 3/n, it the same certain event did not occur in a sample with n subjects. By symmetry, for only successes, the 95% confidence interval is (1-3/n). The numerator value of 3.51 may be used for the 97% confidence interval, the numerator value of 4.61 may be used for the 99% confidence interval and the numerator value 5.3 may be used for 99.5% confidence interval.

2.2.3. The mathematical formula of the causal relationship k

The mathematical formula of the causal relationship k [17] and the chi-square distribution [18] were applied to determine the significance of causal relationship between a Helicobacter pylori infection and human gastric cancer. A one-tailed test makes it much more easier to reject a null hypothesis (no causal relationship) while a two-tailed test makes it more difficult to reject a null hypothesis and is more conservative on this account. For this reason, in causal relationship testing, a two-tailed test is preferred as much as possible. In general, a p value of < 0.05 is considered as significant.

2.2.3. The chi square distribution

The chi-squared distribution [18] is a widely known distribution and used in hypothesis testing, in inferential statistics or in construction of confidence intervals. The critical values of the chi square distribution are visualized by **Table 2**.

Table 2. The critical values of the chi square distribution (degrees of freedom: 1).

	p-Value	One sided X^2	Two sided X^2
The chi square distribution	0,1000000000	1,642374415	2,705543454
	0,0500000000	**2,705543454**	**3,841458821**
	0,0400000000	3,06490172	4,217884588
	0,0300000000	3,537384596	4,709292247
	0,0200000000	4,217884588	5,411894431
	0,0100000000	5,411894431	6,634896601
	0,0010000000	9,549535706	10,82756617
	0,0001000000	13,83108362	15,13670523
	0,0000100000	18,18929348	19,51142096
	0,0000010000	22,59504266	23,92812698
	0,0000001000	27,03311129	28,37398736
	0,0000000100	31,49455797	32,84125335
	0,0000000010	35,97368894	37,32489311
	0,0000000001	40,46665791	41,82145620

2.2.4. Fisher's exact test

A test statistics of independent and more or less normally distributed data which follow a chi-squared distribution is valid as with many statistical tests due to the central limit theorem. Especially, with large samples, a chi-squared distribution can be used. A sample is considered as large when the sample size n is n = 30 or more. With a small sample (n < 30), the central limit theorem does not apply and erroneous results could potentially be obtained from the few observations if the same is applied. Thus far, when the number of observations obtained from a population is too small, a more appropriate test for of analysis of categorical data i. e. contingency tables is R. A. Fisher's exact test [19]. Fisher's exact test is valid for all sample sizes and calculates the significance of the p-value (i. e. the deviation from a null hypothesis) exactly even if in practice it is employed when sample size is small. Fisher's exact test is called exact because the same uses the exact hypergeometric distribution to compute the p-value rather than the approximate chi-square distribution. Still, computations involved in Fisher's exact test can be time consuming to calculate by hand. The formula for the hypergeometric distribution, a discrete probability distribution, is

$$p(x) = \frac{\binom{U}{x} \times \binom{N-U}{n-x}}{\binom{N}{n}} \tag{5}$$

where p(x) is the probability of x successes in n draws, without replacement, from a finite population of size N that contains exactly U successes. Barnard's exact test [20], [21] is another exact test which is useful for the analysis of contingency tables.

3. Results

3.1. Human papillomavirus is a *conditio per quam* of human malignant melanoma

Claims.

Null hypothesis:

The presence of human papillomavirus (HPV) guarantees the presence of malignant melanoma.

($p_{Calc} \geq p_{Crit}$).

Alternative hypothesis:

The presence of human papillomavirus (HPV) does not guarantee the presence of malignant melanoma.

($p_{Calc} < p_{Crit}$).

Significance level (Alpha) below which the null hypothesis will be rejected: 0.05.

Proof.

The data of human papillomavirus in patients and healthy control subjects are viewed in the 2 × 2 table (**Table 1**). In general, the proportion of successes of the conditio per quam relationship p(human papillomavirus → malignant melanoma) is calculated [9] as

$$p(\text{Human papilloma virus} \rightarrow \text{Malignant melanoma}) = \frac{(5+23+6)}{34} = \frac{34}{34} = 1$$

The critical value p_{Crit} (significance level alpha = 0.05) is calculated [9] approximately as

$$p_{Crit} = 1 - \frac{3}{34} = 0{,}9117647058823529$$

The critical value is p_{Crit} = 0,9117647058823529 and is thus far less than the proportion of successes calculated as p(human papillomavirus → malignant melanoma) =1. Consequently, we cannot reject the null hypothesis in favor of the alternative hypotheses. The data as published by Roussaki-Schulze et al. do support our Null hypothesis that *human papillomavirus is a conditio per quam of human malignant melanoma*.

In other words, the presence of human papillomavirus (HPV) in biopsy specimens of human skin guarantees the presence of malignant melanoma.

Q. e. d.

3.2. No significant cause effect relationship between a human papillomavirus and malignant melanoma

Claims.

Null hypothesis: (no causal relationship)

There is no causal relationship between human papillomavirus and malignant melanoma (k=0).

Alternative hypothesis: (causal relationship)

There is a causal relationship between human papillomavirus and malignant melanoma (k<>0).

Conditions.

Alpha level = 5%.

The two tailed critical Chi square value (degrees of freedom = 1) for alpha level 5% is 3.841458821.

Proof.

The data for this hypothesis test are illustrated in the 2 × 2 table (**Table 1**). The causal relationship k(human papillomavirus, malignant melanoma) is calculated [9], [17] as

$$k(\text{Human papilloma virus}, \text{Malignant melanoma}) = \frac{((34\times 5)-(6\times 28))}{\sqrt[2]{(28\times 6)\times(5\times 29)}} = +0{,}192212936$$

The value of the test statistic k=+ 0,192212936 is equivalent to a calculated [9] chi-square value of

$$\chi^2_{\text{Calculated}} = 34\times\frac{((34\times 5)-(6\times 28))}{\sqrt[2]{(28\times 6)\times(5\times 29)}}\times\frac{((34\times 5)-(6\times 28))}{\sqrt[2]{(28\times 6)\times(5\times 29)}}$$

$$\chi^2_{\text{Calculated}} = 34\times 0{,}192212936\times 0{,}192212936$$

$$\chi^2_{\text{Calculated}} = 1{,}256157635$$

The calculated chi-square statistic, uncorrected for continuity, is 1,256157635 and equivalent to a P value of 0,262379652831694. The calculated chi-square statistic does not exceed the critical chi-square value of 3.841458821 (**Table 2**). Consequently, we accept the null hypothesis and reject the alternative hypotheses.

There is not a significant causal relationship between human papillomavirus and human malignant melanoma (k=+0,192212936, p Value = 0,262379652831694).

Q. e. d.

4. Discussion

The statistical technique of sample size calculation and power analysis is beyond the scope of this article. However, the sample size can but must not influences research outcomes and an appropriate sample size is one of the crucial factors which determine any well-planned research investigation. Thus far, in the absence of published sample size and power analysis calculations, the findings of a study should be interpreted with caution. In general, we expect the greater the sample size, the smaller the difference that can be detected. In contrast to a study with greater number of cases, *a study with a small sample often leave the null hypothesis unchallenged*. Very small samples may undermine the validity of a study with the consequence that a small study which obtains a nonsignificant or a negative effect is unlikely to be published.

The sample size of the study of Roussaki-Schulze et al. [5] with n=34 cases is very small and the results should be interpreted with some caution. Roussaki-Schulze et al. [5] found that 23 case of 28 malignant melanoma patients were HPV DNA negative (**Table 1**) which support the hypothesis that there may exist other factors but HPV which determines malignant melanoma. However, due to the PCR-kit and technique used, it is possible that Roussaki-Schulze et al. [5] have underestimated the prevalence of HPV with respect to malignant melanoma because the data as provided by Roussaki-Schulze et al. [5] strongly support the hypothesis that **human papillomavirus is a conditio per quam of malignant melanoma** even if the data of Roussaki-Schulze et al. [5] failed to provide some evidence that there is a cause-effect relationship between human papillomavirus and malignant melanoma. In this context it is appropriate to prospectively study about 1000 patients free of malignant melanoma who were **human papillomavirus positive** at the time of enrollment and equally about 1000 patients free of malignant melanoma who were **human papillomavirus negative** at the time of enrollment. During a follow-up between 1 to 10 years, patients should underwent several standardized investigations at the same time at least by 3 independent investigator to search for malignant melanoma including histologic examination (if melanoma positive), serologic testing, DNA PCR and DNA in situ hybridization et cetera. In toto, the result of this investigation does not justify a position which ignores a possible cause-effect relationship between human papillomavirus and malignant melanoma.

5. Conclusion

Human papillomavirus is a *conditio per quam* of malignant melanoma. A more systematical study with a greater sample size is justified to prove a possible cause effect relationship between human papillomavirus and malignant melanoma.

Acknowledgements

None.

References

[1] Whiteman DC, Green AC, Olsen CM. (2016) The growing burden of invasive melanoma: projections of incidence rates and numbers of new cases in six susceptible populations to 2031. *J. Invest. Dermatol.*, **136**, 1161-1171. http://dx.doi.org/10.1016/j.jid.2016.01.035

[2] Bauer J, Garbe C. (2003). Acquired melanocytic nevi as risk factor for melanoma development. A comprehensive review of epidemiological data, *Pigment. Cell. Res.*, **16**, 297-306. https://doi.org/10.1034/j.1600-0749.2003.00047.x

[3] Holland, F. L. and Harvey I. (2002) Adult vs Childhood Susceptibility to Melanoma: Is There a Difference? *Arch Dermatol.*, **138**,1234-1235. https://doi.org/10.1001/archderm.138.9.1234

[4] Wang S. Q,, Setlow R., Berwick M., Polsky D., Marghoob A. A., Kopf A. W., Bart R. S., (2001) Ultraviolet A and melanoma: a review. *J. Am. Acad. Dermatol.*, **44**,837-846. http://dx.doi.org/10.1067/mjd.2001.114594

[5] Roussaki-Schulze A. V., Kouskoukis C., Rammos C, Rallis E., Kontos F., Zafiriou E., Gross G. (2005) Identification of human papillomavirus DNA in melanoma biopsy specimens of Greek population. *International Journal of Clinical Pharmacology Research*, **25**, 145-150.

[6] Astorga M. L. (2015). Diodorus Cronus and Philo of Megara: Two Accounts of the Conditional. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, **7**, 9-16.

[7] Bochenski, J. M., (1961). A history of formal logic, Translated and edited by Ivo Thomas. Notre Dame: University of Notre Dame Press. pp. 14/15.

[8] Barukčić, I. (1989). *Die Kausalität*, Hamburg: Wissenschaftsverlag, pp 218.

[9] Barukčić, K. and Barukčić, I. (2016) Epstein Barr Virus - The Cause of Multiple Sclerosis. *Journal of Applied Mathematics and Physics*, **4**, 1042-1053. http://dx.doi.org/10.4236/jamp.2016.46109

[10] Agresti, A. and Coull, B. A. (1998) Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician*, **52**, 119–126. http://dx.doi.org/10.2307/2685469

[11] Clopper, C. and Pearson, E. S. (1934) The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, **26**, 404–413. http://dx.doi.org/10.1093/biomet/26.4.404

[12] Leemis, L. M. and Trivedi, K. S. (1996) A Comparison of Approximate Interval Estimators for the Bernoulli Parameter. *The American Statistician*, **50**, 63-68. http://dx.doi.org/10.2307/2685046

[13] Rumke C. L. (1975) Implications of the Statement: No Side Effects Were Observed. *N Engl J Med*, **292**, 372-373. http://dx.doi.org/10.1056/NEJM197502132920723

[14] Hanley, J. A. and Lippman-Hand, A. (1983) If Nothing Goes Wrong, Is Everything All Right? *The Journal of the American Medical Assn.*, **249**, 1743-1745. http://dx.doi.org/10.1001/jama.1983.03330370053031

[15] Jovanovic, B. D. and Levy, P.S. (1997) A Look at the Rule of Three. *The American Statistician*, **51**, 137-139. http://dx.doi.org/10.1080/00031305.1997.10473947

[16] Louis, T. A. (1981) Confidence Intervals for a Binomial Parameter After Observing No Successes. *The American Statistician*, **35**, 154. http://dx.doi.org/10.1080/00031305.1981.10479337

[17] Barukčić, I. (2016) The Mathematical Formula of the Causal Relationship k. *International Journal of Applied Physics and Mathematics*, **6**, 45-65. http://dx.doi.org/10.17706/ijapm.2016.6.2.45-65

[18] Pearson, K. (1900) On the Criterion That a Given System of Deviations from the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling. Philosophical Magazine Series, **5**, 157-175. http://dx.doi.org/10.1080/14786440009463897

[19] Fisher, R. A. (1922). On the interpretation of X^2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society*, 85, 87–94. http://dx.doi.org/10.2307/2340521

[20] Barnard G.A. (1945). A New Test for 2 × 2 Tables. *Nature*, 156, 783-784. http://dx.doi.org/10.1038/156783b0

[21] Barnard G.A. (1947). Significance Tests for 2 × 2 Tables. *Biometrika*, 34, 123-138. https://dx.doi.org/10.1093/biomet/34.1-2.123