HUMAN PAPILLOMAVIRUS IS THE CAUSE OF HUMAN PROSTATE CANCER

Ilija Barukčić[1]

[1] Internist, Horandstrasse, DE-26441 Jever, Germany.

ARTICLE INFO

ABSTRACT

Background: Human papillomavirus (HPV) has an important role in the oncogenesis of several malignant diseases. Some observational studies demonstrated the presence of HPV even in human prostate cancer (PC) while other studies failed on this point. The relationship between HPV infection and PC remains unclear. The aim of the present meta-analysis study is to investigate whether HPV serves as a cause or as the cause of PC.

Methods: The PubMed database was searched for suitable articles. Previously published expert reviews and systematic meta-analysis were used as an additional source to identify appropriate articles. Articles selected for this meta-analysis should fulfill the following inclusion criteria: (a) no data access barrier, (b) polymerase chain reaction (PCR) DNA based identification of HPV. The method of the conditio sine qua non relationship was used to prove the hypotheses whether being married is a necessary condition (a conditio sine qua non) of PC. In other words, *without* being married *no* PC. The method of the conditio per quam relationship (sufficient condition) was used to prove the hypotheses if HPV is present in human prostate tissues then PC is present too. The mathematical formula of the causal relationship k was used to prove the hypothesis, whether there is a cause effect relationship between HPV and PC. Significance was indicated by a p-value (two sided) of less than 0.05.

Results: In toto *more than 136 000 000 cases and controls* were re-analysed while more than 33 studies were considered for a meta-analysis. Several studies support the hypotheses without being married no PC. All the studies considered for a re-analyses support the null-hypotheses if HPV then PC, while the cause effect relationship between HPV and PC was highly significant.

Conclusions: Human papillomavirus is the cause of human prostate cancer.

INTRODUCTION

Human papillomaviruses are frequent pathogens of sexually transmitted diseases and have been implicated in the pathogenesis of a variety of malignancies including the cervical cancer [1,2] and prostate cancer [3]. Several different studies have investigated HPV in relation to prostate cancer with mixed results and the role of HPV infection in the development of prostate cancer is still not yet clarified. Despite great research efforts and an increasing number of studies conducted to evaluate the relationship between HPV infections and prostate cancer, the results remain uncertain. Whether HPV is involved in the pathogenesis of prostate cancer has been a subject of great controversy, the etiology of prostate cancer is still not known in detail. To date, some risk factors [4] for prostate cancer are established and limited to certain genetic polymorphisms, family history of prostate cancer, race, age, height, physical activity, BMI, total energy consumption, intakes of calcium, tomato sauce and alpha-linolenic acid and cigarette smoking history while evidence is conflicting [5].

Prostate cancer is one of the major causes of disease and mortality among men and a growing concern in global public health. Each year more than 1.6 million cases are diagnosed annually, and the mortality burden has risen to over 360,000 deaths per year [6]. Even if studies of HPV infections in sex partners are limited, Human papillomavirus infection is estimated to be one of the most common sexually transmitted infections. In heterosexually active couples, up to a total of 72.9% of their male partners are HPV positive [7]. A discovery of an infectious agent as the cause or a cause of prostate cancer would be of great medical importance. Studies agree on the fact that Human papillomavirus (HPV) is most commonly transmitted through sexual activity. Thus far, marriage could have influence on prostate cancer. Several studies analysed the impact of marital status (single, married, divorced/separated, and widowed) on prostate cancer with contradictory results. The study of Newell[8] et al. does not support any association between an incidence of prostate cancer and marital status while Liu[9] et al. found evidence that marital status was associated with better outcomes for the survival of prostate cancer patients.

Badar et al. [10] reported no evidence of human prostate cancer in very young and thus far sexually inactive male children. These data provide some biological support for HPV transmission between sex partners as the route to prostate cancer. In particular, most HPV infections are asymptomatic or subclinical and become undetectable over time while more than 200 types [11, 12] of human papillomaviruses have been identified. About 40 types infect the anogenital region and have been further classified into low-risk types (e.g., 6 and 11) and high-risk types (e.g., 16, 18, 31, and 45). Several expert reviews published investigated whether HPV infection is a risk factor for PC but opposing reports were stated. A 2011 systematic review of 14 articles by Lin et al. (Lin et al., 2011) documented an increased prevalence of HPV-16 DNA in PCa tissues. An expert review published in 2013 by Hrbacek [13] et al. concluded that there was no evidence to support an association between Human papillomavirus infection and prostate cancer. Yang [14] et al. investigated the prevalence of HPV in prostate cancer by pooling data of 46 studies with 4919 prostate cancer cases and concluded that HPV infections may contribute to the risk of prostate cancer. The results of the meta-analysis of Bae[15] suggest that Human papillomavirus 16 infection is relatively frequent in prostate carcinoma. The thirty studies review of Russo [16] et al. 2018 suggested that HPV-16 infection could represent a risk factor for PC. To clarify the contradictory results of these investigations, we have carried out meta-analysis with updated data to obtain a more precise picture of the association between HPV and PC.

MATERIAL AND METHODS

Search strategy

No electronic database is able to contain all the necessary information needed to perform a review. In order to choose an appropriate database for a review and to ensure to retrieve the type of information required, it is necessary to consider what subjects are covered in a particular database. In this context, the electronic database PubMed is associated with several and serious shortcomings. PubMed does not necessarily index (i.e. cover) every type of a high quality medical publication available nor does the same database cover every language available. Still, for the questions addressed in this paper, PubMed was searched for appropriate studies conducted in any country which investigated the relationship between HPV and PC. The search in PubMed was performed while using some medical key words like "prostate cancer" and "human papillomavirus" and "PCR". The articles found where saved as a *.txt file while using PubMed's support (Menu: Send to, Choose Destination (Button): File, Choose Format: Abstract (text), choose sort by: publication date, click bottom "create file"). The created *.txt file was converted into a *.pdf file.

The abstracts where studied within the *.pdf file. Considered articles for a review were those which provided access to data without any data access barrier; no data access restrictions were accepted. In assessing the shortcomings of PubMed, additionally, appropriate review articles and references published within the same were checked. Furthermore, studies were excluded if data published were self-contradictory or insufficient to calculate the necessary causal relationship.
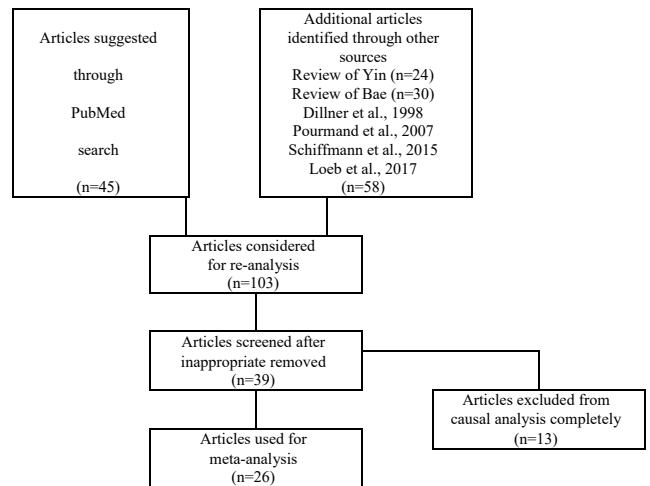

Figure 1. Studies identification in search strategy. Adopted from PRISMA [17,18] 2009 Flow Diagram

Statistical analysis

All statistical analyses were performed with Microsoft Excel version 14.0.7166.5000 (32-Bit) software (Microsoft GmbH, Munich, Germany). An error found in a printed work (*corrigendum*) has to be corrected even after peer-review processes and after publication. In order to increase the transparency, to correct some of the misprints of former publications and to simplify the understanding of this article several of the following lines are repeated sometimes word by word and taken from my former publications.

The Data of the Studies Analysed

The studies reviewed in this publication investigated the relationship between HPV and PC while using the highly sensitive PCR technique. The data of the studies reviewed in this publication are presented in more detail by several tables (Table 1, Table 3, Table 4, Table 5, Table 6, Table 7).

The Data of the Studies not analysed

The data which were self-contradictory [19, 20] are viewed by Table 2 and have not been considered for a review of the causal relationship. The reason for the contradiction is highlighted with bold letters. Still, the majority of these studies (i. e. 12/20) support the hypotheses that HPV is a sufficient condition of PC since X^2 (IMP) is less than 3.841458821.

Martial status and HPV positivity

The Iranian study of Pourmand [21] et al. and Ghasemian [22] et al. provided detailed information about the martial status and HPV positivity. The data on the relationship between martial status and HPV positivity are viewed by the Table 3. The data of study of Pourmand [21] et al. and Ghasemian [22] et al. are viewed in detail by Table 4.

Martial status and prostate cancer

Any debate concerning the relationship between a heterosexual marital status of a men and prostate cancer may not extend into a deep debate about what marriage fundamentally, is. Historically, in many human societies marriage functioned primarily as an institution to control inheritance, to share resources and labour and to create kinship bonds. Nowadays marriage involves a relationship of friendship, of love or companionship and is still the main social institution to regulate sex, the sexual relationship between a man and a woman, their reproduction goals and family life.

A heterosexual marriage as such does not exclude pre-, post- or extra-marital sex activity completely, but is able to contribute to the reduction of the non-marital sexual activity significantly. On this approach, it is of course very difficult to treat the marital status as such as a cause or as the cause of prostate cancer. Still, it is reasonable to assume that a group of men, which is not married, should suffer from prostate cancer to the same amount as a group of men which is married. More or less, the martial status is an indicator of a frequent and regular sexual activity and is of use to prove a hypotheses whether an infectious agent is a cause or the cause of prostate cancer in heterosexually active couples.

The Iranian study of Ghasemian [22] et al. provided information about the martial status and prostate cancer. The data on the relationship between martial status and prostate cancer as published by Ghasemian [22] et al. are viewed by the Table 5.

The Iranian data on the relationship between martial status and prostate cancer were compared with the US data as published by the study of Dillner [23] et al. The data as published by Dillner et al. are viewed by the Table 5.

Schiffmann [24] et al. examined the impact of marital status on prostate cancer in Hamburg. The total population of Hamburg 2016 is about 1 787 000. In Germany, the total population 2018 was 82 521 653 inhabitants while 49.3 % were male. Assuming that ~49.3% of Hamburg's population were male too, we obtain a sample size Hamburg's male population about (1 787 000 * 0.493) 880 991 inhabitants. The controls = (male population) - (prostate cancer cases) are given as (880 991 - 8 088) = 872903. In Germany, there are (82 521 653 * 0.493) about 40 683 174 male inhabitants while (21 667 700/40 683 174) or 53.2596105 % are married or divorced or separated or widowed. We assume that about 53.2596105 % of Hamburg's male population are married or divorced or separated or widowed too. We obtain the number of Hamburg's married or divorced or separated or widowed as (880 991*0.532596105) = 469 212. Furthermore, in Hamburg the number of married or divorced or separated or widowed which were without cancer (*controls and positive*) in about (469 212 - 8 088) = 461 124. The data of Schiffmann [24] et al. are viewed by the Table 5.

In the year 2009, the US population was about 306 771 529 (https://www.cdc.gov/nchs/data/dvs/national_marriage_divorce_rates_00-16.pdf) while **116 666 000** were men (https://www.census.gov/data/tables/time-series/demo/families/marital.html). In the year 2009 from about 116 666 000 men **39 052 000** never were married (https://www.census.gov/data/tables/time-series/demo/families/marital.html). or (116 666 000 – 39 052 000) = **77 614 000** were married or divorced or separated or widowed. In the year 2009, based on the statistics of the U.S. population presented by SEER Cancer Statistics Review stat facts (https://seer.cancer.gov/statfacts/html/prost.html) there were about 155.54 new Prostate cancer cases per 100000 males and females. In other words, in the year 2009 there were (306 771 529 / 100 000)*155.54 = **477 152** new U. S. cancer cases. According to the study of Huang [25] et al. 9 072 men from 95 846 prostate cancer cases never were married, while about 28% of the entire population in the United States was considered. We calculate that ((9 072/958 46)*477 152) = **45 163** of all new prostate cancer cases in the U. S. in the year 2009 never were married. In about 477 152 – 45 163 = **431 989** U. S. men with new prostate cancer cases were married or divorced or separated or widowed. In the following, about (77 614 000 – 431 989) = **77 182 011** of all U.S. men were married or divorced or separated or widowed but without prostate cancer.

The control group is calculated as (116 666 000 – 477 152) = **116 188 848**. The approximate data about the relationship between martial status and prostate cancer in the U.S. in the year 2009 are view by the table 6.

Table 6: The U.S. Population 2009 and prostate cancer

| | | Prostate cancer <B> | | |
		Yes	No	Men total
Married	Yes	431 989	77 182 011	77 614 000
<A>	No	45 163	39 006 837	39 052 000
	Total	477 152	116 188 848	116 666 000

The 2x2 Table

The meaning of the abbreviations a_t, b_t, c_t, d_t, N_t of the data table used are explained by a 2 by 2-table (Table 8).

Table 8. The sample space of a contingency table.

| | | Conditioned B_t (Outcome) | | |
		Yes = 1	Not = +0	Total
Condition A_t	**Yes =+1**	a_t	b_t	A_t
(risk factor)	**Not = +0**	c_t	d_t	$\underline{A_t}$
	Total	B_t	$\underline{B_t}$	N_t

In general it is $(a_t+b_t) = A_t$, $(c_t+d_t) = \underline{A_t}$, $(a_t+c_t) = B_t$, $(b_t+d_t) = \underline{B_t}$ and $a_t+b_t+c_t+d_t=N_t$. Equally, it is $B_t+\underline{B_t} = A_t + \underline{A_t} = N_t$. In this context, it is $p(a_t)=p(A_t \cap B_t)$, $p(A_t) = p(a_t)+p(b_t)$ or $p(A_t)= p(A_t \cap B_t)+ p(b_t) =p(A_t \cap B_t)+p(A_t \cap \underline{B_t})$ while $p(A_t)$ is not defined as $p(a_t)$. In the same context, it is $p(B_t) = p(a_t)+p(c_t) = p(A_t \cap B_t) +p(c_t)$ and equally in the same respect $p(\underline{B_t}) = 1- p(B_t) =p(b_t)+p(d_t)$. Furthermore, the joint probability of A_t and B_t is denoted in general by $p(A_t \cap B_t)$. Thus far, it is $p(A_t \cap B_t) = p(A_t) - p(b_t) = p(B_t) - p(c_t)$ or in other words it follows clearly that $p(B_t) + p(b_t) - p(c_t) = p(A_t)$. In general, it is $p(a_t)+p(c_t)+p(b_t)+p(d_t) = 1$.

Independence
In the case of independence [26, 27] of A_t and B_t it is generally valid that

$$p(A_t \cap B_t) \equiv p(A_t) \times p(B_t) \qquad (1)$$

Sufficient Condition (Conditio per Quam)

The mathematical formula of the *sufficient* condition relationship [2, 19, 20, 28-32] (conditio per quam) of a population is defined as

$$
\begin{aligned}
p(A_t \to B_t) &\equiv \frac{a_t + c_t + d_t}{N_t} \\
&\equiv p(a_t)+p(c_t)+p(d_t) \\
&\equiv p(d_t)+p(B_t) \\
&\equiv +1
\end{aligned}
\qquad (2)
$$

and used to prove the hypothesis: *if* A_t *then* B_t or is taken to express that the occurrence of an event A_t is a sufficient condition [33, 34] for existence or occurrence of an event B_t.

The occurrence of an event A_t is a sufficient condition for occurrence of the event B_t or B_t is a necessary condition for A_t. In other words, sufficient and necessary conditions [33, 34] converse relations.

Self-contradictory data I

Let $p(A_t)$ denote the probability of the condition (i.e. risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t$ and $B_t)$ denote the joint probability of A_t and B_t. Under conditions were the relationship between A_t and B_t is determined by *a sufficient condition* we obtain the formula $p(A_t \rightarrow B_t) = p(A_t$ and $B_t) + (1 - p(A_t)) = 1$ and it is equally $p(A_t$ and $B_t) = p(A_t)$. In general, under these circumstances it is

$$
\begin{array}{ccc}
k(A_t \cap B_t) & \geq & 0 \\
& \text{Proof.} & \\
+1 & \geq & p(B_t) \\
p(A_t) & \geq & p(A_t) \times p(B_t) \\
\text{If } p(A_t \rightarrow B_t) \equiv p(A_t \cap B_t) + (1 - p(A_t)) \equiv 1 \quad \text{then} & p(A_t \cap B_t) = p(A_t) \\
p(A_t \cap B_t) & \geq & p(A_t) \times p(B_t) \\
p(A_t \cap B_t) - p(A_t) \times p(B_t) & \geq & 0 \\
\dfrac{p(A_t \cap B_t) - p(A_t) \times p(B_t)}{\sqrt[2]{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}} & \geq & \dfrac{0}{\sqrt[2]{\ldots}} \\
k(A_t \cap B_t) & \geq & 0 \\
& \text{Q.e.d.} &
\end{array}
\tag{3}
$$

In many problems, data gained from some observations provide an opportunity to increase the degree of confidence, when a decision is made to *either* accept the null hypotheses *or* accept the alternative hypothesis. Clearly, the null hypotheses and the alternative hypotheses are mutually exclusive thus that exactly one of the hypotheses must be true. Still, the quality of data varies and data as such do not assure an exact and true picture of reality with the consequence that a decision of an investigator can be wrong in principle.

An investigator can accept null hypotheses as true even if the same is wrong and vice versa. It is possible to accept alternative hypotheses as true even if the same is wrong. Data which provide evidence that A_t is a sufficient condition of B_t must not in the same respect provide evidence that there is a significant cause effect relationship. In fact, our ability to recognize conditions or risk factors might be seriously endangered by treating a cause as being identical with a condition. A cause is a condition too but not vice versa. A condition must not be a cause. Therefore and due to mathematical requirements, a significant cause effect relationship is not necessary to establish a significant sufficient condition relationship. The analysis of alleged examples can show, among other things, how sufficient conditions should be understood, especially with relation to causation. For example there might be wet and dry conditions of a street while the relationship between raining and the state of a street is measured or investigated in a case control study. Rain (A_t) is generally known to be a sufficient condition for wet streets (B_t). In other words, rain as such guarantees that the event 'the street is wet' occurs. *If* it is raining *then* the street is wet (n=1000). Every time it is raining, the street gets wet, which was measured n=4 times. It isn't raining and the street isn't wet was documented n=500 times. It is raining and the street isn't wet was not measured at all (n=0).

However, the presence of a street which is wet is not enough to conclude that was raining. In point of fact, there are also other possible factors (n=496) which are able to make the street wet. The neighbor might have poured water on the street; a lorry may have lost oil et cetera.

Table 9: Raining and street

		The street is wet $<B_t>$		
		Yes	No	Total
It is raining $<A_t>$	Yes	4	**0**	4
	No	496	500	996
	Total	500	500	1000

$$k = \mathbf{0.0634}$$
$$\text{p value (k)} = \mathbf{0.06212481}$$

The chi square value can be calculated as

$$
\chi^2(A_t \rightarrow B_t) = \frac{(b^2)}{(N_t - B_t)} = \frac{(b^2)}{(A_t)}
\tag{4}
$$

or as

$$
\chi^2(A_t \rightarrow B_t) = \frac{(0^2)}{(500)} = \frac{(0^2)}{(4)} = 0
\tag{5}
$$

Independent of the study design, both methods provide the same Chi square value. Depending upon study design and other factors, it is possible that data support the null-hypotheses that A_t is a sufficient condition of B_t while the causal relationship is not significant. Such data are not self-contradictory and can be used for the analysis of conditions or risk factors, but not for causal analysis. In particular, as proofed before, the cause effect relationship should be greater or equal to zero or $k \geq 0$. A study group investigated once again the relationship between the risk factor rain and the outcome "street is wet" and obtained the following data.

Table 10: Antidot

		The street is wet $<B>$		
		Yes	No	Total
It is raining $<A>$	Yes	4	**20**	24
	No	496	480	976
	Total	500	500	1000

k =	**-0.1045**
p value (k) =	**0.00056237**
Odds ratio =	0.1935
95% CI (Odds ratio) =	(0.0657; 0.5704)
IF $<A_t>$	THEN $<B_t>$
p (IMP) =	0.9800
X² (IMP) =	**15.8438**

Even if the relationship between rain and the state of a street is clear, it is necessary to consider the following case. Different conditions of investigation can have an impact on the quality of conclusions based on data gained by studies. The data presented before would support a null hypotheses that the rain has a protective effect against the wetness of a street. It is *raining* and *the street isn't wet* was measured **n=20** times. How is such a result possible? One reason for such a fundamental error can be an incorrect definition of cases and controls. It is possible that the street was wet but not recognized as being wet or not recorded as being wet although it has rained. In other studies, the controls may have been contaminated et cetera. A mismatch of cases and controls excluded, it is possible that the control group possess an *antidote against the effect of the rain on the street.* In other words, it is possible that the measurements were performed under conditions were the street was protected against the effect of the rain i.e. by a great (transparent) tent or something similar thus that the street could not become wet even if it was raining. The conditions under which investigations are performed can have influence on the quality of data and the validity of the conclusions drawn. Truth is one of the central subjects in scientific inquiry. And yet, despite a long history of debate in its own right going back for more than thousands of years the truth was, is and stays relative. Narrowly speaking, the truth or falsity of a scientific conclusion is based on many factors, among them the quality of data and the circumstances of investigation and has the potential to vary, sometimes extensively. In addition to a careful systematic observation and experiments, any scientific success achieved requires appropriate methods of scientific inference which enable us to infer beyond what is known by observation.

The X² Test of Goodness of Fit of a Sufficient Condition

A random sample of observations can come from a particular distribution (sufficient condition distribution) but must not. The X² test of goodness-of-fit is one appropriate method for testing the null hypotheses that a random sample of observations comes from a specific distribution (i.e. sufficient condition) against the alternative hypotheses that the data have some other distribution. The additive property of X² distribution may sometimes be used as an additional test of significance. In this case, the continuity correction should be omitted from each X² value. Under conditions where the chi-square goodness of fit test cannot be used it is possible to use an approximate and conservative (one sided) confidence interval known as *the rule of three.*

The X² distribution is a particular type of a gamma distribution and widely applied in the field of mathematical statistics. The applicability of using the Pearson chi-squared statistic in cases where the cell frequencies of 2× 2 contingency tables are not greater than five is widely discussed [35] in literature and the use of Yate's continuity correction [36] is proposed.

However, studies provided evidence that incorporating Yate's continuity correction is not essential [37, 38]. Still, using *the continuity correction* [36], the chi-square value of a conditio per quam distribution is derived [32] as

$$\chi^2\left(A_t \to B_t\right) \equiv \frac{\left(b_t - \left(\frac{1}{2}\right)\right)^2}{\left(A_t\right)} + 0 = 0 \tag{6}$$

or alternatively as

$$\chi^2\left(A_t \to B_t\right) \equiv \frac{\left(b_t - \left(\frac{1}{2}\right)\right)^2}{\left(\underline{B}_t\right)} + 0 = 0 \tag{7}$$

Self-contradictory data II

As long as the whole population is not investigated, the study design of a case-control or of another study should assure that the same chi square value can be achieved from the data recorded. This condition is seldom provided by studies published. Many times, it is

$$A_t < \left(N_t - B_t\right) \tag{8}$$

Multiplying by $b^2/(A_t \times (N_t - B_t))$ it is

$$A_t \times \frac{\left(b^2\right)}{\left(\left(A_t\right) \times \left(N_t - B_t\right)\right)} < \left(N_t - B_t\right) \times \frac{\left(b^2\right)}{\left(\left(A_t\right) \times \left(N_t - B_t\right)\right)} \tag{9}$$

Simplifying, we obtain

$$\frac{\left(b^2\right)}{\left(N_t - B_t\right)} < \frac{\left(b^2\right)}{\left(A_t\right)} \tag{10}$$

or

$$\chi^2\left(A_t \to B_t \mid \left(N_t - B_t\right)\right) < \chi^2\left(A_t \to B_t \mid \left(A_t\right)\right) \tag{11}$$

If **$A_t < (N_t - B_t)$** then X²($A_t{\to}B_t \mid (N_t - B_t)$) < X²($A_t{\to}B_t \mid (A_t)$) and the question arises, which X² should be used. Statistical tests primary handle samples and not populations. Still, an appropriate sample should assure that something insignificant stays significant and that a test correctly rejects a false null hypothesis. Circumstances were the inequality

$$A_t > \left(N_t - B_t\right) \tag{12}$$

leads to another point of view. Multiplying by $b^2/(A_t \times (N_t - B_t))$ it is

$$A_t \times \frac{\left(b^2\right)}{\left(\left(A_t\right) \times \left(N_t - B_t\right)\right)} > \left(N_t - B_t\right) \times \frac{\left(b^2\right)}{\left(\left(A_t\right) \times \left(N_t - B_t\right)\right)} \tag{13}$$

Simplifying, we obtain

$$\frac{\left(b^2\right)}{\left(N_t - B_t\right)} > \frac{\left(b^2\right)}{\left(A_t\right)} \tag{14}$$

or

$$\chi^2\left(A_t \to B_t \mid \left(N_t - B_t\right)\right) > \chi^2\left(A_t \to B_t \mid \left(A_t\right)\right) \tag{15}$$

Again, a study design which is grounded on the assumption that $A_t > (N_t - B_t)$ leads to $X^2(A_t \to B_t \mid (N_t - B_t)) > X^2(A_t \to B_t \mid (A_t))$ and the question arises again, which X^2 is valid and which X^2 should be used. Thus far, suppose exact and ideal theoretical experimental conditions and that there is for sure a conditio per quam relationship between A_t and B_t. Under such conditions any study design will not be able to produce any bias and both different Chi square values are equivalent to each other or it is

$$\chi^2\left(A_t \to B_t \mid \left(N_t - B_t\right)\right) = \chi^2\left(A_t \to B_t \mid \left(A_t\right)\right) \tag{16}$$

Even under conditions were a study design assures the relationship before, the basic relationship will be recognized while the bias will be reduced. Thus far, an ongoing controversy raised by the issues discussed before how strong is the support of a sample for the hypotheses, whether we may rely on the hypotheses in our decisions at all, and so on can be shortened by an appropriate study design. Under conditions were a study design demands that

$$\chi^2\left(A_t \to B_t\right) = \chi^2\left(A_t \to B_t \mid \left(N_t - B_t\right)\right) = \chi^2\left(A_t \to B_t \mid \left(A_t\right)\right) \tag{17}$$

it is equally

$$\chi^2\left(A_t \to B_t\right) = \frac{\left(b^2\right)}{\left(A_t\right)} = \frac{\left(b^2\right)}{\left(N_t - B_t\right)} \tag{18}$$

or

$$\frac{\left(b^2\right)}{\left(A_t\right)} = \frac{\left(b^2\right)}{\left(N_t - B_t\right)} \tag{19}$$

while N_t is the sample size. In this example, we have not used the *continuity correction*. Rearranging this equation, we obtain

$$\frac{1}{\left(A_t\right)} = \frac{1}{\left(N_t - B_t\right)} \tag{20}$$

If $b_t=0$, we set $b_t=1$. Simplifying it is

$$A_t = N_t - B_t \tag{21}$$

or

$$A_t + B_t = N_t \tag{22}$$

Before considering the definition of an *index of unfairness* (IOF) we normalize the relationship between A_t and B_t. We get

$$\frac{A_t + B_t}{N_t} = \frac{N_t}{N_t} = 1 \tag{23}$$

and *the index of unfairness* (IOU) follows as

$$IOF \equiv \left(\left(\frac{A_t + B_t}{N_t}\right) - 1\right) = 0 \tag{24}$$

The range of *the index of unfairness* is [-1;+1]. Let us assume our *null hypotheses* is: *without A_t no B_t*. An IOU = -1 indicates an extremely unfair study design and provides an unfair advantage to the party which tries to reject the null hypothesis. An IOU = +1 indicates an extremely unfair study design too because such a study design provides an unfair advantage to the party which tries to accept the null hypothesis. In this context let IOF=0 denote a *fair* study design, let $0 < |IOF| \leq 0.25$ denote an *unfair* study design, let $0.25 < |IOF| \leq 0.5$ denote a *very unfair* study design, let $0.5 < |IOF| \leq 0.75$ denote a *highly unfair* study design and let $0.75 < |IOF| \leq 1$ denote an *extremely unfair* study design.

The principle of equality of scientific arms

The relation between data and hypotheses is of key importance in almost all empirical research. The foundations of statistical methods should be logically and mathematically correct. Statistical methods which are relating hypotheses in the light of empirical facts may enable us even to extrapolate from data to predictions and general facts. Data have an impact on a hypothesis, but the impact should depend on the data themselves and not just on the study design of the researcher. The guarantee of a fair study design is fundamental in any empirical scientific research and of every modern medical investigation. The framework of a fair study design should obey especially *the principle of equality of arms* which is a central feature of every scientific combat to ensure completely only the discovery of the truth.

The principle of equality of arms leaves no room for defending material interest, ideological position or wishful thinking but requires that *advocates of a special null hypotheses* and *opponents of a the same null hypotheses* have the same data or scientific means at their disposal. One could sum up the principle of equality of scientific arms by saying that no party should have an unfair advantage over the other party especially due to study design. Put in other terms, any scientific research is not complete without the notion of fairness. Ignoring the historical origins and theoretical foundations of the principle of equality of scientific arms a fair and careful study design directed to the goal that *a correct null-hypotheses has to be accepted* and that *a false null-hypotheses has to be rejected* is the core of the evaluations to determine how believable a hypotheses is. The question arises, therefore, how can such a goal be achieved? Under conditions were the data are analyzed by a Chi-square goodness of fit test, the equality of scientific arms is given, if the index of unfairness is IOF = 0.

Self-contradictory data III

Necessary Condition (Conditio Sine Qua Non)

Among the many generally valid natural laws and principles under which nature or matter itself assures its own self-organization, a relationship between events denoted as a necessary condition [2, 19, 20, 28-32] (a conditio sine qua non) is one of the most important. A necessary (or an essential) event or condition A_t for some event B_t is a condition that must be satisfied in order to obtain B_t. In this respect, to say that an event A_t with its own probability $p(A_t)$ is at the same (period of) time a necessary condition for another event B_t with its own probability $p(B_t)$ is equivalent to say that it is impossible to have B_t without A_t. In other words, *without* A_t *no* B_t or the absence of A_t guarantees the absence of B_t. The mathematical formula of the *necessary* condition relationship (conditio sine qua non) of a population is defined as

$$
\begin{aligned}
p\left(A_t \leftarrow B_t\right) &\equiv \frac{a_t + b_t + d_t}{N_t} \\
&\equiv p\left(a_t\right) + p\left(b_t\right) + p\left(d_t\right) \\
&\equiv p\left(a_t\right) + \left(1 - p\left(B_t\right)\right) \\
&\equiv p\left(d_t\right) + \left(p\left(a_t\right) + p\left(b_t\right)\right) \\
&= p\left(d_t\right) + p\left(A_t\right) \\
&\equiv +1
\end{aligned}
\tag{25}
$$

The X² Test of Goodness of Fit of a Necessary Condition

Under conditions where the chi-square goodness of fit test cannot be used it is possible to use an approximate and conservative (one sided) confidence interval known as *the rule of three*. Using *the continuity correction,* the chi-square value of a *conditio sine qua non* [2, 19, 20, 28-32] distribution before changes to

$$
\chi^2\left(A_t \leftarrow B_t\right) \equiv \frac{\left(c_t - \left(\frac{1}{2}\right)\right)^2}{\left(B_t\right)} + 0 = 0
\tag{26}
$$

Depending upon the study design, another method to calculate the chi-square value of a *conditio sine qua non* distribution is defined as

$$
\chi^2\left(A_t \leftarrow B_t\right) \equiv \frac{\left(d_t - \left(\frac{1}{2}\right)\right)^2}{\left(\underline{B}_t\right)} + 0 = 0
\tag{27}
$$

Self-contradictory data IV

Let $p(A_t)$ denote the probability of the condition (i.e. risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t$ and $B_t)$ denote the joint probability that A_t and B_t will occur/has occurred.

Under conditions were the relationship between A_t and B_t is determined by *a necessary condition* $p(A_t \leftarrow B_t) = p(A_t$ and $B_t)$ $+(1 - p(B_t)) = 1$ it is equally $p(A_t$ and $B_t) = p(B_t)$ and in general

$$
\begin{array}{ccc}
k\left(A_t \cap B_t\right) & \geq & 0 \\
& \text{Pr oof.} & \\
+1 & \geq & p\left(A_t\right) \\
p\left(B_t\right) & \geq & p\left(A_t\right) \times p\left(B_t\right) \\
\text{If } p\left(A_t \leftarrow B_t\right) \equiv p\left(A_t \cap B_t\right) + \left(1 - p\left(B_t\right)\right) \equiv 1 \quad \text{then} & p\left(A_t \cap B_t\right) = p\left(B_t\right) \\
p\left(A_t \cap B_t\right) & \geq & p\left(A_t\right) \times p\left(B_t\right) \\
p\left(A_t \cap B_t\right) - p\left(A_t\right) \times p\left(B_t\right) & \geq & 0 \\
\dfrac{p\left(A_t \cap B_t\right) - p\left(A_t\right) \times p\left(B_t\right)}{\sqrt[2]{p\left(A_t\right) \times \left(1 - p\left(A_t\right)\right) \times p\left(B_t\right) \times \left(1 - p\left(B_t\right)\right)}} & \geq & \dfrac{0}{\sqrt[2]{\ldots}} \\
k\left(A_t \cap B_t\right) & \geq & 0 \\
& \text{Q.e.d.} &
\end{array}
\tag{28}
$$

In many problems, data gained from some observations provide an opportunity to increase the degree of confidence, when a decision is made to *either* accept the null hypotheses *or* accept the alternative hypothesis. Clearly, the null hypotheses and the alternative hypotheses are mutually exclusive thus that either the null hypothesis is false and the alternative hypothesis is true or the null hypothesis is true and the alternative hypothesis is false.

Exclusion (A_t Excludes B_t and Vice Versa Relationship)

The mathematical formula of the *exclusion* relationship (A_t excludes [2, 19, 20, 28-32] B_t and vice versa) of a population was defined as

$$
\begin{aligned}
p\left(A_t \mid B_t\right) &\equiv \frac{b_t + c_t + d_t}{N_t} \\
&\equiv 1 - p\left(a_t\right) \\
&\equiv p\left(b_t\right) + p\left(c_t\right) + p\left(d_t\right) \\
&\equiv p\left(c_t\right) + \left(1 - p\left(B_t\right)\right) \\
&\equiv p\left(b_t\right) + \left(1 - p\left(A_t\right)\right) \\
&\equiv +1
\end{aligned}
\tag{29}
$$

and used to prove the hypothesis: A_t *excludes* B_t and vice versa.

The X² Test of Goodness of Fit of the Exclusion Relationship

The chi square value with degree of freedom 2-1=1 of the exclusion [2, 19, 20, 28-32] relationship with a *continuity correction* can be calculated as

$$
\chi^2\left(A_t \mid B_t\right) = \frac{\left(-\left(a_t\right) - 0,5\right)^2}{A_t} + \frac{\left(-\left(a_t\right) - 0,5\right)^2}{B_t}
\tag{30}
$$

The chi square Goodness of Fit Test of the exclusion relationship examines how well observed data compare with the expected theoretical distribution of an exclusion relationship.

Self-contradictory data V

Let $p(A_t)$ denote the probability of the condition (i.e. risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t$ and $B_t)$ denote the joint probability that A_t and B_t will occur/has occurred. Under conditions were the relationship between A_t and B_t is determined by *a necessary condition* $p(A_t \leftarrow B_t) = p(A_t$ and $B_t) + (1 - p(B_t)) = 1$ it is equally $p(A_t$ and $B_t) = p(B_t)$ and in general

$$
\begin{aligned}
k(A_t \cap B_t) &< 0 \\
&\text{Proof.} \\
p(A_t \mid B_t) = 1 - p(a_t) &= 1 \\
p(a_t) = p(A_t \cap B_t) &= 1 - 1 = 0 \\
k(A_t, B_t) &= \frac{p(A_t \cap B_t) - p(A_t) \times p(B_t)}{\sqrt[2]{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}} \\
&\text{Under conditions were} \quad p(A_t \cap B_t) = p(a_t) = 0 \text{ it is} \\
k(A_t, B_t) &= \frac{0 - p(A_t) \times p(B_t)}{\sqrt[2]{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}} \\
&\text{or} \\
k(A_t, B_t) &< 0 \\
&\text{Q.e.d.}
\end{aligned}
\tag{31}
$$

In other words, data which provide significant evidence that A_t excludes B_t and vice versa should equally demand that the causal relationship should be $k(A_t, B_t) < 0$, otherwise the data should be treated as self contradictory.

The Mathematical Formula of the Causal Relationship k

The mathematical formula of the causal [2, 19, 20, 28-32] relationship k is defined *at every single event, at every single Bernoulli trial t,* as

$$
k(A_t, B_t) \equiv \frac{\left(p(A_t \cap B_t) - \left(p(A_t) \times p(B_t) \right) \right)}{\sqrt[2]{\left(p(A_t) \times p(\underline{A}_t) \right) \times \left(p(B_t) \times p(\underline{B}_t) \right)}}
\tag{32}
$$

where A_t denotes the cause and B_t denotes the effect. The chi-square distribution can be applied to determine the significance of causal relationship k.

Pearson's concept of correlation is not identical with causation. Causation as such is not identical with correlation. This has been prove many times and is widely discussed in many publications.

The 95% Confidence Interval of the Causal Relationship k

A confidence interval (CI) of the causal relationship k calculated from the statistics of the observed data can help to estimate the true value of an unknown population parameter with a certain probability. Under some conditions, the 95% interval for the causal relationship k is derived as

$$
\left\{ k(A_t, B_t) - \sqrt[2]{\frac{5}{n}}, \, k(A_t, B_t) + \sqrt[2]{\frac{5}{n}} \right\}
\tag{33}
$$

Hypergeometric distribution

The hypergeometric distribution with its own history [39, 40, 41] is defined by the parameters population size, event count in population, sample size and can be used to calculate the exact probability of an event even for small samples which are drawn from relatively small populations, without replacement.

The hypergeometric distribution differs from the binomial distribution. In contrast to the hypergeometric distribution, the probability of a binomially distributed random variable is the same from trial to trial. While the chi square distribution is of limited value for samples drawn from relatively small populations, the hypergeometric distribution can be used to calculate the exact probabilities for samples drawn from relatively small populations and without replication and for large populations too. The probability of having randomly exactly a_t (Table 1) successes in N_t hypergeometric trials or the significance of the causal relationship k can be tested under conditions of sampling without replacement by the hypergeometric distribution too.

The probability of having exactly a_t successes by chance in N_t hypergeometric experimental trials is given by

$$
p(a_t) = \frac{\binom{A_t}{a_t} \times \binom{N_t - A_t}{B_t - a_t}}{\binom{N_t}{B_t}}
\tag{34}
$$

Odds Ratio

The odds [42, 43, 44, 45] ratio (OR) is a very commonly used measure of association for 2×2 contingency tables and given by

$$
OR(A_t, B_t) \equiv \frac{a_t / b_t}{c_t / d_t} = \frac{a_t \times d_t}{c_t \times b_t}
\tag{35}
$$

In addition, researchers are regularly relying on Odds ratio to gain some new knowledge. Still, we need to address some different aspect of Odds ratio itself to find out the straightforward contradictions and the deep theoretical inconsistency which is associated with Odds ratio. It turns out that we are ill-advised if we believe blindly, uncritically in Odds ratio.

Case $c_t = 0$.

Under conditions were $c_t = 0$, there is *a conditio sine qua non relationship* between A_t and B_t while the Odds ratio collapses. To date, it is not generally accepted to divide by zero. The Odds ratio cannot speak about one of the natural, profound and far reaching relationships (i.e conditio sine qua non) but must pass over in silence on this relationship. Pagano & Gauvreau [46] are quietly returning through the back door to circumvent this fundamental problem of Odds ratio by adding 0.5 to the cells [46] a_t, b_t, c_t, d_t.

This simple way to circumvent the inconsistency and spectacular methodological incompleteness of Odds ratio is fundamentally misleading. To date, a substantial amount of research is analyzed by the Odds ratio. The more serious difficulty of this point of view is that it appears to be impossible to rely on Odds ratio in principle.

Case $b_t = 0$.

Furthermore, under conditions were $b_t=0$, *a conditio per quam relationship* between A_t and B_t is given while the Odds ratio collapses again. For this reason, the Odds ratio is overshadowed by a deep theoretical inconsistency and appears not to be grounded on a seemingly sound piece of reasoning. More likely, the Odds ratio (OR) is nothing more but *Yule's coefficient of association* [47] Q re-written [48] in a non-normalized form and given by

$$Q(A_t, B_t) \equiv \frac{OR(A_t, B_t) - 1}{OR(A_t, B_t) + 1}$$

$$Q(A_t, B_t) = \frac{\dfrac{(a_t \times d_t)}{(b_t \times c_t)} - 1}{\dfrac{(a_t \times d_t)}{(b_t \times c_t)} + 1} \qquad (36)$$

$$Q(A_t, B_t) = \frac{\dfrac{(a_t \times d_t) - (b_t \times c_t)}{(b_t \times c_t)}}{\dfrac{(a_t \times d_t) + (b_t \times c_t)}{(b_t \times c_t)}}$$

$$Q(A_t, B_t) = \frac{(a_t \times d_t) - (b_t \times c_t)}{(a_t \times d_t) - (b_t \times c_t)}$$

Under conditions where Yule's coefficient [47] of association Q=0, there is no association. Although severely and justifiably criticized especially by Karl Pearson (1857–1925), the long-time and rarely challenged leader of statistical science and Heron [49], Odds ratio is still regularly referred to. The standard error and 95% confidence interval of the Odds ratio (OR) can be calculated according to Altman [50]. Given the severely limited character of odds ratio, the standard error of the log Odds ratio is calculated as

$$SE\left(\ln\left(OR(A_t, B_t)\right)\right) \equiv \sqrt{\frac{1}{a_t} + \frac{1}{b_t} + \frac{1}{c_t} + \frac{1}{d_t}} \qquad (37)$$

where ***ln*** denotes the *logarithmus naturalis*. The 95% confidence interval of the odds ratio is given by

$$95\% \ CI = \exp\left(\ln\left(OR(A_t, B_t)\right) - \left(1.96 \times SE\left(\ln\left(OR(A_t, B_t)\right)\right)\right)\right)$$
$$\text{to} \qquad (38)$$
$$\exp\left(\ln\left(OR(A_t, B_t)\right) + \left(1.96 \times SE\left(\ln\left(OR(A_t, B_t)\right)\right)\right)\right)$$

The unknown population proportion π_{upper}

Tests of hypotheses concerning the sampling distribution of the sample proportion **p** (i. e. conditio sine qua non p(SINE), conditio per quam p(IMP) et cetera) can be performed using the normal approximation. The calculation of the rejection region based on the sample proportion to construct a confidence interval for an unknown population proportion π_{upper} can be performed under conditions of *sampling without replacement* [51] by the formula

$$\pi_{critical\ upper} = \left(p - \frac{1}{2 \times n}\right) - \left(Z \times {}_2\sqrt{\left(\frac{p \times (1-p)}{n}\right) \times \left(\frac{N-n}{N-1}\right)}\right) \qquad (39)$$

while the term $((N-n)/(N-1))$ denotes *the finite population correction* [52].

The Chi Square Distribution

The following critical values[51] of the chi square distribution[53] as visualized by Table 11 are used in this publication.

Table 11. The critical values of the chi square distribution (degrees of freedom: 1)

	p-Value	One sided X²	Two sided X²
	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
The chi square	0.0010000000	9.549535706	10.82756617
distribution	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

The rule of three

The Chi-square goodness of fit test[53] used to test whether a sample distribution is identical with a *theoretical* distribution yields only an approximate p-value and works when the dataset analysed is large enough (n ~ 30 and more). An approximate and conservative (one sided) confidence interval as discussed by Rumke [54], Louis [55], Hanley [56] et al. and Jovanovic [57] and known as the rule of three can be used if the Chi-square goodness of fit test (with a *continuity correction*) cannot be applied. The rule of three is known to be derived as

$$p_{critical} = \left(1 - \frac{3}{N}\right) \qquad (40)$$

And is one way to calculate the probability of events which occur with a probability near 1. Another and a very simple path to calculate the probability of an event can be performed by the following method.

Theorem.

In general, the probability can be calculated approximately as

$$p \equiv \sqrt[N]{e^{-\underline{\lambda}}} \qquad (41)$$

Proof.

In general, it is **+1=+1** (*lex identitatis* [58, 59]). Multiplying by p, we obtain 1×p=1×p or p=p. Let N denote something like the number of trials or the sample size et cetera. Performing the power operation it is

$$p^N \equiv p^N \qquad (42)$$

According to mathematical requirements it is $p = 1 - \underline{p}$ and $\lambda \equiv N \times p$ and $\underline{\lambda} \equiv N \times \underline{p} \equiv N \times (1-p)$. The equation changes to

$$p^N \equiv \left(1 - \underline{p}\right)^N \qquad (43)$$

or to

$$p^N \equiv \left(1 - \left(\left(N \times \underline{p}\right)/N\right)\right)^N \qquad (44)$$

Due to our definition $\underline{\lambda} \equiv N \times \underline{p} \equiv N \times (1-p)$ we obtain finally

$$p^N \equiv \left(1 - \left(\underline{\lambda}/N\right)\right)^N \qquad (45)$$

It is important to note, that not every population goes to infinity. Still, taking the limit as the number of trials as N goes to *positive infinity (N → +∞)*, we obtain

$$\left(1 - \left(\underline{\lambda}/N\right)\right)^N \equiv \lim_{N \to +\infty} es\left(1 - \left(\underline{\lambda}/N\right)\right)^N \qquad (46)$$

or according to elementary [60] calculus

$$\lim_{N \to +\infty} es\left(1 - \left(\underline{\lambda}/N\right)\right)^N = e^{-\underline{\lambda}} \qquad (47)$$

Our equation changes to

$$p^N \equiv e^{-\underline{\lambda}} \qquad (48)$$

In general, the probability of an event can be calculated approximately as

$$p \equiv \sqrt[N]{e^{-\underline{\lambda}}} \qquad (49)$$

Q. e. d.

Example.

Suppose a team of Astronomers has investigated N=10 galaxies and found one black hole inside each galaxy, consequently it is $\underline{\lambda} = 0$. The probability that every possible galaxy has a black hole can be calculated approximately as

$$p \equiv \sqrt[10]{e^{-0}} = 1 \qquad (50)$$

According to the rule of three, the probability that every galaxy do possess a black could be calculated as

$$p_{critical} = \left(1 - \frac{3}{N}\right) = 1 - \frac{3}{10} = 0.7 \qquad (51)$$

Example.

Suppose an investigation is performed with N=100 cases and controls. The probability of an event within the population is assumed to be p=0.95. What is the critical p value?

$$p_{Critical} \equiv \sqrt[100]{e^{-\left(100 \times (1-0.05)\right)}} = p \equiv \sqrt[100]{e^{-5}} = 0.951229425 \qquad (52)$$

The probability found within the study should not be lower than 0.951229425. Otherwise the data do not support the hypotheses that p = 0.95 or even more.

Are such observations appropriate at all to justify some predictions about *observations we have not yet made* and may be even something like general claims that go beyond the observed? The question is of course are we allowed to infer a hypotheses about the general situation based on the observation of such a limited sample? In other words, how can we be sure about *the unknown 'land'*, the unobserved, on what ground and to what extent? One may object that any analysis of the notions of cause and effect is confronted by the unobserved too. On this view, how many galaxies are given within the universe? We do not know for sure. How many of all galaxies do possess a black hole? We do not know for sure, either. Still, even such a small sample of observations justifies the conclusion and provides some degree of support but of course not the ultimate evidence for the truth that about 100 % of all galaxies possess a black hole. It is not the main goal of this paper to solve *the famous philosophical problem of induction* and inductive inference as introduces by *David Hume* [61] in 1739 in his book A Treatise of Human Nature (Book 1, part iii, section 6). However, in order to approach to the solution of this problem it is necessary to point out that under certain circumstances logic, mathematics and statistics is able to provide us with methods of direct inference even about the unknown.

Anti distribution

Suppose that S defines the *sample space* of an experiment completely. Let a real-valued function (*a random variable*) X which is defined on the sample space S assign a real number X(s) to each possible outcome s ∈ S in a particular experiment. The distribution of the random variable of X is defined as the collection of all probabilities p(X ∈ A) for all subsets A of the real numbers. A *discrete random variable* is defined as a random variable X which can take only a finite number of k different values $x_1, \ldots, x_k$ or at most, an infinite sequence of x_1, $x_2, \ldots$ The distribution of a discrete random variable X is defined as the *probability mass function* and abbreviated as p(x) or p. m. f.(x) of X, namely p(x) = p. m. f.(x) = p(X = x) for all x in the set of possible values.

A random variable X which can take every value in an interval is called a *continuous random variable*. A continuous distribution is defined by its own *probability density function* (p.d.f.) of the distribution of X for every interval (a,b) as

$$p\left(a < X \le b\right) \equiv \int_{a}^{b} f\left(x\right)dx \tag{53}$$

Continuous random variables satisfy the condition p(X=x)=0. In practical problems it may sometimes be necessary to consider a distribution as a mixture of a continuous distribution and a discrete distribution. The cumulative distribution function abbreviated as P(x) or as F(x) or as d.f.(x) or c.d.f.(x) of every random variable X, regardless of whether the distribution of X is continuous, discrete or mixed, for each real number x is defined as

$$P\left(x\right) \equiv F\left(x\right) = p\left(X \le x\right) = p\left(X = x\right) + p\left(X < x\right) \tag{54}$$

for -∞ < x < +∞.

Theorem I.
For every value x,

$$p\left(X > x\right) = 1 - P\left(x\right) \tag{55}$$

Proof.
Since

$$p\left(X > x\right) = 1 - P\left(x\right) \equiv 1 - p\left(X = x\right) - p\left(X < x\right) \tag{56}$$

the theorem follows directly from the definition of the cumulative distribution function.
Quod erat demonstrandum.

Theorem II.
For every value x, the *anti distribution of x* is determined as

$$p\left(X \ne x\right) = p\left(X < x\right) + p\left(X > x\right) \tag{57}$$

Proof.
The cumulative distribution function for each real number x regardless of whether the distribution of X is continuous, discrete or mixed, for each real number x is defined as

$$p\left(X < x\right) + p\left(X = x\right) = p\left(X \le x\right) = P\left(x\right) \equiv F\left(x\right) \tag{58}$$

for -∞ < x < +∞. Adding p(X > x), we obtain

$$p\left(X < x\right) + p\left(X = x\right) + p\left(X > x\right) = P\left(x\right) + p\left(X > x\right) \tag{59}$$

The right term equals 1 and the equation simplifies as

$$p\left(X < x\right) + p\left(X = x\right) + p\left(X > x\right) = 1 \tag{60}$$

Rearranging we obtain

$$p\left(X < x\right) + p\left(X > x\right) = 1 - p\left(X = x\right) \tag{61}$$

We define $p(\underline{x}) \equiv p(X<x) + p(X>x)$ as the distribution for every value of *anti x* denoted as $\underline{x}$ or as the *anti distribution of x* and do obtain

$$p\left(\underline{x}\right) = p\left(X = \underline{x}\right) = p\left(X < x\right) + p\left(X > x\right) = 1 - p\left(X = x\right) \tag{62}$$
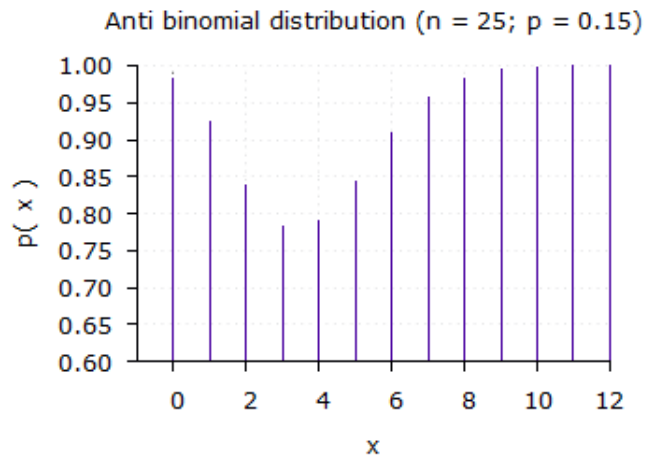
Quod erat demonstrandum.

Example.
The anti binomial distribution can be derived as

$$p\left(X \ne x\right) \equiv 1 - \left(\left(\frac{N!}{x! \times \left(N-x\right)!}\right) \times p^x \times \left(1-p\right)^{N-x}\right) \tag{63}$$

For n=25 and p=015 we obtain the following figure.

Anti binomial distribution (n = 25; p = 0.15)

The probability density of the *anti normal (or Gaussian or Gauss or Laplace–Gauss) distribution* follows as

$$p\left(X \ne x\right) \equiv 1 - \left(\left(\frac{1}{\sqrt[2]{2 \times \pi \times \sigma\left(x\right)^2}}\right) \times \left(e^{-\left(x-\mu\right)/\left(2\times\sigma\left(x\right)^2\right)}\right)\right) \tag{64}$$

where μ denotes the mean or expectation of the distribution and σ(x)² is the variance.

The normal distribution is useful because of the unofficial sovereign and the foundation of any statistics and probability theory, the *central limit theorem*. Any average of enough independent copies of a random variable will result nearly in a normal (Gaussian) distribution.

Anti Normal PDF with μ=1.0 and sigma =1.5
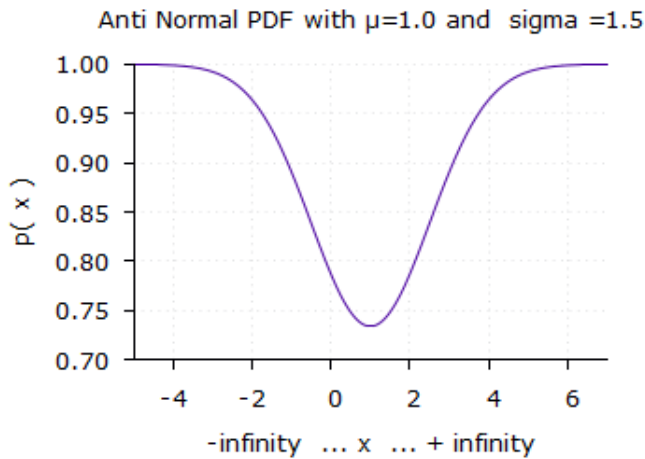
The French-born mathematician Abraham de Moivre (1667 – 1754) while working on "Bernoulli's Law of Large Numbers", the main theorem of Jakob Bernoulli's[62] *Ars conjectandi*, published 1733 the first historical pre-work on the central limit theorem. After the monumental work "Théorie analytique des probabilités" of the famous French mathematician Pierre-Simon Laplace[63] published in 1812 and the very precisely prove provided 1901 by the Russian mathematician Aleksandr Lyapunov[64], the Hungarian born mathematician George Pólya[65] coined in 1920 the German term "zentraler Grenzwertsatz" or *central limit theorem*. In a similar way, anti distributions of other distributions can be derived as demonstrated before.

The Anti Poisson distribution

The Poisson distribution, given previously by Abraham de Moivre [66], is ascribed to Siméon Denis Poisson (1781–1840), a French mathematician, physicist, and engineer who published the same distribution 1837 in his work "*Recherches sur la probabilité des jugements en matière criminelle et en matière civile*" [67]. Ladislaus Bortkiewicz [68] provided in 1898 one of the first practical applications of Poisson's distribution while investigating the number of soldiers in the Prussian army killed accidentally by horse kicks. A discrete random variable X is said to have a Poisson distribution with parameter $\lambda > 0$, if, for x = 0, 1, 2, ..., the probability mass function of X is given by

$$p(X = x) \equiv \frac{\lambda^x}{x!} \times e^{-\lambda} \qquad (65)$$

were x is the number of times an event occurs in an interval and x can take values 0, 1, 2, …., e is Euler's number (the number 2.71828..., the base of the natural logarithms) and x! is the factorial of x or x!=x*(x-1)*(x-2)*...2*1. Many times, the Poisson distribution is applied to experimental conditions or situations with a large number of trials N while the occurrence of each event is very rare. The *anti Poisson distribution* is given by

$$p(X \neq x) \equiv 1 - \left(\frac{\lambda^x}{x!} \times e^{-\lambda} \right) \qquad (66)$$

and is useful too, for events which occur very often and with a probability near 1 or nearly for sure. In other words, if we know that the probability of x very rare p(X=x), we know equally that the probability p(X≠x) of very often events/non-events is p(X≠x) = 1- p(X=x).

Properties.
In general, it is

$$p(\underline{x}) + p(x) = p(X = \underline{x}) + p(X = x) = 1 \qquad (67)$$

The distribution of likely events
Mathematically, the probability that an event will occur is expressed as a number between +0 and +1 and can be defined in many different ways. If $p(X_t)$ equals zero, event X_t will almost definitely not occur and a probability near 0 indicates an unlikely event. A probability $p(X_t)$ near 1 indicates a likely event. Under some circumstances, a binomial distribution can be approximated by the normal distribution. Another extreme of the binomial distribution is the case when **p→ 0** while N goes to infinity. In this case, a binomial distribution can be approximated by the normal distribution. Another extreme of the binomial distribution is the case when **p→ 1** while N goes to infinity. In this case, a binomial distribution can be approximated by the anti Poisson distribution derived from the Poisson distribution and given by

$$p(X \neq x) \equiv 1 - \left(\frac{\lambda^x}{x!} \times e^{-\lambda} \right) \qquad (68)$$

Example.
Suppose that, on the average, 1999 houses in 2000 in a certain district *are free of* fire or not burning during a year. If N = 4000 houses are in that district, what is the probability that exactly 3995 houses will stay free of fire or will not have a fire during the year. We focus on the fact that 1999 houses from 2000 houses will not burn, which is not a Poisson distributed random variable. In turn, it is insightful to point to the fact that 1 out of 2000 houses will have a fire which is a very rare event and Poisson distributed. As is so often the case, it is a matter of personal taste whether a glass is treated as half full or whether the same glass is treated as half empty. The Anti Poisson distribution can be used to calculate the probability. Since 1999 houses have no fire, we know that 1 house in 2000 has fire or it is p = (1/2000) or λ = N×(1-p)=N×p =4000×(1/2000)= 2. The probability that exactly 3995 houses will *have no fire* during a year means that exactly 5 houses or 4000-3995 = 5 houses will have a fire. In other words, we obtain

$$p(X \neq x) \equiv 1 - \left(\frac{\lambda^x}{x!} \times e^{-\lambda} \right) = 1 - \left(\frac{2^5}{5!} \times e^{-2} \right) = 1 - (0.036089408863097) \quad (69)$$

or p = 0.963910591. The probability that exactly 3995 houses in 4000 will *have no fire* during a year is with p= 0.963910591 extremely near 1 and equivalent with the rare event 1 minus the probability that exactly 5 houses in 4000 houses will have a fire (p=0.036089408863097).
Ultimately, under conditions were an event occurs its own complementary event does not occur or it is p + p =1, the two terms are more or less interchangeable and it remains a matter of personal taste what is understood as p and what is taken as p.

In the case of a conditio sine qua non, conditio per quam, exclusion relationship et cetera we expect that the probability p of such a relationship in N Bernoulli trials is extremely near 1 or it is **p~1**. In the same context, the probability, denoted by p, that the same relationship (event) *will not occur* in N Bernoulli is extremely small and will be very near to zero. In other words, since p + p = 1 it is p = 1-p and the expectation value of this very rare event is $\lambda = N \times p = N \times (1-p)$. In an experiment we observed X=N×(p) events. The rate of very rare events which should not have occurred is x=N-X or in detail **x**= N-X= N-(N×(p)) = N×(1-p) = λ.

Anti Poisson distribution with μ = 2.0
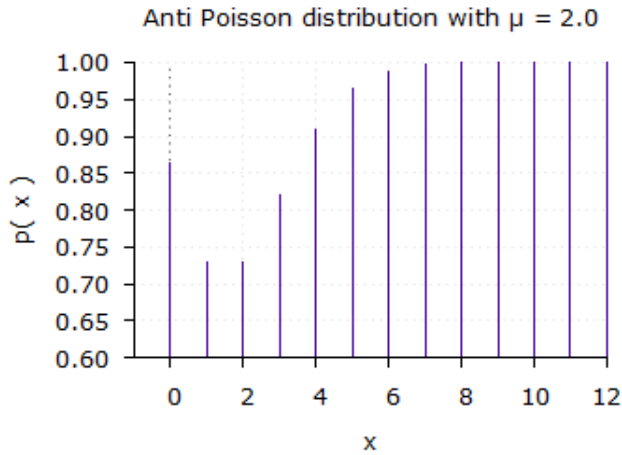
Consequently, the anti Poisson distribution above *under experimental conditions were the expectation value λ is equal to the number of rare events* x, i. e. where **(λ=x) > 0** simplifies as

$$p\left(X \neq x\right) \equiv 1 - \left(\frac{x^x}{x!} \times e^{-x}\right) \tag{70}$$

where x indicates the expected (or average) number of occurrences of a very rare event. This very simplified form of the Poisson distribution can be called *the distribution of the likely*. Under conditions where λ=N the anti Poisson distribution simplifies as

$$p\left(X \neq x\right) \equiv 1 - \left(\frac{N^x}{x!} \times e^{-N}\right) \tag{71}$$

and were **(λ=x=N)** as

$$p\left(X \neq x\right) \equiv 1 - \left(\frac{N^N}{N!} \times e^{-N}\right) \tag{72}$$

The Poisson distribution can be regarded as a limiting form of the binomial distribution and is one of the most widely used distributions in science and industry. The normal distribution is more or less a limiting form of the binomial distribution when p is very near 0.5 and n became very large.
Still, when p is near 0 or 1, is not appropriate to use the normal distribution as an approximation of the binomial distribution. The Chi square distribution is grounded on the normal distribution. Thus far, the Chi-square of goodness of fit test could provide an inappropriate picture when N is very large and p near 1.

Under these experimental conditions, another strategy should be adopted. Suppose an event which occurs with a probability p very near 1 or **p~1**. Under these conditions, *the rate or the average (or expected) number non-occurrence of this event* is very small and very near zero or should be zero. Let us suppose that in an investigation some very rare non-events occurred which should not have occurred. Using the Poisson distribution it is possible to calculate the probability, how likely is to obtain some very rare non-events during a certain period of observation.

The exact probability of a single event
Mathematically, the probability that an event will occur is expressed as a number between +0 and +1 and can be defined in many different ways. For our purposes, the probability of event, which has a value or quantity X_t is represented by $p(X_t)$ and we define the probability that a single event has the value X_t at the Bernoulli trial t by the relationship

$$p\left(X_t\right) \equiv \left(\frac{E\left(X_t\right)}{X_t}\right) \tag{73}$$

where $E(X_t)$ denotes the expectation value of a single event. Such a definition of probability assumes that every single event is associated with its own expectation value even under circumstances where $p(X_t) = 1$. Under these conditions it is equally $E(X_t) = X_t$. In other words, we define

$$E\left(X_t\right) \equiv X_t \times p\left(X_t\right) \tag{74}$$

or

$$E\left(X_t\right)^2 \equiv \left(X_t \times p\left(X_t\right)\right)^2 \equiv X_t^2 \times p\left(X_t\right)^2 \tag{75}$$

or

$$E\left(X_t^2\right) \equiv X_t^2 \times p\left(X_t\right) \equiv X_t \times \left(X_t \times p\left(X_t\right)\right) \equiv X_t \times E\left(X_t\right) \tag{76}$$

while the definitions above are independent of the distribution of X_t. The variance $\sigma(X_t)^2$ of a single event X_t is independent of the distribution of X_t and defined as

$$\sigma\left(X_t\right)^2 \equiv E\left(X_t^2\right) - E\left(X_t\right)^2 \equiv X_t^2 \times p\left(X_t\right) \times \left(1 - p\left(X_t\right)\right) \tag{77}$$

Theorem III.
The expectation value E(x) for each real number x is defined as

$$E\left(x\right) = \left(x\right) \times p\left(x\right) \tag{78}$$

or as

$$E\left(x\right)^2 = \left(x\right)^2 \times p\left(x\right)^2 \tag{79}$$

or

$$E\left(x^2\right) = \left(x^2\right) \times p\left(x\right) \tag{80}$$

In general, the variance $\sigma(x)^2$ for each real number x is defined as $\sigma(x)^2 \equiv E(x^2) - E(x)^2$.

Claim.
The variance for each real number x is defined as

$$\sigma\left(x\right)^2 \equiv E\left(x^2\right) - E\left(x\right)^2 \tag{81}$$

Proof.

It is as $\sigma(x)^2 \equiv E(x^2) - E(x)^2$ or

$$
\begin{aligned}
\sigma(x)^2 &\equiv E(x^2) - E(x)^2 \\
&\equiv (x^2 \times p(x)) - (x^2 \times p(x)^2) \\
&\equiv (x^2) \times (p(x) - p(x)^2) \\
&\equiv (x^2) \times p(x) \times (1 - p(x)) \\
&\equiv (x^2) \times p(x) \times p(\underline{x})
\end{aligned} \tag{82}
$$

Q.e.d.

Theorem IV.

The variance for each real number x can be normalized as

$$
\frac{E(x)^2}{E(x^2)} + \frac{\sigma(x)^2}{E(x^2)} \equiv +1 \tag{83}
$$

Proof.

In general, it is

$$
\begin{aligned}
\sigma(x)^2 &\equiv E(x^2) - E(x)^2 \\
&\text{Re arraing, it is} \\
E(x)^2 + \sigma(x)^2 &\equiv E(x^2) \\
&\text{Dividing, we obtian} \\
\frac{E(x)^2}{E(x^2)} + \frac{\sigma(x)^2}{E(x^2)} &\equiv \frac{E(x^2)}{E(x^2)} \\
\frac{E(x)^2}{E(x^2)} + \frac{\sigma(x)^2}{E(x^2)} &\equiv 1
\end{aligned} \tag{84}
$$

Q.e.d.

Chebyshev's inequality

Let X be a random variable with finite expected value E(x) and finite non-zero variance $\sigma(x)^2$. Then for any real number x > 0, the probability p(x) for each real number x calculated according to the Chebyshev's inequality [69] follows as

$$
p\left(\left|x - E(x)\right| \geq \sqrt[2]{E(x^2)}\right) \leq 1 - \frac{\sigma(x)^2}{E(x^2)} \tag{85}
$$

The Chebyshev's inequality (also called the Bienaymé-Chebyshev inequality) guarantees only an approximate value. In contrast to Chebyshev's inequality, it is possible to calculate the exact probability p(x) for each real number x. Suppose that a random variable X has a certain distribution and can have different single values X_i, i= 1, … each with its own probability $p(X_i)$. Let E(X) denote the expectation value of X.

The number E(X) is also called the mean of X or the expected value of X. The terms mean, expected value or expectation value are used interchangeably. We get

$$
E(X) \equiv p(\overline{X}) \times \overline{X} \equiv E(X_1) + E(X_2) + \ldots \equiv (X_1 \times p(X_1)) + (X_2 \times p(X_2)) + \ldots \tag{86}
$$

or

$$
E(X^2) \equiv p(\overline{X}) \times \overline{X}^2 \equiv (X_1^2 \times p(X_1)) + (X_2^2 \times p(X_2)) + \ldots \tag{87}
$$

The variance of X is defined as

$$
\sigma(X)^2 \equiv E(X^2) - E(X)^2 \tag{88}
$$

Rearranging equation before, we obtain

$$
E(X^2) \equiv E(X)^2 + \sigma(X)^2 \tag{89}
$$

or normalizing the relationship it is

$$
1 \equiv \frac{E(X^2)}{E(X^2)} \equiv \frac{E(X)^2}{E(X^2)} + \frac{\sigma(X)^2}{E(X^2)} \equiv \frac{\overline{X}^2 \times p(\overline{X})^2}{\overline{X}^2 \times p(\overline{X})} + \frac{\sigma(X)^2}{E(X^2)} \tag{90}
$$

Simplifying equation it is

$$
1 \equiv p(\overline{X}) + \frac{\sigma(\overline{X})^2}{E(\overline{X}^2)} \tag{91}
$$

or

$$
p(\overline{X}) = 1 - \frac{\sigma(X)^2}{E(X^2)} \tag{92}
$$

Theorem V.

The exact probability for each real number x can be calculated as

$$
p(x) \equiv 1 - \frac{\sigma(x)^2}{E(x^2)} \tag{93}
$$

Proof.

According to the theorem before, it is

$$
\begin{aligned}
\frac{E(x)^2}{E(x^2)} + \frac{\sigma(x)^2}{E(x^2)} &\equiv 1 \\
&\text{Re arraing, it is} \\
\frac{E(x)^2}{E(x^2)} &\equiv 1 - \frac{\sigma(x)^2}{E(x^2)} \\
&\text{Simplifying, it is} \\
\frac{(x)^2 \times p(x)^2}{(x^2) \times p(x)} &\equiv 1 - \frac{\sigma(x)^2}{E(x^2)} \\
&\text{or} \\
p(x) &\equiv 1 - \frac{\sigma(x)^2}{E(x^2)}
\end{aligned} \tag{94}
$$

Q.e.d.

The p-value

Historically, the evidence of the first use of the p-value in statistics dates back as far the late 17th century. The question of the p-value was addressed especially by John Arbuthnot[70] in 1710. Arbuthnot (1667 – 1735) examined birth records in London for each of the 82 years from 1629 to 1710 and compared the human sex ratio at birth to the null hypothesis of equal probability. About 100 years later, Pierre-Simon Laplace starts the Chapter V of his book "Théorie analytique des probabilités" [63] with the computation of a p-value[71]. In Chapter VI, of his book Laplace provided his famous study on the statistics of almost half a million births and demonstrated an excess of boys compared to girls. Laplace concluded by calculation of a p-value that the excess was a real effect. Formally, it was Karl Pearson who introduced the p-value[53] as capital P. In point of fact, Fisher himself proposed in his influential book "Statistical Methods for Research Workers" [72] the level p-value = 0.05 as a limit for statistical significance. Many times, studies or experiments are investigating whether there is a difference between different experimental set-ups that the researchers are testing. In particular, a sample is drawn from a population, studied and the results are extrapolated to the population from where the sample was drawn. A condition or factor being studied can produces an effect or can makes a difference but must not. In every experiment, the observed difference in the sample data must not reflect a true difference in the populations or in objective reality. To a certain extent, it is possible that a true null hypothesis is incorrectly rejected (type I error (or error of the first kind)). In other words, we falsely infer that something (i.e. *Ho; there is no difference*) is present when it actually it is not present. The probability of rejecting the null hypothesis given that the null hypothesis is true is called type I error rate or significance level, denoted by the Greek letter α (alpha). By convention, statisticians and journals suggest a significance level of α=5% (Type I error) with the consequence (or potential consequence) that the difference observed is not due to chance but equally we have to accept to be fooled by randomness or subjective or objective random errors 1 time out of 20. In particular, the probability of incorrectly rejecting the null hypothesis or p (incorrectly rejecting the null hypothesis) = 5% is defined as being acceptable. A false null hypothesis should be rejected. Theoretically, it is possible fail to reject a false null hypothesis (type II error or error of the second kind, ß error). A false null hypothesis is rejected with the probability 1- ß, denoted by the Greek letter ß (beta). In an investigation, several statements based on the result of hypothesis tests are presented along with the associated p values. A hypothesis test should provide some help to decide whether the results of a study, based on a small sample, provide enough evidence against a claimed null hypothesis (denoted by H_0), with the consequence that it is reasonable to believe that in a larger target population, H_0 is false too. The strength of our evidence against H_0 is measured by the p-value. Still, there are some misunderstandings associated with the interpretation of a p value. In particular, a very small p value does provide strong evidence that H_0 is not true. In contrast to this, even as large p value does not provide real evidence that H_0 is true. In general and depending on the point of view, the p-value is defined as the probability of obtaining a result equal to or more extreme than an actually observed result under the condition that a null hypothesis is valid. Thus, *the p-value for a right tail event* is given by

$$p - \text{value} \equiv p(X > x \mid H_0) \qquad (95)$$

In general, it is

$$1 = 1 - p(X \le x \mid H_0) + p(X \le x \mid H_0) \qquad (96)$$

or

$$1 = p(X > x \mid H_0) + p(X \le x \mid H_0) \qquad (97)$$

It follows that

$$p(X > x \mid H_0) \equiv 1 - p(X \le x \mid H_0) \qquad (98)$$

Under the condition of the validity of the null-hypothesis, the p value can be calculated as

$$p(X \le x \mid H_0) = 1 - p(X > x \mid H_0) \qquad (99)$$

In other words, the calculation of the probability $p(X>x \mid H_0)$ enable us to calculate the p-value.

p value for a Binomial Distribution (Binomial test)

Unfortunately, there is always the possibility that the results of a study may be wrong and sometimes, a differences observed during an investigation is just the result of random subjective or objective errors or random effects. A statistical test is more or less about managing such and similar risks by the tools of probability theory and not about certainty. In point of fact, a true null hypothesis (there is no difference) should be accepted. Thus far we assume that a null hypothesis (H_0) is true.

Example. conditio sine qua non.
Suppose **x = 395 as the number of times the conditio sine qua non relationship occured in n = 400 trials**. This random variable has the binomial distribution where **π is the population parameter** corresponding to the probability of success on any trial. The binomial distribution is used when there are exactly two *mutually exclusive outcomes* of a trial. The formula for the binomial probability mass function of observing exactly x successes in *n* trials, with the probability of success on a single trial denoted by **π** is

$$p(X = x) = \left(\frac{n!}{x!(n-x)!} \right) \times \left(\pi^x \right) \times \left(1 - \pi \right)^{n-x} \qquad (100)$$

The probability of exactly x=395 events out of n=400 trials is p(X = 395) =0.0000412947.
The probability of exactly *not* x=395 events out of n=400 trials is p(X<>395) = 1- p(X= 395) = 0.9999587053.
The probability of exactly, or more than x=395 events out of n=400 trials is calculated as p(X≥395) =0.0000533965.
The probability of less than x=395 events out of n=400 trials is calculated as p(X < 395) = 0.9999466035.
The probability of more than x=395 events out of n=400 trials is calculated as p(X >395) =0.0000121017.
The probability of exactly, or fewer than, x=395 events out of n=400 trials is p = 0.9999878983 or in other words

$$p(X \le 395 \mid H_0) = \sum_{t=0}^{395} \left(\frac{400!}{t!(n-t)!} \right) \times \left(0.95^t \right) \times \left(1 - 0.95 \right)^{n-t} = 0.99998789826754 \qquad (101)$$

In our example, we use the following null and alternative hypotheses:

H_0: π ≤ 0.95; i.e. No conditio sine qua non relationship
H_A: π > 0.95 i.e. conditio sine qua non relationship

Setting α = 0.05, we have the cumulative probability of x=395 out of 400 events as

$$p(X \le x \mid H_0) = p(X = 0) + p(X = 1) + ... + p(X = 395) \qquad (102)$$

or

$$p\left(X \leq x \mid H_0\right) = \sum_{t=0}^{x} \left(\frac{n!}{t!(n-t)!}\right) \times \left(\pi^t\right) \times \left(1-\pi\right)^{n-t} \qquad (103)$$

The probability of exactly, or fewer than, x=395 out of n=400 is p = 0.9999878983 or

$$p\left(X \leq 395 \mid H_0\right) = \sum_{t=0}^{395} \left(\frac{400!}{t!(n-t)!}\right) \times \left(0.95^t\right) \times \left(1-0.95\right)^{n-t} = 0.99998789826754 \qquad (104)$$

It is

$$p\left(X \leq x \mid H_0\right) + p\left(X > x \mid H_0\right) = 1 \qquad (105)$$

or

$$p\left(X > x \mid H_0\right) = 1 - p\left(X \leq x \mid H_0\right) \qquad (106)$$

The probability p(X > 395) follows as

$$p\left(X > x \mid H_0\right) = 1 - \sum_{t=0}^{x=395} \left(\frac{n!}{t!(n-t)!}\right) \times \left(\pi^t\right) \times \left(1-\pi\right)^{n-t} \qquad (107)$$

or

$$p\left(X > x \mid H_0\right) = 1 - \left(p(X=0) + p(X=1) + ... + p(X=395)\right) \qquad (108)$$

or

$$p \text{ value} = p\left(X > 395 \mid H_0\right) = 1 - p\left(X \leq 395 \mid H_0\right) = 0.00001210173246 \qquad (109)$$

If p value > α then accept H_0. If p value < α then reject H_0. We used a one-tailed test with null and alternative hypotheses and conclude with 95% confidence to reject the null hypothesis H_0 and accept the alternative hypothesis H_A since the *p value* = 0.00001210173246 < 0.05 .The data observed support the rejection of the null hypothesis, because the associated p-value is less or equal to the level of significance α. In this context, the p value is thus the smallest level of significance to which the null hypothesis can still be rejected. Under some certain circumstances n×π×(1-π) > 9. Another general rule of thumb demands that the sample size *n* is "sufficiently large" and the binomial distribution can be approximated by the normal distribution if n×π ≥ 5 *and* if n×(1-π)≥5. If these conditions are met, then the binomial distribution can be treated as approximating the normal distribution and a z-test for significance can be performed.

p value for a Poisson Distribution
A binomial distribution is a sum of *n* independent Bernoulli random variables with the probability π. For very high or very low π, a binomial distribution is a very skewed distribution. Under conditions with very low π probability and very large n, the Poisson distribution may be used as an approximation to the binomial distribution. In practice it is possible not to observe a conditio sine qua non relationship within a sample even if within a population, such a relationship is given. Events like these can be accepted only under very limited circumstances and should be extremely small with the consequence that the law of rare events or Poisson limit theorem can be used to test the significance. The Poisson distribution mathematical formula is used to find out the probability of given number of events occurred for the instances of k = 0, 1,2, …, n. In this example, λ is a positive real number or the mean or equal to the expected number of occurrences of the conditio sine qua non relationship and is calculated as

$$\lambda = \sum_{t=0}^{n} \left(\frac{x_t}{n}\right) \qquad (110)$$

The Poisson distribution is defined as

$$p\left(X = x\right) = \left(\frac{\mu^x}{x!}\right) \times e^{-(\mu)} \qquad (111)$$

where x is the number of observed rare events and λ is a positive real number or the mean or equal to the expected number of occurrences of an event (i. e. the conditio sine qua non relationship). Assuming that $\lambda(A_t \leftarrow B_t) = p(A_t \leftarrow B_t)$ is the relative frequency of a conditio sine qua non relationship within sample data in *n* trials, the relative frequency that a conditio sine qua non relationship within sample data of *n* trials *will not be observed* is $\underline{\lambda}(A_t \leftarrow B_t)$ = 1- $p(A_t \leftarrow B_t)$. In other words, we are more or less sure that we will not observe x=0. The probability of x = 0 while the relative frequency is $\underline{\lambda}(A_t \leftarrow B_t)$ = 1- $p(A_t \leftarrow B_t)$ can be calculated as

$$p\left(X = 0\right) = \left(\frac{\mu^0}{0!}\right) \times e^{-\left(1-p(A_t \leftarrow B_t)\right)} = e^{-\left(1-p(A_t \leftarrow B_t)\right)} \qquad (112)$$

where x=0 denotes that no rare events were observed. Under conditions of x = 0, the cumulative distribution function of the Poisson distribution is defined as

$$p\left(X \leq 0 \mid H_0\right) = \sum_{t=0}^{t=0} \left(\frac{\lambda^t}{t!}\right) \times \left(e^{-\lambda}\right) = p\left(X = 0\right) \qquad (113)$$

It is

$$p\left(X \leq 0 \mid H_0\right) + p\left(X > 0 \mid H_0\right) = 1 \qquad (114)$$

or

$$p\left(X > 0 \mid H_0\right) = 1 - p\left(X \leq 0 \mid H_0\right) \qquad (115)$$

The probability p(X > 0) follows as

$$p\left(X > 0 \mid H_0\right) = 1 - e^{-\underline{\lambda}} \qquad (116)$$

or

$$p\left(X > 0 \mid H_0\right) = 1 - \left(p(X = 0)\right) \qquad (117)$$

or

$$p \text{ value} = p\left(X > 0 \mid H_0\right) = 1 - p\left(X \leq 0 \mid H_0\right) = 1 - p\left(X = 0 \mid H_0\right) \qquad (118)$$

Under these conditions, the null and alternative hypotheses are as follows:
H_0 *(Null hypothesis):* p(X = 0) ≤ 1 - α (i.e. p(X = 0) ≤ 0.95)
There is no significant conditio sine qua non relationship.
H_A *(Alternative hypothesis):* p(X = 0) > 1 - α
(i.e. p(X = 0) > 0.95)
There is a significant conditio sine qua non relationship.
If the p value > α then we will accept H_0. If the p value < α then we will reject H_0. Under conditions of the null-hypothesis where the probability p(X < x|H_0) it is x=0 and μ= (1- p(A_t <- B_t)). We obtain based on the data of to the example before

$$p\left(X = 0\right) = \frac{\lambda^x}{x!} \times e^{-\lambda} = \frac{0.0125^0}{0!} \times e^{-0.0125} = 0.9875778 \qquad (119)$$

while the p-value is calculated as

$$p\left(X > 0\right) = 1 - \frac{\lambda^x}{x!} \times e^{-\lambda} = 1 - \frac{0.0125^0}{0!} \times e^{-0.0125} = 0.0124222 \qquad (120)$$

RESULTS

In this publication different [21-25, 73-108] kind of studies have been considered for a meta-analysis.

Without being married no HPV positivity of an Iranian man

Claims.
Null hypothesis:
Being married is a necessary condition (a conditio sine qua non) of HPV positivity of an Iranian man.
In other words, *withou*t being married *no* HPV positivity of an Iranian man.
Alternative hypothesis:
Being married is not a necessary condition (a conditio sine qua non) of a HPV positivity of an Iranian man.
In other words, *without* being married a HPV positivity of an Iranian man is possible.
The significance level (Alpha) below which the null hypotheses will be rejected is alpha=0.05.

Proof.
The conditio sine qua non relationship between being married and HPV positivity of an Iranian man was investigated by Ghasemian et al. (Table 3). Null-Hypotheses due to Pourmand et al. 2007: *without* being married *no* HPV positivity. The critical probability is p(Critical) = 0.9512 (N = 205; Table 3) while the probability calculated is p(SINE) = 0.960976. Hence, accept null-hypothesis: *without* being married *no* HPV positivity because p(SINE) > p(Critical). Null-Hypotheses according to Ghasemian et al. 2013: without being married *no* HPV positivity. The critical probability is p(Critical) = 0.9512 (N = 196; Table 3) while the probability calculated is p(SINE) = 0.994898. Hence, accept null-hypothesis: *without* being married *no* HPV positivity because p(SINE) > p(Critical). The data as presented (Table 3) support the null-hypotheses ***without*** being married ***no*** HPV positivity of an Iranian men.
Q. e. d.

Without being married no prostate cancer
Marital status can be a risk factor of prostate cancer. To evaluate the influence of marriage on men diagnosed with prostate cancer a sample size of N = 136402861 prostate cancer cases and controls was examined. Our expectation was that married men should not have a higher risk of prostate cancer compared unmarried men.

Claims.
Null hypothesis:
Being married is a necessary condition (a conditio sine qua non) of prostate cancer.
In other words, ***without*** being married *no* prostate cancer.
Alternative hypothesis:
Being married is not a necessary condition (a conditio sine qua non) of prostate cancer.
In other words, *without* being married a prostate cancer can develop.
The significance level (Alpha) below which the null hypotheses will be rejected is alpha=0.05.

Proof.
The conditio sine qua non relationship between being the martial status and prostate cancer was investigated by several studies (Table 5).
Null-Hypotheses due to Dillner et al. 1998: without being married no prostate cancer. The critical probability calculated according to the Anti Poisson distribution is equal to p(Critical) = 0.874889964 (N = 452; Table 5) while the probability calculated is p(SINE) = 0.97787611. Hence, accept null-hypothesis: without being married no prostate cancer because p(SINE) > p(Critical).
In particular, following Ghasemian et al. 2013, the Null-Hypotheses is: *without* being married *no* prostate cancer. The critical probability calculated according to the Anti Poisson distribution is equal to p(Critical) = 0.729329434 (N = 196; Table 5) while the probability calculated is p(SINE) = 0.98979592. Hence, accept null-hypothesis: without being married no prostate cancer because p(SINE) > p(Critical).
Null-Hypotheses based on the data of Schiffmann et al. 2015: *without* being married *no* prostate cancer. The critical probability calculated according to the Anti Poisson distribution is equal to p(Critical) = 0.98514433 (N = 880991; Table 5) while the probability calculated is p(SINE) = 0.9991816. Hence, accept null-hypothesis: without being married no prostate cancer because p(SINE) > p(Critical).
Null-Hypotheses according to Loeb et al. 2017: *without* being married *no* prostate cancer. The critical probability calculated according to the Anti Poisson distribution is equal to p(Critical) = 0.996485624 (N = 18855222; Table 5) while the probability calculated is p(SINE) = 0.99931658. Hence, accept null-hypothesis: without being married no prostate cancer because p(SINE) > p(Critical).
Null-Hypotheses with reference to this publication: *without* being married *no* prostate cancer. The critical probability calculated according to the Anti Poisson distribution is equal to p(Critical) = 0.998122768 (N = 116666000; Table 5) while the probability calculated is p(SINE) = 0.99961289. Hence, accept null-hypothesis: without being married no prostate cancer because p(SINE) > p(Critical).
The studies re-analysed with a sample size N = 136402861 support the Null-hypotheses ***without*** being married **no** prostate cancer.
Q. e. d.

If HPV PCR DNA is detected in human prostate tissues ***then*** prostate cancer

Claims.
Null hypothesis:
HPV is a sufficient condition (a conditio per quam) of prostate cancer. In other words, **if** HPV PCR DNA is positive **then** prostate cancer.

Alternative hypothesis:
HPV is not a sufficient condition (a conditio per quam) of prostate cancer. The significance level (Alpha) below which the null hypotheses will be rejected is alpha=0.05.

Proof.
The conditio per quam relationship between HPV and prostate cancer was investigated by more than 22 PCR based studies (Table 1) with a sample size of N = 2260. The studies of Whitaker et al. were re-analysed according to the rule of three. All studies meta-analysed (Table 1) support the Null-hypothesis: **if** HPV PCR DNA is detected in human prostate tissues **then** prostate cancer is present too.
Q. e. d.

The causal relationship between HPV PCR DNA and prostate cancer

Claims.
Null hypothesis:
HPV and prostate cancer are not causally related, both are independent of each other. k = 0.

Alternative hypothesis:
HPV and prostate cancer are causally related, both are not independent of each other k ≠ 0.
The significance level (Alpha) below which the null hypotheses will be rejected is alpha=0.05.

Proof.
The causal relationship between HPV and prostate cancer was investigated by 13 PCR based studies (Table 7) with a sample size of n=1369. The studies meta-analysed do not support the Null-hypothesis. According to the studies analysed HPV and prostate cancer are not independent of each other. In opposite, there is a significant i. e. highly significant cause effect relationship between HPV and prostate cancer.
Q. e. d.

DISCUSSION

Human papilloma viruses have been implicated in the pathogenesis of a variety of malignancies, especially in carcinomas of the female genital tract like human cervical cancer [2] but also in the development of human prostate cancer. Whether oncogenic human papilloma viruses (HPVs) are involved in the pathogenesis of prostate cancers is still a subject of some controversy. The study purpose was to clarify the contradictory results of investigations of the association of human papillomavirus (HPV) infection with prostate cancer.
In summary, several epidemiological studies have suggested that sexual behaviours such as or larger numbers of sexual partners and an early age at first intercourse are also related to an increased risk of prostate cancer. However, the results of sero-epidemiological studies [23, 109-113] in estimating the potential role of human papillomavirus infection in prostate carcinogenesis have produced differing results and remain inconclusive. There is much conflicting data surrounding sero-epidemiological studies and prostatic cancer. These studies were not considered for a review.

Several publication discussed the relationship between sexual behaviour and prostate cancer. In fact, it is known that human prostate cancer in *sexually inactive* male[10] children has not been reported. In contrast to young and male children, HPV infection is reported to be highly prevalent in *sexually active* men [7]. Thus far, a significant relationship between the martial status and the development of prostate cancer is able strengthen the degree of evidence of this relationship. Still, the results of investigations of the effect of marital status on the development of PC must not be consistent since the sexual culture of a human society and the individual behaviour can have great impact on the results of such studies. Especially in human societies were premarital sexual intercourse or extramarital sexual intercourse is common such investigations can be of limited value and possess the potential to lead to conflicting results. To date, in opposite to USA, in Iran premarital or extramarital sexual intercourse between members of the opposite sex is assiduously avoided and sex is more or less restricted to the marital bed. Besides of these fundamental differences in sexual behaviour between USA and Iran findings of Dillner[23] et al. and Ghasemian[22] et al. agree both on the relationship between martial status and the development of prostate cancer. The US data as published by Dillner[23] et al. support the hypotheses without being married no prostate cancer (p (*Without* HPV *no* PC) = (442/452) =0.9779; X²(Without HPV no PC)=0.5503). In line with the data of Dillner[23] et al. the Iranian study group of Ghasemian[22] et al. provided evidence that without being married no prostate cancer (p (without HPV no PC) = (194/196) =0.9898; X²(SINE)=0.0776).
Several hypothetical models can explain the relationship between marital status and prostate cancer. It is of course not reasonable to assume that the martial status as such is the cause of prostate cancer. Enjoying sexuality at martial sexual intercourse is not automatically accompanied with practising safer sex or to avoid getting a sexually transmitted infection like HPV. Human papillomavirus (HPV) is known to be the most common sexually transmitted infection. In the United States approximately 80% of all women acquire[114] an HPV infection by the age of 50. Furthermore, Ghasemian[22] et al. was able to document that HPV infection is transmitted by martial sexual intercourse. According to the data as published by Ghasemian[22] et al., without being married no HPV infection of an Iranian male (p (without being married no HPV infection of an Iranian man) = (194/196) = 0.9898; X²(without being married no HPV infection of an Iranian man)=0.0776). In particular, the studies of Dillner[23] et al. and Ghasemian[22] et al. and other studies (Table 5) provide strict evidence that HPV infection is related to prostate cancer.
To goal of this study was not to re-evaluate again the conventional risk[112] factors for prostate cancer which were already established[5] by publications but to investigate exclusively the relationship between HPV and PC based on PCR based methodology.
Yet, even after years of HPV DNA analysis in malignant and benign prostate samples, the causal involvement of HPV in prostate carcinogenesis is still a matter of controversial debate.

Investigations evaluating the presence of human papillomavirus (HPV) in prostatic tissue by polymerase chain reaction (PCR) technology have yielded very different detection rates between 0% and 100% and the negative or reduced HPV status demonstrated by some studies was used to provide strong arguments against an etiological role of HPV infection in the development and progression of prostate cancer. For instance, view studies [76, 81, 88, 92, 97] provided data against an etiological role of HPV infection in the development of prostate cancer.

Several factors have been suggested to explain the discrepancies observed. HPV DNA was detected by polymerase chain reaction (PCR) using different primers on different DNA regions (L1 region, E6/E7 region et cetera) and the discrepancies in HPV detection[116] published may be solely due to the differences in primer sets utilized. Terris & Peehl[85] were able to provide evidence that Human papillomavirus detection by polymerase chain reaction in benign and malignant prostate tissue is dependent more or less on the primer set utilized. A contamination by viral DNA i. e. from prostatic urethral colonization was not systematically excluded, less than optimal laboratory conditions and other factors must be considered too. In principle, the various studies have searched for different segments of the HPV genome and not the whole HPV DNA while utilizing different specific oligonucleotide primers for amplification was analysed. Many times, there was no systematic testing whether the material analysed was adequate (no evidence of DNA found by beta-globin testing prior to investigation) and gave variable and unsatisfactory results. Furthermore, the quality of the paraffin-embedded archival samples differs from the quality of the fresh frozen samples. But even if the investigations which evaluated the presence of human papillomavirus DNA in prostatic tissue by polymerase chain reaction (PCR) technology have yielded different detection rates the evidence is convincing and cannot be ignored. One objective of this study was to address these differences too. In this context, 22 studies (Table 1) with as sample size of N = 2260 support the null-hypotheses *if* HPV infection of human prostate *then* prostate cancer (Table 1). Even if 13 studies with as sample size of N = 1054 were self-contradictory (Table 2) and not considered for a causal meta-analysis, the evidence is convincing. The causal relationship between HPV and PC was at the same time significant or highly significant (13 studies, N = 1369, Table 7). In other words, there is a highly significant cause effect relationship between a HPV infection of human prostate and PC (13 studies, N = 1369, Table 7). Arguably, the following conclusion is inescapable since the studies presented demand us to accept the null-hypothesis: **without** being married **no** prostate cancer (5 studies, N = 136402861, Table 5).

According to several studies, **without** being married **no** HPV positivity of a men (Table 3). In the same context, **without** being married **no** prostate cancer (Table 5). All studies analysed support the null-hypothesis: **if** HPV in prostate tissues **then** prostate cancer (Table 1) while the cause effect relationship (Table 7) was significant/highly significant. Given this, it is scarcely not surprising that the data presented in this publication necessarily and inescapably demand us to articulate the need for something like the following conclusion.

CONCLUSION

Human papillomavirus is the cause of human prostate cancer.

References

001. IARC. Human papillomaviruses. *IARC monographs on the evaluation of carcinogenic risks to humans* 2007; 90: 1–636. https://www.ncbi.nlm.nih.gov/pubmed/18354839.

002. Barukčić I. Human Papillomavirus—The Cause of Human Cervical Cancer. *Journal of Biosciences and Medicines* 2018d; 06 (04): 106–125. https://doi.org/10.4236/jbm.2018.64009.

003. Rabkin CS, Biggar RJ, Melbye M & Curtis RE. Second primary cancers following anal and cervical carcinoma: evidence of shared etiologic factors. *American journal of epidemiology* 1992; 136 (1): 54–58. https://doi.org/10.1093/oxfordjournals.aje.a116420 https://www.ncbi.nlm.nih.gov/pubmed/1329500.

004. Giovannucci E, Liu Y, Platz EA, Stampfer MJ & Willett WC. Risk factors for prostate cancer incidence and progression in the health professionals follow-up study. *International journal of cancer* 2007; 121 (7): 1571–1578. https://doi.org/10.1002/ijc.22788 https://www.ncbi.nlm.nih.gov/pubmed/17450530.

005. van Dong H, Lee AH, Nga NH, Quang N, Le Chuyen V & Binns CW. Epidemiology and Prevention of Prostate Cancer in Vietnam. *Asian Pacific Journal of Cancer Prevention* 2014; 15 (22): 9747–9751. https://doi.org/10.7314/APJCP.2014.15.22.9747.

006. Pernar CH, Ebot EM, Wilson KM & Mucci LA. The Epidemiology of Prostate Cancer. *Cold Spring Harbor perspectives in medicine* 2018 https://doi.org/10.1101/cshperspect.a030361 https://www.ncbi.nlm.nih.gov/pubmed/29311132.

007. Bleeker MCG, Hogewoning CJA, Berkhof J, Voorhorst FJ, Hesselink AT, van Diemen PM, van den Brule AJC, Snijders PJF & Meijer CJLM. Concordance of specific human papillomavirus types in sex partners is more prevalent than would be expected by chance and is associated with increased viral loads. *Clinical infectious diseases an official publication of the Infectious Diseases Society of America* 2005; 41 (5): 612–620. https://doi.org/10.1086/431978 https://www.ncbi.nlm.nih.gov/pubmed/16080082.

008. Newell GR, Pollack ES, Spitz MR, Sider JG & Fueger JJ. Incidence of prostate cancer and marital status. *Journal of the National Cancer Institute* 1987; 79 (2): 259–262. https://www.ncbi.nlm.nih.gov/pubmed/3474458.

009. Liu Y, Xia Q, Xia J, Zhu H, Jiang H, Chen X, Zheng Y, Zhang F & Li S. The impact of marriage on the overall survival of prostate cancer patients: A Surveillance, Epidemiology, and End Results (SEER) analysis. *Canadian Urological Association journal = Journal de l'Association des urologues du Canada* 2018 https://doi.org/10.5489/cuaj.5413 https://www.ncbi.nlm.nih.gov/pubmed/30332597.

010. Badar F & Mahmood S. Epidemiology of cancers in Lahore, Pakistan, among children, adolescents and adults, 2010-2012: a cross-sectional study part 2. *BMJ open* 2017; 7 (12): e016559 https://doi.org/10.1136/bmjopen-2017-016559 https://www.ncbi.nlm.nih.gov/pubmed/29273649.

011. Bernard H-U, Burk RD, Chen Z, van Doorslaer K, Zur Hausen H & Villiers E-M de. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 2010; 401 (1): 70–79. https://doi.org/10.1016/j.virol.2010.02.002 https://www.ncbi.nlm.nih.gov/pubmed/20206957.

012. Doorbar J, Egawa N, Griffin H, Kranjec C & Murakami I. Human papillomavirus molecular biology and disease association. *Reviews in medical virology* 2015; 25 Suppl 1: 2–23. https://doi.org/10.1002/rmv.1822 https://www.ncbi.nlm.nih.gov/pubmed/25752814.

013. Hrbacek J, Urban M, Hamsikova E, Tachezy R, Eis V, Brabec M & Heracek J. Serum antibodies against genitourinary infectious agents in prostate cancer and benign prostate hyperplasia patients: a case-control study. *BMC cancer* 2011; 11: 53. https://doi.org/10.1186/1471-2407-11-53 https://www.ncbi.nlm.nih.gov/pubmed/21291519.

014. Yang L, Xie S, Feng X, Chen Y, Zheng T, Dai M, Zhou CK, Hu Z, Li N & Hang D. Worldwide Prevalence of Human Papillomavirus and Relative Risk of Prostate Cancer: A Meta-analysis. *Scientific reports* 2015; 5: 14667. https://doi.org/10.1038/srep14667 https://www.ncbi.nlm.nih.gov/pubmed/26441160.

015. Bae J-M. Human papillomavirus 16 infection as a potential risk factor for prostate cancer: an adaptive meta-analysis. *Epidemiology and health* 2015; 37: e2015005 https://doi.org/10.4178/epih/e2015005 https://www.ncbi.nlm.nih.gov/pubmed/25687950.

016. Russo GI, Calogero AE, Condorelli RA, Scalia G, Morgia G & La Vignera S. Human papillomavirus and risk of prostate cancer: a systematic review and meta-analysis. *The aging male the official journal of the International Society for the Study of the Aging Male* 2018: 1–7. https://doi.org/10.1080/13685538.2018.1455178 https://www.ncbi.nlm.nih.gov/pubmed/29571270.

017. Moher D, Liberati A, Tetzlaff J & Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine* 2009; 151 (4): 264-9, W64 https://www.ncbi.nlm.nih.gov/pubmed/19622511.

018. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J & Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine* 2009; 6 (7): e1000100 https://doi.org/10.1371/journal.pmed.1000100 https://www.ncbi.nlm.nih.gov/pubmed/19621070.

019. Barukčić I. Gastric Cancer and Epstein-Barr Virus Infection. *Modern Health Science* 2018; 1 (2): 1-18 https://doi.org/10.30560/mhs.v1n2p1.

020. Barukčić I. Human Cytomegalovirus is the Cause of Glioblastoma Multiforme. *Modern Health Science* 2018; 1 (2): p19 https://doi.org/10.30560/mhs.v1n2p19.

021. Pourmand G, Salem S, Mehrsai A, Lotfi M, Amirzargar MA, Mazdak H, Roshani A, Kheirollahi A, Kalantar E, Baradaran N, Saboury B, Allameh F, Karami A, Ahmadi H & Jahani Y. The risk factors of prostate cancer: a multicentric case-control study in Iran. *Asian Pacific journal of cancer prevention APJCP* 2007; 8 (3): 422–428. https://www.ncbi.nlm.nih.gov/pubmed/18159981.

022. Ghasemian E, Monavari SHR, Irajian GR, Jalali Nodoshan MR, Roudsari RV & Yahyapour Y. Evaluation of human papillomavirus infections in prostatic disease: a cross-sectional study in Iran. *Asian Pacific journal of cancer prevention APJCP* 2013; 14 (5): 3305–3308. https://www.ncbi.nlm.nih.gov/pubmed/23803120.

023. Dillner J, Knekt P, Boman J, Lehtinen M, Af Geijersstam V, Sapp M, Schiller J, Maatela J & Aromaa A. Sero-epidemiological association between human-papillomavirus infection and risk of prostate cancer. *International journal of cancer* 1998; 75 (4): 564–567. https://doi.org/10.1002/(SICI)1097-0215(19980209)75:4%3C564:AID-IJC12%3E3.0.CO;2-9 https://www.ncbi.nlm.nih.gov/pubmed/9466657.

024. Schiffmann J, Beyer B, Tennstedt P, Boehm K, Mehring G, Schlomm T, Salomon G, Karakiewicz P & Graefen M. Oncological outcome after radical prostatectomy: Marital status does not make a difference. *International journal of urology official journal of the Japanese Urological Association* 2015; 22 (5): 484–489.

https://doi.org/10.1111/iju.12717
https://www.ncbi.nlm.nih.gov/pubmed/25781055.
025. Huang T-B, Zhou G-C, Dong C-P, Wang L-P, Luan Y, Ye J-T, Gu X, Yao X-D, Zheng J-H & Ding X-F. Marital status independently predicts prostate cancer survival in men who underwent radical prostatectomy: An analysis of 95,846 individuals. *Oncology letters* 2018; 15 (4): 4737–4744. https://doi.org/10.3892/ol.2018.7964
https://www.ncbi.nlm.nih.gov/pubmed/29552113.
026.Moivre Ad. *The Doctrine of Chances or a Method of Calculating the Probability of Events in Play*. London: printed by W. Pearson for the author, 1718. pp. 175. p. 7
https://doi.org/10.3931/e-rara-10420 https://www.e-rara.ch/download/pdf/3043288?name=
The%20Doctrine%20of%20Chances%20or%20a%20Meth
od%20of%20Calculating%20the%20Probability%20of%20
Events%20in.pdf.
027. Kolmogoroff A. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1933 https://doi.org/10.1007/978-3-642-49888-6. *ISBN: 978-3-642-49596-0*.
028- Barukčić I. *Die Kausalität*, edn 1. Hamburg: Wiss.-Verl., 1989. 218 p. *ISBN: 3980221601*.
029. Barukčić I. *Theoriae causalitatis principia mathematica*. Norderstedt: Books on Demand, 2017. Online-Ressourcen. *ISBN: 9783744815932*.
030. Barukčić I. The Mathematical Formula of the Causal Relationship k. *International Journal of Applied Physics and Mathematics* 2016; 6 (2): 45–65.
https://doi.org/10.17706/ijapm.2016.6.2.45-65.
031. Barukčić K & Barukčić I. Epstein Barr Virus—The Cause of Multiple Sclerosis. *Journal of Applied Mathematics and Physics* 2016; 04 (06): 1042–1053.
https://doi.org/10.4236/jamp.2016.46109.
032. Barukčić I. Epstein-barr virus is the cause of multiple sclerosis. *International Journal of Current Medical and Pharmaceutical Research* 2018; 4 9 (A): 3674-3682
https://doi.org/10.24327/23956429.ijcmpr20180538
http://journalcmpr.com/sites/default/files/issue-files/1420-A-2018.pdf.
033. Wertheimer R. Conditions. *Journal of Philosophy* 1968; 65: 355–364. http://www.jstor.org/stable/2023797.
034. Gomes G. Are Necessary and Sufficient Conditions Converse Relations. *Australasian Journal of Philosophy* 2009; 87: 375–387. https://doi.org/10.1080/00048400802587325.
035. Fisher RA. On the Interpretation of χ 2 from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society* 1922; 85 (1): 87.
https://doi.org/10.2307/2340521.
036. Yates F. Contingency Tables Involving Small Numbers and the χ 2 Test. *The Journal of the Royal Statistical Society (Supplement)* 1934; 1 (2): 217.
https://doi.org/10.2307/2983604.
037. Grizzle JE. Continuity Correction in the χ 2 -Test for 2 × 2 Tables. *The American Statistician* 1967; 21 (4): 28.
https://doi.org/10.2307/2682103.
038. Conover WJ. Some Reasons for Not Using the Yates Continuity Correction on 2×2 Contingency Tables. *Journal of the American Statistical Association* 1974; 69 (346): 374–376.
https://doi.org/10.1080/01621459.1974.10482957.
039. Huygens C(1-1) & van Schooten F(1-1). *De ratiociniis in ludo alae: In: Exercitationum mathematicarum liber primus [-quintus]*. Lugdunum Batavorum (Leiden, The Netherlands): ex officina Johannis Elsevirii, 1657. 521-534
https://doi.org/10.3931/e-rara-8813 https://www.e-rara.ch/download/pdf/2486116?name=
Francisci%20%C3%A0%20Schooten%20Exercitationum%
20mathematicarum%20liber%20primus%20-
%20quintus.pdf.

040. Pearson K. XV. On certain properties of the hypergeometrical series, and on the fitting of such series to observation polygons in the theory of chance. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 1899; 47 (285): 236–246.
https://doi.org/10.1080/14786449908621253.
041. Gonin HT. XIV. The use of factorial moments in the treatment of the hypergeometric distribution and in tests for regression. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 1936; 21 (139): 215–226.
https://doi.org/10.1080/14786443608561573.
042. Fisher RA. The Logic of Inductive Inference. *Journal of the Royal Statistical Society* 1935; 98 (1): 39.
https://doi.org/10.2307/2342435.
043. CORNFIELD J. A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast, and cervix. *Journal of the National Cancer Institute* 1951; 11 (6): 1269–1275.
https://www.ncbi.nlm.nih.gov/pubmed/14861651.
044. Edwards AWF. The Measure of Association in a 2 × 2 Table. *Journal of the Royal Statistical Society. Series A (General)* 1963; 126 (1): 109. https://doi.org/10.2307/2982448.
045. Mosteller F. Association and Estimation in Contingency Tables. *Journal of the American Statistical Association* 1968; 63 (321): 1.
https://doi.org/10.2307/2283825.
046. Pagano M & Gauvreau K. *Principles of Biostatistics, Second Edition*, edn 2. Milton: CRC Press, 2018. 1585 p. *ISBN: 9781138593145* https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=5301930.
047. Yule GU. On the Association of Attributes in Statistics: With Illustrations from the Material of the Childhood Society, &c. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 1900; 194 252-261: 257–319. https://doi.org/10.1098/rsta.1900.0019.
048. Warrens MJ. On Association Coefficients for 2x2 Tables and Properties That Do Not Depend on the Marginal Distributions. *Psychometrika* 2008; 73 (4): 777–789.
https://doi.org/10.1007/s11336-008-9070-3
https://www.ncbi.nlm.nih.gov/pubmed/20046834.
049. Pearson K & Heron D. On Theories of Association. *Biometrika* 1913; 9 1-2: 159–315.
https://doi.org/10.1093/biomet/9.1-2.159.
050. Altman DG. *Practical statistics for medical research*. Boca Raton, Fla: Chapman & Hall/CRC, 1999. 611 p. *ISBN: 0412276305*.
051. Sachs L. *Angewandte Statistik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992. *ISBN: 978-3-540-52085-6*.
052. Isserlis L. On the Value of a Mean as Calculated from a Sample. *Journal of the Royal Statistical Society* 1918; 81 (1): 75.
https://doi.org/10.2307/2340569.
053. Pearson K. X. 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. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 1900; 50 (302): 157–175.
054. Rumke CL. Implications of the Statement: No Side Effects Were Observed. *The New England journal of medicine* 1975; 292 (7): 372–373.
055. Louis TA. Confidence Intervals for a Binomial Parameter after Observing No Successes. *The American Statistician* 1981; 35 (3): 154.
056. Hanley JA. If Nothing Goes Wrong, Is Everything All Right? *JAMA* 1983; 249 (13): 1743.
057. Jovanovic BD & Levy PS. A Look at the Rule of Three. *The American Statistician* 1997; 51 (2): 137–139.
058. Barukčić I. The Physical Meaning of the Wave Function. *Journal of Applied Mathematics and Physics* 2016; 04 (06): 988–1023. https://doi.org/10.4236/jamp.2016.46106.

059. Barukčić I. Unified Field Theory. *Journal of Applied Mathematics and Physics* 2016; 04 (08): 1379–1438. https://doi.org/10.4236/jamp.2016.48147.

060. DeGroot MH, Schervish MJ, Fang X, Lu L & Li D. *Probability and Statistics*. Beijing (China): Higher Education Press, 2005. p. 195. *ISBN: 9787040167658*.

061. Hume D. *David Hume: A Treatise of Human Nature (Second Edition)*: Oxford University Press, 1739 [Reprinted: 1978] https://doi.org/10.1093/actrade/9780198245872.book.1. *ISBN: 9780198245872* https://people.rit.edu/wlrgsh/HumeTreatise.pdf.

062.Bernoulli J. *Ars conjectandi, Opus posthumus: Accedit Tractatus de seriebus infinitis ; et epistola Gallice scripta De Ludo Pilae Reticularis*. Basileae (Basel, Suisse): Impensis Thurnisiorum [Tournes], fratrum, 1713. 306 p. https://doi.org/10.3931/e-rara-9001 https://www.e-rara.ch/download/pdf/2600091?name=Jacobi_Bernoulli_ars_coniectandi_opus_posthumus.pdf.

063. LaPlace, Pierre Simon de. *Théorie analytique des probabilités*. Paris (France): Courcier, 1812. 464 p. https://doi.org/10.3931/e-rara-9457 https://www.e-rara.ch/download/pdf/2802935?name=Th%C3%A9orie%20analytique%20des%20probabilit%C3%A9s.pdf.

064. Lyapunov AM. Nouvelle forme du théorème sur la limite de probabilité. *Mémoires de l'Académie Impériale des Sciences de St.-Pétersbourg. Série VIIIe. Classe Physico-Mathématique* 1901; 12: 1–24.

065. Pólya G. Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem: [In English: On the central limit theorem of probability calculation and the problem of moments]. *Mathematische Zeitschrift* 1920; 8 3-4: 171–181. https://doi.org/10.1007/BF01206525.

066. Moivre A de. *Approximatio ad summam terminorum binomii $(a+b)^n$ in seriem expansi*. London: Privately (Publisher not identified), 1733. 7 p. https://www.york.ac.uk/depts/maths/histstat/demoivre.pdf.

067. Poisson SD. *Recherches sur la Probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilitiés*. Paris, France: Bachelier, 1837. p. 206 https://www-liphy.ujf-grenoble.fr/pagesperso/bahram/Phys_Stat/Biblio/Poisson_Proba_1838.pdf.

068. Bortkiewicz L von, Ed. *Das Gesetz der kleinen Zahlen: [Transl. into English: The law of small numbers]*. Leipzig (Germany): B.G. Teubner, 1898.

069. Tchébychef, Pafnuty Lvovich. Des valeurs moyennes. *Journal de Mathématiques Pures et Appliquées* 1867; 2 (12): 177–184.

070. Arbuthnott J. An Argument for Divine Providence, Taken from the Constant Regularity Observ'd in the Births of Both Sexes. By Dr. John Arbuthnott, Physitian in Ordinary to Her Majesty, and Fellow of the College of Physitians and the Royal Society. Philosophical Transactions of the Royal Society of London 1710; 27 325-336: 186–190. https://doi.org/10.1098/rstl.1710.0011

071. Du Prel J-B, Hommel G, Röhrig B & Blettner M. Confidence Interval or P-Value?: Part 4 of a Series on Evaluation of Scientific Publications. Deutsches Ärzteblatt International 2009; 106 (19): 335–339. https://doi.org/10.3238/arztebl.2009.0335 https://www.ncbi.nlm.nih.gov/pubmed/19547734 .

072. Fisher RA. Statistical Methods for Research Workers.Edinburgh: Oliver and Boyd, 1925. 145 p. http://www.haghish.com/resources/materials/Statistical_Methods_for_Research_Workers.pdf

073. Loeb S, Folkvaljon Y, Damber J-E, Alukal J, Lambe M & Stattin P. Testosterone Replacement Therapy and Risk of Favorable and Aggressive Prostate Cancer. *Journal of clinical oncology official journal of the American Society of Clinical Oncology* 2017; 35 (13): 1430–1436. https://doi.org/10.1200/JCO.2016.69.5304 https://www.ncbi.nlm.nih.gov/pubmed/28447913.

074. Anwar K, Nakakuki K, Shiraishi T, Naiki H, Yatani R & Inuzuka M. Presence of ras oncogene mutations and human papillomavirus DNA in human prostate carcinomas. *Cancer research* 1992; 52 (21): 5991–5996. https://www.ncbi.nlm.nih.gov/pubmed/1382850.

075. McNicol PJ & Dodd JG. High prevalence of human papillomavirus in prostate tissues. *The Journal of urology* 1991; 145 (4): 850–853. https://www.ncbi.nlm.nih.gov/pubmed/1848641.

076. Masood S, Rhatigan RM, Powell S, Thompson J & Rodenroth N. Human papillomavirus in prostatic cancer: no evidence found by in situ DNA hybridization. *Southern medical journal* 1991; 84 (2): 235–236. https://www.ncbi.nlm.nih.gov/pubmed/1703667.

077. Ibrahim GK, Gravitt PE, Dittrich KL, Ibrahim SN, Melhus O, Anderson SM & Robertson CN. Detection of human papillomavirus in the prostate by polymerase chain reaction and in situ hybridization. *The Journal of urology* 1992; 148 (6): 1822–1826. https://www.ncbi.nlm.nih.gov/pubmed/1279224.

078. Effert PJ, Frye RA, Neubauer A, Liu ET & Walther PJ. Human papillomavirus types 16 and 18 are not involved in human prostate carcinogenesis: analysis of archival human prostate cancer specimens by differential polymerase chain reaction. *The Journal of urology* 1992; 147 (1): 192–196. https://www.ncbi.nlm.nih.gov/pubmed/1309581.

079. Rotola A, Monini P, Di Luca D, Savioli A, Simone R, Secchiero P, Reggiani A & Cassai E. Presence and physical state of HPV DNA in prostate and urinary-tract tissues. *International journal of cancer* 1992; 52 (3): 359–365. https://www.ncbi.nlm.nih.gov/pubmed/1328067.

080. Dodd JG, Paraskevas M & McNicol PJ. Detection of human papillomavirus 16 transcription in human prostate tissue. *The Journal of urology* 1993; 149 (2): 400–402. https://www.ncbi.nlm.nih.gov/pubmed/7678873.

081. Tu H, Jacobs SC, Mergner WJ & Kyprianou N. Rare incidence of human papillomavirus types 16 and 18 in primary and metastatic human prostate cancer. *Urology* 1994; 44 (5): 726–731. https://www.ncbi.nlm.nih.gov/pubmed/7974946.

082. Moyret-Lalle C, Marçais C, Jacquemier J, Moles JP, Daver A, Soret JY, Jeanteur P, Ozturk M & Theillet C. ras, p53 and HPV status in benign and malignant prostate tumors. *International journal of cancer* 1995; 64 (2): 124–129. https://www.ncbi.nlm.nih.gov/pubmed/7542226.

083. Suzuki H, Komiya A, Aida S, Ito H, Yatani R & Shimazaki J. Detection of human papillomavirus DNA and p53 gene mutations in human prostate cancer. *The Prostate* 1996; 28 (5): 318–324. https://doi.org/10.1002/(SICI)1097-0045(199605)28:5<318:AID-PROS8>3.0.CO;2-7.

084. Wideroff L, Schottenfeld D, Carey TE, Beals T, Fu G, Sakr W, Sarkar F, Schork A, Grossman HB & Shaw MW. Human papillomavirus DNA in malignant and hyperplastic prostate tissue of black and white males. *The Prostate* 1996; 28 (2): 117–123. https://doi.org/10.1002/(SICI)1097-0045(199602)28:2<117:AID-PROS7>3.0.CO;2-D.

085. Terris MK & Peehl DM. Human papillomavirus detection by polymerase chain reaction in benign and malignant prostate tissue is dependent on the primer set utilized. *Urology* 1997; 50 (1): 150–156. https://doi.org/10.1016/S0090-4295(97)00126-X.

086. Anderson M, Handley J, Hopwood L, Murant S, Stower M & Maitland NJ. Analysis of prostate tissue DNA for the presence of human papillomavirus by polymerase chain reaction, cloning, and automated sequencing. *Journal of

medical virology 1997; 52 (1): 8–13.
https://www.ncbi.nlm.nih.gov/pubmed/9131451.
087. Noda T, Sasagawa T, Dong Y, Fuse H, Namiki M & Inoue M. Detection of human papillomavirus (HPV) DNA in archival specimens of benign prostatic hyperplasia and prostatic cancer using a highly sensitive nested PCR method. *Urological research* 1998; 26 (3): 165–169.
https://www.ncbi.nlm.nih.gov/pubmed/9694597.
088. Strickler HD, Burk R, Shah K, Viscidi R, Jackson A, Pizza G, Bertoni F, Schiller JT, Manns A, Metcalf R, Qu W & Goedert JJ. A multifaceted study of human papillomavirus and prostate carcinoma. *Cancer* 1998; 82 (6): 1118–1125.
https://www.ncbi.nlm.nih.gov/pubmed/9506358.
089. Serth J, Panitz F, Paeslack U, Kuczyk MA & Jonas U. Increased levels of human papillomavirus type 16 DNA in a subset of prostate cancers. *Cancer research* 1999; 59 (4): 823–825. https://www.ncbi.nlm.nih.gov/pubmed/10029070.
090. Carozzi F, Lombardi FC, Zendron P, Confortini M, Sani C, Bisanzi S, Pontenani G & Ciatto S. Association of human papillomavirus with prostate cancer: analysis of a consecutive series of prostate biopsies. *The International journal of biological markers* 2004; 19 (4): 257–261.
https://www.ncbi.nlm.nih.gov/pubmed/15646830.
091. Leiros GJ, Galliano SR, Sember ME, Kahn T, Schwarz E & Eiguchi K. Detection of human papillomavirus DNA and p53 codon 72 polymorphism in prostate carcinomas of patients from Argentina. *BMC urology* 2005; 5: 15.
https://doi.org/10.1186/1471-2490-5-15
https://www.ncbi.nlm.nih.gov/pubmed/16307686.
092. Silvestre RVD, Leal MF, Demachki S, Nahum MCdS, Bernardes JGB, Rabenhorst SHB, Smith MdAC, Mello WAd, Guimarães AC & Burbano RR. Low frequency of human papillomavirus detection in prostate tissue from individuals from Northern Brazil. *Memorias do Instituto Oswaldo Cruz* 2009; 104 (4): 665–667.
https://www.ncbi.nlm.nih.gov/pubmed/19722096.
093. Gazzaz FS & Mosli HA. Lack of detection of human papillomavirus infection by hybridization test in prostatic biopsies. *Saudi medical journal* 2009; 30 (5): 633–637.
https://www.ncbi.nlm.nih.gov/pubmed/19417961.
094. Martinez-Fierro ML, Leach RJ, Gomez-Guerra LS, Garza-Guajardo R, Johnson-Pais T, Beuten J, Morales-Rodriguez IB, Hernandez-Ordoñez MA, Calderon-Cardenas G, Ortiz-Lopez R, Rivas-Estilla AM, Ancer-Rodriguez J & Rojas-Martinez A. Identification of viral infections in the prostate and evaluation of their association with cancer. *BMC cancer* 2010; 10: 326. https://doi.org/10.1186/1471-2407-10-326
https://www.ncbi.nlm.nih.gov/pubmed/20576103.
095. Aghakhani A, Hamkar R, Parvin M, Ghavami N, Nadri M, Pakfetrat A, Banifazl M, Eslamifar A, Izadi N, Jam S & Ramezani A. The role of human papillomavirus infection in prostate carcinoma. *Scandinavian journal of infectious diseases* 2011; 43 (1): 64–69.
https://doi.org/10.3109/00365548.2010.502904
https://www.ncbi.nlm.nih.gov/pubmed/20662618.
096. Chen AC-H, Waterboer T, Keleher A, Morrison B, Jindal S, McMillan D, Nicol D, Gardiner RA, McMillan NAJ & Antonsson A. Human papillomavirus in benign prostatic hyperplasia and prostatic adenocarcinoma patients. *Pathology oncology research POR* 2011; 17 (3): 613–617.
https://doi.org/10.1007/s12253-010-9357-4
https://www.ncbi.nlm.nih.gov/pubmed/21240663.
097. Salehi Z & Hadavi M. Analysis of the codon 72 polymorphism of TP53 and human papillomavirus infection in Iranian patients with prostate cancer. *Journal of medical virology* 2012; 84 (9): 1423–1427.
https://doi.org/10.1002/jmv.23268
https://www.ncbi.nlm.nih.gov/pubmed/22825821.

098. Tachezy R, Hrbacek J, Heracek J, Salakova M, Smahelova J, Ludvikova V, Svec A, Urban M & Hamsikova E. HPV persistence and its oncogenic role in prostate tumors. *Journal of medical virology* 2012; 84 (10): 1636–1645.
https://doi.org/10.1002/jmv.23367
https://www.ncbi.nlm.nih.gov/pubmed/22930513.
099. Mokhtari M, Taghizadeh F & Hani M. Is prostatic adenocarcinoma in a relationship with Human Papilloma Virus in Isfahan -Iran. *Journal of Research in Medical Sciences The Official Journal of Isfahan University of Medical Sciences* 2013; 18 (8): 707–710.
https://www.ncbi.nlm.nih.gov/pubmed/24379849.
100. Whitaker NJ, Glenn WK, Sahrudin A, Orde MM, Delprado W & Lawson JS. Human papillomavirus and Epstein Barr virus in prostate cancer: Koilocytes indicate potential oncogenic influences of human papillomavirus in prostate cancer. *The Prostate* 2013; 73 (3): 236–241.
https://doi.org/10.1002/pros.22562
https://www.ncbi.nlm.nih.gov/pubmed/22851253.
101. Ghasemian E, Monavari SHR, Irajian GR, Jalali Nodoshan MR, Roudsari RV & Yahyapour Y. Evaluation of human papillomavirus infections in prostatic disease: a cross-sectional study in Iran. *Asian Pacific journal of cancer prevention APJCP* 2013; 14 (5): 3305–3308.
https://www.ncbi.nlm.nih.gov/pubmed/23803120.
102. Michopoulou V, Derdas SP, Symvoulakis E, Mourmouras N, Nomikos A, Delakas D, Sourvinos G & Spandidos DA. Detection of human papillomavirus (HPV) DNA prevalence and p53 codon 72 (Arg72Pro) polymorphism in prostate cancer in a Greek group of patients. *Tumour biology the journal of the International Society for Oncodevelopmental Biology and Medicine* 2014; 35 (12): 12765–12773. https://doi.org/10.1007/s13277-014-2604-7
https://www.ncbi.nlm.nih.gov/pubmed/25213701.
103. Yow MA, Tabrizi SN, Severi G, Bolton DM, Pedersen J, Longano A, Garland SM, Southey MC & Giles GG. Detection of infectious organisms in archival prostate cancer tissues. *BMC cancer* 2014; 14: 579.
https://doi.org/10.1186/1471-2407-14-579
https://www.ncbi.nlm.nih.gov/pubmed/25106851.
104. Singh N, Hussain S, Kakkar N, Singh SK, Sobti RC & Bharadwaj M. Implication of high risk human papillomavirus HR-HPV infection in prostate cancer in Indian population--a pioneering case-control analysis. *Scientific reports* 2015; 5: 7822.
https://doi.org/10.1038/srep07822
https://www.ncbi.nlm.nih.gov/pubmed/25592643.
105. Huang L, Wu M-G, He J, Wei Z-S, Lü W-X, Song X-J, Zhang Y, Wu S-X, Yin Y-l & Fan Y-Y. Correlation of high-risk HPV 16/18 infections with prostate cancer. *Zhonghua nan ke xue = National journal of andrology* 2016; 22 (6): 501–505. https://www.ncbi.nlm.nih.gov/pubmed/28963837.
106. Atashafrooz F & Rokhbakhsh-Zamin F. Frequency and Type Distribution of Human Papilloma Virus in Patients with Prostate Cancer, Kerman, Southeast of Iran. *Asian Pacific journal of cancer prevention APJCP* 2016; 17 (8): 3953–3958. https://www.ncbi.nlm.nih.gov/pubmed/27644644.
107. Aydin M, Bozkurt A, Cikman A, Gulhan B, Karabakan M, Gokce A, Alper M & Kara M. Lack of evidence of HPV etiology of prostate cancer following radical surgery and higher frequency of the Arg/Pro genotype in Turkish men with prostate cancer. *International braz j urol official journal of the Brazilian Society of Urology* 2017; 43 (1): 36–46.
https://doi.org/10.1590/S1677-5538.IBJU.2015.0429
https://www.ncbi.nlm.nih.gov/pubmed/28124524.
108. Zhao X, Zhou Z, Chen Y, Chen W, Ma H & Pu J. Role of antibodies to human papillomavirus 16 in prostate cancer: A seroscreening by peptide microarray. *Tumour biology the journal of the International Society for Oncodevelopmental Biology and Medicine* 2017; 39 (6): 1010428317698371

https://doi.org/10.1177/1010428317698371
https://www.ncbi.nlm.nih.gov/pubmed/28618964.

109. Adami H-O, Kuper H, Andersson S-O, Bergström R & Dillner J. Prostate cancer risk and serologic evidence of human papilloma virus infection: a population-based case-control study. *Cancer epidemiology, biomarkers & prevention a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2003; 12 (9): 872–875.
https://www.ncbi.nlm.nih.gov/pubmed/14504197.

110. Rosenblatt KA, Carter JJ, Iwasaki LM, Galloway DA & Stanford JL. Serologic evidence of human papillomavirus 16 and 18 infections and risk of prostate cancer. *Cancer epidemiology, biomarkers & prevention a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2003; 12 (8): 763–768.
https://www.ncbi.nlm.nih.gov/pubmed/12917208.

111. Sutcliffe S, Viscidi RP, Till C, Goodman PJ, Hoque AM, Hsing AW, Thompson IM, Zenilman JM, Marzo AM de & Platz EA. Human papillomavirus types 16, 18, and 31 serostatus and prostate cancer risk in the Prostate Cancer Prevention Trial. *Cancer epidemiology, biomarkers & prevention a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2010; 19 (2): 614–618. https://doi.org/10.1158/1055-9965.EPI-09-1080 https://www.ncbi.nlm.nih.gov/pubmed/20142255.

112. Huang W-Y, Hayes R, Pfeiffer R, Viscidi RP, Lee FK, Wang YF, Reding D, Whitby D, Papp JR & Rabkin CS. Sexually transmissible infections and prostate cancer risk. *Cancer epidemiology, biomarkers & prevention a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2008; 17 (9): 2374–2381.
https://doi.org/10.1158/1055-9965.EPI-08-0173
https://www.ncbi.nlm.nih.gov/pubmed/18768506.

113. Dennis LK, Coughlin JA, McKinnon BC, Wells TS, Gaydos CA, Hamsikova E & Gray GC. Sexually transmitted infections and prostate cancer among men in the U.S. military. *Cancer epidemiology, biomarkers & prevention a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2009; 18 (10): 2665–2671. https://doi.org/10.1158/1055-9965.EPI-08-1167
https://www.ncbi.nlm.nih.gov/pubmed/19755645.

114. Braaten KP & Laufer MR. Human Papillomavirus (HPV), HPV-Related Disease, and the HPV Vaccine. *Reviews in Obstetrics and Gynecology* 2008; 1 (1): 2–10.
https://www.ncbi.nlm.nih.gov/pubmed/18701931 .

115. Giovannucci E, Liu Y, Platz EA, Stampfer MJ & Willett WC. Risk factors for prostate cancer incidence and progression in the health professionals follow-up study. *International journal of cancer* 2007; 121 (7): 1571–1578.
https://doi.org/10.1002/ijc.22788
https://www.ncbi.nlm.nih.gov/pubmed/17450530.

116. Cuzick J. Human papillomavirus infection of the prostate. *Cancer surveys* 1995; 23: 91–95.
https://www.ncbi.nlm.nih.gov/pubmed/7621477.

Table 1 : The HPV PCR Studies considered for a re-analysis

Study Id	Year	Country	Risk Factor	Case_P	Case_T	Con_P	Con_T	k	p-val	IOU	X²(IMP)
Ibrahim et al.	1992	USA	High-risk HPV16/18 PCR	6	24	2	36	0.2802243	0.033141089	-0.47	0.28
Anwar et al.	1992	Japan	High-risk HPV16/18/33 PCR	28	68	0	10	0.28697202	0.008163513	0.23	0.01
Tu et al.	1994	USA	High-risk HPV16/18 PCR	1	43	0	1	0.02325581	0.977272727	0.00	0.25
Moyret-Lalle et al.	1995	France	High-risk HPV16/18 PCR	14	27	8	24	0.18663084	0.094519201	-0.04	2.56
Suzuki et al.	1996	Japan	High-risk HPV16 PCR	8	51	0	51	0.29172998	0.002903682	-0.42	0.03
Wideroff et al.	1996	USA	HPV PCR	7	56	4	42	0.0466577	0.231679548	-0.32	1.11
Terris & Peehl et al.	1997	USA	High-risk HPV16/18 PCR	10	53	5	37	0.07069265	0.185598295	-0.24	1.35
Serth et al.	1999	Germany	HPV16 PCR	10	47	1	37	0.27333482	0.010314777	-0.31	0.02
Carozzi et al.	2004	Italy	High-risk HPV type	14	26	5	25	0.34995662	0.01058851	-0.12	1.07
Leiros et al.	2005	Argentina	HPV PCR	17	41	0	30	0.47995031	1.46345E-05	-0.18	0.01
Silvestre et al.	2009	Brasil	HPV PCR	2	65	0	6	0.05172606	0.837022133	-0.06	0.13
Martinez-Fierro et al.	2010	Mexico	HPV PCR	11	55	4	75	0.22680303	0.008602189	-0.46	0.82
Aghakhani et al.	2011	Iran	HPV PCR	13	104	8	104	0.07978836	0.095738433	-0.40	2.68
Salehi and Hadavi	2012	Iran	HPV PCR	3	68	0	85	0.15811388	0.085627977	-0.54	0.08
Mokhtari et al.	2013	Iran	HPV PCR	3	30	1	90	0.21442251	0.044481939	-0.72	0.06
Whitaker et al.	2013	Australia	HPV PCR	7	10	2	10	0.50251891	0.032150512	-0.05	0.25
Michopoulou et al.	2014	Greece	HPV PCR	8	50	1	30	0.19406961	0.069453811	-0.26	0.03
Singh et al.	2015	India	HPV PCR	39	95	11	55	0.21521103	0.004234054	-0.03	2.21
Huang et al.	2016	China	High-risk HPV16/18 PCR	30	75	0	73	0.49745113	3.80058E-11	-0.29	0.01
Atashafrooz et al.	2016	Iran	HPV PCR	20	100	8	100	0.17291713	0.008230537	-0.36	2.01
Aydin et al.	2017	Turkey	HPV PCR	1	60	0	36	0.07947194	0.625	-0.36	0.25
Zhao et al.	2017	China	High-risk HPV16 PCR	48	75	14	80	0.47434165	2.10403E-09	-0.12	2.94
			Total	300	1223	74	1037	0.23323425	1.27175E-30		14.4445

N =	**2260**
Alpha =	0.05
Degrees of freedom (d. f.) =	22
X² Critical (SINE) =	**33.9244**
X² Calculated (SINE) =	14.4445
Index of unfairness =	-0.29336

Case_P: cases, positive; Case_T: cases, total; Con_P: controls, positive; Con_T: controls, total.

Table 2. Studies not considered for a re-analysis

Study Id	Year	Country	Risk Factor	Case_P	Case_T	Con_P	Con_T	k	p-val	IOU	X²(SINE)	X²(IMP)	X²(IMP^SINE)	X²(EXCL)
McNicol and Dodd	1991	Canada	HPV PCR	14	27	34	56	**-0.08407643**	0.139723165	-0.10	5.79	23.38	29.17	**10.55**
Masood et al.	1991	USA	HPV PCR	0	20	0	20	#DIV/0!	1	-0.50	19.01	#DIV/0!	#DIV/0!	#DIV/0!
Rotola et al.	1992	Italy	HPV PCR	6	8	14	17	**-0.08574929**	0.358366271	0.12	**0.28**	9.11	9.39	5.29
Dodd et al.	1993	Canada	HPV PCR	3	7	5	10	**-0.07042952**	0.362813657	-0.12	1.75	**2.53**	4.28	1.67
Effert et al.	1992	USA	High-risk HPV16/18 PCR	0	30	0	0	#DIV/0!	1	0.00	29.01	#DIV/0!	#DIV/0!	#DIV/0!
Anderson et al.	1997	UK	HPV PCR	0	0	0	0	#DIV/0!	1	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
Noda et al.	1998	Japan	HPV PCR	0	38	3	71	**-0.12307513**	0.272252232	-0.62	37.01	**2.08**	39.09	0.09
Strickler et al.	1998	USA	HPV PCR	0	63	0	61	#DIV/0!	1	-0.49	62.00	#DIV/0!	#DIV/0!	#DIV/0!
Gazzaz and Mosli	2009	Saudi Arabia	HPV PCR	0	6	0	50	#DIV/0!	1	-0.89	5.04	#DIV/0!	#DIV/0!	#DIV/0!
Chen et al.	2011	Australia	HPV PCR	7	51	3	11	**-0.14071179**	0.177670024	-0.02	37.10	**0.63**	37.73	5.05
Tachezy et al.	2012	Czech Republic	HPV PCR	1	51	2	95	**-0.00485537**	0.448187293	-0.63	48.04	**0.75**	48.79	0.09
Ghasemian et al.	2013	Iran	HPV PCR	5	29	8	167	**0.17764904**	0.02231058	-0.79	19.04	4.33	23.37	**2.26**
Yow et al.	2014	Australia	HPV PCR	0	115	0	51	#DIV/0!	1	-0.31	114.00	#DIV/0!	#DIV/0!	#DIV/0!
			Total	36	445	69	609							

N = **1054**

Table 3. The relationship between martial status and HPV positivity

Study Id	Year	Country	Risk Factor	Case_P	Case_T	Con_P	Con_T	p (SINE)	p Critical	IOU	X²(SINE)
Pourmand et al.	2007	Iran	Married	122	130	72	75	0.960976	0.9512	0.58	0.43
Ghasemian et al.	2013	Iran	Married	12	13	167	183	0.994898	0.9512	-0.02	0.02
			Total	134	143	239	258				0.5052

N = **401**

Alpha = 0.05

Degrees of freedom (d. f.) = 2

X² Critical (SINE) = **5.99146**

X² Calculated (SINE) = 0.5052

Index of unfairness = 0.286783

Case_P: cases, positive; Case_T: cases, total; Con_P: controls, positive; Con_T: controls, total.

Table 5. The relationship between martial status and prostate cancer

Study Id	Year	Country	Risk Factor	Case_P	Case_T	Con_P	Con_T	p(SINE)	p(Critical)
Dillner et al.	1998	USA	Married	154	164	259	288	0.97787611	0.874889964
Ghasemian et al.	2013	Iran	Married	27	29	152	167	0.98979592	0.729329434
Schiffmann et al.	2015	Germany	Married	7367	8088	461124	872903	0.9991816	0.98514433
Loeb et al.	2017	Sweden	Married	25684	38570	12607157	18816652	0.99931658	0.996485624
Barukčić	2018	Germany	Married	431989	477152	77182011	116188848	0.99961289	0.998122768
			Total	465221	524003	90250703	135878858		

N = **136402861**

Case_P: cases, positive; Case_T: cases, total; Con_P: controls, positive; Con_T: controls, total.

Table 4: The Study of Ghasemian et al.

		HPV Positive <B>		
		Yes	No	Total
Married	Yes	12	167	179
<A>	No	**1**	16	17
	Total	13	183	196

k = **0.0093**

p value (k) = **0.39106814**

95% CI (k) = [-0.1504; 0.1690]

Index of unfairness = -0.0204 [-1; +1]

WITHOUT <A>   NO <B>.

p (SINE) = **0.994898**

X²(SINE) = **0.0192**

p critical= **0.9512**

p(Anti Poisson) = *0.6321*

Odds ratio = 1.1497

95% CI (Odds ratio) = [0.1403-9.4218]

Table 7. The causal relationship between human papilloma virus and prostate cancer

Study Id	Year	Country	Risk Factor	Case_P	Case_T	Con_P	Con_T	k	p-val	IOU	X²(k)
Huang et al.	2016	China	High-risk HPV16/18 PCR	30	75	0	73	0.49745113	3.80058E-11	-0.29	36.62
Zhao et al.	2017	China	High-risk HPV16 PCR	48	75	14	80	0.47434165	2.10403E-09	-0.12	34.88
Leiros et al.	2005	Argentina	HPV PCR	17	41	0	30	0.47995031	1.46345E-05	-0.18	16.36
Suzuki et al.	1996	Japan	High-risk HPV16 PCR	8	51	0	51	0.29172998	0.002903682	-0.42	8.68
Singh et al.	2015	India	HPV PCR	39	95	11	55	0.21521103	0.004234054	-0.03	6.95
Anwar et al.	1992	Japan	High-risk HPV16/18/33 PCR	28	68	0	10	0.28697202	0.008163513	0.23	6.42
Atashafrooz et al.	2016	Iran	HPV PCR	20	100	8	100	0.17291713	0.008230537	-0.36	5.98
Martinez-Fierro et al.	2010	Mexico	HPV PCR	11	55	4	75	0.22680303	0.008602189	-0.46	6.69
Serth et al.	1999	Germany	HPV16 PCR	10	47	1	37	0.27333482	0.010314777	-0.31	6.28
Carozzi et al.	2004	Italy	High-risk HPV type	14	26	5	25	0.34995662	0.01058851	-0.12	6.25
Whitaker et al.	2013	Australia	HPV PCR	7	10	2	10	0.50251891	0.032150512	-0.05	5.05
Ibrahim et al.	1992	USA	High-risk HPV16/18 PCR	6	24	2	36	0.2802243	0.033141089	-0.47	4.71
Mokhtari et al.	2013	Iran	HPV PCR	3	30	1	90	0.21442251	0.044481939	-0.72	5.52
			Total	241	697	48	672	0.69849758	2.95206E-73	-0.49	150

N =	1369
Alpha =	0.05
Degrees of freedom (d. f.) =	13
X² Critical (k) =	22.36
X² Calculated (k) =	150.37
p value (k) <	0.00001

Case_P: cases, positive; Case_T: cases, total; Con_P: controls, positive; Con_T: controls, total.