Fusobacterium (nucleatum) - The cause of human colorectal cancer

Ilija Barukčić

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

Abstract

Objective: Accumulating evidence indicates that the gut microbiome has an increasingly important role in human disease and health. Fusobacterium nucleatum has been identified in several studies as the leading gut bacterium which is present in colorectal cancer (CRC). Still it is not clear if Fusobacterium plays a causal role.
Methods: To explore the cause-effect relationship between Fusobacterium nucleatum and colorectal cancer, a systematic review and re-analysis of studies published was performed. The method of the conditio sine qua non relationship was used to proof the hypothesis without Fusobacterium nucleatum infection no colorectal cancer. The mathematical formula of the causal relationship k was used to proof the hypothesis, whether there is a cause effect relationship between Fusobacterium nucleatum and colorectal cancer. Significance was indicated by a p-value of less than 0.05
Result: The data analyzed support the Null-hypothesis that without Fusobacterium nucleatum infection no colorectal cancer. In the same respect, the studies analyzed provide highly significant cause effect relationship between Fusobacterium nucleatum and colorectal cancer.
Conclusion: The findings of this study suggest that Fusobacterium (nucleatum) is the cause of colorectal cancer.

Keywords

Fusobacterium nucleatum, human colorectal cancer, causal relationship

1. Introduction

Colorectal cancer has been ranked as the fourth most common cancer cause of death and the third most common cancer worldwide [1], [2]. Almost 694,000 deaths and about 1.4 million new cases occurred in 2012. The mortality of patients with metastatic colorectal cancer disease is very high. In point of fact, the necessity of a good and reliable screening method to detect colorectal cancer at an early operable stage is very important. Several techniques including biochemical tests for colorectal cancer (immunochemical FOBT,

guaiac faecal occult blood test - gFOBT), sigmoidoscopy, colonoscopy, CT colonography, stool DNA, capsule endoscopy and other methods are used to detect even early stages of colorectal cancer. Colonoscopy is currently the most reliable method for detection of CRC. Sometimes, colonoscopy is uncomfortable for the patients, time consuming and very costly. The management of screen-detected or of symptomatic colon rectal cancer includes high-quality surgery, chemotherapy, adjuvant radiotherapy given either pre-operatively or post operatively and other measures. Very often the outcomes are unsatisfactory. The etiology of colorectal cancer is still not fully understood. Over the past few decades, a number of life style and environmental factors contributing to the occurrence of colorectal cancer have been identified, including low vegetable and fruit intake [3] and tobacco smoking [4], family history of colorectal cancer [5], high consumption of red and processed meat [6], excessive alcohol consumption [7], Diabetes mellitus [8], inflammatory bowel disease [9], obesity [10]. However, results remain inconsistent and no single risk factor was identified as being responsible for colorectal cancer. The most of these factors at best confer a very moderate risk for colorectal cancer [11]. Numerous studies have aimed to provide evidence of the presence of infectious agents viral DNA such as human papillomaviruses (HPV), human polyomaviruses, human herpesviruses et cetera [12] in colorectal tumor tissues but the evidence has remained inconclusive and is still very limited. The hypothesis that viral infections are involved in the etiology of colorectal cancer is of some public healthy relevance too. However, an impressive systematic review [13] of studies assessing the association between viral infections and colorectal cancer documented that very inconsistent results were observed across the studies analyzed. Overall, there is no published convincing evidence on the role of viral infections in colorectal cancer. Thus far, viral infections do not contribute to the etiology of colorectal cancer. The human intestinal microbiome encompasses at least 100 trillion (10^{14}) microorganisms and is harbored by more than 1000 species. Some of these species of microorganisms bring about beneficial some other deleterious effects on the host and the gut microbiome is increasingly recognized as having an important role in human health and disease, including colorectal cancer [14], [15], [16]. Recent studies have shown that some harmful microbiota of the huge number of microbial communities which are continuously colonized in the gut may play roles in the development of colorectal cancer [17], [18]. On the whole, accumulating [19] evidence indicates that there is none relationship between a helicobacter pylori infection and colorectal cancer. Regarding the association between the gut microbiome and immunity, a number of studies have shown that Fusobacterium species are somehow related to colorectal cancer. In point of fact, it has been found in former studies that Fusobacterium species particularly Fusobacterium nucleatum as a leading gut bacterium is enriched in colorectal cancer compared to normal tissues or controls [20]-[26]. Fusobacterium species are part of the human oral [27] and intestinal microbiota. But it is still not clear if Fusobacterium nucleatum, an anaerobic gram-negative bacterium, plays an oncogenic role in the development of colorectal cancer.

2. Material and methods

2.1. Search strategy

A systematic literature search and review according to a predefined protocol in PubMed, Google scholar and other sites was conducted to identify relevant studies published while reporting follows the PRISMA statements as much as possible [28]. A combination of different keywords like: review, bacterium, colorectal cancer, virus et cetera has been used in the search filed to search for eligible articles. In addition, the reference lists of the relevant articles including review articles was additionally used as a possible source for identifying studies related to the topic. Titles and abstracts of all identified articles were checked. Studies with potential relevance for the study topic underwent a review only if detailed data information could be extracted without any data access barriers.

2.2.1. Study of Castellarin et al. 2012 (Canada)

Castellarin et al. [29] screened a total of 99 subjects with colorectal carcinoma and matched normal tissue specimens and were able to verify an overabundance of Fusobacterium sequences in tumor versus matched normal control tissue by quantitative PCR analysis. The data as obtained Castellarin et al. are presented by the 2 by 2-table (**Table 1**).

Table 1. Fusobacterium and colorectal cancer due to Castellarin et al. (2012)

		Colorectal cancer		
		yes	**no**	**Total**
Fusobac-terium	**yes**	~78	~21	99
	no	~21	~78	99
	Total	99	99	198

The article does not provide the necessary exact information and was not considered for an statistical analysis. Still, Castellarin et al. found the mean overall abundance of Fusobacterium as being 415 times greater in the tumor samples (n = 99) than in the matched normal samples (n = 99).

2.2.2. Study of Ahn et al. 2013 (USA)

Ahn et al. [30] investigated whether an altered community of gut microbes is related with risk of colorectal cancer in a study of 94 control subjects and 47 colorectal cancer case subjects. Fusobacterium was found positive in 17/47 (36,2 %) cases and in 15/94 (16 %) controls. The data as obtained 2013 by Ahn et al. are presented by the 2 by 2-table (**Table 2**).

Table 2. Fusobacterium and colorectal cancer due to Ahn et al. (2013)

		Colorectal cancer		
		yes	**no**	**Total**
Fusobacterium	**yes**	17	15	32
	no	30	79	109
	Total	47	94	141

2.2.3. Study of Fukugaiti et al. 2015 (Brazil)

Fukugaiti et al. [31] investigated seventeen patients, 7 of whom were diagnosed with colorectal carcinoma, to evaluate the presence of Fusobacterium nucleatum and other intestinal microorganisms in the fecal microbiota of colorectal cancer patients (n=7) and healthy controls (n=10). Fecal samples were collected two days before colonoscopy while patients who had taken antibiotics or with any systemic infection were excluded from the study. Bacterial DNA from feces was obtained using a commercial Kit (QIAGEN, Hilden, Germany). Fusobacterium nucleatum was found in 7/7 (100 %) of the patients with carcinoma and in 9/10 of healthy patients. The data as obtained 2015 by Fukugaiti et al. are presented by the 2 by 2-table (**Table 3**).

Table 3. Fusobacterium and colorectal cancer due to Fukugaiti et al. (2015)

		Colorectal cancer		
		yes	**no**	**Total**
Fusobacterium	**yes**	7	9	16
	no	0	1	1
	Total	7	10	17

2.2.4. Study of Vogtmann et al. 2016 (USA)

Vogtmann et al. [32] investigated fecal samples from 52 matched controls and 52 pretreatment colorectal cancer cases from Washington, DC (USA) to evaluate the relationship between Fusobacterium and colorectal cancer. In point of fact, about 40/52 (76.9 %) of cases and 25/52 (48.1 %) of controls had detectable Fusobacteria. The data as obtained by Vogtmann et al. 2016 (USA) are presented by the 2 by 2-table (**Table 4**).

 http://vixra.org/author/ilija_barukcic

Table 4. Fusobacterium and colorectal cancer due to Vogtmann et al. (2016)

		Colorectal cancer		Total
		yes	**no**	**Total**
Fusobac-terium	**yes**	40	25	65
	no	12	27	39
	Total	52	52	104

2.2.5. Study of Li et al 2016 (China)

Li et al. [33] conducted a matched-case control study to investigate Fusobacterium nucleatum (F. nucleatum) abundance in colorectal cancer (CRC) tissues. Adjacent normal tissues 10 cm beyond cancer margins from 101 consecutive patients with resected colorectal cancer were used as matched controls. Fusobacterium nucleatum was detected in CRC and normal tissues by fluorescent quantitative polymerase chain reaction (FQ-PCR). Li et al were able to detect F. nucleatum as over-represented in 88/101 (87.1%) colorectal cancer samples compared to matched non-cancerous controls. The data as obtained 2016 by Li et al. are presented by the 2 by 2-table (**Table 5**).

Table 5. Fusobacterium and colorectal cancer due to Li et al. (2016)

		Colorectal cancer		Total
		yes	**no**	**Total**
Fusobac-terium	**yes**	88	13	101
	no	13	88	101
	Total	101	101	202

2.2.6. Study of Amitay et al 2017 (Germany)

Amitay et al. [34] collected fecal samples prior to bowel preparation from participants of screening colonoscopy in the German BliTz study. The rRNA gene analysis was used to examine the presence and relative abundance of Fusobacterium in fecal samples from 46 individuals with colorectal cancer and from 231 controls. Fusobacterium was positive in 25/46 (54.3%) of the cases and in 58/231 (25.1%) of the controls. The data as obtained 2017 by Amitay et al. are presented by the 2 by 2-table (**Table 6**).

Table 6. Fusobacterium and colorectal cancer due to Amitay et al. (2017)

		Colorectal cancer		Total
		yes	**no**	**Total**
Fusobac-terium	**yes**	25	58	83
	no	21	173	194
	Total	46	231	277

2.2.7. Study of Eklöf et al. 2017 (Sweden)

Eklöf et al. [35] conducted a nested case-control study with 65 control subjects and with 39 cancer cases to explore the relationship between Fusobacterium nucleatum and colorectal cancer. Fusobacterium nucleatum was found high in 27/39 (69.2 %) of the cancer cases compared to 15/65 (24.3 %) of the controls. The data as obtained 2017 Eklöf et al. are presented by the 2 by 2-table (**Table 7**).

Table 7. Fusobacterium and colorectal cancer due to Eklöf et al. (2017)

		Colorectal cancer		
		yes	**no**	**Total**
Fusobacterium high	**yes**	27	15	42
	no	12	50	62
	Total	39	65	104

2.2. Statistical Analysis

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

2.2.1. Bernoulli trials

Among some discrete distributions like the hypergeometric distribution, the Poisson distribution et cetera the binomial distribution is of special interest. Sometimes, the binomial distribution is called the Bernoulli distribution in honor of the Swiss mathematician Jakob Bernoulli (1654-1705), who derived the same. Bernoulli trials are an essential part of the Bernoulli distribution. Thus far, let us assume two fair coins named as $_0W_t$ and as $_RU_t$. In our model, *heads* of such a coin are considered as success T (i.e. true) and labeled as +1 while *tails* may be considered as failure F (i. e. false) and are labeled as +0. Such a coin is called a *Bernoulli-Boole coin*. The probability of success of $_RU_t$ at one single Bernoulli trial t is denoted as

$$p\left(_R U_t = +1\right) \equiv p\left(_R U_t\right) \tag{1}$$

The probability of failure of $_R\underline{U}_t$ at one single Bernoulli trial t is denoted as

$$p\left(_R U_t = +0\right) \equiv p\left(_R \underline{U}_t\right) \equiv 1 - p\left(_R U_t\right) \tag{2}$$

Furthermore, no matter how many times an experiment is repeated, let the probability of a head or the tail remain the same. The trials are independent which implies that no matter

how many times an experiment is repeated, the probability of a single event at a single trial remain the same. Repeated independent trials which are determined by the characteristic that there are always only two possible outcomes, either $+1$ or $+0$ and that the probability of an event (outcome) remain the same at each single trial for all trials are called *Bernoulli trials*. The definition of Bernoulli trials provides a theoretical model which is of further use. However, in many practical applications, we may by confronted by circumstances which may be considered as approximately satisfying Bernoulli trials. Thus far, let us perform an experiment of tossing two fair coins simultaneously. Suppose two fair coins are tossed twice. Then there are $2^2=4$ possible outcomes (the sample space), which may be shown as

$([_RU_t = +1], [_0W_t =+1])$, $([_RU_t = +1], [_0\underline{W}_t =+0])$, $([_R\underline{U}_t = +0], [_0W_t =+1])$, $([_R\underline{U}_t = +0], [_0\underline{W}_t =+0])$

This may also be shown as a 2-dimensional sample space in the form of a contingency table (**Table 8**).

Table 8. The sample space of a contingency table

		Conditioned		
		Yes = +1	No = +0	Total
Condition	**Yes =+1**	**$([_RU_t = +1], [_0W_t =+1])$**	$([_RU_t = +1], [_0\underline{W}_t =+0])$	$_RU_t$
	No = +0	$([_R\underline{U}_t = +0], [_0W_t =+1])$	$([_R\underline{U}_t = +0], [_0\underline{W}_t =+0])$	$_R\underline{U}_t$
	Total	$_0W_t$	$_0\underline{W}_t$	$_RW_t$

In the following, the contingency table is defined more precisely (**Table 9**).

Table 9. The sample space of a contingency table

		Conditioned		
		Yes = +1	No = +0	Total
Condition	**Yes =+1**	a	b	$_RU_t$
	No = +0	c	d	$_R\underline{U}_t$
	Total	$_0W_t$	$_0\underline{W}_t$	$N =_RW_t$

In general it is $(a+c) = {}_0W_t$, $(a+b) = {}_RU_t$, $(c+d) = {}_0\underline{W}_t$, $(b+d) = {}_R\underline{U}_t$ and $a+b+c+d=N={}_RW_t$. Equally, it is $_0W_t+_0\underline{W}_t = {}_RU_t + {}_R\underline{U}_t = {}_RW_t = N$. Thus far, if one fair coin is tossed *n* times, we have *n* repeated Bernoulli trials and an *n* dimensional sample space with 2^n sample points is generated. In general, when given *n* Bernoulli trials with *k* successes, the probability to obtain exactly *k* successes in *n* Bernoulli trials is given by

$$p(k) = \binom{n}{k} \times p(_R U_t = +1)^k \times (1 - p(_R U_t = +1))^{n-k} \qquad (3)$$

The random variable k is sometimes called a *binomial variable*. The probability to obtain *k* events or more (*at least k events*) in *n* trials is calculated as

$$p\left(k \geq X\right) = p\left(k = X\right) + p\left(k > X\right) = \sum_{k=X}^{k=n}\left(\binom{n}{k} \times p\left(_R U_t = +1\right)^k \times \left(1 - p\left(_R U_t = +1\right)\right)^{n-k}\right) \quad (4)$$

The probability to obtain less than *k* events in *n* Bernoulli trials is calculated as

$$p\left(k < X\right) = 1 - p\left(k \geq X\right) = 1 - \sum_{k=X}^{k=n}\left(\binom{n}{k} \times p\left(_R U_t = +1\right)^k \times \left(1 - p\left(_R U_t = +1\right)\right)^{n-k}\right) \quad (5)$$

2.2.2. Sufficient condition (conditio per quam)
The formula of the conditio per quam [36]-[50] relationship was derived as

$$p\left(\text{EBV DNA} \rightarrow \text{Hodgkin's lymphoma}\right) \equiv \frac{a + c + d}{N} \quad (6)$$

and used to proof the hypothesis: *if* presence of EBV infection (EBV DNA) *then* presence of Hodgkin's lymphoma.

2.2.3. Necessary condition (conditio sine qua non)
The formula of the conditio per quam [36]-[50] relationship was derived as

$$p\left(\text{EBV DNA} \leftarrow \text{Hodgkin's lymphoma}\right) \equiv \frac{a + b + d}{N} \quad (7)$$

and used to proof the hypothesis: *without* presence of EBV infection (EBV DNA) *no* presence of Hodgkin's lymphoma.

2.2.4. Necessary and sufficient condition
The necessary and sufficient condition relationship was defined [36]-[50] as

$$p\left(\text{EBV DNA} \longleftrightarrow \text{Hodgkin's lymphoma}\right) \equiv \frac{a + d}{N} \quad (8)$$

Scholium.
Historically, the notion sufficient condition is known since thousands of years. Many authors testified original contributions of the notion material implication only for Diodorus Cronus. Still, Philo the Logician (~ 300 BC), a member of a group of early Hellenistic philosophers (the Dialectical school), is the main forerunner of the notion material impli-

http://vixra.org/author/ilija_barukcic

cation and has made some groundbreaking contributions [51] to the basics of this relationship. As it turns out, it is very hard to think of the "conditio per quam" relationship without considering the historical background of this concept. Remarkable as it is, Philo's concept of the material implications came very close to that of modern concept material implication. In propositional logic, a conditional is generally symbolized as "**p → q**" or in spoken language "if p then q". Both q and p are statements, with q the consequent and p the antecedent. Many times, the logical relation between the consequent and the antecedent is called a **material implication**. In general, a conditional "if p then q" is false only if p is true and q is false otherwise, in the three other possible combinations, the conditional is always true. In other words, to say that p is a sufficient condition for q is to say that the presence of p guarantees the presence of q. In particular, it is impossible to have p without q. *If p is present, then q must be present too*. To show that p is not sufficient for q, we come up with cases where p is present but q is not. It is well-known that the notion of a necessary condition can be used in defining what a sufficient condition is (and vice versa). In general, p is a necessary condition for q if it is impossible to have q without p. In fact, the absence of p guarantees the absence of q. **Example (Condition: Our earth).** Without oxygen no fire. The following table (**Table 10**) may demonstrate this relationship.

Table 10. Without Oxygen no fire (on our planet earth).

		Fire		
		Yes = +1	No = +0	Total
Oxygen	Yes =+1	a	b	$_RU_t$
	No = +0	**0**	d	$_R\underline{U}_t$
	Total	$_0W_t$	$_0\underline{W}_t$	$N = _RW_t$

In contrast to such a point of view, the opposite point of view is correct too. Thus far, there is a straightforward way to give a precise and comprehensive account of the meaning of the term necessary or sufficient condition itself. In other words, **if** fire is present **then** oxygen is present too. The following table (**Table 11**) may demonstrate this relationship.

Table 11. If fire is present **then** oxygen is present too (on our planet earth).

		Oxygen		
		Yes = +1	No = +0	Total
Fire	Yes =+1	a	**0**	$_RU_t$
	No = +0	c	d	$_R\underline{U}_t$
	Total	$_0W_t$	$_0\underline{W}_t$	$N = _RW_t$

Especially, necessary and sufficient conditions are converses of each other. Still, *the fire is not the cause of oxygen* and vice versa. *Oxygen is note the cause of fire*. In this example before, oxygen is a necessary condition, *a conditio sine qua non*, of fire. A necessary con-

dition is sometimes also called "an essential condition" or a conditio sine qua non. In propositional logic, a necessary condition, a condition sine qua non, is generally symbolized as "**p ← q**" or in spoken language "**without** p **no** q". Both q and p are statements, with p the antecedent and q the consequent. To show that p is not a necessary condition for q, it is necessary to find an event or circumstances where q is present (i. e. an illness) but p (i. e. a risk factor) is not. On any view, (classical) logic has as one of its goals to characterize the most basic, the most simple and the most general laws of objective reality. Especially, in classical logic, the notions of necessary conditions, of sufficient conditions of necessary and sufficient conditions et cetera are defined very precisely for a single event, for a single Bernoulli trial t. In point of fact, no matter how many times an experiment is repeated, the relationship of the conditio sine qua or of the conditio per quam which is defined for every single event will remain the same. Under conditions of independent trials this implies that no matter how many times an experiment is repeated, the probability of the conditio sine qua or of the conditio per quam of a single event at a single trial *t* remain the same which transfers the relationship of the conditio sine qua or of the conditio per quam et cetera into the sphere of (Bio-) statistics. Consequently, (Bio) statistics generalizes the notions of a sufficient or of a necessary condition from one single Bernoulli trial to N Bernoulli trials. However, in many practical applications, we may by confronted by circumstances which may be considered as approximately satisfying the notions of a sufficient or of a necessary condition. Thus far, under these circumstances, we will need to perform some tests to investigate, can we rely on our investigation.

2.2.5. The central limit theorem

Many times, for some reason or other it is not possible to study exhaustively a whole population. Still, sometimes it is possible to draw a sample from such a population which itself can be studied in detail and used to convince us about the properties of the population. Roughly speaking, statistical inference derived from a randomly selected subset of a population (a sample) can lead to erroneous results. The question raised is how to deal with the uncertainty inherent in such results?   The concept of confidence intervals, closely related to statistical significance testing, was formulated to provide an answer to this problem. Confidence intervals, introduced to statistics by Jerzy Neyman in a paper published in 1937 [52], specifies a range within a parameter, i. e. the population proportion π, with a certain probability, contain the desired parameter value. Most commonly, the 95% *confidence interval* is used.   Interpreting a confidence interval involves a couple of important but subtle issues. In general, a 95% confidence interval for the value of a random number means that there is a 95% probability that the "true" value of the value of a random number is within the interval. Confidence intervals for proportions or a population mean of random variables which are not normally distributed in the population can be constructed while relying on the central limit theorem as long as the sample sizes and counts are big enough (i. e. a sample size of *n=30* and more). A formula, justified by the central limit theorem, is known as

         http://vixra.org/author/ilija_barukcic

$$p_{Crit} = p_{Calc} \pm \left(z_{Alpha/2} \times \left(\sqrt[2]{\frac{1}{N} \times p_{Calc} \times \left(1 - p_{Calc}\right)} \right) \right) \qquad (9)$$

where p$_{Calc}$ is the sample proportion of successes in a Bernoulli trial process with N trials yielding X successes and N-X failures and z is i.e the 1 – (Alpha/2) quantile of a standard normal distribution corresponding to the significance level alpha. For example, for a 95% confidence level alpha = 0.05 and z is z = 1.96. A very common technique for calculating binomial confidence intervals was published by Clopper-Pearson [53]. Agresti-Coull proposed another different method [54] for calculating binomial confidence intervals. A faster and an alternative way to determine the lower and upper "exact" confidence interval is justified by the F distribution [55].
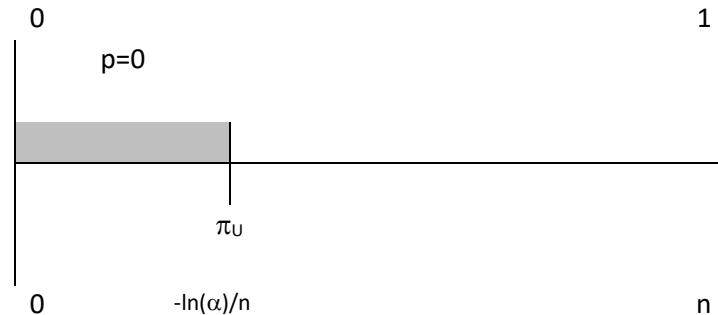
2.2.6. The rule of three

Furthermore, an approximate and conservative (one sided) confidence interval was developed by Louis [56], Hanley et al. [57] and Jovanovic [58] known as the rule of three. Briefly sketched, the rule of three can be derived from the binomial model. Let π_U denote the upper limit of the one-sided $100 \times (1 - \alpha)$% confidence interval for the unknown proportion when in N independent trials **no events occur** [58]. Then π_U is the value such that

$$\pi_U = \left(\frac{-\ln(\alpha)}{n} \right) \approx \left(\frac{3}{n} \right) \qquad (10)$$

assuming that α =0,05. In other words, an one-sided approximate upper 95% confidence bound for the true binomial population proportion π, the rate of occurrences in the population, based on a sample of size n where no successes are observed (p=0) is 3/n [58] or given approximately by $[0 < \pi < (3/n)]$. The rule of three is a useful tool especially in the analysis of medical studies. The following table (**Table 12**) will illustrate this relationship.

Table 12. The one-sided approximate upper $100 \times (1 - \alpha)$% confidence bound where **no successes** (p=0) are observed.
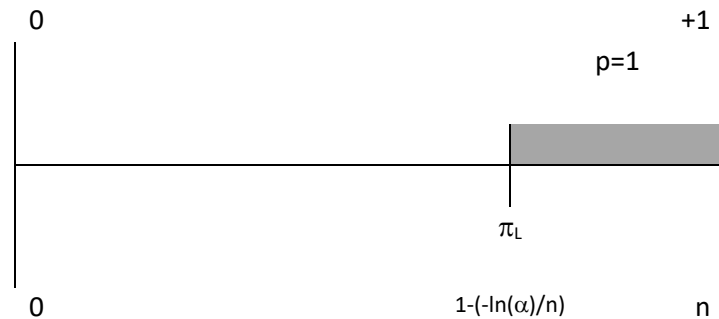
Under conditions where **a certain event did not occur** [56] in a sample with n subjects (i. e. **p=0**) the interval from 0 to $(-\ln(\alpha)/n)$ is called a $100 \times (1 - \alpha)\%$ confidence interval for the binomial parameter for the rate of occurrences in the population.

Another special case of the binomial distribution is based on a sample of size n where **only successes are observed (p=1)**. Accordingly, the lower limit of a one-sided $100 \times (1- \alpha)\%$ confidence interval for a binomial probability π_L, the rate of occurrences in the population, based on a sample of size n where only successes are observed is given approximately by $[(1-(-\ln(\alpha)/n)) < \pi < +1]$ or (assuming $\alpha=0{,}05$)

$$\pi_L = 1-\left(\frac{-\ln(\alpha)}{n}\right) \approx 1-\left(\frac{3}{n}\right) \tag{11}$$

The following table (**Table 13**) may illustrate this relationship.

Table 13. The one-sided approximate upper $100 \times (1 - \alpha)\%$ confidence bound where **only successes** are observed.

To construct a two-sided $100 \times (1 - (\alpha))\%$ interval according to the rule of three, it is necessary to take a one-sided $100 \times (1 - (\alpha/2))\%$ confidence interval. In this study, we will use the rule of three [59] too, to calculate the confidence interval for the value of a random number.

2.2.7. Fisher's exact test

A test statistics of independent and more or less normally distributed data which follow a chi-squared distribution is valid as with many statistical tests due to the central limit theorem. Especially, with large samples, a chi-squared distribution can be used. A sample is considered as large when the sample size n is n = 30 or more. With a small sample (n < 30), the central limit theorem does not apply and erroneous results could potentially be obtained from the few observations if the same is applied. Thus far, when the number of observations obtained from a population is too small, a more appropriate test for of analysis of categorical data i. e. contingency tables is R. A. Fisher's exact test [60]. Fisher's exact

http://vixra.org/author/ilija_barukcic

test is valid for all sample sizes and calculates the significance of the p-value (i. e. the deviation from a null hypothesis) exactly even if in practice it is employed when sample size is small. Fisher's exact test is called exact because the same uses the exact hypergeometric distribution to compute the p-value rather than the approximate chi-square distribution. Still, computations involved in Fisher's exact test can be time consuming to calculate by hand.

2.2.8. Hypergeometric distribution

The hypergeometric distribution, illustrated in a table (**Table 14**), is a discrete probability distribution which describes the probability of a events/successes in a sample with the size $_0W_t$, without replacement, from a finite population of the size N which contains exactly $_RU_t$ objects with a certain feature while each event is either a success or a failure. The formula for the hypergeometric distribution, a discrete probability distribution, is

$$p(a) = \frac{\binom{_RU_t}{a} \times \binom{N - {}_RU_t}{{}_0W_t - a}}{\binom{N}{{}_0W_t}} \tag{12}$$

Table 14. The hypergeometric distribution

		Conditioned		
		Yes = +1	No = +0	Total
Condition	Yes =+1	**a**	b=($_RU_t$ –a)	$_RU_t$
	No = +0	c=($_0W_t$ –a)	N-$_RU_t$ -$_0W_t$ +a	N-$_RU_t$
	Total	$_0W_t$	N - $_0W_t$	N

The hypergeometric distribution has a wide range of applications. The Hypergeometric distribution can be approximated by a Binomial distribution. The elements of the population being sampled are classified into one of two mutually exclusive categories: **either** conditio sine qua non **or** no conditio sine qua non relationship. We are sampling without replacement from a finite population. How probable is it to draw specific c events/successes out of $_0W_t$ total draws from an aforementioned population of the size N? The hypergeometric distribution, as shown in a table (**Table 15**) is of use to calculate how probable is it to obtain c=($_0W_t$–a) events out of N events.

Table 15. The hypergeometric distribution and conditio sine qua non

		Conditioned		
		Yes = +1	No = +0	Total
No Condition	Yes =+1	c=($_0W_t$ –a)	N-$_RU_t$ -$_0W_t$ +a	N-$_RU_t$
	No = +0	a	b=($_RU_t$ –a)	$_RU_t$
	Total	$_0W_t$	N - $_0W_t$	N

http://vixra.org/author/ilija_barukcic

2.2.9. Statistical hypothesis testing

A statistical hypothesis test is a method to extract some inferences from data. A hypothesis is compared as an alternative hypothesis. Under which conditions does the outcomes of a study lead to a rejection of the null hypothesis for a pre-specified level of significance. According to the rules of a proof by contradiction, a null hypothesis (H_0) is a statement which one seeks to disproof. The related specific alternative hypothesis (H_A) is opposed to the null hypothesis such that if   null hypothesis (H_0) is true, the alternative hypothesis (H_A) is false and vice versa. If the alternative hypothesis (HA) is true then the null hypothesis (H0) is false. In principle, a null hypothesis that is true can be rejected (type I error) which lead us to falsely infer the existence of something which is not given. The significance level, also denoted as α (alpha) is the probability of rejecting a null hypothesis when the same is true. A type II error is given, if we falsely infer the absence of something which in reality is given.   A null hypothesis can be false but a statistical test may fail to reject such a false null hypothesis. The probability of accepting a null hypothesis when the same is false (type II error), is denoted by the Greek letter ß (beta) and related to the power of a test (which equals 1-ß). The power of a test indicates *the probability by which the test correctly rejects the null hypothesis (H_0) when a specific alternative hypothesis (H_A) is true*. Most investigator assess the power of a tests using **1-ß** = 0.80 as a standard for adequacy. A tabularized relation between truth/falseness of the null hypothesis and outcomes of the test are shown precisely within a table (**Table 16**).
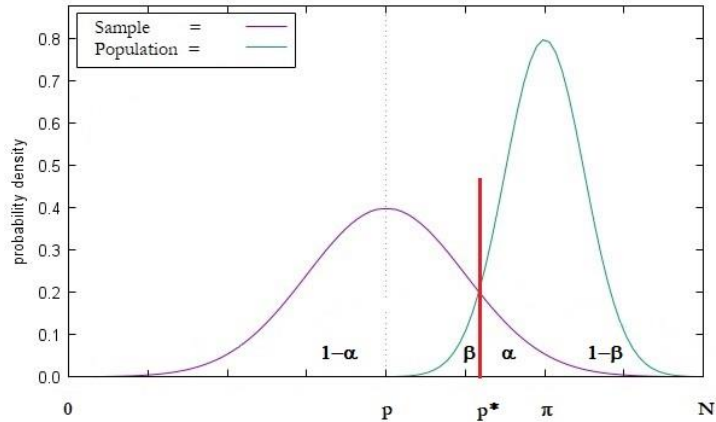
Table 16. Table of error types

Null Hypothesis (H_0)		Null Hypothesis (H_0) is		
		True	**False**	**Total**
	Accepted	1-α	β	1-α+ß
	Rejected	α	**1-ß**	1+α-ß
	Total	1	1	**2**

In general, it is 1- α + α = 1 or (1- α -β) + α = **1- β**. The following figure may illustrate these relationships (**Figure 1**). The relationships can be normalized in the following way which is shown schematically in a table (**Table 17**).

Table 17. Table of error types

Null Hypothesis (H_0)		Null Hypothesis (H_0) is		
		True	**False**	**Total**
	Accepted	(1-α)/2	β/2	(1-α+ß)/2
	Rejected	α/2	**(1-ß)/2**	(1+α-ß)/2
	Total	1/2	1/2	**(2/2)=1**

Figure 1 .   The relationship between error types.

2.2.10. The mathematical formula of the causal relationship k

The mathematical formula of the causal relationship k [36]-[50] defined as

$$k\left(_R U_t, _0 W_t\right) \equiv \frac{\left((N \times a) - \left(_R U_t \times _0 W_t\right)\right)}{\sqrt[2]{\left(_R U_t \times _R \underline{U}_t\right) \times \left(_0 W \times _0 \underline{W}_t\right)}} \tag{13}$$

and the chi-square distribution [61] were applied to determine the significance of causal relationship between a EBV and HL. A one-tailed test makes it much more easier to reject a null hypothesis (no causal relationship) while a two-tailed test makes it more difficult to reject a null hypothesis and is more conservative on this account. For this reason, in causal relationship testing, a two-tailed test is preferred. In general, a p value of less than 0.05 is considered as significant. In this context, what is the necessary connection between a cause and effect? What ties a cause and its own effect together? Is there a necessary connection between a cause and effect at all? Theoretically, it is neither justified nor necessary to reduce causation as such to an act of observation or measurement. Sill, case-control studies, experiments, observations et cetera can help us to recognize cause effect relationships. In this context it is necessary to stress out that **every single event (effect) has its own cause**, which is the logical foundation of the mathematical formula of the causal relationship k. It is therefore entirely clear that this is the fundamental difference to Pearson's methodological approach. Obviously, although under some certain specified circumstances Pearson's product-moment correlation coefficient [62] or Pearson's Phi [63] coefficient can yield the same numerical result as the mathematical formula of the causal relationship k, there is nothing truly exciting about such a coincidence. Nevertheless, when conducting experiments and analyzing data, views in which correlation and causation are brought very close together are incorrect and worthless. The mathematical formula of the causal relationship k is neither identical nor can the same mathematical formula be reduced to Pearson's product-moment correlation coefficient [62] or to Pearson's Phi [63] Coefficient (Mean Square

Contingency Coefficient). In contrast to Pearson's product-moment correlation coefficient and to Pearson's Phi Coefficient (Mean Square Contingency Coefficient) the mathematical formula of the causal relationship k is defined and valid at every single Bernoulli trial *t* or at every single event. Sir Austin Bradford Hill (1897 - 1991), an English epidemiologist, proposed 1965 a set of nine criteria (Strength, Consistency, Specificity, Temporality, Biological gradient, Plausibility, Coherence, Experiment, Analogy) [64] to establish epidemiologic evidence of a causal relationship (Bradford Hill criteria). In point of fact, Bredford's "fourth characteristic is the temporal relationship of the association" [64] and in last consequence the "*post hoc ergo propter hoc*" logical fallacy. Causation cannot be derived from the "*post hoc ergo propter hoc*" [49] logical fallacy. Consequently, the Mathematical Formula of the causal relationship k can neither be reduced to the Bradford Hill criteria nor is the same just a mathematization of Bradford Hill criteria.

2.2.11. The chi square distribution

The chi-squared distribution [61] is a widely known distribution and used in hypothesis testing, in inferential statistics or in construction of confidence intervals. The critical values of the chi square distribution are visualized by **Table 18**.

Table 18. The critical values of the chi square distribution (degrees of freedom: 1).

	p-Value	One sided X^2	Two sided X^2
	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

2.2.12. The X² goodness of fit test

A chi-square goodness of fit test can be applied to determine whether sample data are consistent with a hypothesized distribution. *The chi-square goodness of fit test is appropriate when some conditions are met. A view of these conditions are simple random sampling, categorical variables and an expected value of the number of sample observations which is at least 5*. If the expectation value is less than 5, then *the rule of three* is of use too. The null hypothesis (H_0) and its own alternative hypothesis (H_A) are stated in such a way that they are mutually exclusive. In point of fact, if the null hypothesis (H_0) is true, the other, alternative hypothesis (H_A), must be false; and vice versa. For a chi-square goodness of fit test, the hypotheses can take the following form.

H_0: The <u>sample distribution</u> agrees with the hypothetical (<u>theoretical</u>) <u>distribution</u>.
H_A: The sample distribution does not agree with the hypothetical (theoretical) distribution.

The X² Goodness-of-Fit Test can be shown schematically as

$$\chi^2 \equiv \sum_{t=+1}^{t=+N} \left(\frac{\left(\text{Observed}_t - \text{Expected}_t \right)^2}{\text{Expected}_t} \right) \qquad (14)$$

The degrees of freedom are calculated as N-1. If there is no discrepancy between an observed and a theoretical distribution, then X²=0. As the discrepancy between an observed and a theoretical distribution becomes larger, the X² becomes larger. This X² values are evaluated by the known X² distribution.

The original X² values are calculated from an original theoretical distribution, which is continuous, whereas the approximation by the X² Goodness of fit test we are using is discrete. Thus far, there is a tendency to underestimate the probability, which means that the number of rejections of the null hypothesis can increase too much and must be corrected downward. Such an adjustment (*Yate's correction for continuity*) is used only when there is one degree of freedom. When there is more than one degree of freedom, the same adjustment is not used. Applying this to the formula above, we find the X² Goodness-of-Fit Test *with continuity correction* shown schematically as

$$\chi^2 \equiv \sum_{t=+1}^{t=+N} \left(\frac{\left(\left| \text{Observed}_t - \text{Expected}_t \right| - \left(\frac{1}{2} \right) \right)^2}{\text{Expected}_t} \right) \qquad (15)$$

When the term ($|\text{Observed}_t - \text{Expected}_t|$) is less than ½, the continuity correction should be omitted.

2.2.12.1. The X² goodness of fit test of a sufficient condition

The theoretical (hypothetical) distribution of a sufficient condition [36]-[50] is shown schematically by the 2x2 table (**Table 19**).

Table 19. The theoretical distribution of a sufficient condition (conditio pre quam).

		Conditioned		
		Yes = +1	No = +0	Total
Condition	Yes =+1	a	**b=0**	(a+b)
	No = +0	c	d	(c+d)
	Total	(a+c)	(b+d)	(a+b+c+d)

The theoretical distribution of a sufficient condition (conditio pre quam) is determined by the fact that **b=0**. The X² Goodness-of-Fit Test *with continuity correction* of a sufficient condition (conditio per quam) is calculated as

$$\chi^2(\text{IMP}) \equiv \left(\frac{\left(|a-(a+b)| - \left(\frac{1}{2}\right) \right)^2}{(a+b)} \right) + \left(\frac{\left(||(c+d)-(c+d)| - \left(\frac{1}{2}\right) \right)^2}{(c+d)} \right) = \left(\frac{\left(|a-(a+b)| - \left(\frac{1}{2}\right) \right)^2}{(a+b)} \right) + 0 \quad (16)$$

or more simplified as

$$\chi^2(\text{IMP}) \equiv \left(\frac{\left(|-b| - \left(\frac{1}{2}\right) \right)^2}{(a+b)} \right) + 0 \quad (17)$$

Under these circumstances, the degree of freedom is d.f. = N-1=2-1=1.

2.2.12.2. The X² goodness of fit test of a necessary condition

The theoretical (hypothetical) distribution of a necessary condition is shown schematically by the 2x2 table (**Table 20**).

Table 20. The theoretical distribution of a necessary condition (conditio sine qua non).

		Conditioned		
		Yes = +1	No = +0	Total
Condition	Yes =+1	a	b	(a+b)
	No = +0	**c=0**	d	(c+d)
	Total	(a+c)	(b+d)	(a+b+c+d)

The theoretical distribution of a necessary condition (conditio sine qua non) is determined by the fact that **c=0**. The X² Goodness-of-Fit Test *with continuity correction* of a necessary condition (conditio sine qua non) is calculated as

$$\chi^2(\text{SINE}) \equiv \left(\frac{\left(||(a+b)-(a+b)| - \left(\frac{1}{2}\right) \right)^2}{(a+b)} \right) + \left(\frac{\left(||(d)-(c+d)| - \left(\frac{1}{2}\right) \right)^2}{(c+d)} \right) = 0 + \left(\frac{\left(|d-(c+d)| - \left(\frac{1}{2}\right) \right)^2}{(c+d)} \right) \quad (18)$$

 http://vixra.org/author/ilija_barukcic

or more simplified as

$$\chi^2(SINE) \equiv \left(\frac{\left(|-c| - \left(\frac{1}{2} \right) \right)^2}{(c+d)} \right) + 0 \tag{19}$$

Under these circumstances, the degree of freedom is d.f. = N-1=2-1=1. For example, without water and oxygen, there would be no human life on this planet; hence water and oxygen are necessary conditions for the existence of human beings on this planet.

2.2.12.3. The X² goodness of fit test of a necessary and sufficient condition

Like other fundamental concepts, the concepts of necessary and sufficient conditions is not specified uniquely. It is well-known that the notion of sufficient condition is of use to define what a necessary condition is (and vice versa) but a generally accepted and straightforward concept to give a precise and comprehensive account of the meaning of the term "necessary (or sufficient) condition" itself has not met with success. Thus far, what then is a necessary and a sufficient condition? Central to this problem is the question under which circumstances can certain phenomena truly be said to be necessary and sufficient conditions.

Especially J. L. Mackie used the terminology of necessary and sufficient conditions to define a cause of some particular event as an INUS condition which is an "insufficient, but necessary part of an unnecessary but sufficient condition" [65] of an effect.

The theoretical (hypothetical) distribution of a necessary and sufficient condition is shown schematically by the 2x2 table (**Table 21**).

Table 21. The theoretical distribution of a necessary and sufficient condition.

		Conditioned		
		Yes = +1	No = +0	Total
Condition	Yes =+1	a	**b=0**	(a+b)
	No = +0	**c=0**	d	(c+d)
	Total	(a+c)	(b+d)	(a+b+c+d)

The theoretical distribution of a necessary and sufficient condition is determined by the fact that **b=0** and that **c=0**. The X² Goodness-of-Fit Test *with continuity correction* of a necessary and sufficient condition is calculated as

$$\chi^2(\text{Necessary AND Sufficient}) \equiv \left(\frac{\left(|(a)-(a+b)| - \left(\frac{1}{2} \right) \right)^2}{(a+b)} \right) + \left(\frac{\left(|(d)-(c+d)| - \left(\frac{1}{2} \right) \right)^2}{(c+d)} \right) \tag{20}$$

or more simplified as

$$\chi^2(\text{Necessary AND Sufficient}) \equiv \left(\frac{\left(|-b| - \left(\frac{1}{2} \right) \right)^2}{(a+b)} \right) + \left(\frac{\left(|-c| - \left(\frac{1}{2} \right) \right)^2}{(c+d)} \right) \tag{21}$$

Under these circumstances, the degree of freedom is d.f. = N-1=2-1=1.

2.2.12.4. The X² goodness of fit test of either a necessary or a sufficient condition

The theoretical (hypothetical) distribution of a neither necessary nor a sufficient condition is shown schematically by the 2x2 table (**Table 22**).

Table 22. The theoretical distribution of **either** a necessary dondition **or** a sufficient condition.

		Conditioned		Total
		Yes = +1	No = +0	
Condition	Yes =+1	a=0	b	(a+b)
	No = +0	c	d=0	(c+d)
	Total	(a+c)	(b+d)	(a+b+c+d)

The theoretical distribution of **either** a necessary condition **or** a sufficient condition is determined by the fact that **a=0** and that **d=0**. The X² Goodness-of-Fit Test *with continuity correction* of either a necessary condition or a sufficient condition is calculated as

$$\chi^2\left(\text{Either Necessary or Sufficient}\right) \equiv \left(\frac{\left(\left|(b)-(a+b)\right|-\left(\frac{1}{2}\right)\right)^2}{(a+b)}\right) + \left(\frac{\left(\left|(c)-(c+d)\right|-\left(\frac{1}{2}\right)\right)^2}{(c+d)}\right) \quad (22)$$

or more simplified as

$$\chi^2\left(\text{Either Necessary or Sufficient}\right) \equiv \left(\frac{\left(\left|-a\right|-\left(\frac{1}{2}\right)\right)^2}{(a+b)}\right) + \left(\frac{\left(\left|-d\right|-\left(\frac{1}{2}\right)\right)^2}{(c+d)}\right) \quad (23)$$

Under these circumstances, the degree of freedom is d.f. = N-1=2-1=1.

2.2.12.5. The X² goodness of fit test of exclusion

The theoretical (hypothetical) distribution of exclusion is shown schematically by the 2x2 table (**Table 23**).

Table 23. The theoretical distribution of coindicence.

		Conditioned		Total
		Yes = +1	No = +0	
Condition	Yes =+1	a=0	b	(a+b)
	No = +0	c	d	(c+d)
	Total	(a+c)	(b+d)	(a+b+c+d)

The theoretical distribution of exclusion is determined by the fact that **a=0**. The X² Goodness-of-Fit Test *with continuity correction* of exclusion is calculated as

$$\chi^2\left(\text{Exclusion}\right) \equiv \left(\frac{\left(\left|b-(a+b)\right|-\left(\frac{1}{2}\right)\right)^2}{(a+b)}\right) + \left(\frac{\left(\left|(c+d)-(c+d)\right|-\left(\frac{1}{2}\right)\right)^2}{(c+d)}\right) = \left(\frac{\left(\left|b-(a+b)\right|-\left(\frac{1}{2}\right)\right)^2}{(a+b)}\right) + 0 \quad (24)$$

         http://vixra.org/author/ilija_barukcic

or more simplified as

$$\chi^2\left(\text{Exclusion}\right) \equiv \left(\frac{\left(\left| -a \right| - \left(\frac{1}{2} \right) \right)^2}{\left(a + b \right)} \right) + 0 \tag{25}$$

Under these circumstances, the degree of freedom is d.f. = N-1=2-1=1.

2.2.12.6. The X² goodness of fit test of disjunction

The theoretical (hypothetical) distribution of disjunction is shown schematically by the 2x2 table (**Table 24**).

Table 24. The theoretical distribution of disjunction.

		Conditioned		
		Yes = +1	No = +0	Total
Condition	**Yes =+1**	a	b	(a+b)
	No = +0	c	**d=0**	(c+d)
	Total	(a+c)	(b+d)	(a+b+c+d)

The theoretical distribution of disjunction is determined by the fact that **d =0**. The X² Goodness-of-Fit Test *with continuity correction* of disjunction can be calculated as

$$\chi^2\left(\text{Disjunction}\right) \equiv \left(\frac{\left(\left| (a+b) - (a+b) \right| - \left(\frac{1}{2} \right) \right)^2}{\left(a + b \right)} \right) + \left(\frac{\left(\left| (c) - (c+d) \right| - \left(\frac{1}{2} \right) \right)^2}{\left(c + d \right)} \right) = 0 + \left(\frac{\left(\left| c - (c+d) \right| - \left(\frac{1}{2} \right) \right)^2}{\left(c + d \right)} \right) + 0 \tag{26}$$

or more simplified as

$$\chi^2\left(\text{Disjuntion}\right) \equiv 0 + \left(\frac{\left(\left| -d \right| - \left(\frac{1}{2} \right) \right)^2}{\left(c + d \right)} \right) \tag{27}$$

Under these circumstances, the degree of freedom is d.f. = N-1=2-1=1.

2.2.12.7. The X² goodness of fit test of a neither nor relationship

The theoretical (hypothetical) distribution of a neither nor relationship is shown schematically by the 2x2 table (**Table 25**).

Table 25. The theoretical distribution of a neither nor relationship.

		Conditioned		
		Yes = +1	**No = +0**	**Total**
Condition	**Yes =+1**	a=0	b=0	(a+b)
	No = +0	c=0	d	(c+d)
	Total	(a+c)	(b+d)	(a+b+c+d)

The theoretical distribution of a neither nor relationship is determined by the fact that **a=0 and b=0 and c=0**. The X² Goodness-of-Fit Test *with continuity correction* of a neither nor relationship is calculated as

$$\chi^2(\text{IMP}) \equiv \left(\frac{\left(\left| d-(c+d) \right| - \left(\frac{1}{2} \right) \right)^2}{(c+d)} \right) + \left(\frac{\left(\left| d-(a+b+c+d) \right| - \left(\frac{1}{2} \right) \right)^2}{(a+b+c+d)} \right) = \left(\frac{\left(\left| -c \right| - \left(\frac{1}{2} \right) \right)^2}{(c+d)} \right) + \left(\frac{\left(\left| -(a+b+c) \right| - \left(\frac{1}{2} \right) \right)^2}{(a+b+c+d)} \right) \quad (28)$$

Under these circumstances, the degree of freedom is d.f. = N-1=2-1=1.

3. Results

3.1. *Without* an infection by Fusobacterium *no* colorectal cancer

Claims.
Null hypothesis:
An infection by Fusobacterium <u>is</u> a conditio sine qua non of colorectal cancer.
Alternative hypothesis:
An infection by Fusobacterium <u>is not</u> a conditio sine qua non of colorectal cancer.
Significance level (Alpha) below which the null hypothesis will be rejected: 0.05.

Proof.
The data of Fukugaiti et al. [31] of an infection by Fusobacterium and colorectal cancer are viewed in the 2×2 table (**Table 3**). The proportion of successes in the sample of a conditio sine qua non relationship p(Fusobacterium $\leftarrow$ Colorectal cancer) is calculated [36]-[50] as

$$p\left(\text{Fusobacterium} \leftarrow \text{Colorectal cancer}\right) = \frac{(7+9+1)}{17} = \frac{17}{17} = 1$$

The chi-square goodness of fit test can be used to test the significance of this result if some conditions are met. A view of these conditions are simple random sampling, categorical variables and an expected value of the number of sample observations which is at least 5. In the study of Fukugaiti et al. [31] (**Table 3**), one expectation value is less than 5. Therefore, we use the rule of three to prove the significance of the result above. The critical value p_{Crit} (significance level alpha = 0.05) calculated [36]-[50] according to the rule of three is

$$p_{\text{Crit}} = 1 - \frac{3}{17} = 0{,}823780454$$

The critical value is p_{Crit} = 0,823780454 and is less than the proportion of successes calculated as p(Fusobacterium $\leftarrow$ Colorectal cancer) =1. Due to this evidence, we do not reject the null hypothesis in favor of the alternative hypotheses. The data as published by Fukugaiti et al. [31] do support our Null hypothesis that an infection by Fusobacterium is a conditio sine qua non of colorectal cancer. In other words, *without* an infection by Fusobacterium *no* colorectal cancer.

Q. e. d.

         http://vixra.org/author/ilija_barukcic

3.2. *Without* an infection by Fusobacterium *no* colorectal cancer

Claims.
Null hypothesis:
An infection by Fusobacterium <u>is</u> a conditio sine qua non of colorectal cancer.
Alternative hypothesis:
An infection by Fusobacterium <u>is not</u> a conditio sine qua non of colorectal cancer.
Significance level (Alpha) below which the null hypothesis will be rejected: 0.05.

Proof.
The data of Vogtmann et al. [32] of an infection by Fusobacterium and colorectal cancer are viewed in the 2×2 table (**Table 4**). The X² Goodness-of-Fit Test *with continuity correction* of a necessary condition (conditio sine qua non) known to be defined as p(Fusobacterium ← Colorectal cancer) is calculated [36]-[50] as

$$\chi^2\left(\text{SINE}\right) \equiv \left(\frac{\left(|-c| - \left(\frac{1}{2}\right)\right)^2}{(c+d)} \right) + 0 = \left(\frac{\left(|-12| - \left(\frac{1}{2}\right)\right)^2}{(12+27)} \right) = 3{,}391025641$$

Under these circumstances, the degree of freedom is d. f. = N-1=2-1=1. The two sided critical X² (significance level alpha = 0.05) is known to be 3,841458821 (**Table 18**). The calculated X² value = 3,391025641and less than the critical X² = 3,841458821. Hence, our calculated X² value = 3,391025641 is not significant and we accept our null hypothesis. Due to this evidence, we do not reject the null hypothesis in favor of the alternative hypotheses. In other words, the sample distribution agrees with the hypothetical (theoretical) distribution. Our hypothetical distribution was the distribution of the necessary condition. Thus far, the data as published by Vogtmann et al. [32] do support our null hypothesis that an infection by Fusobacterium is a conditio sine qua non of colorectal carcinoma. In other words, *without* an infection by Fusobacterium *no* colorectal carcinoma.

Q. e. d.

3.3. *Without* an infection by Fusobacterium *no* colorectal cancer

Claims.

Null hypothesis:

An infection by Fusobacterium <u>is</u> a conditio sine qua non of colorectal cancer.

Alternative hypothesis:

An infection by Fusobacterium <u>is not</u> a conditio sine qua non of colorectal cancer.

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

Proof.

The data of Li et al. [33] of an infection by Fusobacterium and colorectal cancer are viewed in the 2×2 table (**Table 5**). The X² Goodness-of-Fit Test *with continuity correction* of a necessary condition (conditio sine qua non) known to be defined as p(Fusobacterium ← Colorectal cancer) is calculated [36]-[50] as

$$\chi^2 (SINE) \equiv \left(\frac{\left(|-c| - \left(\frac{1}{2} \right) \right)^2}{(c+d)} \right) + 0 = \left(\frac{\left(|-13| - \left(\frac{1}{2} \right) \right)^2}{(13+88)} \right) = 1,547029703$$

Under these circumstances, the degree of freedom is d. f. = N-1=2-1=1. The one sided critical X² (significance level alpha = 0.05) is known to be 2,705543454 (**Table 18**). The calculated X² value = 1,547029703 and less than the critical X² = 2,705543454. Hence, our calculated X² value = 1,547029703 is not significant and we accept our null hypothesis. Due to this evidence, we do not reject the null hypothesis in favor of the alternative hypotheses. In other words, the sample distribution agrees with the hypothetical (theoretical) distribution. Our hypothetical distribution was the distribution of the necessary condition. Thus far, the data as published by Li et al. [33] do support our null hypothesis that an infection by Fusobacterium is a conditio sine qua non of colorectal carcinoma. In other words, *without* an infection by Fusobacterium *no* colorectal carcinoma.

Q. e. d.

3.4. *Without* an infection by Fusobacterium *no* colorectal cancer

Claims.

Null hypothesis:

An infection by Fusobacterium <u>is</u> a conditio sine qua non of colorectal cancer.

Alternative hypothesis:

An infection by Fusobacterium <u>is not</u> a conditio sine qua non of colorectal cancer.

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

Proof.

The data of Amitay et al. [34] of an infection by Fusobacterium and colorectal cancer are viewed in the 2 × 2 table (**Table 6**). The X² Goodness-of-Fit Test *with continuity correction* of a necessary condition (conditio sine qua non) known to be defined as p(Fusobacterium ← Colorectal cancer) is calculated [36]-[50] as

$$\chi^2\left(SINE\right) \equiv \left(\frac{\left(|-c|-\left(\frac{1}{2}\right)\right)^2}{\left(c+d\right)}\right) + 0 = \left(\frac{\left(|-21|-\left(\frac{1}{2}\right)\right)^2}{\left(21+173\right)}\right) + 0 = 2{,}166237113$$

Under these circumstances, the degree of freedom is d. f. = N-1=2-1=1. The one sided critical X² (significance level alpha = 0.05) is known to be 2,705543454 (**Table 18**). The calculated X² value = 2,166237113 and less than the critical X² = 2,705543454. Hence, our calculated X² value = 2,166237113 is not significant and we accept our null hypothesis.

Due to this evidence, we do not reject the null hypothesis in favor of the alternative hypotheses. In other words, the sample distribution agrees with the hypothetical (theoretical) distribution. Our hypothetical distribution was the distribution of the necessary condition. Thus far, the data as published by Amitay et al. [34] do support our null hypothesis that an infection by Fusobacterium is a conditio sine qua non of colorectal carcinoma. In other words, *without* an infection by Fusobacterium *no* colorectal carcinoma.

Q. e. d.

3.5. *Without* an infection by Fusobacterium *no* colorectal cancer

Claims.

Null hypothesis:

An infection by Fusobacterium is a conditio sine qua non of colorectal cancer.

Alternative hypothesis:

An infection by Fusobacterium is not a conditio sine qua non of colorectal cancer.

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

Proof.

The data of Eklöf et al. [35] of an infection by Fusobacterium and colorectal cancer are viewed in the 2×2 table (**Table 7**). The X² Goodness-of-Fit Test *with continuity correction* of a necessary condition (conditio sine qua non) known to be defined as p(Fusobacterium ← Colorectal cancer) is calculated [36]-[50] as

$$\chi^2\left(\text{SINE}\right) \equiv \left(\frac{\left(\left|-c\right|-\left(\frac{1}{2}\right)\right)^2}{\left(c+d\right)}\right)+0 = \left(\frac{\left(\left|-12\right|-\left(\frac{1}{2}\right)\right)^2}{\left(12+50\right)}\right)+0 = 2{,}166237113$$

Under these circumstances, the degree of freedom is d. f. = N-1=2-1=1. The one sided critical X² (significance level alpha = 0.05) is known to be 2,705543454 (**Table 18**). The calculated X² value = 2,133064516and less than the critical X² = 2,705543454. Hence, our calculated X² value = 2,133064516 is not significant and we accept our null hypothesis.

Due to this evidence, we do not reject the null hypothesis in favor of the alternative hypotheses. In other words, the sample distribution agrees with the hypothetical (theoretical) distribution. Our hypothetical distribution was the distribution of the necessary condition. Thus far, the data as published by Eklöf et al. [35] do support our null hypothesis that an infection by Fusobacterium is a conditio sine qua non of colorectal carcinoma. In other words, *without* an infection by Fusobacterium *no* colorectal carcinoma.

Q. e. d.

3.6. Fusobacterium is the cause of colorectal cancer

Claims.

Null hypothesis: (**no causal relationship**)

There is no significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k=0).

Alternative hypothesis: (**causal relationship**)

There is a significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k<>0).

Conditions.

Alpha level = 5%.

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

Proof.

The data for this hypothesis test were provided by Ahn et al. [30] and are illustrated in the 2 × 2 table (**Table 2**). The causal relationship k(Fusobacterium , Colorectal cancer) is calculated [36]-[50] as

$$\text{k}\left(\text{Fusobacterium, Colorectal carcinoma}\right)=\frac{\left(\left(141\times17\right)-\left(32\times47\right)\right)}{\sqrt[2]{\left(47\times94\right)\times\left(32\times109\right)}}=+0{,}227483743$$

The value of the test statistic k=+0,227483743 is equivalent to a calculated [36]-[50] chi-square value of

$$\chi^{2}{}_{\text{Calculated}}=141\times\left(\frac{\left(\left(141\times17\right)-\left(32\times47\right)\right)}{\sqrt[2]{\left(47\times94\right)\times\left(32\times109\right)}}\right)\times\left(\frac{\left(\left(141\times17\right)-\left(32\times47\right)\right)}{\sqrt[2]{\left(47\times94\right)\times\left(32\times109\right)}}\right)$$

$$\chi^{2}{}_{\text{Calculated}}=141\times0{,}227483743\times0{,}227483743$$

$$\chi^{2}{}_{\text{Calculated}}=7{,}296588303$$

The chi-square statistic, uncorrected for continuity, is calculated as X² = 7,296588303 and thus far equivalent to a P value of 0,00690856692290541000. The calculated chi-square statistic exceeds the critical chi-square value of 3.841458821 (**Table 18**). Consequently, we reject the null hypothesis and accept the alternative hypotheses. There is a significant causal relationship between an infection Fusobacterium and colorectal cancer (k=+0,227483743, p Value = 0,00690856692290541000). The result is significant at p < 0.05.

Q. e. d.

3.7. Fusobacterium is the cause of colorectal cancer

Claims.

Null hypothesis: (**no causal relationship**)

There is no significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k=0).

Alternative hypothesis: (**causal relationship**)

There is a significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k<>0).

Conditions.

Alpha level = 5%.

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

Proof.

The data for this hypothesis test were provided by Vogtmann et al. [32] and are illustrated in the 2×2 table (**Table 4**). The causal relationship k(Fusobacterium , Colorectal cancer) is calculated [36]-[50] as

$$k\left(\text{Fusobacterium, Colorectal carcinoma}\right) = \frac{\left(\left(104 \times 40\right) - \left(65 \times 52\right)\right)}{\sqrt[2]{\left(52 \times 52\right) \times \left(65 \times 39\right)}} = +0,0023798165$$

The value of the test statistic k=+ 0,297921796 is equivalent to a calculated [36]-[50] chi-square value of

$$\chi^2_{\text{Calculated}} = 104 \times \left(\frac{\left(\left(104 \times 40\right) - \left(65 \times 52\right)\right)}{\sqrt[2]{\left(52 \times 52\right) \times \left(65 \times 39\right)}}\right) \times \left(\frac{\left(\left(104 \times 40\right) - \left(65 \times 52\right)\right)}{\sqrt[2]{\left(52 \times 52\right) \times \left(65 \times 39\right)}}\right)$$

$$\chi^2_{\text{Calculated}} = 104 \times 0,0023798165 \times 0,0023798165$$

$$\chi^2_{\text{Calculated}} = 9,230769231$$

The chi-square statistic, uncorrected for continuity, is calculated as X² = 9,230769231 and thus far equivalent to a P value of 0,0023798164638097. The calculated chi-square statistic exceeds the critical chi-square value of 3.841458821 (**Table 18**). Consequently, we reject the null hypothesis and accept the alternative hypotheses. There is a significant causal relationship between an infection Fusobacterium and colorectal cancer (k=+0,297921796, p Value = 0,0023798164638097). The result is significant at p < 0.05.

Q. e. d.

3.8. Fusobacterium is the cause of colorectal cancer

Claims.

Null hypothesis: (**no causal relationship**)

There is no significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k=0).

Alternative hypothesis: (**causal relationship**)

There is a significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k<>0).

Conditions.

Alpha level = 5%.

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

Proof.

The data for this hypothesis test were provided by Li et al. [33] and are illustrated in the 2 × 2 table (**Table 5**). The causal relationship k(Fusobacterium , Colorectal cancer) is calculated [36]-[50] as

$$k\left(\text{Fusobacterium, Colorectal carcinoma}\right) = \frac{\left(\left(202 \times 88\right) - \left(101 \times 101\right)\right)}{\sqrt[2]{\left(101 \times 101\right) \times \left(101 \times 101\right)}} = +0,742574257$$

The value of the test statistic k=+ 0,742574257 is equivalent to a calculated [36]-[50] chi-square value of

$$\chi^2_{\text{Calculated}} = 202 \times \left(\frac{\left(\left(202 \times 88\right) - \left(101 \times 101\right)\right)}{\sqrt[2]{\left(101 \times 101\right) \times \left(101 \times 101\right)}}\right) \times \left(\frac{\left(\left(202 \times 88\right) - \left(101 \times 101\right)\right)}{\sqrt[2]{\left(101 \times 101\right) \times \left(101 \times 101\right)}}\right)$$

$$\chi^2_{\text{Calculated}} = 202 \times 0,742574257 \times 0,742574257$$

$$\chi^2_{\text{Calculated}} = 111,3861386$$

The chi-square statistic, uncorrected for continuity, is calculated as X² = 111,3861386 and thus far equivalent to a P value of 0,000000000000000000000000048699. The calculated chi-square statistic exceeds the critical chi-square value of 3.841458821 (**Table 18**). Consequently, we reject the null hypothesis and accept the alternative hypotheses. There is a highly significant causal relationship between an infection Fusobacterium and colorectal cancer (k=+ 0,742574257, p Value = 0,000000000000000000000000048699). The result is significant at p < 0.001.

Q. e. d.

3.9. Fusobacterium is the cause of colorectal cancer

Claims.

Null hypothesis: (**no causal relationship**)

There is no significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k=0).

Alternative hypothesis: (**causal relationship**)

There is a significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k<>0).

Conditions.

Alpha level = 5%.

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

Proof.

The data for this hypothesis test were provided by Amitay et al. [34] and are illustrated in the 2 × 2 table (**Table 6**). The causal relationship k(Fusobacterium , Colorectal cancer) is calculated [36]-[50] as

$$k\left(\text{Fusobacterium, Colorectal carcinoma}\right) = \frac{\left(\left(277 \times 25\right) - \left(83 \times 46\right)\right)}{\sqrt[2]{\left(46 \times 231\right) \times \left(83 \times 194\right)}} = +0{,}237528888$$

The value of the test statistic k=+0,237528888 is equivalent to a calculated [36]-[50] chi-square value of

$$\chi^2_{\text{Calculated}} = 277 \times \left(\frac{\left(\left(277 \times 25\right) - \left(83 \times 46\right)\right)}{\sqrt[2]{\left(46 \times 231\right) \times \left(83 \times 194\right)}}\right) \times \left(\frac{\left(\left(277 \times 25\right) - \left(83 \times 46\right)\right)}{\sqrt[2]{\left(46 \times 231\right) \times \left(83 \times 194\right)}}\right)$$

$$\chi^2_{\text{Calculated}} = 277 \times 0{,}237528888 \times 0{,}237528888$$

$$\chi^2_{\text{Calculated}} = 15{,}62833247$$

The chi-square statistic, uncorrected for continuity, is calculated as X² =15,62833247 and thus far equivalent to a P value of 0,0000770908214559830. The calculated chi-square statistic exceeds the critical chi-square value of 3.841458821 (**Table 18**). Consequently, we reject the null hypothesis and accept the alternative hypotheses. There is a highly significant causal relationship between an infection Fusobacterium and colorectal cancer (k=+0,237528888, p Value = 0,0000770908214559830). The result is significant at p < 0.001.

Q. e. d.

3.10. Fusobacterium is the cause of colorectal cancer

Claims.

Null hypothesis: (**no causal relationship**)

There is no significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k=0).

Alternative hypothesis: (**causal relationship**)

There is a significant causal relationship between an infection by Fusobacterium and colorectal cancer.

(k<>0).

Conditions.

Alpha level = 5%.

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

Proof.

The data for this hypothesis test were provided by Eklöf et al. [35] and are illustrated in the 2 × 2 table (**Table 7**). The causal relationship k(Fusobacterium , Colorectal cancer) is calculated [36]-[50] as

$$k\left(\text{Fusobacterium, Colorectal carcinoma}\right)=\frac{\left(\left(104\times27\right)-\left(42\times39\right)\right)}{\sqrt[2]{\left(39\times65\right)\times\left(42\times62\right)}}=+0,455382556$$

The value of the test statistic k=+ 0,455382556 is equivalent to a calculated [36]-[50] chi-square value of

$$\chi^2_{\text{Calculated}}=104\times\left(\frac{\left(\left(104\times27\right)-\left(42\times39\right)\right)}{\sqrt[2]{\left(39\times65\right)\times\left(42\times62\right)}}\right)\times\left(\frac{\left(\left(104\times27\right)-\left(42\times39\right)\right)}{\sqrt[2]{\left(39\times65\right)\times\left(42\times62\right)}}\right)$$

$$\chi^2_{\text{Calculated}}=104\times0,455382556\times0,455382556$$

$$\chi^2_{\text{Calculated}}=21,56682028$$

The chi-square statistic, uncorrected for continuity, is calculated as X² = 21,56682028 and thus far equivalent to a P value of 0,00000341712543966276. The calculated chi-square statistic exceeds the critical chi-square value of 3.841458821 (**Table 18**). Consequently, we reject the null hypothesis and accept the alternative hypotheses. There is a highly significant causal relationship between an infection Fusobacterium and colorectal cancer (k =+0,455382556, p Value = 0,00000341712543966276). The result is significant at p < 0.001.

Q. e. d.

4. Discussion

There are a few limitations of this study. Firstly, several studies where not considered. Secondly, the studies analyzed were very heterogeneous. On the whole, the results of this study reflects the main features of other studies. The studies of Fukugaiti et al. [31], Vogtmann et al. [32] (two sided), Li et al. [33], Amitay et al. [34], Eklöf et al. [35], where able to provide evidence that Fusobacterium is a necessary condition, a conditio sine qua non, of colon cancer. The presence of Fusobacterium nucleatum does not simply represent an opportunistic infection at an immuno-compromised site. In contrast to such a view, Fusobacterium nucleatum is a necessary condition of colon cancer. In point of fact, *without* an infection by Fusobacterium nucleatum *no* colon cancer will develop. Several recent studies observed a highly significant over-representation of Fusobacterium nucleatum in colorectal tumor specimens. Consistent with this finding, the current study using data published by Ahn et al. [30], Vogtmann et al. [32] documented a significant cause effect relationship between Fusobacterium and colon cancer while the studies of Li et al. [33], Amitay et al. [34], Eklöf et al. [35] were able to provide evidence of *a highly significant cause effect relationship* between Fusobacterium and colon cancer. This study is for the first time to comprehensively identify Fusobacterium nucleatum as the cause of colorectal cancer. Since *without* an infection by Fusobacterium (nucleatum) *no* colon cancer will develop and since there is a highly significant cause effect relationship between Fusobacterium (nucleatum) we are authorized to deduce that Fusobacterium (nucleatum) is not only a cause of colon cancer, *Fusobacterium (nucleatum) is the cause of human colon cancer*. Still, the sample size of the studies investigated was small. Thus far, larger and very systematic studies are justified and needed to examine and to clarify the association of Fusobacterium and colon cancer definitely while every trace of Fusobacterium in human body should be documented and treated as Fusobacterium positive.

5. Conclusion

In summary, this study represents a systematic review of studies on the relationship between Fusobacterium and colorectal cancer. This report reinforces the notion that F. nucleatum is the cause of colorectal cancer.

Acknowledgements

         http://vixra.org/author/ilija_barukcic

References

[1] Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., Bray, F. GLOBOCAN 2012 v1.0, cancer incidence and mortality world-wide: IARC Cancer Base no. 11: Lyon, International Agency for Research on Cancer; 2013. http://globocan.iarc.fr/Default.aspx

[2] Jemal, A., Bray, F., Center, M.M., Ferlay J, Ward E, Forman D. (2011) Global cancer statistics. *CA Cancer. J. Clin.*, **61**, 69-90. https://doi.org/10.3322/caac.20107

[3] Riboli, E., Norat, T. (2003) Epidemiologic evidence of the protective effect of fruit and vegetables on cancerrisk. *The American Journal of Clinical Nutrition*, **78**, 559S - 69S. http://ajcn.nutrition.org/content/78/3/559S.long

[4] Liang, P.S., Chen, T.Y., Giovannucci, E. (2009) Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. *International Journal of Cancer*, **124**, 2406-15. https://doi.org/10.1002/ijc.24191

[5] Taylor, D.P., Burt, R.W., Williams, M.S., Haug, P.J., Cannon-Albright, L.A. (2010) Population based family history-specific risks for colorectal cancer: a constellation approach. *Gastroenterology*, **138**, 877 - 85. https://doi.org/10.1053/j.gastro.2009.11.044

[6] Chan, D.S., Lau, R., Aune, D., Vieira, R., Greenwood, D.C., Kampman, E., Norat, T. (2011) Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. *PLoS One*, **6**, e20456. https://doi.org/10.1371/journal.pone.0020456

[7] Fedirko, V., Tramacere, I., Bagnardi, V., Rota, M., Scotti, L., Islami, F., Negri, E., Straif, K., Romieu, I., La Vecchia, C., Boffetta, P., Jenab, M. (2011) Alcoholdrinking and colorectal cancer risk: an overall and dose-response metaanalysis of published studies. *Annals of Oncology*, **22**, 1958 - 72. https://doi.org/10.1093/annonc/mdq653

[8] Jiang, Y., Ben, Q., Shen, H., Lu, W., Zhang, Y., Zhu, J. (2011) Diabetes mellitus and incidence and mortality of colorectal cancer: a systematic review and metaanalysis of cohort studies. *Eur J Epidemiol.*, **26**, 863 - 76. https://doi.org/10.1007/s10654-011-9617-y

[9] Jess, T., Rungoe, C., Peyrin-Biroulet, L. (2012) Risk of colorectal cancer in patients with ulcerative colitis: a meta-analysis of population-based cohort studies. *Clin Gastroenterol Hepatol*, **10**, 639 -45. https://doi.org/10.1016/j.cgh.2012.01.010

[10] Ma, Y., Yang, Y., Wang, F., Zhang, P., Shi, C., Zou, Y., et al. (2013) Obesity and risk of colorectal cancer: a systematic review of prospective studies. *PLoS One*, **8**, e53916. https://doi.org/10.1371/journal.pone.0053916

[11] Brenner. H., Kloor, M., Pox, C.P. (2014) Colorectal cancer. *Lancet*, **383**, 1490-502. http://dx.doi.org/10.1016/S0140-6736(13)61649-9

[12] De Flora, S., La Maestra, S. (2015) Epidemiology of cancers of infectious origin and prevention strategies. Journal of Preventive Medicine and Hygiene, **56**, E15-20. http://dx.doi.org/10.15167/2421-4248/jpmh2015.56.1.470

[13] Chen, H., Chen, X.Z., Waterboer, T., Castro, F.A., Brenner, H. (2015) Viral infections and colorectal cancer: a systematic review of epidemiological studies. *Int J Cancer*, **137**, 12-24. https://doi.org/10.1002/ijc.29180

[14] Schwabe, R.F., Jobin, C. (2013) The microbiome and cancer. *Nat Rev Cancer*, **13**, 800-812. https://doi.org/10.1038/nrc3610

[15] Dickson, R.P., Martinez, F.J., Huffnagle, G.B. (2014) The role of the microbiome in exacerbations of chronic lung diseases. **Lancet**, *384*, 691-702 https://doi.org/10.1016/S0140-6736(14)61136-3

[16] Kahn, S.E., Cooper, M.E., Del Prato, S. (2014) Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future. *Lancet*, **383**, 1068-1083 https://doi.org/10.1016/S0140-6736(13)62154-6

[17] Rowland, I.R. (2009) The role of the gastrointestinal microbiota in colorectal cancer. *Curr Pharm Des.*, **15**, 1524-1527. https://doi.org/10.2174/138161209788168191

[18] Warren, R.L., Freeman, D.J., Pleasance, S., Watson, P., Moore, R.A., Cochrane, K., Allen-Vercoe, E., Holt, R.A. (2013) Co-occurrence of anaerobic bacteria in colorectal carcinomas.

Microbiome, **1**, 16. https://doi.org/10.1186/2049-2618-1-16

[19] Wu, Q., Yang, Z.P., Xu, P., Gao, L.C., Fan, D.M. (2016) Association between Helicobacter pylori infection and the risk of colorectal neoplasia: a systematic review and meta-analysis. *Colorectal Dis.*, **15**, e352-64. https://doi.org/10.1111/codi.12284

[20] Castellarin, M., Warren, R.L., Freeman, J.D., Dreolini, L., Krzywinski, M., Strauss, J., et al. (2012) Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Research*, **22**, 299-306. https://dx.doi.org/10.1101/gr.126516.111

[21] Kostic, A.D., Gevers, D., Pedamallu, C.S., Michaud, M., Duke, F., Earl, A.M., et al. (2012) Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Research*, **22**, 292-8. https://dx.doi.org/10.1101/gr.126573.111

[22] Kostic, A.D., Chun, E., Robertson, L., Glickman, J.N., Gallini, C.A., Michaud, M., Clancy, T.E., Chung, D.C., Lochhead P., Hold, G.L., El-Omar, E.M., Brenner, D., Fuchs, C.S., Meyerson, M., Garrett, W.S. (2013) Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe*, **14**, 207-215. https://doi.org/10.1016/j.chom.2013.07.007

[23] Zeller, G., Tap, J., Voigt, A.Y., Sunagawa, S., Kultima, J.R., Costea, P.I., Amiot, A., Böhm, J., Brunetti, F., Habermann, N., Hercog, R., Koch, M., Luciani, A., Mende, D.R., Schneider, M.A., Schrotz-King, P., Tournigand, C., Tran Van Nhieu, J., Yamada, T., Zimmermann, J., Benes, V., Kloor, M., Ulrich, C.M., von Knebel Doeberitz, M., Sobhani, I., Bork, P. (2014) Potential of fecal microbiota for early-stage detection of colorectal cancer. *Molecular Systems Biology*, **10**, 766. https://doi.org/10.15252/msb.20145645

[24] Zackular, J.P., Rogers, M.A., Ruffin, M.T. 4th, Schloss, P.D. (2014) The human gut microbiome as a screening tool for colorectal cancer. *Cancer Prevention Research (Phila.)*, **7**, 1112-1121. https://doi.org/10.1158/1940-6207.CAPR-14-0129

[25] Baxter, N.T., Ruffin, M.T. 4th, Rogers, M.A., Schloss PD. (2016) Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions. *Genome Medicine*, **8**, 37. https://doi.org/10.1186/s13073-016-0290-3

[26] Yu, J., Feng, Q., Wong, S.H., Zhang, D., Liang, Q.Y., Qin, Y., Tang, L., Zhao, H., Stenvang, J., Li, Y., Wang, X., Xu, X., Chen, N., Wu, W.K., Al-Aama, J., Nielsen, H.J., Kiilerich, P., Jensen, B.A., Yau, T.O., Lan, Z., Jia, H., Li, J., Xiao, L., Lam, T.Y., Ng, S.C., Cheng, A.S., Wong, V.W., Chan, F.K., Xu, X., Yang, H., Madsen, L., Datz, C., Tilg, H., Wang, J., Brünner, N., Kristiansen, K., Arumugam, M., Sung, J.J., Wang, J. (2017) Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut*, **66**, 70-78. https://doi.org/10.1136/gutjnl-2015-309800

[27] Bolstad, A.I., Jensen, H.B., Bakken V. (1996) Taxonomy, biology, and periodontal aspects of Fusobacterium nucleatum. *Clin Microbiol Rev.*, **9**, 55-71. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC172882/

[28] Panic, N., Leoncini, E., de Belvis, G., Ricciardi, W., Boccia, S. (2013) Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PLoS One*, **8**, e83138. https://doi.org/10.1371/journal.pone.0083138

[29] Castellarin, M., Warren, R.L., Freeman, D., Dreolini, L., Krzywinski, M., Strauss, J., Barnes, R., Watson, P., Allen-Vercoe, E., Moore, R.A., Holt, R.A. (2012) Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Research*, **22**, 299-306. https://dx.doi.org/10.1101/gr.126516.111

[30] Ahn, J., Sinha, R., Pei, Z., Dominianni, C., Wu, J., Shi, J., Goedert, J.J., Hayes, R.B., Yang, L. (2013) Human gut microbiome and risk for colorectal cancer. *Journal of the National Cancer Institute*, **105**, 1907-11. https://doi.org/10.1093/jnci/djt300

[31] Fukugaiti, M.H., Ignacio, A., Fernandes, M.R., Ribeiro Júnior, U., Nakano, V., Avila-Campos, M.J. (2015) High occurrence of Fusobacterium nucleatum and Clostridium difficile in the intestinal microbiota of colorectal carcinoma patients. *Brazilian Journal of Microbiology*, **46**, 1135-40. https://doi.org/10.1590/S1517-838246420140665

[32] Vogtmann, E., Hua, X., Zeller, G., Sunagawa, S., Voigt, A.Y., Hercog, R., Goedert, J.J., Shi, J., Bork, P., Sinha, R. (2016) Colorectal Cancer and the Human Gut Microbiome: Reproducibility with Whole-Genome Shotgun Sequencing. *PLoS One*, **11**, e0155362.

https://doi.org/10.1371/journal.pone.0155362

[33] Li, Y.Y., Ge, Q.X., Cao, J., Zhou, Y.J., Du, Y.L., Shen, B., Wan, Y.J., Nie, Y.Q. (2016) Association of Fusobacterium nucleatum infection with colorectal cancer in Chinese patients. *World J Gastroenterol.*, **22**, 3227-33. https://doi.org/10.3748/wjg.v22.i11.3227

[34] Amitay, E.L., Werner, S., Vital, M., Pieper, D.H., Höfler, D., Gierse, I.J., Butt, J., Balavarca, Y., Cuk, K., Brenner, H. (2017) Fusobacterium and colorectal cancer: causal factor or passenger? Results from a large colorectal cancer screening study. *Carcinogenesis*, **38**, 781-788. https://doi.org/10.1093/carcin/bgx053

[35] Eklöf, V., Löfgren-Burström, A., Zingmark, C., Edin, S., Larsson, P., Karling, P, Alexeyev, O., Rutegård, J., Wikberg, M.L., Palmqvist R. (2017) Cancer-associated fecal microbial markers in colorectal cancer detection. *Int J Cancer.*, **141**, 2528-2536. https://doi.org/10.1002/ijc.31011

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

[37] Barukčić, I. (1997) Die Kausalität. Wilhelmshaven: Scientia, 1997. pp. 374.

[38] Barukčić, I. (2005) Causality. New Statistical Methods. Hamburg - Norderstedt: Books on Demand, pp. 488.

[39] Barukčić, I. (2006) Causality. New Statistical Methods, Second English Edition. Hamburg-Norderstedt: Books on Demand, pp. 488.

[40] Barukčić, I. (2006) New method for calculating causal relationships. Proceeding of XXIIIrd International Biometric Conference. July 16-21; McGill University, Montréal, Québec, Canada. p. 49.

[41] Barukčić, I. (2011) Causality I. A Theory of Energy, Time and Space. Morrisville: Lulu, pp. 648.

[42] Barukčić, I. (2011) Causality II. A Theory of Energy, Time and Space. Morrisville: Lulu, pp. 376.

[43] Barukčić, I. (2012) The deterministic relationship between cause and effect. *International International Biometric Conference*, Kobe, JAPAN, 26 - 31 August 2012. https://www.biometricsociety.org/conference-abstracts/2012/programme/p1-5/P-1/249-P-1-30.pdf

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

[45] Barukčić, K., Barukčić, I. (2016) Epstein Barr Virus - The Cause of Multiple Sclerosis. *Journal of Applied Mathematics and Physics*. **4**, 1042-53. https://doi.org/10.4236/jamp.2016.46109

[46] Barukčić, I. (2016) Unified Field Theory. *Journal of Applied Mathematics and Physics*. **4**, 1379-1438. https://doi.org/10.4236/jamp.2016.48147

[47] Barukčić, I. (2017) Helicobacter pylori-The Cause of Human Gastric Cancer. *Journal of Biosciences and Medicines.* **5**, 1-19. https://doi.org/10.4236/jbm.2017.52001

[48] Barukčić, I. (2017) Anti Bohr - Quantum Theory and Causality. *International Journal of Applied Physics and Mathematics*. **7**, 93-111. https://doi.org/10.17706/ijapm.2017.7.2.93-111

[49] Barukčić, I. (2017) Theoriae causalitatis principia mathematica. Hamburg - Norderstedt: Books on Demand, pp. 244. https://www.bod.de/buchshop/theoriae-causalitatis-principia-mathematica-ilija-barukcic-9783744815932

[50] Barukčić, I. (2018) Epstein Barr Virus - The Cause of Hodgkin's Lymphoma. *Journal of Biosciences and Medicines,* **1**, …………………

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

[52] Neyman, J. (1937) Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability. *Philosophical Transactions of the Royal Society A,* **236**, 333-380. https://doi.org/10.1098/rsta.1937.0005

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

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

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

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

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

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

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

[60] Fisher, RA. (1922) On the interpretation of X² from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society,* **85**, 87 - 94. https://doi.org/10.2307/2340521

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

[62] Pearson, K. (1896) VII. Mathematical contributions to the theory of evolution.-III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Ser. A*, **187**, 253-18. https://doi.org/10.1098/rsta.1896.0007

[63] Pearson, K. (1904) Mathematical contributions to the theory of evolution. -XIII. On the Theory of Contingency and Its Relation to Association and Normal Correlation. London, Dulau and Co., pp. 1-35. https://archive.org/details/cu31924003064833

[64] Hill, AB. (1965) The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine*, **58**, 295-300. https://doi.org/10.1177/0141076814562718

[65] Mackie, J. (1965) Causes and Conditions. *American Philosophical Quarterly*, **2**, 245-264.