Journal of Biosciences and Medicines, 2018, 6, 31-69
http://www.scirp.org/journal/jbm
ISSN Online: 2327-509X
ISSN Print: 2327-5081

Fusobacterium nucleatum—The Cause of Human Colorectal Cancer

Ilija Barukčić

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

How to cite this paper: Barukčić, I. (2018) *Fusobacterium nucleatum*—The Cause of Human Colorectal Cancer. *Journal of Biosciences and Medicines*, **6**, 31-69.
https://doi.org/10.4236/jbm.2018.63004

Received: January 29, 2018
Accepted: March 11, 2018
Published: March 14, 2018

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 were 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 etc. [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 are 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] [29]. 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.

2.1.1. Study of Castellarin *et al.* 2012 (Canada)

Castellarin *et al.* [20] [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).

The article does not provide the necessary exact information and was not considered for a 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).

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

		Colorectal cancer		Total
		Yes	No	
Fusobacterium	Yes	~78	~21	99
	No	~21	~78	99
	Total	99	99	198

2.1.3. 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).

2.1.4. Study of Vogtmann *et al.* 2016 (USA)

Vogtmann *et al.* [32] investigated fecal samples from 52 matched controls and 52 pre-treatment 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).

2.1.5. Study of Li *et al.* 2016 (China)

Li *et al.* [33] conducted a matched-case control study to investigate Fusobacterium

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

		Colorectal cancer		Total
		Yes	No	
Fusobacterium	Yes	17	15	32
	No	30	79	109
	Total	47	94	141

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

		Colorectal cancer		Total
		Yes	No	
Fusobacterium	Yes	7	9	16
	No	0	1	1
	Total	7	10	17

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

		Colorectal cancer		Total
		Yes	No	
Fusobacterium	Yes	40	25	65
	No	12	27	39
	Total	52	52	104

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).

2.1.6. 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).

2.1.7. 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).

2.2. Statistical Analysis

All statistical analyses were performed with Microsoft Excel version 14.0.7166.5000 (32-Bit) software (Microsoft GmbH, Munich, Germany). In order to simplify the understanding of this article and to increase the transparency for the reader

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

		Colorectal cancer		Total
		Yes	No	
Fusobacterium	Yes	88	13	101
	No	13	88	101
	Total	101	101	202

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

		Colorectal cancer		Total
		Yes	No	
Fusobacterium	Yes	25	58	83
	No	21	173	194
	Total	46	231	277

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

		Colorectal cancer		Total
		Yes	No	
Fusobacterium high	Yes	27	15	42
	No	12	50	62
	Total	39	65	104

several of the following lines are *repeated word by word* and taken form Barukčić [36].

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

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

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

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

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

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

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

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

Table 8. The sample space of a contingency table.

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

Table 9. The sample space of a contingency table.

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

In general it is $(a+c) = {}_0 W_t$, $(a+b) = {}_R U_t$, $(c+d) = {}_0 \underline{W}_t$, $(b+d) = {}_R \underline{U}_t$ and $a+b+c+d = N = {}_R W_t$. Equally, it is ${}_0 W_t + {}_0 \underline{W}_t = {}_R U_t + {}_R \underline{U}_t = {}_R W_t = N$. 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

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

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

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

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

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

2.2.2. Sufficient Condition (Conditio Per Quam)

The formula of the conditio per quam [36]-[54] relationship was derived as

$$p(\text{Fusobacerium nucleatum} \rightarrow \text{Colorectal cancer}) \equiv \frac{a+c+d}{N} \tag{6}$$

and used to proof the hypothesis: *if* presence of *Fusobacterium nucleatum* infection *then* presence of colorectal cancer.

2.2.3. Necessary Condition (Conditio Sine Qua Non)

The formula of the conditio per quam [36]-[54] relationship was derived as

$$p\left(\text{Fusobacerium nucleatum} \leftarrow \text{Colorectal cancer}\right) \equiv \frac{a+b+d}{N} \tag{7}$$

and used to proof the hypothesis: *without* presence *Fusobacterium nucleatum no* presence of colorectal cancer.

2.2.4. Necessary and Sufficient Condition

The necessary and sufficient condition relationship was defined [36]-[54] as

$$p\left(\text{Fusobacerium nucleatum} \longleftrightarrow \text{Colorectal cancer}\right) \equiv \frac{a+d}{N} \tag{8}$$

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 implication and has made some groundbreaking contributions [55] 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 "$\mathbf{p} \rightarrow \mathbf{q}$" 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.

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

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

		Fire		Total
		Yes = +1	No = +0	
Oxygen	Yes =+1	a	b	${}_{R}U_{t}$
	No = +0	0	d	${}_{R}\underline{U}_{t}$
	Total	${}_{0}W_{t}$	${}_{0}\underline{W}_{t}$	$N = {}_{R}W_{t}$

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

		Oxygen		Total
		Yes = +1	No = +0	
Fire	Yes =+1	a	0	${}_{R}U_{t}$
	No = +0	c	d	${}_{R}\underline{U}_{t}$
	Total	${}_{0}W_{t}$	${}_{0}\underline{W}_{t}$	$N={}_{R}W_{t}$

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.

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 condition 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 [56], 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

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

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 [57]. Agresti-Coull proposed another different method [58] 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 [59].

2.2.6. The Rule of Three

Furthermore, an approximate and conservative (one sided) confidence interval was developed by Louis [60], Hanley *et al.* [61] and Jovanovic [62] 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 × (1 − α)% confidence interval for the unknown proportion when in N independent trials no events occur [62]. Then π_U is the value such that

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

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 $3/n$ [62] or given approximately by $\left[0 < \pi < (3/n) \right]$. The rule of three is a useful tool especially in the analysis of medical studies. The follow-

ing table (Table 12) will illustrate this relationship.

Under conditions where a certain event did not occur [60] 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 $\left[\left(1-\left(-\ln\alpha/n\right)\right) < p < +1\right]$ 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.

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 [63] 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

Table 12. The one-sided approximate upper $100 \times (1 - \alpha)\%$ confidence bound where no successes ($p = 0$) are observed.

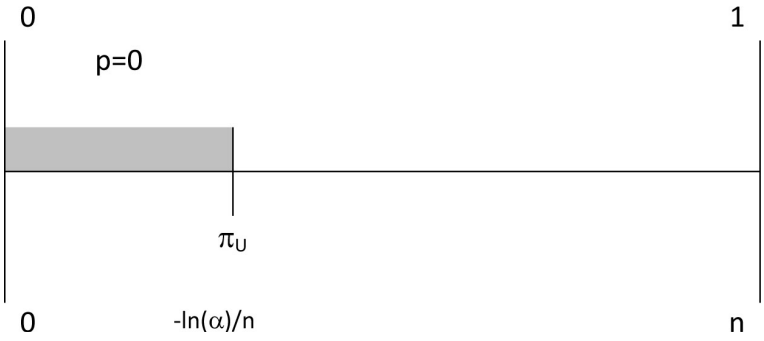
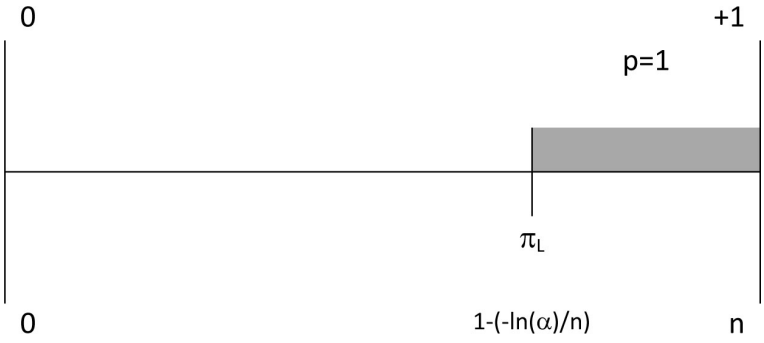


Table 13. The one-sided approximate upper $100 \times (1 - \alpha)\%$ confidence bound where only successes are observed.



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

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

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

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? In 2016, 537 people drowned [65] in Germany as shown in a table (**Table 15**), while a total of 911,000 died [66]. The hypergeometric distribution is of use to calculate how probable is it to obtain $c = ({}_0W_t - a) = 0$ events out of $N = 83{,}431{,}000$ events.

Table 14. The hypergeometric distribution.

		Conditioned		Total
		Yes = +1	No = +0	
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

Table 15. The hypergeometric distribution and conditio sine qua non.

		Human being alive (on Dec. 31, 2016 in Germany)		Total
		Yes = +1	No = +0	
Gaseous oxygen	Yes = +1	82,520,000	910,463	83,430,463
	No = +0	0	537	537
	Total	82,520,000	911,000	83,431,000

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 outcome 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 (H_A) is true then the null hypothesis (H_0) 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 $1 - \beta$). 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 $1 - \beta = 0.80$ 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).

In general, it is $1-\alpha+\alpha=1$ or $(1-\alpha-\beta)+\alpha=1-\beta$. 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).

2.2.10. The Mathematical Formula of the Causal Relationship *k*

The mathematical formula of the causal relationship k [36]-[54] defined as

$$k\left({}_{R}U_{t}, {}_{0}W_{t}\right) \equiv \frac{\left(\left(N \times a\right)-\left({}_{R}U_{t} \times {}_{0}W_{t}\right)\right)}{\sqrt[2]{\left({}_{R}U_{t} \times {}_{R}\underline{U}_{t}\right) \times \left({}_{0}W \times {}_{0}\underline{W}_{t}\right)}} \tag{13}$$

and the chi-square distribution [67] were applied to determine the significance of causal relationship between *Fusobacterium nucleatum* and colorectal cancer. A one-tailed test makes it much 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

Table 16. Table of error types.

		Null hypothesis (H_0) is		Total
		True	False	
Null hypothesis (H_0)	Accepted	$1-\alpha$	β	$1-\alpha+\beta$
	Rejected	α	$\mathbf{1-\beta}$	$1+\alpha-\beta$
	Total	1	1	**2**

Table 17. Table of error types.

		Null hypothesis (H_0) is		Total
		True	False	
Null hypothesis (H_0)	Accepted	$(1-\alpha)/2$	$\beta/2$	$(1-\alpha+\beta)/2$
	Rejected	$\alpha/2$	$\mathbf{(1-\beta)/2}$	$(1+\alpha-\beta)/2$
	Total	1/2	1/2	**(2/2) = 1**

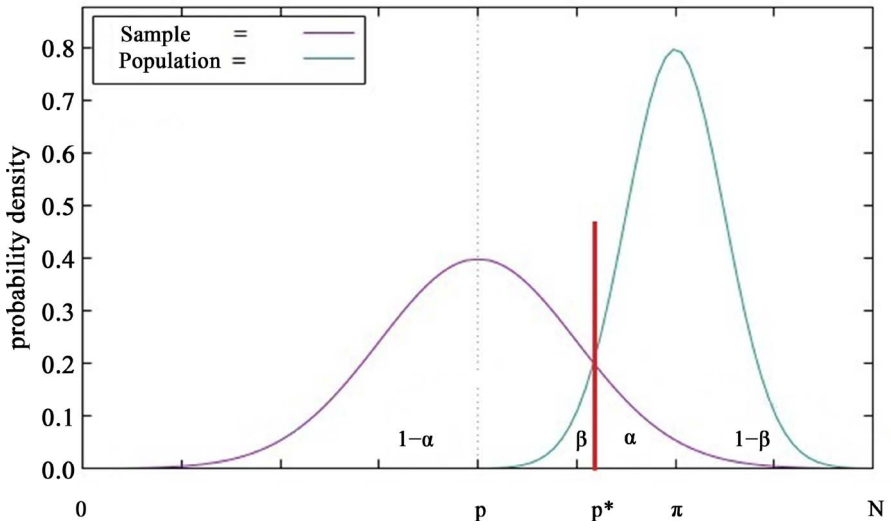


Figure 1. The relationship between error types.

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 [68] or Pearson's Phi [69] 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 [68] or to Pearson's Phi [69] 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) [70] 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" [70] 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.

Medical practice should be based on evidence as much as possible. In this context, the increasing volume of research by the medical community is highly welcomed. Still, the increasing medical literature supported by statistics leads sometimes to an increasing number of contradictory conclusions and findings and the question is justified under which conditions can we accept a significant cause effect relationship as given? Often, studies are performed on a small [51], [52] heterogeneous or limited number of specimens or patients. Thus far, a difference observed can be a true difference or is likely attributable to chance alone. The sample size [53] [54] has an essential influence on the relevance of a finding, on the variability of the data, the level of statistical significance, the power et cetera. In general, a low statistical power is able to undermine the purpose of scientific research while reducing the chance of detecting a true effect. Under these conditions, a statistically significant result must not reflect a true effect. Improving reproducibility of studies to assure the quality of data obtained is a key priority and requires great attention to some well-established methodological principles. In this context it is necessary to consider that *every cause is a kind of a condition too but not vice versa. Every condition must not be equally a cause too.* Considering that other relationships like the condition sine qua non, the conditio per quam, the necessary and sufficient condition et cetera are not statistically significant, a statistically significant cause effect relationship alone should be treated with great care especially under conditions where the sample size is small. Is it possible to generalize these findings? Again there is need for great care in determining whether we can talk about a (highly) significant causal relationship. In summary, a statistically highly significant causal relationship

alone must not guarantee a real causal relationship. Finally, it should be noted that neither a highly significant necessary condition nor a highly necessary sufficient condition guarantees a cause effect relationship. In point of fact, a significant necessary and sufficient condition supports the claims of a study which was able to provide evidence of a statistically (highly) significant causal relationship too. Thus far, in this context it is appropriate to consider the significance of a necessary condition, the significance of a sufficient condition and the significance of a necessary and sufficient condition too.

2.2.11. The Chi Square Distribution

The chi-squared distribution [67] 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**.

2.2.12. The X^2 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 is 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.

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

	P-value	One sided X^2	Two sided X^2
The chi square distribution	0.1000000000	1.642374415	2.705543454
	0.0500000000	**2.705543454**	**3.841458821**
	0.0400000000	3.06490172	4.217884588
	0.0300000000	3.537384596	4.709292247
	0.0200000000	4.217884588	5.411894431
	0.0100000000	5.411894431	6.634896601
	0.0010000000	9.549535706	10.82756617
	0.0001000000	13.83108362	15.13670523
	0.0000100000	18.18929348	19.51142096
	0.0000010000	22.59504266	23.92812698
	0.0000001000	27.03311129	28.37398736
	0.0000000100	31.49455797	32.84125335
	0.0000000010	35.97368894	37.32489311
	0.0000000001	40.46665791	41.82145620

H_0: The sample distribution agrees with the hypothetical (theoretical) distribution.

H_A: The sample distribution does not agree with the hypothetical (theoretical) distribution.

The X^2 Goodness-of-Fit Test can be shown schematically as

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

The degrees of freedom are calculated as $N-1$. If there is no discrepancy between an observed and a theoretical distribution, then $X^2 = 0$. As the discrepancy between an observed and a theoretical distribution becomes larger, the X^2 becomes larger. This X^2 values are evaluated by the known X^2 distribution.

The original X^2 values are calculated from an original theoretical distribution, which is continuous, whereas the approximation by the X^2 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 X^2 Goodness-of-Fit Test *with continuity correction* shown schematically as

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

When the term ($|\text{Observed}_t - \text{Expected}_t|$) is less than 1/2, the continuity correction should be omitted.

1) The X^2 goodness of fit test of a sufficient condition

Gaseous oxygen which is colorless, odorless, tasteless, and nonflammable is necessary to support human life. But it is not the cause of human life. The theoretical (hypothetical) distribution of a sufficient condition [36]-[54] is shown schematically by the 2 × 2 table (Table 19).

The theoretical distribution of a sufficient condition (conditio per quam) is determined by the fact that $b = 0$. The X^2 Goodness-of-Fit Test *with continuity correction* of a sufficient condition (conditio per quam) is calculated as

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

		Gaseous oxygen		Total
		Yes = +1	No = +0	
Human being alive (Germany Dec. 31, 2016)	Yes = +1	82,520,000	0	82,520,000
	No = +0	910,463	537	911,000
	Total	83,430,463	537	83,431,000

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

or more simplified as

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

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

2) The X^2 goodness of fit test of a necessary condition

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

The theoretical distribution of a necessary condition (conditio sine qua non) is determined by the fact that $c = 0$. The X^2 Goodness-of-Fit Test *with continuity correction* of a necessary condition (conditio sine qua non) is calculated as

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

or more simplified as

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

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

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

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.

3) The X^2 goodness of fit test of a necessary and sufficient condition

Like other fundamental concepts, the concept 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" [71] of an effect.

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

The theoretical distribution of a necessary and sufficient condition is determined by the fact that $b = 0$ and that $c = 0$. The X^2 Goodness-of-Fit Test *with continuity correction* of a necessary and sufficient condition is calculated as

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

or more simplified as

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

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

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

		Conditioned		
		Yes = +1	**No = +0**	**Total**
Condition	**Yes = +1**	a	$\boldsymbol{b = 0}$	$(a + b)$
	No = +0	$\boldsymbol{c = 0}$	d	$(c + d)$
	Total	$(a + c)$	$(b + d)$	$(a + b + c + d)$

4) The X^2 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 2 × 2 table (Table 22).

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 X^2 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. = $N-1=2-1=1$.

5) The X^2 goodness of fit test of exclusion

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

The theoretical distribution of exclusion is determined by the fact that $a = 0$. The X^2 Goodness-of-Fit Test *with continuity correction* of exclusion is calculated as

Table 22. The theoretical distribution of either a necessary condition 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	$(a + c)$	$(b + d)$	$(a + b + c + d)$

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	$(a + c)$	$(b + d)$	$(a + b + c + d)$

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

or more simplified as

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

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

6) The X^2 goodness of fit test of disjunction

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

The theoretical distribution of disjunction is determined by the fact that $d = 0$. The X^2 Goodness-of-Fit Test *with continuity correction* of disjunction can be calculated as

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

or more simplified as

Table 24. The theoretical distribution of disjunction.

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

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

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

7) The X^2 goodness of fit test of a neither nor relationship

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

The theoretical distribution of a neither nor relationship is determined by the fact that $a = 0$ and $b = 0$ and $c = 0$. The X^2 Goodness-of-Fit Test *with continuity correction* of a neither nor relationship is calculated as

$$\begin{aligned} \chi^2\left(\text{Neither nor}\right) &\equiv \left(\frac{\left(\left| \left(d \right) - \left(a+b+c+d \right) \right| - \left(\frac{1}{2} \right) \right)^2}{\left(a+b+c+d \right)} \right) + 0 \\ &= \left(\frac{\left(\left| -\left(a+b+c \right) \right| - \left(\frac{1}{2} \right) \right)^2}{\left(a+b+c+d \right)} \right) + 0 \end{aligned} \tag{28}$$

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 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 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

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

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

in the sample of a conditio sine qua non relationship p (Fusobacterium ← Colorectal cancer) is calculated [36]-[54] as

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

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 is 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]-[54] according to the rule of three is

$$p_{Crit}=1-\frac{3}{17}=0.823780454$$

The critical value is p_{Crit} = 0.823780454 and is less than the proportion of successes calculated as p (Fusobacterium ← 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.

3.2. *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 Vogtmann *et al.* [32] of an infection by Fusobacterium and colorectal cancer are viewed in the 2 × 2 table (**Table 4**). The X^2 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]-[54] as

$$\chi^2\left(\text{SINE}\right)\equiv\left(\frac{\left(\left|-c\right|-\left(\frac{1}{2}\right)\right)^2}{\left(c+d\right)}\right)+0=\left(\frac{\left(\left|-12\right|-\left(\frac{1}{2}\right)\right)^2}{\left(12+27\right)}\right)=3.391025641$$

Under these circumstances, the degree of freedom is d.f. = $N-1=2-1=1$. The two sided critical X^2 (significance level alpha = 0.05) is known to be 3.841458821 (**Table 18**). The calculated X^2 value = 3.391025641 and less than the

critical X^2 = 3.841458821. Hence, our calculated X^2 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 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 Li *et al.* [33] of an infection by Fusobacterium and colorectal cancer are viewed in the 2 × 2 table (Table 5). The X^2 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]-[54] as

$$\chi^2(\text{SINE}) \equiv \left(\frac{\left(|-c| - \left(\frac{1}{2} \right) \right)^2}{(c+d)} \right) + 0 = \left(\frac{\left(|-13| - \left(\frac{1}{2} \right) \right)^2}{(13+88)} \right) = 1.547029703$$

Under these circumstances, the degree of freedom is d.f. = $N-1=2-1=1$. The one sided critical X^2 (significance level alpha = 0.05) is known to be 2.705543454 (Table 18). The calculated X^2 value = 1.547029703 and less than the critical X^2 = 2.705543454. Hence, our calculated X^2 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. *Fusobacterium Nucleatum* Is a Necessary and Sufficient Condition of Human Colorectal Cancer

Claims.

Null hypothesis:

An infection by Fusobacterium is a necessary and sufficient of human colorectal cancer.

Alternative hypothesis:

An infection by Fusobacterium is not a necessary and sufficient of human 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 X^2 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]-[54] as

$$\begin{aligned}&\chi^2(\text{Necess. and Suffic.})\\ &\equiv\left(\frac{\left(\left|-b\right|-\left(\frac{1}{2}\right)\right)^2}{(c+d)}\right)+\left(\frac{\left(\left|-c\right|-\left(\frac{1}{2}\right)\right)^2}{(c+d)}\right)\\ &=\left(\frac{\left(\left|-13\right|-\left(\frac{1}{2}\right)\right)^2}{(13+88)}\right)+\left(\frac{\left(\left|-13\right|-\left(\frac{1}{2}\right)\right)^2}{(13+88)}\right)\\ &=3.094059406\end{aligned}$$

Under these circumstances, the degree of freedom is d. f. = $N-1=2-1=1$. The two sided critical X^2 (significance level alpha = 0.05) is known to be 3.841458821 (Table 18). The calculated X^2 value = 3.094059406 and less than the critical X^2 = 3.841458821. Hence, our calculated X^2 value = 3.094059406 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 and sufficient condition. Thus far, the data as published by Li *et al.* [33] do support our null hypothesis that an infection by Fusobacterium is a necessary and sufficient condition of colorectal carcinoma. In other words, *without* an infection by Fusobacterium *no* colorectal carcinoma and *if* an infection by Fusobacterium *then* 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 Amitay *et al.* [34] of an infection by Fusobacterium and colorectal cancer are viewed in the 2 × 2 table (**Table 6**). The X^2 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]-[54] as

$$\chi^2(\text{SINE}) \equiv \left(\frac{\left(|-c| - \left(\frac{1}{2} \right) \right)^2}{(c+d)} \right) + 0 = \left(\frac{\left(|-21| - \left(\frac{1}{2} \right) \right)^2}{(21+173)} \right) + 0 = 2.166237113$$

Under these circumstances, the degree of freedom is d. f. = $N-1=2-1=1$. The one sided critical X^2 (significance level alpha = 0.05) is known to be 2.705543454 (**Table 18**). The calculated X^2 value = 2.166237113 and less than the critical X^2 = 2.705543454. Hence, our calculated X^2 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.

3.6. *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 X^2 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]-[54] as

$$\chi^2(\text{SINE}) \equiv \left(\frac{\left(|-c| - \left(\frac{1}{2} \right) \right)^2}{(c+d)} \right) + 0 = \left(\frac{\left(|-12| - \left(\frac{1}{2} \right) \right)^2}{(12+50)} \right) + 0 = 2.166237113$$

Under these circumstances, the degree of freedom is d. f. = $N-1=2-1=1$. The one sided critical X^2 (significance level alpha = 0.05) is known to be 2.705543454 (Table 18). The calculated X^2 value = 2.133064516and less than the critical X^2 = 2.705543454. Hence, our calculated X^2 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.

3.7. 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.

($k\neq 0$).

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]-[54] as

$$k(\text{Fusobacterium, Colorectal carcinoma}) = \frac{((141\times 17)-(32\times 47))}{\sqrt[2]{(47\times 94)\times(32\times 109)}} = +0.227483743$$

The value of the test statistic k = +0.227483743 is equivalent to a calculated [36]-[54] chi-square value of

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

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

$$\chi^2_{\text{Calculated}} = 7.296588303$$

The chi-square statistic, uncorrected for continuity, is calculated as X^2 = 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.

3.8. 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.

($k \neq 0$).

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]-[54] as

$$k(\text{Fusobacterium, Colorectal carcinoma}) = \frac{\left((104\times40)-(65\times52)\right)}{\sqrt[2]{(52\times52)\times(65\times39)}} = +0.0023798165$$

The value of the test statistic k = +0.297921796 is equivalent to a calculated [36]-[54] chi-square value of

$$\chi^2_{\text{Calculated}} = 104\times\left(\frac{\left((104\times40)-(65\times52)\right)}{\sqrt[2]{(52\times52)\times(65\times39)}}\right)\times\left(\frac{\left((104\times40)-(65\times52)\right)}{\sqrt[2]{(52\times52)\times(65\times39)}}\right)$$

$$\chi^2_{\text{Calculated}} = 104\times0.0023798165\times0.0023798165$$

$$\chi^2_{\text{Calculated}} = 9.230769231$$

The chi-square statistic, uncorrected for continuity, is calculated as X^2 = 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.

3.9. 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.

($k \neq 0$).

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]-[54] 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]-[54] 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 X^2 = 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 sig-

nificant causal relationship between an infection Fusobacterium and colorectal cancer (k =+0.742574257, p Value = 0.00000000000000000000000048699). 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.

($k \neq 0$).

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]-[54] 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]-[54] 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 X^2 = 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.

3.11. 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.

($k \neq 0$).

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]-[54] as

$$k\left(\text{Fusobacterium, Colorectal carcinoma}\right) = \frac{\left(\left(104\times 27\right)-\left(42\times 39\right)\right)}{\sqrt[2]{\left(39\times 65\right)\times\left(42\times 62\right)}} = +0.455382556$$

The value of the test statistic k = +0.455382556 is equivalent to a calculated [36]-[54] chi-square value of

$$\chi^2_{\text{Calculated}} = 104\times\left(\frac{\left(\left(104\times 27\right)-\left(42\times 39\right)\right)}{\sqrt[2]{\left(39\times 65\right)\times\left(42\times 62\right)}}\right)\times\left(\frac{\left(\left(104\times 27\right)-\left(42\times 39\right)\right)}{\sqrt[2]{\left(39\times 65\right)\times\left(42\times 62\right)}}\right)$$

$$\chi^2_{\text{Calculated}} = 104\times 0.455382556\times 0.455382556$$

$$\chi^2_{\text{Calculated}} = 21.56682028$$

The chi-square statistic, uncorrected for continuity, is calculated as X^2 = 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

An important objective of scientific research is to identify causes for disease. Depending on the particular question being asked, several types of studies are conducted including cohort studies, case-control studies, cross-sectional studies et cetera. The study design of a study with its own inclusion and exclusion crite-

ria determines to a major extent scientific informativeness and value of a medical study and can have a decisive influence on the analysis and the results of a study. In general, it is difficult to correct errors in study design afterwards. The sample can be highly representative of the study population if it is properly selected. A sample can but must not produce true description of a population. A sample properly selected can but must not be highly representative of the study population itself. In point of fact, the sample size itself has an effect on the measures to be calculated and the analyses to be performed. Very often, the size of medical studies is too small, and the power of such studies is small too with the consequence that a real difference is described imprecisely or unidentified. Often, inaccuracies are caused by measuring or recording errors too. Therefore, the findings of several types of studies, comparative therapeutic trials and some diagnostic tests must be interpreted accurately for a correct understanding of these findings. According to such opinions, generalizing about the properties of a population based on some number of observations of particular instances can ultimately be associated with fallacious reasoning.

In this context, imprecision seems to be a defect which impedes scientific progress. However, efforts in the form of studies which have been made to take up a challenge can provide a contribution to scientific development by the addition of new truths to a stock of already known and secured old truths and in the odd case to the correction of past errors. In particular, great scientist may accelerate scientific progress, but in one way or another, scientific progress itself is an evolutionary process too and remains associated with errors. In line with thoughts like these it is important to note that a calculated significant causal relationship does not guarantee a cause effect relationship within the population and we must keep searching for an explanation. In response to this worry and while characterizing the relationship between cause and effect it is important address a couple of other preliminary issues. It is nowadays regarded as good scientific practice to study the efficacy of a drug in a randomized, placebo-controlled and double-blinded trial. Still, such verum and placebo controlled (clinical) trials provide more questions than definitive answers. A general assumption in almost all placebo-controlled (drug) trials is that a suitable sample size should be determined more or less by a number of subjects where verum = placebo. Still, meta-analyses and re-analyses of such (trial) data are achieving conflicting results. Many times, it is assumed that the placebo response rates in the placebo arm equals to the placebo response rates in the drug arm [72] [73] without a general proof. Despite an increase in scientific knowledge based on placebo response trials, such an approach poses the serious challenge whether such trials can be taken as evidence of drug effects in clinical routine at all. In point of fact, clinical trials of new drugs are many time multiples trials in multiple locations (multicenter trials) and often, it is time consuming and very expensive to conduct such clinical trials. Even after a drug has been approved and marketed manufacturers are forced to continue testing after approval to demonstrate that

a special drug indeed provides benefit to the patient. In the past, *i.e.* due data generated according to an insufficient study design, already approved drugs have been withdrawn from the market. Thus far, it is necessary speed up the development of new drugs which promise significant benefit over existing therapies while reducing the costs for the manufacturers and increasing the security for patients and the manufacturers. A more appropriate study design is necessary to increase the quality of data and to contribute to a more accurate picture of reality. In this context, for example studies grounded on a sample size where the number of ((exposed to a specific risk factor) and (disease developed)) = ((not exposed to a specific risk factor) and (disease not developed)) will reduce the cost, speed up the development of drugs and are more appropriate in clinical trial studies or causal analysis. A simple, very impressive and preliminary example of such an approach is the Chinese study conducted by Li *et al.* [33]. With these general considerations in place, let us examine once again the relationship between conditions and causes. The mathematical formula of causal relationship k is defined for one single event. While analyzing the data of a study we are extrapolating form one single event to n events and erroneous results are possible. Thus far, a significant cause effect relationship alone is to less to rely upon until the problems as associated with study design and generating data are not solved in an appropriate manner. In this context, Gottfried Wilhelm Leibniz (1646-1716), a well-known proponent of the unrestricted version the principle of causality and closely associated with the Principle of Sufficient Reason [74] has been stipulating that everything what is (effect) must have a cause too. In other words, without a cause no effect. In the same respect it is assumed that a cause must have an effect. Consequently, *without* a cause, *no* effect and equally *if* cause, *then* effect. In other words, a cause is a sufficient condition of an effect too. In order to fully understand the relationship between cause and effect it is worth noting *that a cause is equally a necessary and sufficient conditions of an effect too*.

Additionally, we should use the conditio sine qua non relationship, the conditio per quam relationship, the necessary and sufficient condition relationship et cetera in order to achieve reproducible results. A significant (conditio sine qua non relationship or conditio per quam relationship) and a significant causal relationship provides some evidence that there is a cause effect relationship within the population too. A significant necessary and sufficient condition relationship and equally a significant causal relationship make it much more probable that there is a cause effect relationship within the population.

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 reflect 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. Con-

sistent 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. In particular, we need to bear in mind that Li *et al.* [33] were able within certain limits to provide evidence that *Fusobacterium nucleatum* is a necessary and sufficient condition of human colorectal cancer. As it turns out, among other, the study of Li *et al.* [33] provided evidence of *a highly significant cause effect relationship* between Fusobacterium and colon cancer too. Thus far, the presence of *Fusobacterium nucleatum* does not simply represent an opportunistic infection at an immuno-compromised site. In contrast to such a view, due to Li *et al.* [33], *Fusobacterium nucleatum* is a necessary and sufficient condition of colon cancer. In point of fact, *without* an infection by *Fusobacterium nucleatum no* colon cancer will develop. Furthermore, *if* infection by *Fusobacterium nucleatum* is present *then* colon cancer is present too. In toto, several recent studies observed a highly significant over-representation of *Fusobacterium nucleatum* in colorectal tumor specimens. 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*.

However, the sample size of the studies analyzed was (very) small and can have influence on the quality of the conclusions drawn. Nonetheless, the studies must be regarded as being independent of each other and have provided the same or similar results. In other words, *without* an infection by Fusobacterium (nucleatum) *no* colon cancer. 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

The public domain software GnuPlot was used to draw the figure.

References

[1] Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., Rebelo, M.,

Parkin, D.M., Forman, D. and Bray, F. (2013) GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 11: Lyon. International Agency for Research on Cancer. http://globocan.iarc.fr/Default.aspx

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

[3] Riboli, E. and Norat, T. (2003) Epidemiologic Evidence of the Protective Effect of Fruit and Vegetables on Cancerrisk. *The American Journal of Clinical Nutrition*, **78**, 559S-569S. http://ajcn.nutrition.org/content/78/3/559S.long https://doi.org/10.1093/ajcn/78.3.559S

[4] Liang, P.S., Chen, T.Y. and Giovannucci, E. (2009) Cigarette Smoking and Colorectal Cancer Incidence and Mortality: Systematic Review and Meta-Analysis. *International Journal of Cancer*, **124**, 2406-2415. https://doi.org/10.1002/ijc.24191

[5] Taylor, D.P., Burt, R.W., Williams, M.S., Haug, P.J. and Cannon-Albright, L.A. (2010) Population Based Family History-Specific Risks for Colorectal Cancer: A Constellation Approach. *Gastroenterology*, **138**, 877-885. 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. and 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. and Jenab, M. (2011) Alcohol-drinking and Colorectal Cancer Risk: An Overall and Dose-Response Metaanalysis of Published Studies. *Annals of Oncology*, **22**, 1958-1972. https://doi.org/10.1093/annonc/mdq653

[8] Jiang, Y., Ben, Q., Shen, H., Lu, W., Zhang, Y. and Zhu, J. (2011) Diabetes Mellitus and Incidence and Mortality of Colorectal Cancer: A Systematic Review and Metaanalysis of Cohort Studies. *European Journal of Epidemiology*, **26**, 863-876. https://doi.org/10.1007/s10654-011-9617-y

[9] Jess, T., Rungoe, C. and Peyrin-Biroulet, L. (2012) Risk of Colorectal Cancer in Patients with Ulcerative Colitis: A Meta-Analysis of Population-Based Cohort Studies. *Clinical Gastroenterology and Hepatology*, **10**, 639-645. 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. and Pox, C.P. (2014) Colorectal Cancer. *Lancet*, **383**, 1490-1502. http://doi.org/10.1016/S0140-6736(13)61649-9

[12] De Flora, S. and La Maestra, S. (2015) Epidemiology of Cancers of Infectious Origin and Prevention Strategies. *Journal of Preventive Medicine and Hygiene*, **56**, E15-E20. http://doi.org/10.15167/2421-4248/jpmh2015.56.1.470

[13] Chen, H., Chen, X.Z., Waterboer, T., Castro, F.A. and 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. and Jobin, C. (2013) The Microbiome and Cancer. *Nature Reviews Cancer*, **13**, 800-812. https://doi.org/10.1038/nrc3610

[15] Dickson, R.P., Martinez, F.J. and 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. and 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. *Current Pharmaceutical Design*, **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. and 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. and 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-e364. 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://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**, 292-298. https://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. and 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. and 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 IV, M.T. and Schloss, P.D. (2014) The Human Gut Microbiome as a Screening Tool for Colorectal Cancer. *Cancer Prevention Research*, **7**, 1112-1121. https://doi.org/10.1158/1940-6207.CAPR-14-0129

[25] Baxter, N.T., Ruffin IV, M.T., Rogers, M.A. and Schloss, P.D. (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. and 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. and Bakken, V. (1996) Taxonomy, Biology, and Periodontal Aspects of *Fusobacterium nucleatum*. *Clinical Microbiology Reviews*, **9**, 55-71. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC172882/

[28] Panic, N., Leoncini, E., de Belvis, G., Ricciardi, W. and 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] Lijmer, J.G., Mol, B.W., Heisterkamp, S., Bonsel, G.J., Prins, M.H., van der Meulen, J.H. and Bossuyt, P.M. (1999) Empirical Evidence of Design-Related Bias in Studies of Diagnostic Tests. *Journal of the American Medical Association*, **282**, 1061-1066. https://doi.org/10.1001/jama.282.11.1061

[30] Ahn, J., Sinha, R., Pei, Z., Dominianni, C., Wu, J., Shi, J., Goedert, J.J., Hayes, R.B. and Yang, L. (2013) Human Gut Microbiome and Risk for Colorectal Cancer. *Journal of the National Cancer Institute*, **105**, 1907-1911. https://doi.org/10.1093/jnci/djt300

[31] Fukugaiti, M.H., Ignacio, A., Fernandes, M.R., Ribeiro Jr., U., Nakano, V. and 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**, 1135-1140. 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. and 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. and Nie, Y.Q. (2016) Association of *Fusobacterium nucleatum* Infection with Colorectal Cancer in Chinese Patients. *World Journal of Gastroenterology*, **22**, 3227-3233. 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. and 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. and 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. (2018) Epstein Barr Virus: The Cause of Hodgkin's Lymphoma. *Journal of Biosciences and Medicines*, **6**, 75-100. https://doi.org/10.4236/jbm.2018.61008

[37] Barukčić, I. (1989) Die Kausalität. Wissenschaftsverlag, Hamburg, 218.

[38] Barukčić, I. (1997) Die Kausalität. Scientia, Wilhelmshaven, 374.

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

[40] Barukčić, I. (2006) Causality. New Statistical Methods, 2nd English Edition, Books on Demand, Hamburg-Norderstedt, 488.

[41] Barukčić, I. (2006) New Method for Calculating Causal Relationships. *Proceeding of XXIIIrd International Biometric Conference*, 16-21 July 2006, McGill University, Montréal, Québec, Canada, 49.

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

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

[44] Barukčić, I. (2012) The Deterministic Relationship between Cause and Effect. *International Biometric Conference*, 26-31 August 2012, Kobe, Japan. https://www.biometricsociety.org/conference-abstracts/2012/programme/p1-5/P-1/249-P-1-30.pdf

[45] 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

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

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

[48] 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

[49] 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

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

[51] Barukčić, I. (2017) Human Papilloma Virus: A Cause of Malignant Melanoma. *viXra*, **10**, 1-13. http:///vixra.org/abs/1710.0311

[52] Barukčić, I. (2017) Human Papillomavirus: A Cause of Human Prostate Cancer. *viXra*, **11**, 1-57. http://vixra.org/abs/1711.0437

[53] Barukčić, I. (2017) Human Papillomavirus: The Cause of Human Cervical Cancer. *viXra*, **11**, 1-56. http://vixra.org/abs/1711.0339

[54] Barukčić, I. (2017) *Helicobacter pylori* Is the Cause of Human Gastric Cancer. *viXra*, **12**, 1-50. http://vixra.org/abs/1712.0018

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

[56] 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

[57] 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://doi.org/10.1093/biomet/26.4.404

[58] 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://doi.org/10.2307/2685469

[59] 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://doi.org/10.2307/2685046

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

[61] 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://doi.org/10.1001/jama.1983.03330370053031

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

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

[64] Fisher, R.A. (1922) On the Interpretation of X^2 from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society*, **85**, 87-94. https://doi.org/10.2307/2340521

[65] Statisitisches Bundesamt (2016) Anzahl der Todesfälle durch Ertrinken in Deutschland von 1993 bis 2016. https://de.statista.com/statistik/daten/studie/5256/umfrage/anzahl-der-jaehrlichen-todesfaelle-durch-ertrinken/

[66] Statisitisches Bundesamt (2016) Sterbefälle, Lebenserwartung. https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Sterbefaelle/Sterbefaelle.html

[67] Pearson, K. (1900) 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*, **50**, 157-175. http://doi.org/10.1080/14786440009463897

[68] Pearson, K. (1896) VII. Mathematical Contributions to the Theory of Evolution.-III. Regression, Heredity, and Panmixia. *Philosophical Transactions of the Royal Society A*, **187**, 253-318. https://doi.org/10.1098/rsta.1896.0007

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

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

[71] Mackie, J. (1965) Causes and Conditions. *American Philosophical Quarterly*, **2**, 245-264. http://joelvelasco.net/teaching/5330(fall2013)/mackie65-causesconditions.pdf

[72] Kirsch, I. and Weixel, L.J. (1988) Double-Blind Versus Deceptive Administration of a Placebo. *Behavioral Neuroscience*, **102**, 319-323. https://doi.org/10.1037/0735-7044.102.2.319

[73] Kirsch, I. (2000) Are Drug and Placebo Effects in Depression Additive? *Biological Psychiatry*, **47**, 733-735. https://doi.org/10.1016/S0006-3223(00)00832-5

[74] Belot, G. (2001) The Principle of Sufficient Reason. *The Journal of Philosophy*, **98**, 55-74. https://doi.org/10.2307/2678482