Glyphosate and Non-Hodgkin lymphoma: No causal relationship

Ilija Barukčić[1]

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

Correspondence: Ilija Barukčić, Horandstrasse, DE-26441 Jever, Germany. Tel: 0049-(0)4466-333. E-mail: Barukcic@t-online.de

Received: May 19, 2019; Accepted: May 19, 2019; Published: May 19, 2019

Abstract

Objective: Herbicides are used worldwide by both residential and agricultural users. Due to the statistical analysis of some epidemiologic studies the International Agency for Research on Cancer classified the broad-spectrum herbicide glyphosate (GS) in 2015, as potentially carcinogenic to humans especially with respect to non-Hodgkin lymphoma (NHL). In this systematic review and re-analysis, the relationship between glyphosate and NHL was re- investigated.

Methods: A systematic review and re-analysis of studies which investigated the relationship between GS and NHL was conducted. The method of the conditio sine qua non relationship, the method of the conditio per quam relationship, the method of the exclusion relationship and the mathematical formula of the causal relationship k were used to proof the hypothesis. Significance was indicated by a p-value of less than 0.05.

Results: The studies analyzed do not provide any direct and indirect evidence that NHL is caused GS.

Conclusion: In this re-analysis, no causal relationship was apparent between glyphosate and NHL and its subtypes.

Keywords: Glyphosate, Non-Hodgkin lymphoma, no causal relationship

1. Introduction

Historically, Marcell Malpighi (1628-1694) described in 1666 as one of the first authors Hodgkin lymphoma (HL) in his publication: *De viscerum structura exercitatio anatomica* (Malpighi, 1666). Centuries later, the English physician Thomas Hodgkin (1798–1866) of Guy's Hospital, London, published 1832 a remarkable paper entitled as "On some morbid cases of the absorbent glands and spleen" (Hodgkin, 1832) and described a new disease, in medical literature known through the use of the term 'Hodgkin's disease' (Wilks, 1865). Lymphomas are traditionally divided into non-Hodgkin lymphoma and Hodgkin's lymphoma, which are responsible for about 10% of all lymphomas (Armitage, Gascoyne, Lunning, & Cavalli, 2017) and known since centuries too. Independently of Hodgkin, the non-Hodgkin lymphoma i. e. leukaemia were described by Virchow (Virchow, 1845), Bennett (Bennett, 1845) and by Cohnheim (Cohnheim, 1865) under the descriptive term `pseudoleukaemia. Non-Hodgkin lymphoma (NHL) is a group of blood cancers with a wide range of histological appearances and clinical features at presentation which includes all different types of lymphoma but Hodgkin's lymphomas. The first systematic and widely accepted classification of lymphomas other than Hodgkin was proposed by Henry Rappaport in 1956 (Rappaport, 1966). Meanwhile, NHL is the leading hematological malignancy worldwide. Non-Hodgkin lymphoma (also known as non-Hodgkin's lymphoma, NHL, or sometimes just lymphoma) starts when white blood cells called (B- or T-) lymphocytes begin to grow out of control. NHL can start anywhere in the body but is usually found in lymph nodes or other lymph tissues (spleen, bone marrow, thymus, adenoids and tonsils, digestive tract). Several NHL risk factors like age, gender, family history, weakened immune system, radiation exposure, exposure to certain chemicals and drugs and glyphosate too have been discussed in literature, but the cause or a cause of NHL has not been identified. Finally, in 2015, the International Agency for Research on Cancer (IARC, 2017) Working Group published limited evidence of increased risk of non-Hodgkin lymphoma (NHL) in some

epidemiologic studies. Glyphosate [N-(phosphonomethyl)glycine], sold in the commercial as Roundup (R) (Monsanto Company, St. Louis, MO), was registered in the U.S. in 1974 and re-registrated 1993 by the US Environmental Protection Agency (EPA, 1993). Since its introduction in the 1970s Glyphosate has been frequently (Williams, Kroes, & Munro, 2000) used in forestry, in cropland and noncropland areas like gardens and lawns et cetera to control vegetation. Especially after genetically engineered glyphosate-tolerant crops were introduced, the use of glyphosate increased dramatically in the late-1990s and 2000s. Glyphosate inhibits the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (Steinrücken & Amrhein, 1980), which is responsible via a mechanism specific to plants for the biosynthesis of aromatic amino acids like phenylalanine, tyrosine, and tryptophan. Questions regarding the safety of glyphosate, its major breakdown product aminomethylphosphonic acid (AMPA) and the predominant surfactant polyethoxylated tallow amine (POEA) have been periodically raised (Olorunsogo, Bababunmi, & Bassir, 1979) (Hietanen, Linnainmaa, & Vainio, 1983) (Yousef et al., 1995) (Bolognesi et al., 1997) (Lioi, Scarfi, et al., 1998) (Lioi, Scarfi, et al., 1998) (Peluso, Munnia, Bolognesi, & Parodi, 1998) (Walsh, McCormick, Martin, & Stocco, 2000) (Daruich, Zirulnik, & Gimenez, 2001) (El-Demerdash, Yousef, & Elagamy, 2001) raised. In the following, different studies have been conducted by several regulatory agencies and scientific institutions worldwide to re-evaluate the relationship between glyphosate and some parameters. Glyphosate had no effects on fertility or reproductive parameters, there was no convincing evidence for direct DNA damage in vitro or in vivo, and neither AMPA nor glyphosate bioaccumulates in any animal tissue (Williams et al., 2000). Nevertheless, the question whether Glyphosate does pose a health risk to humans has not been finally answered. Thus far, considering use of glyphosate in both the United States and the rest of the world, an ongoing risk assessment is necessary. Here we have re-investigated the relationship between GS and NHL by some new statistical methods.

2. Material and Methods

In one way or another, testing hypotheses and theories about the natural world is not completely free of errors. Still, when all goes well, systematic observation and experimentation should assure that different scientists at different times and places are able to generate the same scientific knowledge.

2.1 Definitions

Definition 2.1.1. (The sample space)

Let the sample space denote a set or a collection of all different possible outcomes of an experiment. Each possible single outcome x_t of the experiment is said to be a member of the sample space, or to belong to the space S(X). A single outcome x_t of an experiment S(X) is a member of S(X) and denoted symbolically by the relation $x_t \in S(X)$. A set Y is contained in another set X if every element of the set Y also belongs to the set X. This relation is expressed symbolically by the expression $Y \subset X$, which is the set-theoretic expression for saying that Y is a subset of X. A subset of X that contains no elements is called an empty set, or null set, and it is denoted by the symbol $\emptyset$. In a given experiment, a number $p(x_t)$ is assigned to each event x_t in the sample space S which indicates the probability that x_t will occur. If the event x_t is certain to occur, then the probability of that event is $p(x_t)=1$.

Definition 2.1.2. (Independence)

Let A_t denote random variable at a Bernoulli trial (period of time) t. Let B_t denote another random variable at the same Bernoulli trial (period of time) t. Let $p(A_t)$ denote the probability of A_t. Let $p(B_t)$ denote the probability of B_t. Let $p(A_t \cap B_t)$ denote the joint probability of A_t and B_t. In the case of independence (Moivre, 1718) (Kolmogoroff, 1933) of A_t and B_t it is generally valid that

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

Definition 2.1.3. (A two-way or contingency table)

In this context, let us define that $p(A_t) = p(a_t)+p(b_t)$ or $p(A_t) = p(A_t \cap B_t)+ p(b_t)$ or $p(A_t) = p(A_t \cap B_t)+p(A_t \cap \underline{B}_t)$ while $p(A_t)$ is not identical with $p(a_t)$. Thus far, it is $p(B_t) = p(a_t)+p(c_t)$ or $p(B_t) = p(A_t \cap B_t) +p(c_t)$ and equally $p(\underline{B}_t) = 1- p(B_t)$ or $p(\underline{B}_t) = p(b_t)+p(d_t)$. Since the joint probability of A_t and B_t is denoted in general by $p(A_t \cap B_t)$, it is $p(A_t \cap B_t) = p(A_t) - p(b_t)$ or $p(A_t \cap B_t)=p(B_t) - p(c_t)$ or $p(B_t) + p(b_t) - p(c_t) = p(B_t) + p(\Lambda_t) = p(A_t)$. There may exist circumstances where Λ_t is identical or associated with Einstein's cosmological 'constant'. In general, it is $p(a_t)+p(c_t)+p(b_t)+p(d_t) = +1$. The following table may show the relationship in more details.

Table 1. The probabitlities of a contingency table

		Conditioned B		
		Yes = +1	No = +0	Total
Condition A	Yes =+1	$p(a_t)$	$p(b_t)$	$p(A_t)$
	No = +0	$p(c_t)$	$p(d_t)$	$p(\underline{A}_t)$
	Total	$p(B_t)$	$p(\underline{B}_t)$	1

Consider the case of Bernoulli trials (period of time) with probability $p(a_t)$ for success. Let $a_t = 1$ if the t-th outcome is a success and 0 if it is a failure. Then $a = (a_1 + a_2 + ... + a_n)$ is the number of successes in n trials (period of time) t. It is $p(a_t)= p(A_t \cap B_t)$ the joint probability of A_t and B_t and

$$a \equiv (a_1 + a_2 + \cdots + a_n) \equiv \sum_{t=1}^{t=n} a_t \tag{2}$$

Let $b_t = 1$ if the t-th outcome is a success and 0 if it is a failure. Then $b = (b_1 + b_2 + ... + b_n)$ is the number of successes in n Bernoulli trials (period of time) t. It is $p(b_t)= p(A_t \cap \underline{B}_t)$ the joint probability of (A_t and $\underline{B}_t$) and

$$b \equiv (b_1 + b_2 + \cdots + b_n) \equiv \sum_{t=1}^{t=n} b_t \tag{3}$$

Let $c_t = 1$ if the t-th outcome is a success and 0 if it is a failure. Then $c = (c_1 + c_2 + ... + c_n)$ is the number of successes in n Bernoulli trials (period of time) t. It is $p(c_t)= p(\underline{A}_t \cap B_t)$ the joint probability of ($\underline{A}_t$ and B_t) and

$$c \equiv (c_1 + c_2 + \cdots + c_n) \equiv \sum_{t=1}^{t=n} c_t \tag{4}$$

Let $d_t = 1$ if the t-th outcome is a success and 0 if it is a failure. Then $d = (d_1 + d_2 + ... + d_n)$ is the number of successes in n Bernoulli trials (period of time) t. It is $p(d_t)= p(\underline{A}_t \cap \underline{B}_t)$ the joint probability of ($\underline{A}_t$ and $\underline{B}_t$) and

$$d \equiv (d_1 + d_2 + \cdots + d_n) \equiv \sum_{t=1}^{t=n} d_t \tag{5}$$

Let A denote another binomial random variable with the probability $p(A_t)$. It is $A_t = (a_t + b_t)$ at the same Bernoulli trial (period of time) t and

$$A \equiv \left((a_1 + b_1) + (a_2 + b_2) + \cdots + (a_n + b_n)\right) \equiv \sum_{t=1}^{t=n} A_t \tag{6}$$

Let $\underline{A}$ denote the complementary random variable of the binomial random variable A with the probability $p(\underline{A}_t)$. It is $\underline{A}_t = (c_t + d_t)$ at the same Bernoulli trial (period of time) t and

$$\underline{A} \equiv \left((c_1 + d_1) + (c_2 + d_2) + \cdots + (c_n + d_n)\right) \equiv \sum_{t=1}^{t=n} \underline{A}_t \tag{7}$$

Let B denote another binomial random variable with the probability $p(B_t)$. It is $B_t = (a_t + c_t)$ at the same Bernoulli trial (period of time) t and

$$B \equiv \left((a_1 + c_1) + (a_2 + c_2) + \cdots + (a_n + c_n)\right) \equiv \sum_{t=1}^{t=n} B_t \quad (8)$$

Let $\underline{B}$ denote the complementary random variable of the binomial random variable B with the probability $p(\underline{B}_t)$. It is $\underline{B}_t = (c_t + d_t)$ at the same Bernoulli trial (period of time) t and

$$\underline{B} \equiv \left((b_1 + d_1) + (b_2 + d_2) + \cdots + (b_n + d_n)\right) \equiv \sum_{t=1}^{t=n} \underline{B}_t \quad (9)$$

At each Bernoulli trial it is

$$n_t \equiv (a_t + b_t + c_t + d_t) \equiv A_t + \underline{A}_t \equiv B_t + \underline{B}_t \quad (10)$$

and the sample size n itself equal to

$$n \equiv \sum_{t=1}^{n} (a_t + b_t + c_t + d_t) \equiv \sum_{t=1}^{n} A_t + \underline{A}_t \equiv \sum_{t=1}^{n} B_t + \underline{B}_t \quad (11)$$

The meaning of the abbreviations a, b, c, d, n et cetera are explained by following 2 by 2-table.

Table 2. The sample space of a contingency table

		Conditioned B (Outcome)		
		Yes = +1	No = +0	Total
Condition A (risk factor)	Yes =+1	a	b	A
	No = +0	c	d	$\underline{A}$
	Total	B	$\underline{B}$	n

Definition 2.1.4. (Index of unfairness)

The index of unfairness (IOU) is defined as

$$IOU \equiv \left(\left(\frac{A + B}{n}\right) - 1\right) \quad (12)$$

The range of A is $0 \leq A \leq n$, while the range of B is $0 \leq B \leq n$. A study design based on A=B=0 leads to an index of unfairness of IOU = (((0+0)/n)-1) = -1. A study design which demands that A=B=n leads to an index of unfairness of IOU = (((n+n)/n)-1) = +1. The index of unfairness is of use and valid too, if data are investigated for a causal relationship k. In particular, the range of the index of unfairness is [-1;+1]. In this context let us define the following.

Let IOU = 0 denote a fair study design

Let $0 < |IOU| \leq 0.25$ denote an unfair study design

Let $0.25 < |IOU| \leq 0.5$ denote a very unfair study design.

Let $0.5 < |IOU| \leq 0.75$ denote a highly unfair study design.

Let $0.75 < |IOU| \leq 1$ denote an extremely unfair study design.

Definition 2.1.5. (Sufficient condition)

The definition of the sufficient condition relationship can be found in literature. The mathematical formula of the sufficient (I. Barukčić, 2018d, 1989, 2017, 2018b, 2018c, 2018a, 2018b, 2019a, 2019b; K. Barukčić & Barukčić, 2016) condition relationship of a population is defined as

$$\begin{aligned} p(A_t \rightarrow B_t) &\equiv \frac{(a\ ) + (c\ ) + (d\ )}{n} = 1 \\ &\equiv p(a_t) + p(c_t) + p(d_t) \\ &\equiv \Big(p(B_t) + p(d_t)\Big) \\ &= \Big(\big(1 - p(A_t)\big) + p(a_t)\Big) \\ &\equiv \big(1 - p(b_t)\big) \\ &\equiv +1. \end{aligned} \tag{13}$$

Example.

If a *candle is burning* then *gaseous oxygen is present*. The following table may illustrate this relationship.

Table 3. The relationship between a burning candle and gaseous oxygen.

		Gaseous oxygen is present (Outcome)		
		Yes = +1	No = +0	Total
A candle is burning (Risk factor)	Yes =+1	1	0	A
	No = +0	1	1	A
	Total	B	B	n

Even if we are allowed to conclude that *if* a candle is burning *then* gaseous oxygen is present, a burning candle as such cannot be treated as a cause or as the cause of the presence of gaseous oxygen. In a slightly different way, a conditio per quam relationship cannot be reduced to a causal relationship, both are different.

Definition 2.1.6. (Necessary condition)

The definition of the necessary condition relationship can be found in literature too. The mathematical formula of the necessary condition relationship of a population is defined (I. Barukčić, 2018d, 1989, 2017, 2018b, 2018c, 2018a, 2018b, 2019a, 2019b; K. Barukčić & Barukčić, 2016) as

$$\begin{aligned} p(A_t \leftarrow B_t) &\equiv \frac{(a\ ) + (b\ ) + (d\ )}{n} = 1 \\ &\equiv p(a_t) + p(b_t) + p(d_t) \\ &\equiv \left(p(A_t) + p(d_t)\right) \\ &= \left(\left(1 - p(B_t)\right) + p(a_t)\right) \\ &\equiv \left(1 - p(c_t)\right) \\ &\equiv +1. \end{aligned} \tag{14}$$

Example.

Without *gaseous oxygen present* no *burning candle*. The following table may illustrate this relationship.

Table 4. The relationship between gaseous oxygen and a candle.

		A candle is burning (Outcome)		
		Yes = +1	No = +0	Total
Gaseous oxygen is present (Risk factor)	Yes =+1	1	1	A
	No = +0	0	1	$\underline{A}$
	Total	B	$\underline{B}$	n

Definition 2.1.7. (Necessary and sufficient condition)

The definition of the sufficient condition relationship, the necessary condition and the exclusion relationship can be found in literature. The concept of necessary and sufficient conditions, like other fundamental concepts, is determined by its own parts too, the necessary conditions and the sufficient conditions, which are under some circumstance's converses of each other. An event A_t which is a necessary and sufficient condition of another event B_t, is more than just a necessary condition of an event B_t. The same event A_t is equally more than just a sufficient condition, sometimes referred to as *material implication*, of the same event B_t. Such an event A_t is at the same Bernoulli trial *t*, both, a sufficient and a necessary condition of an event B_t. The account of necessary and sufficient conditions just outlined before is in contrast to the well-known and premature insight of J. L. Mackie that causes are at least INUS conditions, that is, "the so-called cause is, and is known to be, an *insufficient* but *necessary* part of a condition which is itself *unnecessary* but *sufficient* for the result" (Mackie, 1965). In a slightly different way, besides of Mackie's premature generalization and undeniably an oversimplification of the necessary and sufficient condition relationship, how then, can such a necessary and sufficient condition be mathematized? In this respect, let an event A_t with its own probability $p(A_t)$ at the same (period of) time *t* be a necessary and sufficient condition

for another event B_t with its own probability $p(B_t)$. In other words, *without* A_t *no* B_t or the absence of A_t guarantees the absence of B_t and in the same respect *if* A_t is given *then* B_t is given too. The mathematical formula of the necessary and sufficient condition relationship of a population is defined as

$$\begin{aligned} p(A_t \leftrightarrow B_t) &\equiv \frac{(a\ ) + (d\ )}{n} = 1 \\ &\equiv p(a_t\ ) + p(d_t\ ) \\ &\equiv \Big(p(A_t\ ) - p(b_t\ )\Big) + \Big(\big(1 - p(A_t\ )\big) - p(c_t\ )\Big) \\ &\equiv \Big(1 - p(b_t\ ) - p(c_t\ )\Big) \\ &\equiv +1. \end{aligned} \tag{15}$$

Definition 2.1.8. (Either A_t or B_t relationship)

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

$$\begin{aligned} p(A_t >-< B_t) &\equiv \frac{(b\ ) + (c\ )}{n} = 1 \\ &\equiv p(b_t\ ) + p(c_t\ ) \\ &\equiv \Big(p(A_t\ ) - p(a_t\ )\Big) + \Big(\big(1 - p(A_t\ )\big) - p(d_t\ )\Big) \\ &\equiv \Big(1 - p(a_t\ ) - p(d_t\ )\Big) \\ &\equiv +1. \end{aligned} \tag{16}$$

Definition 2.1.9. (The Chi-square goodness-of fit test)

A Chi-Square goodness-of fit test is one of commonly used methods of statistical inference and was originally proposed by Karl Pearson (Karl Pearson, 1900). Given some conditions (simple random sampling, categorical random variable, expected value of the number of sample observations is at least 5 et cetera), the chi-square goodness of fit test can be applied to determine whether (sample distribution) data observed are consistent with (theoretical distribution) hypothesized data. The degrees of freedom (d.f.) of a chi-square goodness of fit test is equal to the number of levels (k) of the categorical variable minus 1. In general, the chi-square goodness of fit test is given by

$$X^2 \equiv \sum_{t=1}^{k} \frac{\Big((x_t\ ) - \big(n \times p(x_t\ )\big)\Big)^2}{\big(n \times p(x_t\ )\big)} \tag{17}$$

Example.

Suppose, a coin, assumed to be fair, is tossed 100 times with the results given in **Table 5**.

Table 5. A fair coin.

Event	Observed (x_t)	Expected ($n \times p(x_t)$)	((x_t)- ($n \times p(x_t)$))	($((x_t)$- ($n \times p(x_t)))^2$)/ ($n \times p(x_t)$)
Heads	40	50	-10	$(-10)^2/50 = 2$
Tails	60	50	+10	$(+10)^2/50 = 2$
n	100	100		$X^2 = 4$

In this context, the chi-square goodness of fit test (Sachs, 1992), p. 421 requires to state a null hypothesis (H_0) and an alternative hypothesis (H_A) too. In point of fact, it is **p=p(Heads) and q=p(Tails)** and **(p +q) = 1** or **(p(Heads) + p(Tails)) = 1 or p(Tails) = 1 – p(Heads).** In our present case ($\alpha = 0.05$), for a chi-square goodness of fit test of this example, the hypotheses take the following form.

Null hypothesis: The data are consistent with a specified distribution or p(Heads) =0.5

The null hypothesis claims equally that p(Heads) = 1 –p(Tails) = 0.5

Alternative hypothesis: The data are not consistent with a specified distribution.

The Null hypothesis is not true.

The value of the test statistics as calculated before is

$$X^2 \equiv \sum_{t=1}^{k} \frac{\left((x_t) - \left(n \times p(x_t)\right)\right)^2}{\left(n \times p(x_t)\right)} = \frac{(40-50)^2}{50} + \frac{(60-50)^2}{50} = = \frac{100}{50} + \frac{100}{50} = 2 + 2 = 4 \quad (18)$$

with d. f. = k-1=2-1 = 1. Unfortunately, the p-value of $X^2=4$ is less than the significance level (0.05). We accept the alternative hypothesis and reject the null-hypothesis. The sample data do not provide support for the hypothesis that the coin tossed is fair. In general, it is not necessary that p = q, to be able use the chi square goodness-of fit test which is the mathematical the foundation of the chi square goodness of fit test of the necessary condition, of a sufficient condition et cetera with d. f. = k-1=2-1 = 1.

Definition 2.1.10. (The X^2 Test of Goodness of Fit of a Sufficient Condition)

A random sample of observations can come from a particular distribution (sufficient condition distribution) but must not. The X^2 test of goodness-of-fit is an appropriate method for testing the *null hypothesis* that *a random sample of observations comes from a specific distribution* (i.e. the distribution of a sufficient condition) against the *alternative hypothesis* that *the data have some other distribution*. The additive property of X^2 distribution may sometimes be used as an additional test of significance. In this case, the continuity correction should be omitted from each X^2 value. Under conditions where the chi-square goodness of fit test cannot be used it is possible to use an approximate and conservative (one sided) confidence interval known as the rule of three. The X^2 distribution is a particular type of a gamma distribution and widely applied in the field of mathematical statistics. The applicability of using the Pearson chi-squared statistic in cases where the cell frequencies of a 2× 2 contingency table are not greater than five is widely discussed (Fisher, 1922) in literature and the use of Yate's continuity correction (Yates, 1934) is proposed. However, studies provided evidence that incorporating Yate's continuity correction is not essential (Grizzle, 1967) (Conover, 1974). Still, using the *continuity correction*, the chi-square value of a conditio per quam relationship is derived (I. Barukčić, 2019b) as

$$X^2\left((A \rightarrow B)|A\right) \equiv \frac{\left((b) - (1/2)\right)^2}{A} + 0 = 0 \tag{19}$$

or alternatively as

$$X^2\left((A \rightarrow B)|\underline{B}\right) \equiv \frac{\left((b) - (1/2)\right)^2}{\underline{B}} + 0 = 0 \tag{20}$$

Definition 2.1.11. (The X² Test of Goodness of Fit of a Necessary Condition)

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

$$X^2\left((A \leftarrow B)|B\right) \equiv \frac{\left((c) - (1/2)\right)^2}{B} + 0 = 0 \tag{21}$$

Depending upon the study design, another method to calculate the chi-square value of a conditio sine qua non distribution (while using the continuity correction) is defined as

$$X^2\left((A \leftarrow B)|\underline{A}\right) \equiv \frac{\left((c) - (1/2)\right)^2}{\underline{A}} + 0 = 0 \tag{22}$$

Definition 2.1.12. (The X² Test of Goodness of Fit of the Exclusion Relationship)

The chi square value with degree of freedom 2-1=1of the exclusion relationship *with a continuity correction* can be calculated as

$$X^2\left((A|B)|A\right) \equiv \frac{\left((a) - (1/2)\right)^2}{A} + 0 = 0 \tag{23}$$

Depending upon the study design, another method to calculate the chi-square value of the exclusion relationship is defined as

$$X^2\left((A|B)|B\right) \equiv \frac{\left((a) - (1/2)\right)^2}{B} + 0 = 0 \tag{24}$$

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

Definition 2.1.13. (The Mathematical Formula of the Causal Relationship k)

The mathematical formula of the causal relationship (I. Barukčić, 2018d, 1989, 2017, 2018b, 2018c, 2018a, 2018b, 2019a, 2019b; K. Barukčić & Barukčić, 2016) k is defined at every single event, at every single Bernoulli trial t, as

$$k(A_t, B_t) \equiv \frac{p(A_t \cap B_t) - \left(p(A_t) \times p(B_t)\right)}{\sqrt[2]{p(A_t) \times \left(1 - p(A_t)\right) \times p(B_t) \times \left(1 - p(B_t)\right)}} \tag{25}$$

where A_t denotes the cause and B_t denotes the effect. Under some certain circumstances, the chi-square distribution can be applied to determine the significance of causal relationship k. Pearson's concept of correlation is not identical with causation. Causation as such is not identical with correlation. This has been proved many times and is widely discussed in many publications.

Definition 2.1.14. (The 95% Confidence Interval of the Causal Relationship k)

A confidence interval (CI) of the causal relationship k calculated from the statistics of the observed data can help to estimate the true value of an unknown population parameter with a certain probability. In the following, let the sample mean S be

$$\left\{S = \overline{k(A_t,B_t)} = \frac{k(A_1,B_1)+k(A_2,B_2)+\cdots+k(A_n,B_n)}{n} = \frac{\sum_{t=1}^{n} k(A_i,B_i)}{n}\right\} \quad (26)$$

The causal relationship $k(A_t,B_t)$ at every single Bernoulli trial is Bernoulli(p) distributed. In this context, we consider that $E(k(A_t,B_t)) = (1\times p(k(A_t,B_t))) + (0\times(1- p(k(A_t,B_t))) = p(k(A_t,B_t))$ where $E(k(A_t,B_t))$ denotes the expected value of $k(A_t,B_t)$. Thus far, it is

$$\left\{E(S) = p\big(k(A_t,B_t)\big) \quad and \quad \sigma(S)^2 = \frac{p\big(k(A_t,B_t)\big)\times\Big(1-p\big(k(A_t,B_t)\big)\Big)}{n}\right\} \quad (27)$$

where $\sigma(S)^2$ denotes the variance of the sampling distribution of $p(k(A_t,B_t))$. When the sample size is not too small, the central limit theorem based normal approximation can be used to estimate the confidence interval (CI) as it is

$$\left\{p\big(k(A_t,B_t)\big) \mp \left(Z\times\sqrt[2]{\frac{p\big(k(A_t,B_t)\big)\times\Big(1-p\big(k(A_t,B_t)\big)\Big)}{n}}\right) = p\big(k(A_t,B_t)\big) \mp \left(\sqrt[2]{\frac{Z^2}{n}\times p\big(k(A_t,B_t)\big)\times\Big(1-p\big(k(A_t,B_t)\big)\Big)}\right)\right\} \quad (28)$$

where $p(k(A_t,B_t))$ denotes the proportion of successes in a Bernoulli trial process and Z is the $(1–(\alpha/2))$ quantile of a standard normal distribution. For a 95% confidence level Z is about $Z \sim 1.96$. For an unknown standard deviation the Student's t distribution t can be used as the critical value. Still, it is known that $\sigma(S)^2$ has the maximum value $(1/(4\times n))$ when p=1/2 and approximately we obtain

$$\left\{p\big(k(A_t,B_t)\big) \mp \left(\sqrt[2]{\frac{Z^2}{n\times 4}}\right) = p\big(k(A_t,B_t)\big) \mp \left(\sqrt[2]{\frac{1.96^2}{n\times 4}}\right) \approx p\big(k(A_t,B_t)\big) \mp \left(\sqrt[2]{\frac{1}{4}}\right)\right\} \quad (29)$$

The proposed approximation is of use even under circumstances where p(…) = 0.9999 … 999 ~ p=1. In this context, we obtain the critical value $p_{critical}$ approximately as $p_{critical} = 1 – (1/(n))^{1/2}$. In particular, the concept of Chebyshev's inequality is profound because the same inequality is true for every distribution even if the distribution isn't normal.

Thus far, Chebyshev's inequality allows calculating the 95% confidence of the causal relationship k and so by the Chebyshev inequality it is

$$p\left\{p\big(k(A_t,B_t)\big) - \left(c\times\sqrt[2]{\sigma(S)^2}\right) < S < p\big(k(A_t,B_t)\big) + \left(c\times\sqrt[2]{\sigma(S)^2}\right)\right\} \geq 1-\frac{1}{c^2} \quad (30)$$

were the right side has the value 0.95 when $c = (20)^{1/2}$. This is the case since $(1-(1/c^2))=0.95$ or $0.05 = (1/c^2)$ or $c^2= (1/0.05)$ or $c^2 = (100/5)$ or $c^2 = 20$ or $c = (20)^{1/2}$. Thus far, if S does lie in the interval

$$\left\{p\left(k(A_t,B_t)\right) \;-\; \left(\sqrt[2]{20\times\sigma(S)^2}\right) \;,\; p\left(k(A_t,B_t)\right) \;+\; \left(\sqrt[2]{20\times\sigma(S)^2}\right)\right\} \tag{31}$$

then $p(k(A_t,B_t))$ itself must be in the interval

$$\left\{S \;-\; \left(\sqrt[2]{20\times\sigma(S)^2}\right) \;,\; S \;+\; \left(\sqrt[2]{20\times\sigma(S)^2}\right)\right\} \tag{32}$$

which is equally the 95% confidence interval for an unknown parameter $p(k(A_t,B_t))$. Again, $\sigma(S)^2$ has the maximum value $(1/(4\times n))$ when p=1/2, so we have

$$\left\{S \;-\; \left(\sqrt[2]{\frac{20\times 1}{4\times n}}\right) \;,\; S \;+\; \left(\sqrt[2]{\frac{20\times 1}{4\times n}}\right)\right\} \tag{33}$$

or **the 95% interval for the causal relationship k** approximately as

$$\left\{k\left(A_t\,,B_t\,\right)-\sqrt[2]{\frac{5}{N}} \;,\; k\left(A_t\,,B_t\,\right)+\sqrt[2]{\frac{5}{N}}\right\} \tag{34}$$

Definition 2.1.15. (Hypergeometric distribution)

The hypergeometric distribution (Huygens & van Schooten, 1657) (Karl Pearson, 1899) (Gonin, 1936) is defined by the parameters population size, event count in population, sample size and can be used to calculate the exact probability of an event even for small samples which are drawn from relatively small populations, *without replacement*. The hypergeometric distribution differs to some extent from the binomial distribution. In contrast to the hypergeometric distribution, the probability of a binomially distributed random variable from trial to trial is the same. While the chi square distribution is of limited value for samples drawn from relatively small populations, the hypergeometric distribution can be used to calculate the exact probabilities for samples drawn from relatively small populations and without replication and for large populations too. The probability of having randomly exactly **a** (**Table 1**) successes in *n* hypergeometric trials or the significance of the causal relationship k can be tested under conditions of sampling without replacement by the hypergeometric distribution too. The probability of having exactly *a* successes by chance in *n* hypergeometric experimental trials is given by

$$\mathrm{p}\left(X=a\;\right) \;=\; \frac{\binom{A}{a}\times\binom{n\;-\;A}{B\;-\;a}}{\binom{n}{B}} \tag{35}$$

Definition 2.1.16. (Odds ratio (OR))

The odds ratio (CORNFIELD, 1951; Edwards, 1963; Fisher, 1935; Mosteller, 1968) abbrivated as OR(A,B), is a very commonly used measure of association for 2× 2 contingency tables (**Table 1**) and given by

$$OR(A\ ,B\ ) \equiv \frac{a\ /b}{c\ /d} \equiv \frac{a\ \times d}{c\ \times b} \tag{36}$$

In addition, researchers are regularly relying on Odds ratio to gain some new knowledge. Still, we need to address some different aspect of Odds ratio itself to find out the straightforward contradictions and the deep theoretical inconsistency which is associated with Odds ratio. It turns out that we are ill-advised if we believe blindly, uncritically in Odds ratio. More likely, the Odds ratio (OR) is nothing more but *Yule's coefficient of association* (Yule, 1900) *Q*(A,B) re-written (Warrens, 2008) in a non-normalized form and given by

$$\begin{aligned} Q(A\ ,B\ ) &\equiv \frac{OR(A\ ,B\ )-1}{OR(A\ ,B\ )+1} \\ &\equiv \frac{\left(\frac{a\ \times d}{b\ \times c}\right)-1}{\left(\frac{a\ \times d}{b\ \times c}\right)+1} \\ &\equiv \frac{\left(\frac{(a\ \times d\ )-(b\ \times c\ )}{(b\ \times c\ )}\right)}{\left(\frac{(a\ \times d\ )+(b\ \times c\ )}{(b\ \times c\ )}\right)} \\ &\equiv \frac{\left((a\ \times d\ )-(b\ \times c\ )\right)}{\left((a\ \times d\ )+(b\ \times c\ )\right)} \end{aligned} \tag{37}$$

Under conditions where Yule's coefficient of association Q=0, there is no association. Although severely and justifiably criticized especially by Karl Pearson (1857–1925), the long-time and rarely challenged leader of statistical science and Heron (K. Pearson & Heron, 1913), Odds ratio is still regularly referred to. The standard error and 95% confidence interval of the Odds ratio (OR) can be calculated according to Altman (Altman, 1999). Given the severely limited character of odds ratio, the standard error of the log Odds ratio is calculated as

$$SE\left(ln\left(OR(A\ ,B\ )\right)\right) \equiv \sqrt[2]{\left(\frac{1}{a}\right)+\left(\frac{1}{b}\right)+\left(\frac{1}{c}\right)+\left(\frac{1}{d}\right)} \tag{38}$$

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

$$95\ \%\ CI \equiv exp\left(ln\left(OR(A\ ,B\ )\right)-\left(1.96\times SE\left(ln\left(OR(A\ ,B\ )\right)\right)\right)\right)$$

$$to \tag{39}$$

$$exp\left(ln\left(OR(A\ ,B\ )\right)+\left(1.96\times SE\left(ln\left(OR(A\ ,B\ )\right)\right)\right)\right)$$

Definition 2.1.17. (The unknown population proportion π_{upper})

Tests of hypotheses concerning the sampling distribution of the sample proportion **p** (i. e. conditio sine qua non p(SINE), conditio per quam p(IMP) et cetera) can be performed using the normal approximation. The calculation of the rejection region based on the sample proportion to construct a confidence interval for an unknown population proportion π_{upper} can be performed under conditions of sampling without replacement (Sachs, 1992) by the formula

$$p_{critical\ upper} = \left(p - \frac{1}{2 \times n}\right) - \left(Z \times \sqrt[2]{\left(\frac{p \times (1-p)}{n}\right) \times \left(\frac{N-n}{N-1}\right)}\right) \quad (40)$$

while the term ((N-n)/(N-1)) denotes *the finite population correction* (Isserlis, 1918).

Definition 2.1.18. (The Chi Square Distribution)

The following critical values of the chi square distribution as visualized by **Table 6** are used in this publication.

Table 6. The critical values of the chi square distribution (degrees of freedom: 1)

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

2.2 Material

2.2.1 Search Strategy

Systematic reviews and meta-analyses are becoming increasingly more important. To answer the questions addressed in this paper, the literature search, the collection and analyzes of data, the flow of information through the different phases of a systematic review was supported by Preferred Reporting Items for Systematic Reviews and Meta - analysis (PRISMA) (Liberati et al., 2009; Moher, Liberati, Tetzlaff, & Altman, 2009). The screening process and results are shown in Fig. 1.

To answer the questions addressed in this paper, the electronic database PubMed was searched for appropriate studies conducted in any country which investigated the relationship between glyphosate and NHL. The search in PubMed was performed while using some medical key words. The articles found where saved as a *.txt file while using the support of PubMed. The created *.txt file was converted into a *.pdf file. The abstracts where studied within the *.pdf file. Those articles were considered for a re-view which provided access to data without any data access barrier. Additionally, the reference list of identified articles was used as a potential source of articles appropriate for this study.

1. Identification of records	Size	Total
Records identified by searching in the databases		
PubMed	9	
Google Scholar	0	
Web of Science	0	
Additional records identified from other sources	2	11
2. Clean-up of search (Screening)		
Records removed after verifying duplication	0	
Records excluded by title	2	
Records excluded due to other reasons	2	
(Articles outside the inclusion criteria)		
3. Eligibility		
Articles evaluated for eligibility		7
Articles excluded for various reasons		
- Language	0	
- Data access barriers	0	
4. Included		
Articles included in the meta-analysis		7

Figure 1.

Flow Diagram of the article selection process. Adopted from PRISMA 2009 (Liberati et al., 2009; Moher et al., 2009).

The study of (L. Hardell & Eriksson, 1999) published (4/404) positive cases and (3/741) positive controls but was not considered for a re-analyses. The data of this study are extremely self-contradictory. The index of unfairness is IOU = -0.64 and highly unfair. At the same time, the exclusion relationship between GS and NHL is positive (p (EXCL) = 0,99650655, X^2 (EXCL) =0,04 and X^2 (EXCL) =2,29) while equally the conditio per quam relationship is significant too (p (IMP) =0,997379913. X^2 (IMP) =0,01. X^2 (IMP) =1,29). This is a contradiction. Mathematically, it is not possible **GS excludes NHL** and at the same time that **if GS then NHL**.

Leon et al. (Leon et al., 2019) investigated the relationship of ever use of glyphosate and non-Hodgkin lymphoid malignancies (NHL) in a pooled analysis of three large agricultural worker cohorts of 316 270 farmers. A control group has not been provided. During follow-up, 2430 NHL cases were diagnosed while 1131 of these cases ever

used glyphosate. Besides of a missing control group, a fair study design assumed, it is possible to calculate the significance of a conditio sine qua non relationship between GS and NHL as X^2(SINE) = ((2430-1131)*(2430-1131))/2430 = 694,41, a highly significant result. In other words, the study of Leon et al. has provided striking evidence that GS is not a necessary condition of NHL. In other words, it is possible to suffer from NHL without GS. According to Leon et al. (Leon et al., 2019) the Null-hypothesis: without GS no NHL must be rejected. The consequence is, that the use of GS must imply that people will suffer from NHL, which is not the case either.

2.2.2 Statistical Analysis

All statistical analyses were performed with Microsoft® Excel® for Mac® version 16.2 (181208) software (© 2018, Microsoft GmbH, Munich, Germany). The level of significance was set to 0.05.

3. Results

Theorem 3.1. (Glyphosate is neither a cause nor the cause of Non-Hodgkin Lymphoma.)

McDuffie et al. (McDuffie et al., 2001) conducted a Canadian multicenter population-based incident, case (n = 517)-control (n = 1506) study to investigate the putative associations of specific pesticides with non-Hodgkin's Lymphoma.

Claim.

Null Hypothesis:

Glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma. In other words, k = 0.

Alternative Hypothesis:

Glyphosate is either the cause or a cause of Non-Hodgkin Lymphoma. In other words, k >0.

Proof.

McDuffie et al. investigated the relationship between exposure to glyphosate of humans with respect to the development of Non-Hodgkin Lymphoma. The data as obtained by McDuffie et al. (McDuffie et al., 2001) are view by **table 7**.

Table 7.

The study of McDuffie et al., 2001.				
		NHL		
		YES	NO	
Glyphosate	YES	51	133	184
	NO	466	1373	1839
		517	1506	2023

Statistical analysis.	IOU =**-0,65**
Causal relationship k =+0,0156.	95% CI: -0,034 to +0,07
p-value (k | HGD) =0,05402	X^2 (k) =0,50
Odds ratio (OR) =1,13	95% CI: (0,80 -1,59)
p (SINE) =0,76965	X^2 (SINE) =420,03
p (IMP) =0,93426	X^2 (IMP) =11,75
p (SINE ^ IMP) =0,70391	X^2 (SINE ^IMP) =431,78

The study of McDuffie et al. (McDuffie et al., 2001) is potentially biased since the index of unfairness of the study IOU is IOU = -0,65 and indicates a highly unfair study design. The relative frequency of the conditio per quam relationship between GS and NHL is p (IMP) = 0,93426 and not significant ($X^2(IMP|A_t)$ = ((133×133)/184) + 0 = 96,14 or $X^2(IMP|\underline{B}_t)$ = ((133×133)/1506) + 0 = 11,746). The data of McDuffie et al. do not support the hypothesis *without* GS *no* NHL (p(SINE) = 0,76965; X^2 (SINE) = 420,03). In the same context, there is not a significant causal relationship between GS and NHL (k = +0,0156 with 95% CI: -0,034 to +0,07). The p value of the causal relationship k calculated according to the hypergeometric distribution is *p-value (k | HGD)* = 0,05402 and not significant. Thus far, according to the data of McDuffie et al., Glyphosate is neither a cause nor the cause of Non-Hodgkin Lymphoma.

Quod erat demonstrandum.

Theorem 3.2. (Glyphosate is neither a cause nor the cause of Non-Hodgkin Lymphoma.)

Hardell, Eriksson, & Nordstrom (Lennart Hardell, Eriksson, & Nordstrom, 2002) investigated the importance of glyphosate and other factors in the etiology of NHL by a pooled analysis performed on two case-control studies. Hardell, Eriksson, & Nordstrom reported that they were not able to find an association between glyphosate and non-Hodgkin lymphoma.

Claim.

Null Hypothesis:

Glyphosate is not a cause of Non-Hodgkin Lymphoma. In other words, k = 0.

Alternative Hypothesis:

Glyphosate is a cause of Non-Hodgkin Lymphoma. In other words, k >0.

Proof.

The data as obtained by Hardell, Eriksson, & Nordstrom (Hardell, Eriksson, & Nordstrom, 2002) are viewed by **table 8**.

Table 8.

The study of Hardell et al., 2002.				
		NHL		
		YES	NO	
Glyphosate	YES	8	**8**	16
	NO	507	1133	1640
		515	1141	1656

Statistical analysis.	IOU = **-0,68**
Causal relationship k = +0,0403.	95 % CI: (-0,015 : 0,10)
p-value (k | HGD) = 0,05682	X^2 (k) = 2,69
Odds ratio (OR) = 2,23	95 % CI: (-0,83 : 5,99)
p (SINE) = 0,69384	X^2 (SINE) = 499,12
p (IMP) = 0,99517	X2 (IMP) = 0,06
p (SINE ^ IMP) = 0,68901	X^2 (SINE ^IMP) = 499,18

It is highly probable that the study of Hardell, Eriksson, & Nordstrom (Hardell, Eriksson, & Nordstrom, 2002) is biased since the *index of unfairness* (I. Barukčić, 2019a) of the study is IOU = -0,68 and indicates a highly unfair study design. The relative frequency of the conditio per quam relationship between GS and NHL is p (IMP) = 0,99517 and significant (X^2(IMP|$\underline{B}_t$) = ((8×8)/1141) + 0 = **0,06**). The data of the same study support the hypothesis that there is not a significant conditio per quam relationship between GS and NHL (X^2(IMP| A_t) = ((8×8)/16) + 0 = **4,00**) too, which is a contradiction. The data of Hardell, Eriksson, & Nordstrom (Hardell, Eriksson, & Nordstrom, 2002) do not support the hypothesis *without* GS *no* NHL (p(SINE) = 0,69384; X^2 (SINE) = 499,12). In the same context, there is a significant causal relationship between GS and NHL (k = +0,04033with 95 % CI: --0,0147 to +0,10). The p value of the causal relationship k calculated according to the hypergeometric distribution is *p-value (k | HGD)* = 0,05682 and not significant. The data of Hardell, Eriksson, & Nordstrom (Hardell, Eriksson, & Nordstrom, 2002) are self-contradictory and do not support a cause effect relationship between GS and NHL.

Quod erat demonstrandum.

Theorem 3.3. (Glyphosate is not a cause of Non-Hodgkin Lymphoma.)

De Roos et al. (A. J. De Roos et al., 2003) examined whether an increased rate of non-Hodgkin's lymphoma (NHL) observed among farmers (Cantor, 1982) is due to pesticide exposures in farming. The term pesticide denotes a wide variety of chemicals used to destroy weeds (herbicides), insects (insecticides), and mold (fungicides).

Claim.

Null Hypothesis:

Glyphosate is not a cause of Non-Hodgkin Lymphoma. In other words, k = 0.

Alternative Hypothesis:

Glyphosate is a cause of Non-Hodgkin Lymphoma. In other words, k >0.

Proof.

De Roos et al. investigated the potential health effects of glyphosate in humans with respect of the development of Non-Hodgkin Lymphoma. The data as obtained by De Roos et al. (De Roos et al., 2003) are view by **table 9**.

Table 9.

The study of De Roos et al., 2003.				
		NHL		
		YES	NO	
Glyphosate	YES	36	61	97
	NO	614	1872	2486
		650	1933	2583

Statistical analysis.	IOU =-0,710801394
Causal relationship k =+0,05439.	95 % CI: (-0,01 to +0,0983)
p-value (k | HGD) =0,00254	X^2 (k) =7,641030004
Odds ratio (OR) =1,8	95 % CI: (1,18- 2,74)
p (SINE) =0,76229	X^2 (SINE) =579,9938462
p (IMP) =0,97638	X^2 (IMP) =1,924987067
p (SINE ^ IMP) =0,73868	X^2 (SINE ^IMP) =581,91

It is highly probable that the study of De Roos et al. (De Roos et al., 2003) is biased since the index of unfairness of the study is IOU = -0,710801394 and indicates a highly unfair study design. The relative frequency of the conditio per quam relationship between GS and NHL is p (IMP) = 0,97638 and significant (X^2(IMP|$\underline{B_t}$) = ((61×61)/1993) + 0 = 1,924987067). The data of the same study support the hypothesis that there is not a significant conditio per quam relationship between GS and NHL (X^2(IMP| A_t) = ((61×61)/97) + 0 = 38,36) too, which is a contradiction. The data of De Roos et al. do not support the hypothesis *without* GS *no* NHL (p(SINE) = 0,76229; X^2 (SINE) = 579,993). In the same context, there is a significant causal relationship between GS and NHL (k = +0,05439 with 95 % CI: -0,01 to +0,0983). The p value of the causal relationship k calculated according to the hypergeometric distribution is *p-value (k | HGD)* = 0,00254 and significant. Formally, according to the data of De Roos et al. it is very difficult to conclude that glyphosate is at least a cause of Non-Hodgkin Lymphoma.

Quod erat demonstrandum.

Theorem 3.4. (Glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma.)

De Roos et al. (Anneclaire J. De Roos et al., 2005) evaluated the associations between the exposure to the broad-spectrum herbicide glyphosate and cancer incidence in a prospective cohort study of 57,311 applicators in the U.S.

Claim.

Null Hypothesis:

Glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma. In other words, k = 0.

Alternative Hypothesis:

Glyphosate is either the cause or a cause of Non-Hodgkin Lymphoma. In other words, k >0.

Proof.

De Roos et al. investigated the potential health effects of glyphosate in humans with respect of the development of Non-Hodgkin Lymphoma. The data as obtained by De Roos et al. (De Roos et al., 2005) are view by **table 10**.

Table 10.

The study of De Roos et al., 2005.				
		NHL		
		YES	NO	
Glyphosate	YES	71	40964	41035
	NO	21	13259	13280
		92	54223	54315

Statistical analysis. IOU =-0,242805855

Causal relationship k =+0,00156. 95 % CI: (-0,0080 to +0,011150)

p-value (k | HGD) =0,09238 X^2 (k) =0,131558657

Odds ratio (OR) =1,09 95 % CI: (0,67 - 1,78)

p (SINE) =0,99961 X^2 (SINE) =4,793478261

p (IMP) =0,24581 X^2 (IMP) =30947,18654

p (SINE ^ IMP) =0,24542 X^2 (SINE ^IMP) =30951,98002

The study of De Roos et al. (De Roos et al., 2005) is potentially biased because the index of unfairness of the study is IOU = -0,242805855 and thus far unfair. The relative frequency of the conditio sine qua non relationship between GS and NHL with p (SINE) =0,99961 is very high, but statistically not significant ($X^2(SINE|B_t)$ = =((21×21)/92) + 0 = **4,793478261**). The data of the same study support the hypothesis that there is a conditio sine qua non relationship between GS and NHL ($X^2(SINE|\underline{A}_t)$ = ((21×21)/13280) + 0 = **0,033207831**) too, which is a contradiction. The data of De Roos et al. do not support the hypothesis *if* use of glyphosate *then* development of Non-Hodgkin Lymphoma (p(IMP) =0,24581; X^2 (IMP) = 30947,19. In the same context, the there is no causal relationship between GS and NHL (k =+0,00156 with 95% CI: -0,0080 to +0,011150). The p value of the causal relationship k calculated according to the hypergeometric distribution is *p-value (k | HGD)* = 0,09238 and not significant. In other words, Glyphosate is neither a necessary condition for the development of Non-Hodgkin Lymphoma nor a sufficient condition. Furthermore, the Null-hypothesis above cannot be rejected. According to the data of De Roos et al., glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma.

Quod erat demonstrandum.

Theorem 3.5. (Glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma.)

Eriksson et al. (Eriksson, Hardell, Carlberg, & Akerman, 2008) evaluated the associations between the exposure to the broad-spectrum herbicide glyphosate and cancer incidence in a prospective cohort study of 57,311 applicators in the U.S.

Claim.

Null Hypothesis:

Glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma. In other words, k = 0.

Alternative Hypothesis:

Glyphosate is either the cause or a cause of Non-Hodgkin Lymphoma. In other words, k >0.

Proof.

Eriksson et al. (Eriksson, Hardell, Carlberg, & Akerman, 2008) investigated the potential health effects of glyphosate in humans with respect of the development of Non-Hodgkin Lymphoma. The data as obtained by Eriksson et al. (Eriksson, Hardell, Carlberg, & Akerman, 2008) are view by **table 11**.

Table 11.

The study of Eriksson et al., 2008.				
		NHL		
		YES	NO	
Glyphosate	YES	29	**18**	47
	NO	881	998	1879
		910	1016	1926

Statistical analysis. IOU =-0,50

Causal relationship k =+0,04579. 95 % CI: (-0,0051 :0,10)

p-value (k | HGD) =0,0159 X^2 (k) =4,04

Odds ratio (OR) =1,83 95 % CI: (1,01 : 3,31)

p (SINE) =0,54258 X^2 (SINE) =852,92

p (IMP) =0,99065 X^2 (IMP) =0,32

p (SINE ^ IMP) =0,53323 X^2 (SINE ^IMP) =853,24

The study of Eriksson et al. is potentially biased because the index of unfairness of the study is IOU = -0,5. In this context, the study design is very unfair. The discrepancy between the number of cases (n=910) and the number of exposed to glyphosate (n=47) is too great. The data of Eriksson et al. do not support the hypothesis *if* use of glyphosate *then* development of Non-Hodgkin Lymphoma (p(IMP) =0,99065; X^2 (IMP|A_t) = ((18×18)/47) + 0 = **6,89** and at the same time the data of the same study do support the hypothesis that there is a significant conditio per quam relationship between GS and NHL (X^2(IMP|B_t) = ((18×18)/1016) + 0 **= 0,32**) too, which is a contradiction. The data of Eriksson et al. are self-contradictory. The same study do not support the hypothesis that there is not a conditio sine qua non relationship between GS and NHL (X^2(SINE|B_t) = ((881×881)/910) + 0 = **852,92**). In the same context, the there is a significant **positive** causal relationship k between GS and NHL (k =+0,04579 with 95 % CI: (-0,0051 : 0,10) and a hypergeometric distribution based *p-value (k | HGD)* = 0,0159). In other words, glyphosate is neither a necessary condition nor a sufficient condition for the development of Non-Hodgkin Lymphoma. Furthermore, since the data of Eriksson et al. are self-contradictory, the significant causal relationship is potentially biased and not of any use. According to the data of Eriksson et al., glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma.

Quod erat demonstrandum.

Theorem 3.6. (Glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma.)

Orsi et al. (Orsi et al., 2009) conducted a hospital-based case-control study in France between 2000 and 2004 to investigate the relationship between occupational exposure to pesticides and the risk of lymphoid neoplasms in men.

Claim.

Null Hypothesis:

Glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma. In other words, k = 0.

Alternative Hypothesis:

Glyphosate is either the cause or a cause of Non-Hodgkin Lymphoma. In other words, k >0.

Proof.

The study of Orsi et al. (Orsi et al., 2009) investigated the potential health effects of glyphosate in humans with respect of the development of Non-Hodgkin Lymphoma. The data as obtained by Orsi et al. (Orsi et al., 2009) are view by **table 12**.

Table 12.

The study of Orsi et al., 2009.				
		NHL		
		YES	NO	
Glyphosate	YES	12	**24**	36
	NO	232	412	644
		244	436	680

Statistical analysis.	IOU = -0,59
Causal relationship k = - 0,013.	95 % CI: (-0,0983 to +0,07)
p-value (k | HGD) =0,13606.	X^2 (k) = 0,11
Odds ratio (OR) =0,89.	95 % CI: (0,44 - 1,81)
p (SINE) =0,65882	X^2 (SINE) = 220,59
p (IMP) =0,96471	X^2 (IMP) = 1,32
p (SINE ^ IMP) =0,62353	X^2 (SINE ^IMP) = 221,91

The study of Orsi et al. (Orsi et al., 2009) is biased because the index of unfairness of the study is IOU = -0,59 and thus far highly unfair. The relative frequency of the conditio sine qua non relationship between GS and NHL with p (SINE) = 0,65882 is not very high and statistically not significant ($X^2(SINE|B_t)$ = ((232×232)/ 244) + 0 = 220,59). The data of Orsi et al. do not support the hypothesis *if* use of glyphosate *then* development of Non-Hodgkin Lymphoma (p(IMP) =0,96471; X^2 $(IMP|A_t)$ = ((24×24)/36) + 0 =**16,00**. The data of the same study support the hypothesis that there is a significant conditio per quam relationship between GS and NHL ($X^2(IMP|\underline{B}_t)$ = ((24×24)/436) + 0 =**1,32**) too, which is a contradiction. In the same context, a non significant **negative** causal relationship between GS and NHL (k = -0,0126 with 95% CI: -0,0983 to +0,07) is documented. The p value of the causal relationship k calculated according to the hypergeometric distribution is *p-value(k | HGD)* = 0,13606 and not significant. In other words, according to the data of Orsi et al. glyphosate is neither a necessary condition for the development of Non-Hodgkin Lymphoma nor a sufficient condition. Furthermore, the Null-hypothesis above cannot be rejected. According to the data of Orsi et al., the use of glyphosate **and** Non-Hodgkin Lymphoma are not causally related.

Quod erat demonstrandum.

Theorem 3.7. (Glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma.)

In the large, prospective cohort study of Andreotti et al. (Andreotti et al., 2018) the previous (De Roos et al., 2005) evaluation of glyphosate with cancer incidence was updated and again no association was apparent between glyphosate and any solid tumors including NHL and its subtypes.

Claim.

Null Hypothesis:

Glyphosate is neither the cause nor a cause of Non-Hodgkin Lymphoma. In other words, k = 0.

Alternative Hypothesis:

Glyphosate is either the cause or a cause of Non-Hodgkin Lymphoma. In other words, k >0.

Proof.

The study of Andreotti et al. investigated the potential health effects of glyphosate in humans with respect of the development of Non-Hodgkin Lymphoma. The data as obtained by De Roos et al. (De Roos et al., 2005) are view by **table 13**.

Table 13.

The study of Andreotti et al., 2018.				
		NHL		
		YES	NO	
Glyphosate	YES	440	43952	44392
	NO	135	9724	9859
		575	53676	54251

Statistical analysis.	IOU =-0,171130486
Causal relationship k =-0,0142.	95 % CI: (-0,023 to -0,004639)
p-value (k | HGD) =0,00023	X^2 (k) = 11,00011393
Odds ratio (OR) =0,72.	95 % CI (OR): (0,59 to 0,88)
p (SINE) =0,99751	X^2 (SINE) =31,69565217
p (IMP) =0,18984	X^2 (IMP) =35989,60996
p (SINE ^ IMP) =0,18735	X^2 (SINE ^IMP) =36021,30561

The study of Andreotti et al. (Andreotti et al., 2018) is potentially biased because the index of unfairness of the study is IOU = -0,171130486 and thus far unfair. The relative frequency of the conditio sine qua non relationship between GS and NHL with p (SINE) =0,99751 is very high, but statistically not significant ($X^2(SINE|B_t)$ = ((135×135)/575) + 0 = 31,69565217). The data of the same study support the hypothesis that there is a significant conditio sine qua non relationship between GS and NHL ($X^2(SINE|\underline{A}_t)$ = ((135×135)/ 9859) + 0 = **1,848564763**) too, which is a contradiction. The data of Andreotti et al. do not support the hypothesis *if* use of glyphosate *then* development of Non-Hodgkin Lymphoma (p(IMP) =0,18984; X^2 (IMP) = 35989,60996. In the same context, a highly significant **negative** causal relationship between GS and NHL (k = -0,0142 with 95% CI: -0,023 to -0,004639) is documented. The p value of the causal relationship k calculated according to the hypergeometric distribution is *p-value(k | HGD)* = 0,00023 and significant. In other words, according to the data of Andreotti et al. glyphosate is neither a necessary condition for the development of Non-Hodgkin Lymphoma nor a sufficient condition. Furthermore, the Null-hypothesis above must be rejected. According to the data of Andreotti et al., the use of glyphosate **prevents** from Non-Hodgkin Lymphoma (k = -0,0142).

Quod erat demonstrandum.

Theorem 3.8. (Without Epstein-Barr virus infection no Non-Hodgkin Lymphoma.)

Non-Hodgkin lymphomas differ in several aspects but share some features too. Epstein-Barr virus (EBV) is possibly one of these common features and has been discussed (IARC, 2012) as a cause of non-Hodgkin lymphoma (NHL). However, the role of EBV in non-Hodgkin lymphomas (NHLs) remains unclear. Teras et al. (Teras et al., 2015) examined the association between prospectively-collected plasma EBV antibodies and NHL risk in the Cancer Prevention Study-II (CPS-II) Nutrition Cohort which included 225 NHL cases and 2:1 matched controls and documented an association between EBV serostatus or antibody levels (early antigen) and risk of the three most common types of NHL (diffuse large B-cell lymphoma, follicular lymphoma, chronic lymphocytic leukemia/small lymphocytic lymphoma).

Claim.

Null Hypothesis:

Epstein-Barr virus infection is a necessary condition of Non-Hodgkin Lymphoma.

In other words, *without* an Epstein-Barr virus infection *no* Non-Hodgkin Lymphoma.

Alternative Hypothesis:

Epstein-Barr virus infection is not a necessary condition of Non-Hodgkin Lymphoma.

In other words, a human being can suffer from Non-Hodgkin Lymphoma even if not Epstein-Barr virus positive.

Proof.

The study of Teras et al. investigated the potential role of EBV in non-Hodgkin lymphomas (NHLs). The data as obtained by Teras et al. (Teras et al., 2015) are view by **table 14**.

Table 14.

The study of Teras et al., 2015.				
		NHL		
		YES	NO	
EBV pos.	YES	212	416	628
	NO	13	33	46
		225	449	674

Statistical analysis. IOU =+0,27

Causal relationship k =+0,0294. 95 % CI: (-0,057 : 0,12)

p-value (k | HGD) =0,09937 **X^2 (k) =0,58**

Odds ratio (OR) =1,29. 95 % CI: (0,67 : 2,51)

p (SINE) =0,98071 **X^2 (SINE) =0,75**

p (IMP) =0,38279 X^2 (IMP) =385,43

p (SINE ^ IMP) =0,3635 X^2 (SINE ^IMP) =386,18

The study of Teras et al. (Teras et al., 2015) is potentially biased because the index of unfairness of the study is IOU = +0,27. Thus far, the study of Teras et al. is unfair. The relative frequency of the conditio sine qua non relationship between EBV and NHL with p (SINE) =0,98071 is very high, and statistically significant ($X^2(SINE|B_t) = ((13\times13)/225) + 0 = 0,75$). Moreover, it is not difficult to see from a different perspective that the data of the same study support the hypothesis of a significant conditio sine qua non relationship between EBV and NHL ($X^2(SINE|\underline{A}_t) = ((13\times13)/ 46) + 0 = 3,67$) too, which is not a contradiction. The data of Teras et al. documented a **positive** causal relationship between EBV and NHL (k = +0,0294 with 95% CI: 0,057 : 0,12 and *p-value(k | HGD)* = 0,09937). In other words, according to the data of Teras et al. (Teras et al., 2015) we cannot reject the null-hypothesis: EBV is a necessary condition for the development of Non-Hodgkin Lymphoma. There is another aspect to the characterization of this relationship: *without* EBV infection *no* Non-Hodgkin Lymphoma.

Quod erat demonstrandum.

4. Discussion

NHL consists of more than 40 major subtypes and is a very heterogeneous group of malignant lymphoid tumors. **Historically, people suffered from NHL before the existence or the use of GS**. In other words, historically, it is proven that **the existence or the use of GS is not a necessary condition for the development of NHL**. Independently of this historical fact, todays data proof this hypothesis too. The National Cancer Institute (NCI) reported 2019 about 19,6 new cases of non-Hodgkin lymphoma per 100,000 men and women per year (National Cancer Institute, 2019). The data as reported by NCI are viewed by the table (**Table 15**) below.

Table 15. Percent of New U. S. Cases of Non-Hodgkin Lymphoma by Age Group according to National Cancer Institute 2019 (NCI, 2019).

Percent of New NHL U.S. Cases	1,7 %	3,6 %	5,1 %	11,8 %	21,3 %	26,0 %	20,9 %	9,6 %
Age	< 20	20-34	35-44	45-54	55-64	65-74	75-84	>84

According to National Cancer Institute, NHL can occur at any age and especially in the childhood (Sandlund, 2015). There does not appear to be any justifiable reason to assume, that very small children are working with glyphosate frequently or at all. Therefore, no human reason can provide serious evidence of the hypothesis that *without* GS *no* NHL. **Glyphosate [N-(phosphonomethyl)glycine] is not a necessary condition for the development of Non-Hodgkin Lymphoma**. None of the studies analyzed provided clear evidence of a significant conditio sine qua non relationship (without GS no NHL) between GS and NHL. Two studies (De Roos et al., 2005; Andreotti et al., 2018) were self-contradictory (**Table 16**) on this point.

Table 16. Overview of the results achieved.

Study ID	Year	N	Case_P	Case_T	Con_P	Con_T	IOU	k	X^2(IMP| A_t)	X^2(IMP|$\underline{B}_t$)	X^2(SINE|B_t)	X^2(SINE|$\underline{A}_t$)
McDuffie et al.	2001	2023	51	517	133	1506	-0,65	+0,02	96,14	11,75	420,03	118,08
Hardell et al.	2002	1656	8	515	8	1141	-0,68	+0,04	**4,00**	**0,06**	99,12	156,74
De Roos et al.	2003	2583	36	650	61	1933	-0,71	**+0,05**	**38,36**	**1,92**	579,99	151,65
De Roos et al.	2005	54315	71	92	40964	54223	-0,24	+0,00	40893,12	30947,19	4,79	0,03
Eriksson et al.	2008	1926	29	910	18	1016	-0,50	**+0,05**	**6,89**	**0,32**	852,92	413,07
Orsi et al.	2009	680	12	244	24	436	-0,59	-0,01	**16,00**	**1,32**	220,59	83,58
Andreotti et al.	2018	54251	440	575	43952	53676	-0,17	**-0,01**	43516,36	35989,61	31,70	1,85

N = sample size. Case_P: case, positive. Case_T: number of cases. Con_P: control, positive, Con_T: number of controls.

The studies of McDuffie et al. 2001 (McDuffie et al.2001), De Roos et al. (De Roos et al., 2005), Orsi et al. (Orsi et al., 2009) and Andreotti et al. (Andreotti et al.,2018) were not able to provide evidence of a significant **positive** cause effect relationship between GS and NHL. In contrast to expectation, the study of Andreotti et al. (Andreotti et al.,2018) provided significant evidence of a negative cause effect relationship between GS and NHL. The data of this study are to some extent self-contradictory since the same study support the contradiction that a conditio sine qua no relationship between GS and NHL is given and equally not given. *A negative causal relationship excludes a conditio sine qua no relationship or a conditio per quam and vice versa* (I. Barukčić, 2019a). Thus far, we cannot rely on the data of Andreotti et al. (Andreotti et al.,2018) in this context. The study of the data provided to us by the study of De Roos et al. (De Roos et al., 2005) are self-contradictory too since the same study support both, no conditio sine qua relationship between GS and NHL (X^2(SINE|B_t) = 4,793478261) and equally a conditio sine qua not relationship between GS and NHL (X^2(SINE|$\underline{A}_t$) = 0,033207831), which is a contradiction. The study of De Roos et al. had a prospective design, but is still potentially susceptible to bias of non-exposure reporting which accounts for discrepancies associated with the study results. The group of the non-exposed was 13280 and to small with respect to the group of the exposed (n=41035). According to the US Census Bureau, the estimated 2018 United States population (February 2018) was about 327,16 million. According to the study of De Roos et al., more than 247,1695 million of U.S. inhabitants are using Glyphosate in a high dose and frequently, which is not realistic. In other words, the data of De Roos et al. (De Roos et al., 2005) are only of limited value and do not provide clear evidence in favor or against the relationship between GS and NHL. The only study which was to some extent not self-contradictory was the study of McDuffie et al. (McDuffie et al., 2001) while the validity of

the results based on the data of the study of McDuffie et al. is endangered by an index of unfairness of IOU = -0,65.

The systematic review and meta-analysis by Chang and Delzell (Chang & Delzell, 2016) examined the relationship between glyphosate exposure and among other, the risk of NHL and was not able to establish a causal relationship between glyphosate exposure and the risk of any type of lymphohematopoietic cancer (LHC) including NHL. In contrast to Chang and Delzell, the meta-analysis conducted by Zang et al. (Zhang, Rana, Taioli, Shaffer, & Sheppard, 2019) used published human studies on the relationship between exposures to GS and NHL and reported that GBH exposure is associated with increased risk of NHL. The meta-analysis of Zang et al. is grossly flawed, one-sided and worthless in toto due to several reasons. The data of the most studies considered by Zang et al. (Zhang, Rana, Taioli, Shaffer, & Sheppard, 2019) are self-contradictory and of none or extremely limited value, which was ignored by the study group completely. Other possible factors which are causally related to NHL were not considered to a necessary extent. Statistical methods, far away from being able, to provide anything valuable on the point of issue, were used with the consequence that everything desirable can be proofed as correct, even pure non-sense. The inconsistency of Forest plot (I. Barukčić, 2019a) supported meta-analysis was ignored completely. In toto, I do justifiably believe that adopting an extremely restricted and unfair one-eyed view on such a complicated matter like the relationship between GS and NHL may be a way to lose its own reputation publicly but is for sure not the path to eternal scientific honor.

The results of this systematic review and meta-analysis suggest that EBV and not glyphosate is causally linked with a wider spectrum of NHL subtypes. Still, this cannot be considered as the final proof of the relationship between EBV and NHL and further and better designed studies are needed to confirm and fully understand the etiology of NHL. Besides of all, as long as no better data are available, it is justified, necessary and allowed to deduce the following conclusion.

5. Conclusion

Glyphosate is neither a cause nor the cause of Non-Hodgkin Lymphoma (McDuffie et al., 2001).

Conflict of interest

Author declare no conflict of interests for this article.

Acknowledgement

None.

References

Altman, D. G. (1999). *Practical statistics for medical research*. Boca Raton, Fla: Chapman & Hall/CRC.

Andreotti, G., Koutros, S., Hofmann, J. N., Sandler, D. P., Lubin, J. H., Lynch, C. F., … Beane Freeman, L. E. (2018). Glyphosate Use and Cancer Incidence in the Agricultural Health Study. *Journal of the National Cancer Institute*, *110*(5), 509–516. doi: https://doi.org/10.1093/jnci/djx233 [PMCID: PMC6279255] [PMID: 29136183]

Armitage, J. O., Gascoyne, R. D., Lunning, M. A., & Cavalli, F. (2017). Non-Hodgkin lymphoma. *Lancet (London, England)*, *390*(10091), 298–310. doi: https://doi.org/10.1016/S0140-6736(16)32407-2 [PMID: 28153383]

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

Barukčić, I. (1989). *Die Kausalität* (1. Aufl.). Hamburg: Wiss.-Verl.

Barukčić, I. (2017). *Theoriae causalitatis principia mathematica*. Norderstedt: Books on Demand.

Barukčić, I. (2018a). Epstein Bar Virus—The Cause of Hodgkin's Lymphoma. *Journal of Biosciences and Medicines*, *06*(01), 75–100. doi: https://doi.org/10.4236/jbm.2018.61008

Barukčić, I. (2018b). Epstein-barr virus is the cause of multiple sclerosis. *International Journal of Current Medical and Pharmaceutical Research*, *4*(9 (A)), 3674–3682. doi: https://doi.org/10.24327/23956429.ijcmpr20180538

Barukčić, I. (2018c). Fusobacterium nucleatum —The Cause of Human Colorectal Cancer. *Journal of Biosciences and Medicines*, *06*(03), 31–69. doi: https://doi.org/10.4236/jbm.2018.63004

Barukčić, I. (2019a). Index of Unfairness. *Modern Health Science*, *2*(1), p22. doi: https://doi.org/10.30560/mhs.v2n1p22

Barukčić, I. (2019b). Smoking of tobacco is the cause of human lung cancer. *Journal of Drug Delivery and Therapeutics*, *9*(1-s), 148–160. doi: https://doi.org/10.22270/jddt.v9i1-s.2273

Barukčić, K., & Barukčić, I. (2016). Epstein Barr Virus—The Cause of Multiple Sclerosis. *Journal of Applied Mathematics and Physics*, *04*(06), 1042–1053. doi: https://doi.org/10.4236/jamp.2016.46109

Bennett, J. H. (1845). Two cases of disease and enlargement of thespleen in which death took place from the presence of purulentmatter in the blood. *Edinburgh Medical Surgical Journal*, *64*, 413–423. Retrieved from https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-2141.2000.01988.x

Bolognesi, C., Bonatti, S., Degan, P., Gallerani, E., Peluso, M., Rabboni, R., … Abbondandolo, A. (1997). Genotoxic Activity of Glyphosate and Its Technical Formulation Roundup. *Journal of Agricultural and Food Chemistry*, *45*(5), 1957–1962. doi: https://doi.org/10.1021/jf9606518

Cantor, K. P. (1982). Farming and mortality from non-Hodgkin's lymphoma: a case-control study. *International Journal of Cancer*, *29*(3), 239–247. [PMID: 7040259]

Chang, E. T., & Delzell, E. (2016). Systematic review and meta-analysis of glyphosate exposure and risk of lymphohematopoietic cancers. *Journal of Environmental Science and Health. Part. B, Pesticides, Food Contaminants, and Agricultural Wastes*, *51*(6), 402–434. doi: https://doi.org/10.1080/03601234.2016.1142748 [PMCID: PMC4866614] [PMID: 27015139]

Cohnheim, P. (1865). Ein Fall von Pseudoleukämie. *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin*, *33*(3), 451–454. doi: https://doi.org/10.1007/BF02137492

Conover, W. J. (1974). Some Reasons for Not Using the Yates Continuity Correction on 2×2 Contingency Tables. *Journal of the American Statistical Association*, *69*(346), 374–376. doi: https://doi.org/10.1080/01621459.1974.10482957

CORNFIELD, J. (1951). A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast, and cervix. *Journal of the National Cancer Institute*, *11*(6), 1269–1275. [PMID: 14861651]

Daruich, J., Zirulnik, F., & Gimenez, M. S. (2001). Effect of the herbicide glyphosate on enzymatic activity in pregnant rats and their fetuses. *Environmental Research*, *85*(3), 226–231. doi: https://doi.org/10.1006/enrs.2000.4229 [PMID: 11237511]

De Roos, A. J., Zahm, S. H., Cantor, K. P., Weisenburger, D. D., Holmes, F. F., Burmeister, L. F., & Blair, A. (2003). Integrative assessment of multiple pesticides as risk factors for non-Hodgkin's lymphoma among men. *Occupational and Environmental Medicine*, *60*(9), E11. doi: https://doi.org/10.1136/oem.60.9.e11 [PMCID: PMC1740618] [PMID: 12937207]

De Roos, Anneclaire J., Blair, A., Rusiecki, J. A., Hoppin, J. A., Svec, M., Dosemeci, M., … Alavanja, M. C. (2005). Cancer incidence among glyphosate-exposed pesticide applicators in the Agricultural Health Study. *Environmental Health Perspectives*, *113*(1), 49–54. doi: https://doi.org/10.1289/ehp.7340 [PMCID: PMC1253709] [PMID: 15626647]

Edwards, A. W. F. (1963). The Measure of Association in a 2 × 2 Table. *Journal of the Royal Statistical Society. Series A (General)*, *126*(1), 109. doi: https://doi.org/10.2307/2982448

El-Demerdash, F. M., Yousef, M. I., & Elagamy, E. I. (2001). Influence of paraquat, glyphosate, and cadmium on the activity of some serum enzymes and protein electrophoretic behavior (in vitro). *Journal of Environmental Science and Health. Part. B, Pesticides, Food Contaminants, and Agricultural Wastes*, *36*(1), 29–42. [PMID: 11281253]

EPA, U. S. (1993). U.S. Environmental Protection Agency Reregistration Eligibility Decision (RED) Glyphosate. *Environmental Protection Agency. Washington, DC:U.S.*, *EPA-738-R-93-014*.

Eriksson, M., Hardell, L., Carlberg, M., & Akerman, M. (2008). Pesticide exposure as risk factor for non-Hodgkin lymphoma including histopathological subgroup analysis. *International Journal of Cancer*, *123*(7), 1657–1663. doi: https://doi.org/10.1002/ijc.23589 [PMID: 18623080]

Fisher, R. A. (1922). On the Interpretation of χ 2 from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society*, *85*(1), 87. doi: https://doi.org/10.2307/2340521

Fisher, R. A. (1935). The Logic of Inductive Inference. *Journal of the Royal Statistical Society*, *98*(1), 39. doi: https://doi.org/10.2307/2342435

Gonin, H. T. (1936). XIV. The use of factorial moments in the treatment of the hypergeometric distribution and in tests for regression. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *21*(139), 215–226. doi: https://doi.org/10.1080/14786443608561573

Grizzle, J. E. (1967). Continuity Correction in the χ 2 -Test for 2 × 2 Tables. *The American Statistician*, *21*(4), 28. doi: https://doi.org/10.2307/2682103

Hardell, L., & Eriksson, M. (1999). A case-control study of non-Hodgkin lymphoma and exposure to pesticides. *Cancer*, *85*(6), 1353–1360. [PMID: 10189142]

Hardell, Lennart, Eriksson, M., & Nordstrom, M. (2002). Exposure to pesticides as risk factor for non-Hodgkin's lymphoma and hairy cell leukemia: pooled analysis of two Swedish case-control studies. *Leukemia & Lymphoma*, *43*(5), 1043–1049. [PMID: 12148884]

Hietanen, E., Linnainmaa, K., & Vainio, H. (1983). Effects of phenoxyherbicides and glyphosate on the hepatic and intestinal biotransformation activities in the rat. *Acta Pharmacologica Et Toxicologica*, *53*(2), 103–112. [PMID: 6624478]

Hodgkin, T. (1832). On some Morbid Appearances of the Absorbent Glands and Spleen. *Medico-Chirurgical Transactions*, *17*, 68–114. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2116706/

Huygens, C. (1629-1695), & van Schooten, F. (1615-1660). (1657). *De ratiociniis in ludo alae: In: Exercitationum mathematicarum liber primus [- quintus]*. Lugdunum Batavorum (Leiden, The Netherlands): ex officina Johannis Elsevirii. doi: https://doi.org/10.3931/e-rara-8813

IARC, W. G. on the E. of C. R. to H. (2012). Biological agents. Volume 100 B. A review of human carcinogens. *Iarc Monographs on the Evaluation of Carcinogenic Risks to Humans*, *100*(PT B), 1–441. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4781184/ [PMCID: PMC4781184] [PMID: 23189750]

IARC, W. G. on the E. of C. R. to H. (2017). *IARC monographs on the evaluation of carcinogenic risks to humans. Volume 112* (Vol. 112).

Isserlis, L. (1918). On the Value of a Mean as Calculated from a Sample. *Journal of the Royal Statistical Society*, *81*(1), 75. doi: https://doi.org/10.2307/2340569

Kolmogoroff, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi: https://doi.org/10.1007/978-3-642-49888-6

Leon, M. E., Schinasi, L. H., Lebailly, P., Beane Freeman, L. E., Nordby, K.-C., Ferro, G., … Schüz, J. (2019). Pesticide use and risk of non-Hodgkin lymphoid malignancies in agricultural cohorts from France, Norway and the USA: a pooled analysis from the AGRICOH consortium. *International Journal of Epidemiology*. doi: https://doi.org/10.1093/ije/dyz017 [PMID: 30880337]

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., … Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine*, *6*(7), 1000100. doi: https://doi.org/10.1371/journal.pmed.1000100 [PMID: 19621070]

Lioi, M. B., Scarfi, M. R., Santoro, A., Barbieri, R., Zeni, O., Di Berardino, D., & Ursini, M. V. (1998). Genotoxicity and oxidative stress induced by pesticide exposure in bovine lymphocyte cultures in vitro. *Mutation Research*, *403*(1–2), 13–20. [PMID: 9726001]

Lioi, M. B., Scarfi, M. R., Santoro, A., Barbieri, R., Zeni, O., Salvemini, F., … Ursini, M. V. (1998). Cytogenetic damage and induction of pro-oxidant state in human lymphocytes exposed in vitro to gliphosate, vinclozolin, atrazine, and DPX-E9636. *Environmental and Molecular Mutagenesis*, *32*(1), 39–46. [PMID: 9707097]

Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly*, *2*(4), 245–264. Retrieved from https://www.jstor.org/stable/20009173

Malpighi, M. (1666). *De viscerum structura exercitatio anatomica.Accedit dissertatio eiusdem De polypo cordis*. Bononiae: Iacobi Montij. doi: https://doi.org/10.3931/e-rara-23974

McDuffie, H. H., Pahwa, P., McLaughlin, J. R., Spinelli, J. J., Fincham, S., Dosman, J. A., … Choi, N. W. (2001). Non-Hodgkin's lymphoma and specific pesticide exposures in men: cross-Canada study of pesticides and health. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *10*(11), 1155–1163. [PMID: 11700263]

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Open Medicine : A Peer-Reviewed, Independent, Open-Access Journal*, *3*(3), 123–130. [PMID: 21603045]

Moivre, A. de [1667-1754]. (1718). *The Doctrine of Chances or a Method of Calculating the Probability of Events in Play*. London: printed by W. Pearson for the author. doi: https://doi.org/10.3931/e-rara-10420

Mosteller, F. (1968). Association and Estimation in Contingency Tables. *Journal of the American Statistical Association*, *63*(321), 1. doi: https://doi.org/10.2307/2283825

National Cancer Institute, N. I. of H. (2019). Non-Hodgkin Lymphoma - Cancer Stat Facts. *National Cancer Institute - Surveillance, Epidemiloogy and End Results Program*, *May 18*. Retrieved from https://seer.cancer.gov/statfacts/html/nhl.html

Olorunsogo, O. O., Bababunmi, E. A., & Bassir, O. (1979). Effect of glyphosate on rat liver mitochondria in vivo. *Bulletin of Environmental Contamination and Toxicology*, *22*(3), 357–364. [PMID: 223703]

Orsi, L., Delabre, L., Monnereau, A., Delval, P., Berthou, C., Fenaux, P., … Clavel, J. (2009). Occupational exposure to pesticides and lymphoid neoplasms among men: results of a French case-control study. *Occupational and Environmental Medicine*, *66*(5), 291–298. doi: https://doi.org/10.1136/oem.2008.040972 [PMCID: PMC2728754] [PMID: 19017688]

Pearson, K., & Heron, D. (1913). On Theories of Association. *Biometrika*, *9*(1–2), 159–315. doi: https://doi.org/10.1093/biomet/9.1-2.159

Pearson, Karl. (1899). XV. On certain properties of the hypergeometrical series, and on the fitting of such series to observation polygons in the theory of chance. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *47*(285), 236–246. doi: https://doi.org/10.1080/14786449908621253

Pearson, Karl. (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*(302), 157–175.

Peluso, M., Munnia, A., Bolognesi, C., & Parodi, S. (1998). 32P-postlabeling detection of DNA adducts in mice treated with the herbicide Roundup. *Environmental and Molecular Mutagenesis*, *31*(1), 55–59. [PMID: 9464316]

Rappaport, H. (1966). *Tumors of the Hematopoietic System*. Washington: Armed Forces Institute of Pathology. Retrieved from http://annals.org/article.aspx?doi=10.7326/0003-4819-67-3-686_2

Sachs, L. (1992). *Angewandte Statistik*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Sandlund, J. T. (2015). Non-Hodgkin Lymphoma in Children. *Current Hematologic Malignancy Reports*, *10*(3), 237–243. doi: https://doi.org/10.1007/s11899-015-0277-y [PMID: 26174528]

Steinrücken, H. C., & Amrhein, N. (1980). The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid-3-phosphate synthase. *Biochemical and Biophysical Research Communications*, *94*(4), 1207–1212. [PMID: 7396959]

Teras, L. R., Rollison, D. E., Pawlita, M., Michel, A., Brozy, J., de Sanjose, S., … Gapstur, S. M. (2015). Epstein-Barr virus and risk of non-Hodgkin lymphoma in the cancer prevention study-II and a meta-analysis of serologic studies. *International Journal of Cancer*, *136*(1), 108–116. doi: https://doi.org/10.1002/ijc.28971 [PMID: 24831943]

Virchow, R. (1845). Weisses Blut. Neue Notizen aus dem Gebiet der Naturfund Heilkunde. *Floriep's Neue Notizen*, *36*, 151. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2116706/

Walsh, L. P., McCormick, C., Martin, C., & Stocco, D. M. (2000). Roundup inhibits steroidogenesis by disrupting steroidogenic acute regulatory (StAR) protein expression. *Environmental Health Perspectives*, *108*(8), 769–776. doi: https://doi.org/10.1289/ehp.00108769 [PMCID: PMC1638308] [PMID: 10964798]

Warrens, M. J. (2008). On Association Coefficients for 2x2 Tables and Properties That Do Not Depend on the Marginal Distributions. *Psychometrika*, *73*(4), 777–789. doi: https://doi.org/10.1007/s11336-008-9070-3 [PMID: 20046834]

Wilks, S. (1865). Cases with enlargement of the lymphatic glands and spleen (or Hodgkin's disease) with remarks. *Guy's Hospital Reports*.

Williams, G. M., Kroes, R., & Munro, I. C. (2000). Safety evaluation and risk assessment of the herbicide Roundup and its active ingredient, glyphosate, for humans. *Regulatory Toxicology and Pharmacology: RTP*, *31*(2 Pt 1), 117–165. doi: https://doi.org/10.1006/rtph.1999.1371 [PMID: 10854122]

Yates, F. (1934). Contingency Tables Involving Small Numbers and the χ 2 Test. *The Journal of the Royal Statistical Society (Supplement)*, *1*(2), 217. doi: https://doi.org/10.2307/2983604

Yousef, M. I., Salem, M. H., Ibrahim, H. Z., Helmi, S., Seehy, M. A., & Bertheussen, K. (1995). Toxic effects of carbofuran and glyphosate on semen characteristics in rabbits. *Journal of Environmental Science and Health. Part. B, Pesticides, Food Contaminants, and Agricultural Wastes*, *30*(4), 513–534. doi: https://doi.org/10.1080/03601239509372951 [PMID: 7797819]

Yule, G. U. (1900). On the Association of Attributes in Statistics: With Illustrations from the Material of the Childhood Society, &c. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *194*(252–261), 257–319. doi: https://doi.org/10.1098/rsta.1900.0019

Zhang, L., Rana, I., Taioli, E., Shaffer, R., & Sheppard, L. (2019). Exposure to Glyphosate-Based Herbicides and Risk for Non-Hodgkin Lymphoma: A Meta-Analysis and Supporting Evidence. *Mutation Research/Reviews in Mutation Research*. doi: https://doi.org/10.1016/j.mrrev.2019.02.001