

Predicting the Efficacy of the COVID 19 Vaccine by Analyzing the Death Rate and Confirmation Data of the Corona Virus

Dodi Irwan Siregar*

Department of STIE, Persada Bunda College, Pekanbaru, Indonesia

Abstract

Viruses are important pathogens of high animals and the major cause of mortality, and they also are drivers of global geochemical cycles; yet, biological sciences considered them as entities separate from the realm of life and acting merely as mechanical artifacts that can exchange genes between different species, genera, phyla and even from one ecosystem to another. During the last decade, the increased awareness of the complexity of the immune system and its determinants, including at the host genetic level, indicated that using system biology approaches to assess how various processes and networks interact in response to immunization could prove more illustrative than trying to isolate and characterize a few components of vaccine responses. These immune-related concerns have largely spread into the population, and questions related to the immunological safety of vaccines that is, their capacity for triggering 2 non-antigen specific responses possibly leading to allergy, auto-immunity, or even premature death are being raised. Certain "off-targets effects" of vaccines have also been recognized and call for studies to quantify their impact and identify the mechanisms at play. This study uses multiple linear regression analysis to predict and predict changes in the value of certain variables when other variables change. Correlation is one of the statistical analysis techniques used to find the relationship of how strong the relationship between two or more variables that are quantitative. By using the Linearity Test (Lillie force test) where $F_{count} > F_{table}$ is 787.62 > 3.24. Then, H_0 is rejected. This means that multiple linear regression analysis can be used to predict The COVID 19 vaccines in world by analyzing variable death rate and confirmed the COVID 19 data. Obtained the multiple linear regression equation is $Y = 0.609211 - 3.15845257X_1 + 3.58359X_2$, the correlation value of the variables above is 0.98389 the value is superior correlation scale ranging from 0.76 to 0.99. Is where $t_1 \text{ count} < t_{table}$ is $-12.078 < 2.009$, then H_0 is accepted, meaning there is no significant effect partially between the Variable The COVID 19 vaccine and The death rate. Then the value of the $t_2 \text{ count} > t_{table}$ is $3.62523 > 2.009$. H_0 is rejected; it means that there is a large significant influence partially between the COVID 19 vaccine against confirmed the COVID 19 variable in World on December 2020 until September 2021.

Keywords: COVID 19 vaccine • Death rate • Multiple Linear Regression • Correlations

Introduction

Pharmacokinetics to generate vaccine-mediated protection is a complex challenge. Currently available vaccines have largely been developed empirically, with little or no understanding of how they activate the immune system. Their early protective efficacy is primarily conferred by the induction of antigen-specific anti bodies. However, there is more to antibody mediated protection than the peak of vaccine-induced antibody titers. The quality of such antibodies (e.g., their avidity, specificity, or neutralizing capacity) has been identified as a determining factor in efficacy. Long-term protection requires the persistence of vaccine antibodies above protective thresholds and/or the maintenance of immune memory cells capable of rapid and effective reactivation with subsequent microbial exposure. The determinants of immune memory induction, as well as the relative contribution of persisting antibodies and of immune memory to protection against specific diseases, are essential parameters of long-term vaccine efficacy. The predominant role of B cells in the efficacy of current vaccines should not overshadow the importance of T-cell responses: T cells are essential to the induction of high-affinity antibodies and immune memory, directly contribute to the protection conferred by current vaccines such as Bacilli Calmette-Guérin (BCG), may play a more critical role than previously anticipated for specific diseases like pertussis, and will be the prime effectors against novel vaccine targets with predominant intracellular localization such as tuberculosis. New methods have emerged allowing the assessment of a growing number of vaccine-associated immune parameters,

including in humans. This development raises new questions about the optimal markers to assess and their correlation with vaccine-induced protection. The identification of mechanistic immune correlates or at least surrogates of vaccine efficacy is a major asset for the development of new vaccines or the optimization of immunization strategies using available vaccines. Thus, their determination generates a considerable amount of interest. During the last decade, the increased awareness of the complexity of the immune system and its determinants, including at the host genetic level, indicated that using system biology approaches to assess how various processes and networks interact in response to immunization could prove more illustrative than trying to isolate and characterize a few components of vaccine responses [1]. Delineating the specific molecular signatures of vaccine immunogenicity is beginning to highlight novel correlates of protective immunity and better explain the heterogeneity of vaccine responses in a population. The tailoring of vaccine strategies for specific vulnerable populations, including very young, elderly, and immunosuppressed populations, also largely relies on a better understanding of what supports or limits vaccine efficacy under special circumstances at the population and individual levels. Lastly, the exponential development of new vaccines raises many questions that are not limited to the targeted diseases and the potential impacts of their prevention, but that address the specific and nonspecific impacts of such vaccines on the immune system and, thus, on health in general. These immune-related concerns have largely spread into the population, and questions related to the immunological safety of vaccines that is, their capacity for triggering 2 non-antigen-specific responses possibly leading to allergy, auto-immunity, or even premature death are being raised. Certain "off-targets effects" of vaccines have also been recognized and call for studies to quantify their impact and identify the mechanisms at play. The objective of this chapter is to extract from the complex and rapidly evolving field of immunology the main concepts that are useful to better address these important questions.

*Address for Correspondence: Siregar DI, Department of STIE, Persada Bunda College, Pekanbaru, Indonesia; Tel: +62082171928606; E-mail: dodi.irwan.siregar@gmail.com

Copyright: © 2021 © Siregar DI. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Received 11 October, 2021; **Accepted** 25 October, 2021; **Published** 01 November, 2021

Theoretical Analysis

Populations and samples

Definition population is the set of all the individuals of interest in a particular study. Although research question concerns an entire population, it usually is impossible for a researcher to examine every single individual in the population of interest. Therefore, researchers typically select a smaller, more manageable group from the population and their studies to the individuals in the selected group. In the statistical terms a set of individuals selected from a population is called a sample. A sample intended to be representative of its population, a sample should always be identified in terms of the populations from which it was selected. Definition A sample is a set of individuals selected from a population, usually intended to represent the population in a research study.

Collecting data: Once you decide on the type of data-quantitative or qualitative-appropriate for the problem at hand, you will need to collect the data. Generally, you can obtain data in four different ways:

- Data from a published source
- Data from a designed experiment
- Data from a survey
- Data collected observationally

Sometimes, the data set of interest has already been collected for you and is available in a published source, such as a book, journal, or newspaper. For example, you may want to examine and summarize the divorce rates (i.e., number of divorces per 1,000 populations) in the 50 states of the World. You can find this data set (as well as numerous other data sets) at your library in the statistical abstract of the World, published annually by the U.S. government.

The independent and dependent variables: Definitions the independent variable is the variable that is manipulated by the researcher. In behavioral research, the independent variable usually consists of the two (or more) treatment conditions to which subjects are exposed. The independent variable consists of the antecedent conditions that were manipulated prior to observing the dependent variable. The dependent variable is the one that is observed for changes in order to assess the effect of the treatment.

A comment on stating null and alternative hypothesis: A statistical test procedure cannot prove the truth of a null hypothesis. When we fail to reject a null hypothesis, all the hypothesis test can establish is that the information in a sample of data is compatible with the null hypothesis. On the other hand, a statistical test can lead us to reject the null hypothesis, with only a small probability, of rejecting the null hypothesis when it is actually true. Thus rejecting a null hypothesis is stronger conclusion than failing to reject it. The null hypothesis is usually stated in such stated in such a way that if our theory is correct, then we will reject the null hypothesis. For example, our airplane seat designer has been operating under the assumption (the maintained or null hypothesis) that the population means hip width is 16.5 inches. Casual observation suggests that people are getting larger all the time. If we are larger and if the airplane wants to continue to accommodate the same percentage of the population, then the seat width must be increased. This costly change should be undertaken only if there is statistical evidence that the population hip size indeed larger. When using a hypothesis test we would like to find out whether there is statistical evidence against our current "theory," or whether the data are compatible with it. With this goal, we set up the null hypothesis that the population mean is 16.5inches, $H_0: \mu=16.5$, against the alternative that have shown that there has been a "statistically significant" increase in hip width. You may view the null hypothesis to be too limited in this case, since it is feasible that the population mean hip width is now smaller than 16.5 inches. The hypothesis test of the null hypothesis $H_0: \mu \leq 16.5$ against the alternative hypothesis $H_1: \mu > 16.5$ is exactly the same as the test for $H_0: \mu=16.5$ against the alternative hypothesis $H_1: \mu=16.5$ against the alternative hypothesis $H_1: \mu > 16.5$. The test statistic and rejection region are exactly the same. For a one-tail test you can form the null hypothesis in either of these ways. Finally, it is important to set up the null and alternative hypotheses before you analyze

or even collects the sample of data. Failing to do so can lead to errors in formulating the alternative hypothesis. Suppose that we wish to test whether $\mu > 16.5$ and the sample mean is $\bar{y}=15.5$. Does that mean we should set up the alternative $\mu < 16.5$, to be consistent with estimate? The answer is no, the alternative is formed to state the conjecture that we wish to establish, $\mu > 16.5$ [2].

Relating two measurements: regression: This chapter considers relationships between two measurements that can be represented by a line. Such relationships are called linear and their models are called regression models. As usual, we deal with settings in which there is variability in measurement. Observations in different circumstances but with the same x may give different values of y . Perhaps the true line goes through the first value of y and perhaps it goes through the second-more likely, it goes through neither. Despite there being strength in numbers, any finite sample gives less than perfect information about the line. Since anything uncertain has a probability distribution, the line itself has a probability distribution. Lines are specified by slopes and intercepts and so probabilities for lines are specified by probabilities for b and α . A major focus of this section is on the slope, b , because it indicates the degree of the relationship between x and y : increasing x by 1 unit means increasing y by b units. In particular, $b=0$ means there is no relationship between x and y , at least not a linear relationship [3].

The regression effect arises in every aspect of life and yet few people appreciate its existence. As a result, some people mistakenly attribute change to an ineffective intervention. A golfer has a terrible round and then takes a lesson-she does better the next time and praises her teacher. A baseball player is in a slump and changes bats-the next day he gets two hits and credits the bat. A woman a backache goes to a specialist and her backache improves -she thinks the specialist is a genius. A man is feeling depressed and goes to psychiatrist may be effective. But another explanation is possible and even likely in all this cases: the regression effect, also called regression to the mean. It is simple and ubiquitous: there are ups and downs in any sequence of exchangeable observation; drops tend to follow the ups and rises tend to follow the downs [3].

The first recorded observation of the regression effect was found that seeds of mother sweet pea plants that happen to be larger than average tend to have daughters that are larger than average tend to have daughters that are smaller than average-but larger than themselves. The comparison from one generation to the next is a reversion or regression to the mean. Similarly, Galton noticed that sons of short men tend to be short-but not as short as their fathers (Imagine what would happen to the population if the opposite were true and tall men tended to have sons taller than themselves and short men tended to have sons shorter than themselves!). Sports statistics that show the regression effect are readily [4].

The correlational method: Definition with the correlation method, two variables are observed to see if there is a relationship. The simplest way to look for relationship between variables is to make observations of two variables as they exist naturally. This is called the observational or correlation method.

Parameters in statistics: Definition parameters in statistics a value, usually a numerical value that describes a population. A parameter may be obtained from a single measurement, or it may be derived from a set measurement from the population. A statistic is a value, usually a numerical value that describes a sample. A statistic may be obtained from a single a measurement, or it may be derived from a set of measurement from the sample. We examine the relationship between the mean obtained for a sample and the mean for the population from which the sample was obtained.

Multiple regression analysis: Multiple regression analysis is a statistical technique that can be used to analyze the relationship between a single dependent (creation) variable and several independent variables. The objective of multiple regression analysis is to use the independent variables whose values are known to predict the single dependent value selected by the researcher. Each independent variable is weighted by the regression analysis procedure to ensure maximal prediction from the set of independent variable. The weights denote the relative contribution of the

independent variables to the overall prediction and facilitate interpretation as to the influence of each variable in making the prediction, although correlation among the independent variables complicates the interpretative process. The set of weighted independent variables forms the regression variety, a linear combination of the independent variables that best predicts the dependent variable. The regression variety, also referred to as the regression equation or regression equation or regression model, is the most widely known example of vitiate among the multivariate techniques. Multiple regression analysis is a dependence technique. Thus, to use it you must be able to divide the variables into dependent and independent variables. Regression analysis is also a statistical tool that should be used only when both the dependent and independent variable are metric. However, under certain circumstances it is possible to include non-metric data either as independent variables (by transforming either ordinal or nominal data with dummy variable coding) or dependent variable. In summary, to apply multiple regression analysis: (1) the data must be metric or appropriately transformed, and (2) before deriving the regression equation, the researcher must decide which variable is to be dependent and which remaining variables will be independent [5].

Term used to express the concept that the model processes the properties of additives and homogeneity. In a simple sense, linear models predict values that falls a straight line by having a constant unit change (slope) of the dependent variable for a constant unit change of the independent variable. In the population models, $Y = b_0 + b_1X_1 + \epsilon$, the effect of a change of 1 in X_1 is to add b_1 (a constant) units of Y [6].

Although there are many problems in which one variable can be predicted quite accurately in terms of another, it stands to reason that predictions should improve if one considers additional relevant information. For instance, we should be able to make better predictions of the performance of newly hired teachers if we consider not only their education, but also their years of experience and their personality. Also, we should be able to make better predictions of new textbook's success if we consider not only the quality of the work, but also the potential demand and the competition.

Although there are many different formulas that can be used to express regression relationships among than two variables, most widely used are linear equations of the form;

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k$$

This is partly a matter of mathematical convenience and partly due to the fact that many relationships are actually of this form or can be approximated closely by linear equations. In the equations above, Y is the random variable whose values we want to predict in terms of given values of x_1, x_2, \dots , and x_k , and $\beta_0, \beta_1, \beta_2, \dots$, and β_k , the multiple regression coefficients, are numerical constants that must be determined from observed data [7].

For each variable, the value of observations from one another is not related. Multiple correlations are a correlation consisting of two independent variables (X_1, X_2) or more, and one dependent variable (Y). The relationship between variables can be described as follows:

From the picture above the formulation of the problem consists of three or more problems, and then multiple linear correlations are used (Figure 1).

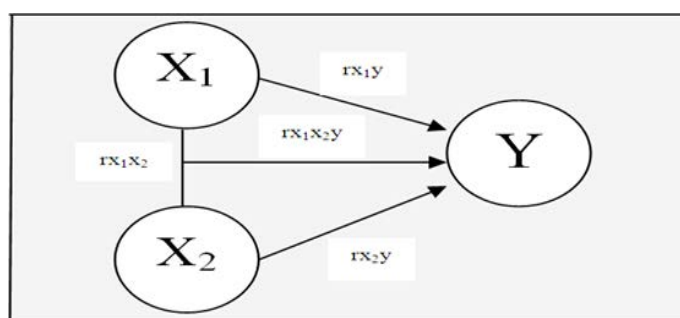


Figure 1. Multiple linear correlation models.

Methodology

Transformation data

Transforms are usually applied so that the data appear to more closely meet the assumptions of a statistical inference procedure that is to be applied, or to improve the interpretability or appearance of graphs. Nearly always, the function that is used to transform the data is invertible, and generally is continuous. The transformation is usually applied to a collection of comparable measurements. For example, if we are working with data on peoples' incomes in some currency unit, it would be common to transform each person's income value by the logarithm function.

Partial F (or t) values: The partial F-test is simply a statistical test for the additional contribution to prediction accuracy of a variable above that of the variables already in the equation. When a variable (X_α) is added to a regression equation after other variables are already in the equation, its contribution may be very small even though it has a high correlation with the dependent variable. The reason is that X_α is highly correlated with the variables already in the equation. The partial F value is calculated for all variables by simply pretending that each variable above all others in the equation indicates its low or insignificant partial F value for a variable not in the equation indicates its low or insignificant contribution to the model as already specified. A t value being approximately the square root of the F value [5].

Partial F values: As discussed earlier, two computational approaches simultaneous and stepwise can be utilized in deriving discriminate functions. When the stepwise method is selected, an additional means of interpreting the relative discriminate power of the independent variables is available through the use of partial F values. This is accomplished by examining the absolute sizes of the significant F values and ranking them. Large F values indicate greater discriminatory power. In practice, rankings using the F values approach are the same as the ranking derived from using discriminate weight, but the F values indicate the associated level of significance for each variable [8].

To test the hypothesis that the amount of variation explained by the regression model is more than the variation explained by the average, the F table is used. The test statistic F table is calculated as:

$$F \text{ table} = (\text{SSE regression} / \text{df regression}) / (\text{SSE total} / \text{df residual})$$

Where

$$\text{df regression} = \text{number of estimated coefficient (including intercept)} - 1$$

$$\text{df residual} = \text{sample size} - \text{number of estimated coefficient (including intercept)}$$

The t-test: The t test assesses the statistical significance of the difference between two independent sample means. If the t value is sufficiently large, then statistically we can say that the difference was not due to sampling variability, but represents a true difference. This is done by comparing the t statistic to the critical value of the t statistic (t_{crit}). If the absolute value of the statistic is greater than the critical value, this leads to rejection of the null hypothesis of no difference in the appeals of the advertising messages between groups. This means that the actual difference due to the appeals is statistically larger than the difference expected from sampling error. We determine the critical value (t_{crit}) for our t statistic and test the statistical significance of the observed differences by the following procedure:

- Compute the t statistic as the table of the difference between samples to the standard error.
- Specify a type I error level (denoted as α , or significance level), which indicates the probability level the researcher will accept in concluding that the group means are difference when in fact they are not.
- Determine the critical value (t_{crit}) by referring to the distribution with $N_1 + N_2 - 2$ degrees of freedom and specified α , where N_1 and N_2 are samples size..

If the absolute value of the computed t statistic exceeds t_{crit} , the researcher

can conclude that the two advertising messages have different levels of appeal (i.e., $\mu_1 \neq \mu_2$), with a type I error probability of α . The researcher can then examine the actual mean values to determine which group is higher on the dependent value.

Correlation coefficient (r): Another measure of the relationship between two variables is the correlation coefficient. This section describes the simple linear correlation, for short linear correlation, which measures the strength of the linear association between two variables. In other words, the linear correlation coefficient measures how closely the points in a scatter diagram are spread around regression line.

Calculated for population data is denoted by ρ (Greek letter rho) and the one calculated for sample data is, denoted by r (Note that the square of the correlation coefficient is equal to the determination) [9]. Coefficient that indicates the strength of the association between any two metric variables. The sign (+ or -) indicates the direction of the relationship. The value can range from -1 to +1, with +1 indicating a perfect positive relationship, 0 indicating no relationship, and -1 indicating a perfect negative or reverse relationship (as one variable grows larger, the other variable grows smaller) [5].

The coefficient of correlation (r) is simply the square root of the coefficient of determination (r^2). Thus, four our, the coefficient of correlation is $r = \sqrt{r^2} = \sqrt{0.982} = 0.991$. As you can see, r is a positive value in this case, but r can also be negative. The algebraic sign for r is always the same as that of b in the regression equation. The coefficient of correlation isn't as useful as the coefficient of determination, since it's an abstract decimal and isn't subject to precise interpretation. (As the square root of a percentage, it cannot itself be interpreted in percentage terms). But r does provide a scale against which the closeness of the relationship between -1.00 and +1.00. When r is zero, there is no correlation, and when $r = -1.00$ or $+1.00$, there is perfect correlation. Thus, the closer r is to its limit of ± 1.00 , the better the correlation, and the closer it is zero, the poorer the relationship between the variables [10] (Table 1).

Result and Discussion

Null hypothesis and alternative hypothesis for F-test

Null Hypothesis (H_0) that is, linear regression analysis cannot be used in analyzing the death rate/Mortality (X_1) and confirmed the COVID 19 (X_2) on The COVID 19 vaccine (Y).

Alternative Hypothesis (H_a) that is, linear regression analysis can be used in analyzing the death rate/Mortality (X_1) and confirmed the COVID 19 (X_2) on The COVID 19 vaccine (Y) (Table 2).

The F-test: The partial F-test is simply a statistical test for the additional contribution to prediction accuracy of a variable above that of the variables already in the equation. When a variable (X_α) is added to a regression equation after other variables are already in the equation, its contribution may be very small even though it has a high correlation with the dependent variable. F value is 787.62 and F table is 3.183. F value > F table is $787.62 > 3.24$ then, H_0 is rejected. So, multiple linear regression analysis can be used in analyzing the death rate/Mortality (X_1) and confirmed the COVID 19 (X_2) on The COVID 19 vaccine (Y) [11-13].

Multiple regression analysis: Although there are many different formulas

Table 1. Correlation pearson value.

Value of Correlation (r)	Interpretation (r)
0.00-0.25	Doubtful correlation
0.26-0.50	Fair correlation
0.51-0.75	Good correlation
0.76-1.00	Superior correlation

Table 2. Data independent and dependent variables.

Date	The death rate/ Mortality COVID 19 (X_1)	Confirmed the COVID 19 (X_2)	The COVID 19 vaccine (Y)
04-Jan-21	79686	5000868	3940793
11-Jan-21	89192	4828430	18725297
18-Jan-21	97667	4244093	27041142
25-Jan-21	100994	3760012	45043901
01-Feb-21	101074	3221158	61855273
08-Feb-21	92656	2730732	79226560
15-Feb-21	85194	2480313	101939133
22-Feb-21	71086	2663764	125494108
08-Mar-21	66300	2961331	189905011
15-Mar-21	62904	3308464	152024730
22-Mar-21	61672	3819085	230804904
29-Mar-21	63835	4060642	277430192
05-Apr-21	69952	4559259	327306417
12-Apr-21	74037	5238407	385434619
19-Apr-21	79986	5704421	455196888
26-Apr-21	87221	5703313	515167425
03-May-21	91414	5476552	567179728
10-May-21	96727	4817080	615591070
17-May-21	93310	4207261	665800483
24-May-21	88499	3563568	718770886
31-May-21	86954	3031537	780273491
07-Jun-21	80330	2667666	848323546
14-Jun-21	74119	2548662	921757497
21-Jun-21	73526	2612568	1627054655
28-Jun-22	64842	2726659	1732610327
05-Jul-21	58554	3075287	1732610327
12-Jul-21	54584	3534912	2185505175
19-Jul-21	56372	3909772	1913584923
22-Jul-21	57229	4118176	1994376343
26-Jul-21	69912	4329532	2078422978
02-Aug-21	64290	4505310	2078422978
09-Aug-21	66127	4559238	2158464818
16-Aug-21	67388	4426234	2551033733
23-Aug-21	68816	4509113	2362262658
30-Aug-21	68274	3999873	2466849992
06-Sep-21	67655	3760926	2550595141
13-Sep-21	64025	3075287	3110927999
	2858421	147374742	41881576167

that can be used to express regression relationships among than two variables, most widely used are linear equations of the form;

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

The results of multiple linear regression analysis obtained the equation The COVID 19 vaccine (Y) by analyzing the death rate/Mortality COVID 19 as (X_1) to the COVID 19 confirmed (X_2) we get the formula for multiple linear regression equation as follows:

$$Y = 0.609211 - 3.15845257X_1 + 3.58359X_2$$

By using the above equation we can predict or predict Life Expectancy in World in the future by using these multiple linear regression equations.

Correlation coefficient (r), partial correlation coefficient: Value that measures the strength of the relationship between the criterion or independent variable and a single independent variable when the effects of other independent variables in the model are held constant. For example, r_{y, x_2, x_1} measures the variation in Y associated with X_2 , when the effect of X_1 on both X_2 and Y is held constant. This value is used in sequential variable selection methods of regression model estimation to identify the independent variable with the greatest incremental predictive power beyond the independent variables already in the regression model [6] (Barnett, V., and T, Lewis, 1984). Value that measures the strength of the relationship between the criterion or independent variable and a single independent variable when the effects of another independent variables in the model are held constant. For r_{y, x_2, x_1} , measures the variation in Y associated with X_2 , when the effect of X_1 on both X_2 and Y is held constant. The value of r_{y, x_2, x_1} is 0.983892 with the interpretation being Superior Correlation which ranges from 0.75 to 0.99. For r_{y, x_1} measures the variation in Y associated with X_1 . The value r_{y, x_1} is 0.98886 with the interpretation being Doubtful Correlation, which ranges from 0.00 to 0.25 [14]. For r_{y, x_2} measures the variation in Y associated with X_2 . The value r_{y, x_2} is 0.9906 with the interpretation being Superior Correlation, which ranges from 0.76 to 0.99. For r_{x_1, x_2} measures the variation in X_1 associated with X_2 . The value of r_{x_1, x_2} is 0.99968 with the interpretation being. Superior Correlation, which ranges from 0.76 to 0.99.

Hypothesis for t-test

Partial influence test (t-test) between X_1 and Y

Null Hypothesis (H_0) that is, there is no significant (very large) effect partially between the death rate/Mortality COVID 19 as variable X_1 on The COVID 19 vaccine as Y.

Alternative Hypothesis (H_a) that is, there is a significant (very large) effect partially between the death rate/Mortality COVID 19 as variable X_1 to The COVID 19 vaccine as Y.

T_1 value is -12.078 and t table is 2.015. So, t_1 value < t table is the value of -12.078 < 2.015; then H_0 is accepted meaning, there is no significant (very large) effect partially between The death rate/Mortality COVID 19 as variable X_1 to The COVID 19 vaccine as Y [12] (Figure 2).

Partial influence test (t-test) between X_2 and Y: Null Hypothesis (H_0) that is, there is no significant (very large) effect partially between the death rate/Mortality COVID 19 (X_1) in World(X_2) on The COVID 19 vaccine in World (Y).

Alternative Hypothesis (H_a), that is, there is a significant (very large) partial effect between the death rate/Mortality corona viruses 19 (X_1) in World (X_2) on The COVID 19 vaccine in World (Y) [15,16].

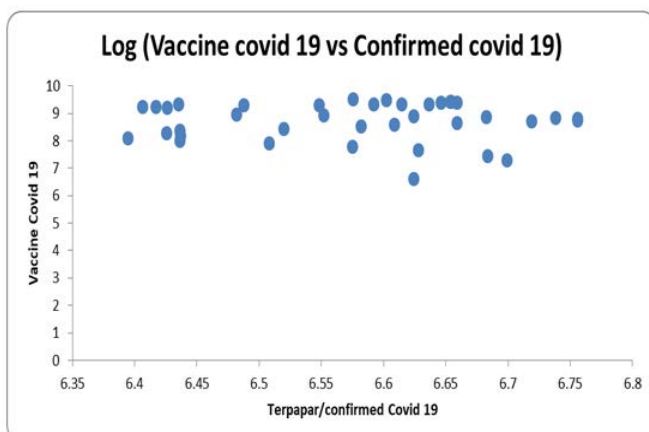


Figure 2. The COVID 19 confirmed plot to the COVID 19 vaccine.

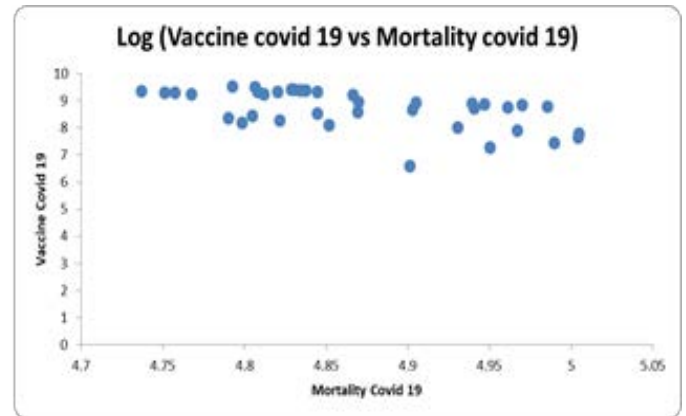


Figure 3. Confirmed COVID 19 plot to the COVID 19 vaccine.

t_2 value is 3.62523 and t table is 2.009. So, for t_2 value > t table that is 3.62523 > 2.009 then H_0 is rejected meaning, there is significant/big correlation partial effect between the death rate/Mortality COVID 19 (X_2) on The COVID 19 vaccine (Y) from December 2020 until September 2021 (Figure 3).

Conclusion

From the results of the study it can be concluded as follows: F value > F table that is 787.62 > 3.24. Then, H_0 is rejected. That is, linear regression analysis can be used in analyzing the influence of the dependent variable (free) Fossil value of fuel energy consumption (% of total) and Confirmed the COVID 19 on the independent variable (dependent) The COVID 19 vaccine. The multiple linear regression equation of these variables is as follows.

$$Y = 0.6092 - 3.15845X_1 + 3.58359X_2$$

Correlation (interrelation) of variables in multiple linear regression (r_{y, x_1, x_2}) that is, independent variable (free); The COVID 19 vaccine is variable (Y) of the dependent variable (free) Fossil value of fuel energy consumption (% of total) (X_1) and Confirmed the COVID 19 (X_2) obtained results (r_{y, x_1, x_2} is 0.983892) with interpretation is Superior Correlation which ranges from 0.75 to 0.99. For t_1 value < t table; that is -12.078 < 2.015, so, H_0 is accepted, that is a significant (large) relationship between The COVID 19 vaccine with the death rate/Mortality COVID 19. For t_2 value > t-table that is 3.6252 > 2.015, then H_0 is rejected, it means, there is a no significant/close relationship between The COVID 19 vaccine with Confirmed the COVID 19.

Acknowledgements

Without reservation, we would like thank you to my parents, my father; Misran Siregar and my Mother; Yurnita Yunus, and thank you to reviewers and the seminar participants at Persada University in Pekanbaru for their helpful comments and suggestions, as well as the World Development Indicators "World Bank Data". The views expressed in this paper are those of the authors.

References

- Bali, Pulendran. "Systems Vaccinology: Probing Humanity's Diverse Immune Systems with Vaccines." *Proc Natl Acad Sci* 111(2014):12300-12306.
- Hill, R Carter, William E Griffiths and Guay C Lim. "Principles of Economics." *J Eco* 1976.
- Donald, A Berry and B W Lindgren. "Statistic: Theory and Methods." *Brooks/Cole* 1990 (1990): 669-670.
- Box, E P George and George C Tiao. "Bayesian Inference in Statistical Analysis." *Wiley* (1973).

5. Cohen, J and P Cohen. "Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences." *Lawrence Erlbaum Associates* 4(1983).
6. Barnett, V and T Lewis. "On Discordance Tests for the Wrapped Cau-chy Distribution." *J Stat* 5(1984).
7. Jhon, E Freund. "Mathematical Statistics 5th Edition." *J Math* (1962).
8. Joseph, FHJR. "Multivariate Data Analysis 5th Edition." (1998).
9. Mann, S Premm. "Introductory Statistics." (1976).
10. Sanders, Donald H and Robert Smidt. "Electric Machinery: A First Course Sixth Edition, California Polytechnic State University." *J Mech* (1995).
11. Acaravci, Ali and Ilhan Ozturk. "Electricity Consumption and Economic Growth Nexus: A Multivariate Analysis for Turkey." *Amfiteatru Economic* 14(2012): 246-257.
12. Louis, A Thomas. "Multivariate Analysis of Variance and Repeated Measures. London: Chapman and Hall. *Wiley Online Library* 8 (1987): 1559-1559.
13. Kelley, William D, Jr. Ratliff Thomas A and Charles Nenadic. "Basic Statistics for laboratories, a primer for laboratory worker." *Van nostran drein hold, new York*. (1992).
14. National Accounts of OECD Countries 2009." OECD, (2009).
15. James, T McClave and Terry Sincich. "Statistics Eight Edition." *Printice Hall* (2000).
16. Ofori, G. "Developing the construction industry to enhance affordability of housing in World." (2013).

How to cite this article: Siregar, Dodi Irwan. "Predicting the Efficacy of the COVID 19 Vaccine by Analyzing the Death Rate and Confirmation Data of the Corona Virus". *J Bioprocess Biotech* 11 (2021): 002