

# COVID19 - A Correlation Study of Infection Fatality Rate vs Age

JAYDIP DATTA (✉ [deepjoy1972@gmail.com](mailto:deepjoy1972@gmail.com))

Indian Institute of Chemical Engineers (CRC), Kolkata, 700032, INDIA <https://orcid.org/0000-0002-0013-5269>

---

## Research Article

**Keywords:** Logistic regression, Covid19, Infection fatality rate, Age, Quadratic regression, Best fit analysis, Epidemiological parameters, Comorbidities, Linear fit

**Posted Date:** February 22nd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-85482/v3>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

In this article one of the most important epidemiological parameter ie Infection fatality rate [ 1 ] is correlated with age of the population through a sigmoid statistics of Logistic model. The IFR is a special case of case fatality rate ( CFR ) . The CFR ( 1 ) is termed as the number of deaths due to symptomatic Covid infection within entire population per unit time . The IFR is a special case of CFR where number of deaths to be considered as total number of deaths due to symptomatic as well as asymptomatic infection within the same population per unit time .The sigmoid fit can also be approximated to modified quadratic fit [ 4-5 ]. CFR can be more specifically correlated to **comorbidities [8 ]**through linear regression analysis. co morbidities due to SARS-COV-2 infection for different chronic diseases like heart , Lung , Kidney , related chronic failure are analysed by a significant Pearson statistics ( 10 ) are discussed here . The IFR can be realised from mild to hospitalisation under ICU , critical care and finally severity to death( 9,12).

## Introduction

The application of statistical data analysis is correlated by using two important epidemiological variables like IFR with Age of Covid19 patients ( 6,7 ) .The best fit regression study like logistic or sigmoid mode is analysed here . Further mathematical analysis on logistic model shows a quadratic regression model as in addition .The scope of the regression analysis ( 4)is applied to comorbities(8) as specific cases of case fatality rate( CFR ) , mortality rate ( MR ) due to Hypertension , Diabetes , Cardiac diseases , Smoking related respiratory disorders and different types of malignancy automatically increasing the CFR and MR . Most of the above chronic cases of co morbidities can be divided into mild SARS-COV-2 infection leading to hospitalised to severe in ICU and ultimately leading to pulmonary sepsis( 12 ) ie death of the patients also termed as fatality of the concerned primary as well as secondary infections.

## Methods

The methods can be sub-divided into following two statistical correlations( 4,10 ) .In each cases the statistical analysis has been carried out by the epidemiological parametric variations of Age , Infection fatality rate , comorbities of fatality frequencies and Fatality percentages(9) . The fatality frequencies include all the cases of high blood pressure, blood sugar, pulmonary disorder, smoking related pulmonary disorder , asthma and the most fatal that is malignancy for long term chronic cases .Each subclasses consists of related data as well as the regression methods of analysis like logistic to linear regression with significant values of correlation coefficient ( r ) with standard error(s) .

### Data Source ( 2,9 )

The mean of the three distribution chart ( REFERENCE [ 2 ] ) is taken as following **Figure –1**.

The Figure – 2 , Figure -3 ( REFERENCE [ 9 ] ) represents the Scattered data , Bar graph , PIE and contour representation of frequency of fatality vs Fatality percentages due to the chronic comorbities due to above mentioned diseases.

**Figure -2 and 3 representing frequency of fatality vs fatality percentages followed by comorbities .**

The data presented in Figure -1 , Figure -2 and Figure -3 are executed by **Standard software –**

**Curve Expert V.1.4** generating the following results. All these datasets showing the trends of comorbities enhanced by the chronic cases of hypertension, diabetes, pulmonary infections including and excluding smoking (COPD) , asthma and different types of carcinomas with age ,infection fatality rate ( IFR ) from mild to severe leading to fatality .

## Results And Discussion

The data analysis of IFR vs Age of the population shows a significant correlation ( r ) of Sigmoid model ( 3-5 ) . Then this mean dataset ( **Figure – 1** ) is analysed by best -fit software like of which Logistic regression shows a significant correlation coefficient ( r ) [ 0 < r < 1.0 ] with tolerable std.error ( s ) , Cov ( Matrix ) and residual table of the Parametric statistics . A part from regression analysis a PIE chart of the dataset is also shown as bellow. The sigmoid function results from the analysis can be represented as follows ( **Figure – 4** ) .

$$Y ( IFR ) = \text{Sigmoid } f [ X ( \text{Age} ) ] = [ a / ( 1 + b * \exp ( -cx ) ] \text{——eqn ( 1 )}$$

a , b , c are the coefficients obtained from best fit analysis [ 5 ]

By expanding [  $e^{ -x }$  ] series we can get the modified equation as follows :

$$\exp ( -cx ) = 1 - cx + c^2 / 2! + c^3 / 3! - \dots \text{approx} = ( 1 - bcx ) \text{——eqn ( 2 )}$$

So **eqn ( 1 )** can be simplified as  $Y = a(1 - bcx)^{-1} = a(1 + bcx - bc^2x^2 + bc^3x^3 + \dots)$

$$\text{approx} = a(1 + bcx - bc^2x^2) \text{——( eqn 3 )}$$

From **eqn ( 3 )** one can simplify the Y ( IFR ) is a second order quadratic polynomial fit of X ( Age ) .

The **eqn (1-3)** represents the desired model equation and the measure of a sigmoid probability (3,5) of fatality rate of Covid-infection with age of the patients. A simplified Excell representation ie Pie chart (Figure-5) also supports the sigmoid distribution. **eqn (3)** also represents the sigmoid probability can be also approximately represented as quadratic fit.

A further statistical analysis on fatality frequency (X) as a result of co morbidity vs fatality percentage (Y) can be expressed as a linear variation ( $Y = a + bX$ ) of the above two parameters like scattered data, bar chart, PIE and contour distribution showing the following results Figure-6, Figure-7, Figure-8 by standard Statistical checkpoint of ChiSq, Covariance matrix, Correlation table.

**Figure – 6,7, 8** an user defined linear model with the following testing of goodness of fit pearson coefficient (r) of 0.9999 with standard error (s) of 0.0390, ChiSq Convergence = 1e-006 tolerance limit, 2x 2 variance matrix correlation, a,b, c are coefficient of variations (**Figure 10,11,12**).

## Conclusion

The above plot shows a significant fit of pre-processed dataset of Age Vs IFR. Naturally the Infection Fatality Ratio of Covid19 of the said population shows just a sharp-S-rise @ Age of 45 Yrs and the infection become more fatal after an old age of 70 yrs for both symptomatic & asymptomatic patient (6,7) of the considered population.

The comorbidities within a specific correlation both for asymptomatic as well as symptomatic patient induced by different types of stages from mild to fatal at infection fatality rate (IFR). The statistical significance (**CI more than 95%, r = 0.9999, Probability p < 0.01**) (11) shows a steeper relationship between number of deaths or frequency of fatality with percentage fatality. Between chronic disease **induced comorbidities (8)** the correlation shows that the **mortality rate or percentage is maximum for heart and cardiovascular disorder lowering towards diabetes minimises for smoking related pulmonary diseases (12)**.

## References

1. *en.wikipedia.org / Case fatality rate*
2. *Mallapaty Smriti - nature.com [ article – d41586-020-02483-2 ]*
3. *en.wikipedia.org / Logistic regression*
4. *Datta J, Preprint at <https://doi.org/10.31219/osf.io/wuqa6>*
5. *Datta J, [https://figshare.com/articles/dataset/Data\\_-\\_INFECTION\\_FATALITY\\_RATE\\_OF\\_COVID19\\_A\\_LOGISTIC\\_MODEL/12979856](https://figshare.com/articles/dataset/Data_-_INFECTION_FATALITY_RATE_OF_COVID19_A_LOGISTIC_MODEL/12979856)*
6. *AYDIP DATTA. (2020, October 5). COVID19 – A CORRELATION STUDY OF INFECTION FATALITY RATE vs AGE (Version <https://semver.org/#semantic-versioning-200>). Zenodo. <http://doi.org/10.5281/zenodo.4066767>*
7. *JAYDIP DATTA. COVID19 - A Correlation Study of Infection Fatality Rate vs Age, 20 October 2020, PREPRINT (Version 2) available at Research Square [<https://doi.org/10.21203/rs.3.rs-85482/v2>]*
8. *<https://en.wikipedia.org/wiki/Comorbidity>*
9. *JAYDIP DATTA. (2021). Supplementary dataset – Comorbidities due to SARS-COV-2 (Version Verson 1.0) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.4553423>*
10. *<https://www.socscistatistics.com/pvalues/pearsondistribution.aspx>*
11. *<https://www.medcalc.org/manual/correlation.php#:~:text=The%20P%2Dvalue%20is%20the,coefficient%20is%20called%20statistically%20significant>*
12. *JAYDIP DATTA. Mortality Rate Due to Pulmonary Fibrosis Associated with Sars-Cov-2 Infection: Scope of Best Fit Regression, 29 December 2020, PREPRINT (Version 1) available at Research Square [<https://doi.org/10.21203/rs.3.rs-137219/v1>]*

## Declarations

I hereby declare that I have no conflict of interest; I also declare that I have no competing interest regarding this submission.

## Figures

Age	IFR
15	0
25	0.01
33.2	0.035
45	0.07
55.52	0.26
63	2.36
78	9.4

Figure 1

The mean of the three distribution chart

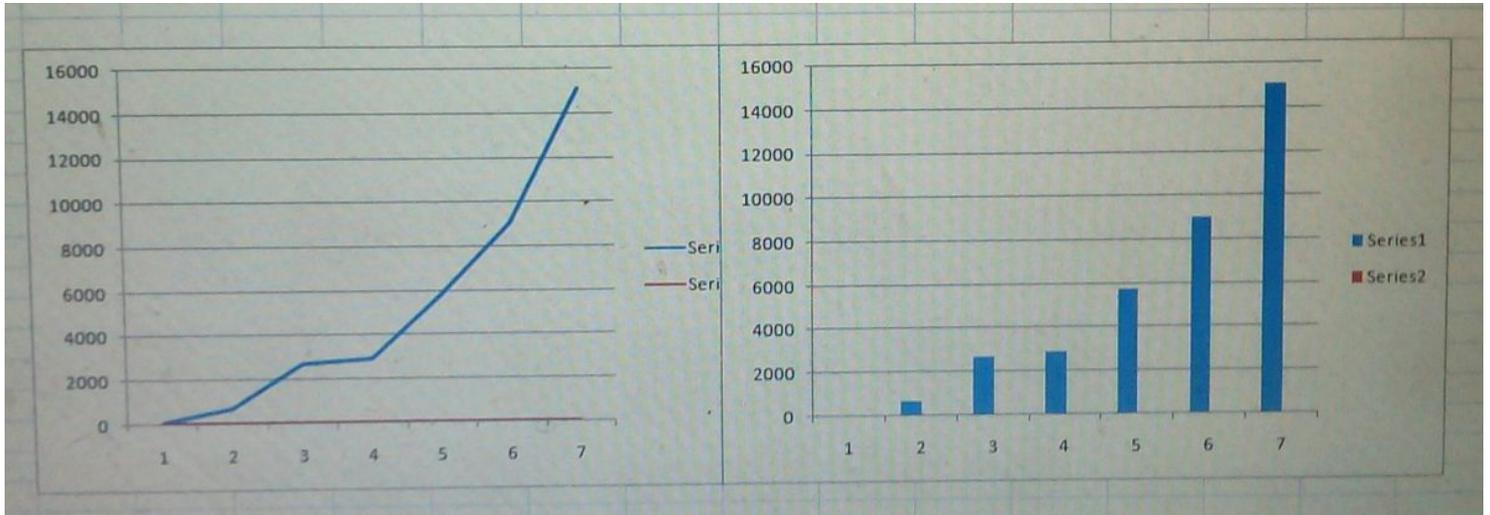


Figure 2

Scattered Data

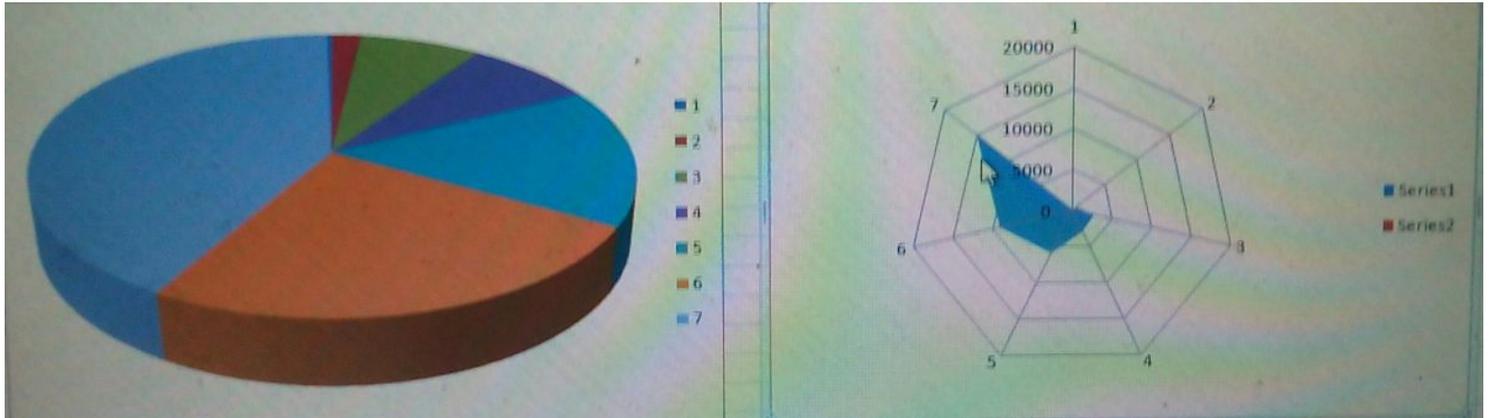


Figure 3

Density variation of Comorbities

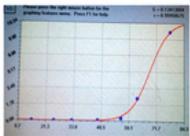


Figure 4

Sigmoid Curve

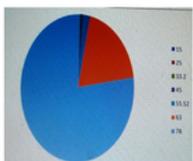


Figure 5

PIE Distribution

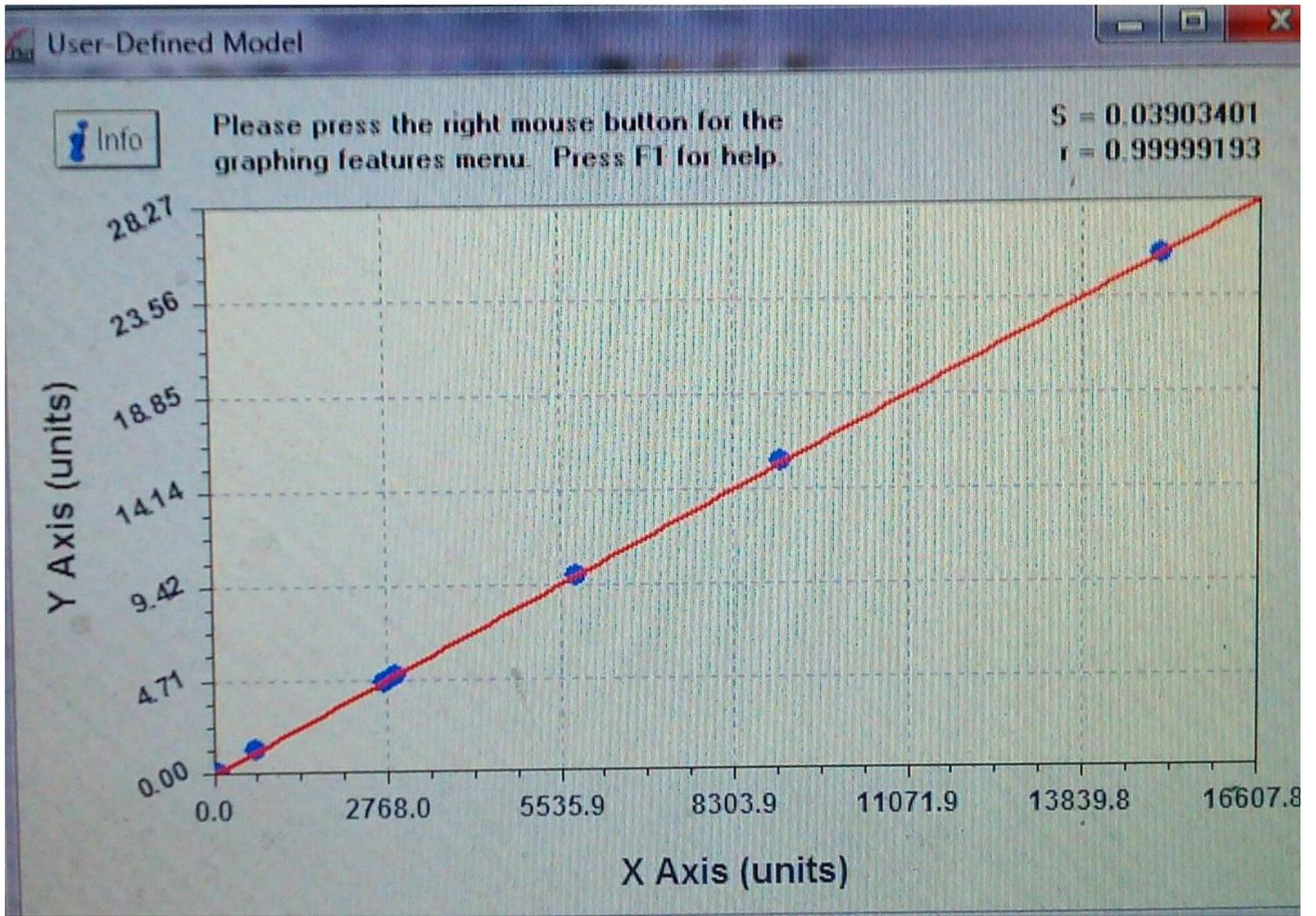


Figure 6

Linear Fit

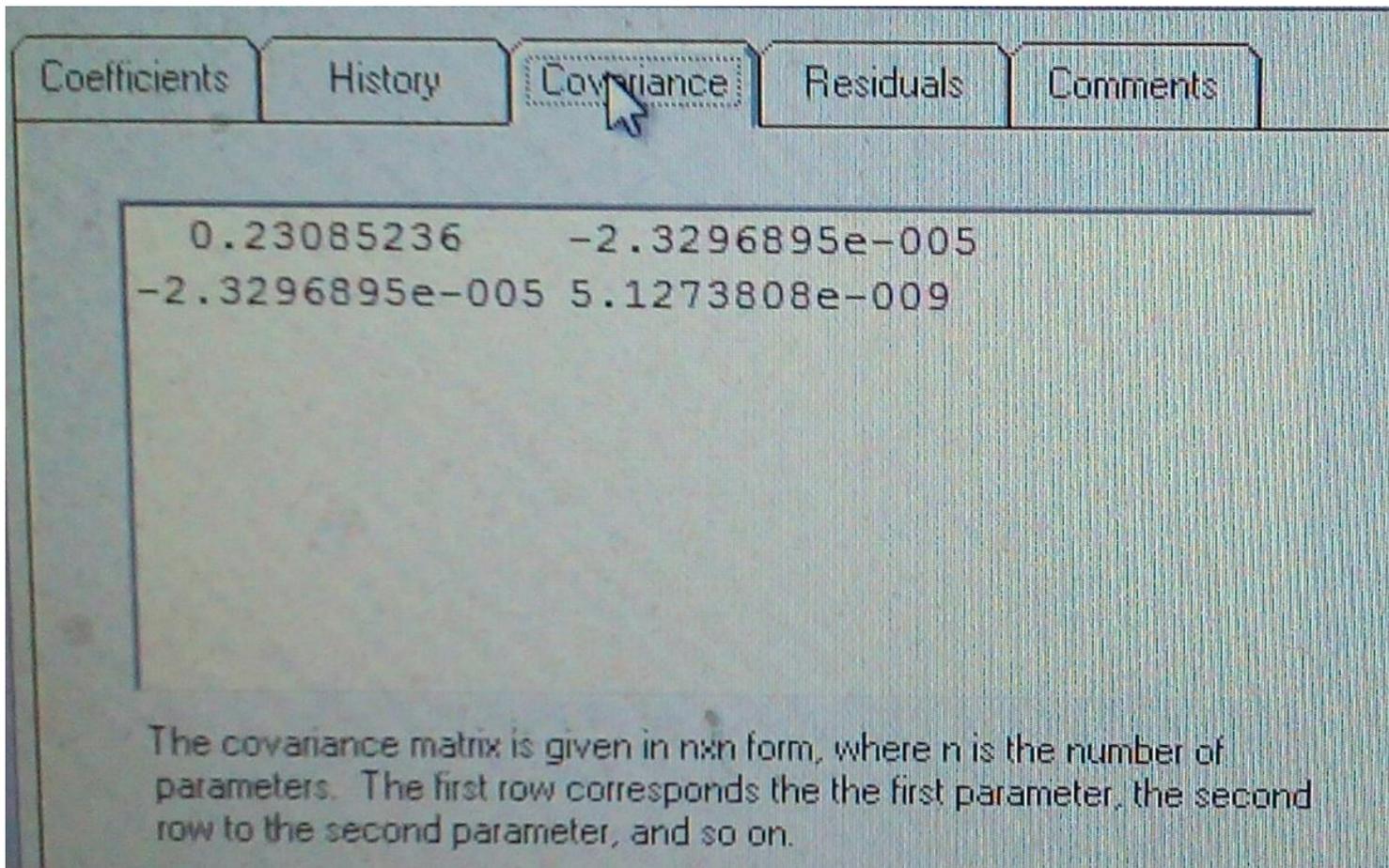


Figure 7

Variance Matrix

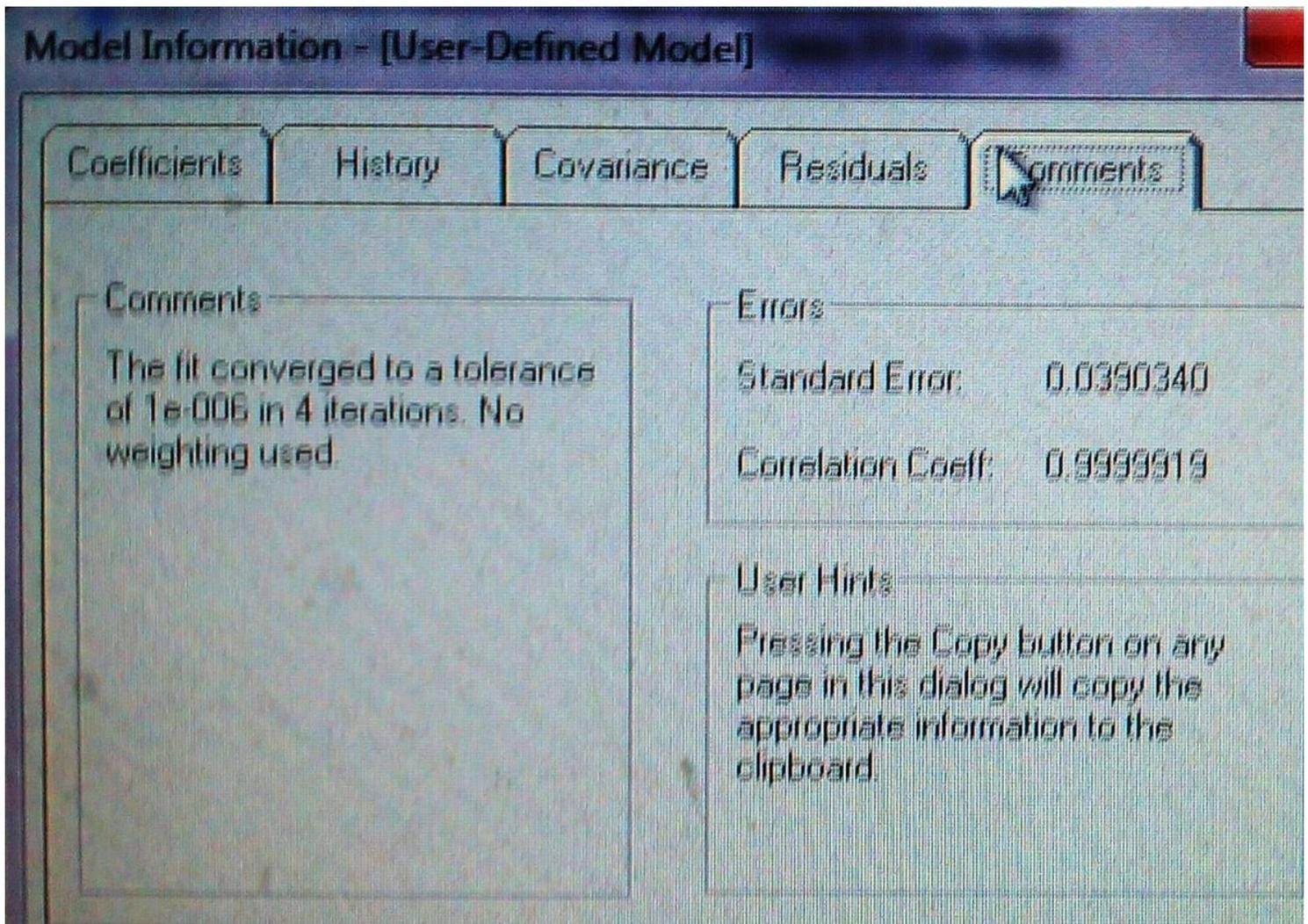


Figure 8

Correlation Table

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [DataCovidIFR.pdf](#)
- [SupplimentaryFileComorbiditiesUpdated.docx](#)