Mathematical and Statistical Study on COVID19- SIR Model

Jaydip Datta (✉ deepjoy1972@gmail.com)

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

Research Article

Keywords: Covid19, SARS-COV-2, Random chain biopolymer, SP3 configuration, Molecular simulation, Exponential, Gaussian, Drug intermediate, Drug modeling, Quantum Chemistry, LCAO, Infection fatality rate, Logistic regression, Sigmoid model, Hydrodynamic modeling, Bio-polymer, First order kinetics, SIR compartmental model, p-value, Confidence limit (CI), Covid-vaccination

Posted Date: April 5th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-89766/v3>

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

MATHEMATICAL AND STATISTICAL STUDY ON COVID19 – SIR MODEL

Jaydip Datta

ABSTRACT -

In this combinatorial study let us try to simulate the four cases starting from viral spreading kinetics , Gaussian distribution of the infectious disease , mortality statistics like infection fatality ratio (IFR) with the distribution of age of the patient through sigmoid regression method approach and finally the most important modeling of remdesivir on the basis of Molecular bonding method (LCAO) as well as vaccination as the traditional method . It is an alternative approach to susceptibility – infectivity – recovery (SIR) model .

Keywords – Covid19 , SARS-COV-2 , Random chain biopolymer , SP3 configuration , Molecular simulation , Exponential , Gaussian , Drug intermediate , Drug modeling , Quantum Chemistry , LCAO , Infection fatality rate , Logistic regression , Sigmoid model , Hydrodynamic modeling , Bio-polymer , First order kinetics , SIR compartmental model , p-value , Confidence limit (CI) , Covid vaccination .

Literature review –

In my earlier work by **Datta J** I have reviewed the three basic mathematical modeling related to Covid19 pointing out exponential series , Gaussian distribution function of one dimension , Probability and Probability density function , Linear combination of atomic orbital starting from wave function and area under the curve of Gaussian as well as flat Gaussian or Pseudo-Gaussian methods . The fourth work is in continuation of infection fatality rate (IFR) vs Age of **Mallapaty S** (Data Source) in different statistical angle of best fit analysis indication logistic or sigmoid probability.

Objective – The main objective of the composite modeling is to correlate the Gaussian , Pseudo-Gaussian function as susceptibility , Sigmoid function as spreading and Drug modelling as Recovery compartment of epidemiological SIR model .

Figure Ref : en.encyclopedia.org

METHODOLOGY -

To start four cases will be discussed as per following order –

Case I – Viral spreading kinetic study :

Most of the microbes including virus should follow four stages namely - lag phase , LOG or exponential phase , stationary phase and death phase . Under this present case of Covid19 the exponential phase is more prominent for spreading of infection as well as sharp one . This exponential phase may be shifted to a flat one by a prolonged lock down may decrease the infection susceptibility. This type of Covid infection then enters to a prolonged static phase of 14 days or more depending on mild to severe infection for symptomatic or asymptomatic patients . The declining or

death phase will be mathematically an asymptotes ie infinite time will be required to irradiate to susceptibility of infection .

Case II – Viral spreading distribution study :

SARS-COV-2 being a virulent viral biopolymer (like other viruses) the distribution probability is a Gaussian one with conventional exponential phase generates a expo-gaussian contour or pseudo Gaussian contour ie why the the infective speed is uncontrollable , epidemic rather pandemic .

The Case I & Case II may be hypothesised to susceptibility study of SARS-COV-2 infection.

Case III – Mortality statistics in terms of infection fatality rate (IFR) with age of the patients . The corresponding best fit analysis shows a logistic or sigmoid regression model. The sharp S rise of the model supports the conventional exponential phase .

The best fit analysis plot as well as excel pie chart shows the infectivity or fatality of infection probability vs Age . Obviously the inflection data changes from 46 yrs to an old age patient of more than 70 yrs are having much more prone to death.

The Case III may be hypothesised to infectivity study of SARS-COV-2 .

Case IV :

Can be sub divided into the following :

(a) It is related to novel drug modelling hypothesis of the most effective use of Covid19 antiviral like Remdesivir . In this hypothesis we will analyse the Linear combination of atomic orbital (LCAO) to overlap S -Orbital and P-Orbital to give rise to a hybrid Sp orbital having 50 % s and 50 % p – Character. Like S–Orbital – spherically symmetrical geometry like spike crown & Remdesivir the most potent synthetic -antiviral may be assumed to dumbbell shaped P-Orbital in geometry. Here Covid – Remdesivir drug – Protein [DP] intermediate may be considered SP hybridisation analogous to molecular bonding approach .The combined approach of pharmaco-kinetic stability - M.O approach can be simulated as Remdesivir – nucleocapsid protein complex as follows .

MOLECULAR SIMULATION STUDY :

A Combined Model for Covid19 (S-orbital) + Remdesivir (P-orbital)----->

□

Covid -Remdesivir (A SP Hybridized Model Hypothesis) .

Case IV(a) may be hypothesised to recovery modeling as a potent antiviral of SARS-COV-2 as mentioned above .

(b) Vaccination study :

Vaccination is a preventive approach rather than antiviral to combat against SARS-COV-2 . During vaccination program the most important variable is population of the said country. China, India the population density is maximum compared to UK, Germany, the minimum population density is observed . The basis of first dosage already vaccinated is 1030 million globally.

RESULTS -

Basically includes model equations & hypothesis in each case.

The exponential viral growth kinetics shows a 1st order kinetics as mentioned –
 $C_t = C_0 * \text{EXP} (- K * t)$ ----- (1) [Case – I]

[Where C_0 = Initial bacterial count, bacterial Colony Count (C_t) at time t & K = 1st order growth const & t = time of growth]

General mathematical expression of Gaussian Probability distribution of biopolymer can be expressed as :

$$P (x) = A \exp (- B * x ^ 2) \text{ (2)}$$

A,B are constants depending on the physic-chemical properties of bio-polymer chain .

considering one dimension Gaussian probability $P (x)$ ranging from - infinity to + infinity with mean distribution is zero .For two /three dimension the expression includes dimension factor . To find best fit analysis through Regression method the Gaussian fit is one of the useful one for hydrodynamic modelling of Covid19 considering the flow of viral infection through our blood .

The Gaussian probability can be represented as randomisation of monomer chain of biopolymers (including virus like Covid19) (N =Number of Monomers vs R_{rms} = An Important Statistical

parameters like Root mean square end to end separation as a $F(X)$ of Radius of gyration (R_g). The representative equation is as follows. The R_g of SARS-COV-2 biopolymer with a tremendous SP3 monomer leads to the virus more and more virulent as well as genetically recombinant.

$$P(N, x) = A \cdot \exp(-0.042 \cdot R_{rms}^2) \dots\dots\dots (3) \text{ [Case – II]}$$

X = distribution chain length of random biopolymer coil

R_{rms} = Root mean square end to end separation of

bio-polymer chain.

CuVer finder V.1.4 run mode gives the best fit logistic

regression as below :

$Y(\text{IFR}) = \text{Sigmoid } f[X(\text{Age})] = [a / (1 + b \cdot \exp(-cx))] \dots\dots\dots (4) \text{ [Case –III]}$

a, b, c are the coefficients obtained from best fit analysis.

The equation (4) represents the desired model equation and the measure of a sigmoid probability of Covid infection with a significant correlation (r) value [$0 < r < 1$]. The Pearson coefficient (r) in this case is 0.9995 with $SE = 0.1341$, $p < 0.00001$, $CI > 95\%$, with $p < 0.05$ makes the sigmoid statistics more significant. The covariance matrix resulting logistic regression of the data of interest is [3x3] square matrix.

Hypothesis regarding Remdesivir antiviral drug-nucleocapsid protein of Covid19 –Quantum Molecular bonding approach of drug-design. (Case –IV .a)

The Theorem of Normalisation - Integration of ψ ($\psi \cdot \psi$) $dV = 1 \dots\dots\dots (5)$

Very useful theorem in quantum chemistry. dV = volume element for a probability $\psi \cdot \psi$

.Considering Covid19 as an uniform sphere we can compare with the spherically symmetrical s -Orbital in Atomic structure. Wave particle duality hypothesis assumes a wave nature for quick viral spreading as well as spike-crown like particle. We can apply the probability density theory on s -Orbital where $dV = \frac{4}{3} \cdot \pi \cdot R^3$.

Applying cylindrical polar symmetry on spike crown (R, θ, ϕ), the cylindrical probability can be calculated as usual.

Now considering Remdesivir as p -orbital, we can assume the [Drug- Protein] Complex as SP hybridisation for linear overlap.

The modeling relationship from LCAO (Linear combination of Atomic Orbital) –

$\psi(\text{Covid19}) + \psi(\text{Remdesivir}) = \psi(\text{Drug- Protein Intermediate})$

$\Rightarrow \psi(\text{Covid19}) [\psi \cdot (\text{Remdesivir})] dV \dots\dots\dots (6)$ represents the mass-probability of the [D-P] complex intermediate. (ψ, ψ^* are wave function and conjugate wave function respectively). The equation (6) correlating the stability of [D-P] complex intermediate through the approach of molecular bonding. On the other hand considering the pharmacokinetic drug receptor stability constant ($K_{\text{stability}}$) = $[D-P] / [D] \times [P] \dots\dots\dots (7)$

So the correlation from (6) & (7) becomes :

$K_{\text{stability}} = \psi(D) \times [\psi^*(P)] = [D-P] / [D] \times [P] \dots\dots\dots (8)$

The equation (8) can be hypothesised as recovery stability of SIR model.

In the Case IV (b). A 4th degree polynomial fit is observed with correlation coefficient 0.9682 signifies the global vaccination with confidence interval ($CI > 95\%$).

CONCLUSION –

Between four cases I – II [eqn (1,2,3)] may be simulated as **Susceptibility**, case-III [eqn(4)] may be simulated as **Infectivity** of Covid19 and case IV [eqn (5,6,7,8)] may be considered as probability of Remdesivir – SARS-COV-2 intermediate complex[IV(a)] and vaccination case [IV (b)] resulting **Recovery stability** of the patient .It is an alternative approach to study SIR model .

REFERENCE :

1. *INDUSTRIAL MICROBIOLOGY* by PRESCOTT & DUNN .
2. Image Source - en.wikipedia.org (microbial growth kinetics curve) , industry.com (Covid19) , can.acs.org (Remdesivir) , en.wikipedia.org (SIR model) .
3. Datta J , *MICROBIAL GROWTH KINETICS : A MATHEMATICAL REVIEW (SPECIAL ADDENDUM - COVID19)* (figshare.com) , May , 2020.
4. Datta J. *Gaussian Probability : Dimensional View* , (figshare.com) , May 2020 .
5. Datta J , *HYDRODYNAMICS: AN EXPERIMENTAL VIEW ON BIO MOLECULE (REVISED)* , osf.io , April , 2020 .
- 6 . Datta J , <https://www.researchgate.net/search> (What are the BASIC mathematical methods specially used in computational chemistry covering Covid19 ? .
- 7 . . https://en.wikipedia.org/wiki/Case_fatality_rate
- 8 . . Mallapaty Smriti - <https://www.nature.com/articles/d41586-020-02483-2>
9. https://en.wikipedia.org/wiki/Logistic_regression#:~:text=Logistic%20regression%20is%20a%20statistical,a%20form%20of%20binary%20regression).
- 10 . Datta J , https://figshare.com/articles/dataset/Data_-_INFECTION_FATALITY_RATE_OF_COVID19_A_LOGISTIC_MODEL/12979856

11. Datta J , at Preprint (<https://doi.org/10.31219/osf.io/wuqa6>) .
12. Datta J , at Preprint (<https://doi.org/10.6084/m9.figshare.12264491>) .
13. https://www.researchgate.net/post/What_are_the_BASIC_mathematical_methods_specially_used_in_COMPUTATIONAL_CHEMISTRY_covering_Covid19
14. Datta J, COVID 19 – A COMPREHENSIVE MATHEMATICAL APPROACH , (<https://osf.io/skzxb> - DOI 10.17605/OSF.IO/SKZXB)
15. https://figshare.com/articles/MICROBIAL_GROWTH_KINETICS_A_MATHEMATICAL_REVIEW_SPECIAL_ADDENDUM_-_COVID19_/12280790 , Datta J , May ,11, 2020 .
16. Datta J , <https://osf.io/78jvr> / (DOI 10.17605/OSF.IO/78JVR) , May ,18 ,2020 .
- 17 . Datta J , Preprint at <https://doi.org/10.6084/m9.figshare.12533348>
- 18 . Datta J , Preprint at figshare.com(<https://doi.org/10.6084/m9.figshare.12579941>)
- 19 JAYDIP 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>.
- 20 .DATTA J, https://www.researchgate.net/post/Justification_of_Remdesivir_in_the_treatment_of_SARS-COV-2.
21. DATTA J, https://www.researchgate.net/post/Vaccinations_or_antivirals_to_treat_SARS-COV-2_infection-Which_options_do_you_prefer
22. Datta J, https://www.researchgate.net/post/Whether_successive_vaccination_will_increase_or_decrease_HERD_IMMUNITY
23. Datta J, https://www.researchgate.net/post/How_best_fit_regression_method_can_be_applied_to_model_SARS-COV-2_epidemiology

CITATION :

- 24 .Jaydip Datta. Mathematical and Statistical Study on Covid19- SIR Model, 09 October 2020, PREPRINT (Version 1) available at Research Square [+<https://doi.org/10.21203/rs.3.rs-89766/v1>].
25. JAYDIP DATTA . COVID19 - A Correlation Study of Infection Fatality Rate vs Age, 05 October 2020, PREPRINT (Version 1) available at Research Square [+<https://doi.org/10.21203/rs.3.rs-85482/v1>].
26. Jaydip Datta. Mathematical and Statistical Study on COVID19- SIR Model, 12 October 2020, PREPRINT (Version 2) available at Research Square [+<https://doi.org/10.21203/rs.3.rs-89766/v2>]
27. Jaydip Datta. A Statistical Analysis on COVID19 Global Vaccination by Best Fit Regression Method, 10 May 2021, PREPRINT (Version 1) available at Research Square [<https://doi.org/10.21203/rs.3.rs-512890/v1>] .

Supplementary Files

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

- [Supplymentaryfile.pdf](#)