Original Research

Forecasting the timeframe of 2019-nCoV and human cells interaction with reverse engineering

Ayesha Sohail [a], [*], Alessandro Nutini [b]

[a] Department of Mathematics, Comsats University Islamabad, Lahore Campus, Lahore, 54000, Pakistan
[b] Centro Studi Attività Motorie - Biomechanics Division, via di Tiglio 94 loc. Arancio, Lucca, 55100 Italy

ABSTRACT

In December 2019, an atypical pneumonia invaded the city of Wuhan, China, and the causative agent of this disease turned out to be a new coronavirus. In January 2020, the World Health Organization named the new coronavirus 2019-nCoV and subsequently it is referred to as SARS-CoV2 and the related disease as CoViD-19 (Lai et al., 2020). Very quickly, the epidemic led to a pandemic and it is now a worldwide emergency requiring the creation of new antiviral therapies and a related vaccine. The purpose of this article is to review and investigate further the molecular mechanism by which the SARS-CoV2 virus infection proceeds via the formation of a hetero-trimer between its protein S, the ACE2 receptor and the B0AT1 protein, which is the "entry receptor" for the infection process involving membrane fusion (Li et al., 2003). A reverse engineering process uses the formalism of the Hill function to represent the functions related to the dynamics of the biochemical interactions of the viral infection process. Then, using a logical evaluation of viral density that measures the rate at which the cells are hijacked by the virus (and they provide a place for the virus to replicate) and considering the "time delay" given by the interaction between cell and virus, the expected duration of the incubation period is predicted. The conclusion is that the density of the virus varies from the "exposure time" to the "interaction time" (virus-cells). This model can be used both to evaluate the infectious condition and to analyze the incubation period.

Background: The ongoing threat of the new coronavirus SARS-CoV2 pandemic is alarming and strategies for combating infection are highly desired. This RNA virus belongs to the β-coronavirus genus and is similar in some features to SARS-CoV. Currently, no vaccine or approved medical treatment is available. The complex dynamics of the rapid spread of this virus can be demonstrated with the aid of a computational framework.

Methods: A mathematical model based on the principles of cell-virus interaction is developed in this manuscript. The amino acid sequence of S proein and its interaction with the ACE-2 protein is mimicked with the aid of Hill function. The mathematical model with delay is solved with the aid of numerical solvers and the parametric values are obtained with the help of MCMC algorithm.

Results: A delay differential equation model is developed to demonstrate the dynamics of target cells, infected cells and the SARS-CoV2. The important parameters and coefficients are demonstrated with the aid of numerical computations. The resulting thresholds and forecasting may prove to be useful tools for future experimental studies and control strategies.

Conclusions: From the analysis, I is concluded that control strategy via delay is a promising technique and the role of Hill function formalism in control strategies can be better interpreted in an inexpensive manner with the aid of a theoretical framework.

---

* Corresponding author.
  E-mail address: ayeshasohail81@gmail.com (A. Sohail).

1. Introduction

Coronavirus (CoV) is a pathogen that infects humans and other vertebrates; more precisely, it infects the respiratory system, the gastrointestinal tract, the hepatic and nervous systems of humans, birds, bats, mice and other animals. The SARS (Severe Acute Respiratory Syndrome) epidemic of 2002 and the MERS (Middle East Respiratory Syndrome) of 2012 demonstrated the "spill over" event of this virus, first in "animal to man" and then "man to man " transmission. In December 2019, atypical pneumonia invaded the city of Wuhan, China, and the causative agent of this disease was identified as a new coronavirus that has been subsequently sequenced by five independent laboratories in China (http://www.virological.org). In January 2020, the World Health Organization named the new coronavirus "2019-nCoV" and subsequently referred to it as SARS-CoV2 with its related disease referred to as CoViD-19 (Lai et al., 2020).

Very soon, the epidemic created a pandemic, which is still progressing due to the human mobility that took place in early 2020. A severe outbreak was subsequently reported in Italy (and now extensive outbreaks are also registered in Iran, France, Germany, Spain, UK, USA and elsewhere), resulting in an emergency situation, to impose containment measures of a kind not recently reported, where civilians have been kept in forced quarantine and the commercial services were "locked down"; this is now a worldwide emergency and for this reason the creation of new antiviral therapies and related vaccines is urgently needed, based upon a possible "functional model" of this virus.

1.1. Structure

To understand the infectious nature of this coronavirus SARS-CoV2 (RNA virus), it is necessary to understand its structure. The virus belongs to the beta-coronavirus genus and is closely related to another virus of the same family known as SARS-CoV but, unlike this, SARS-CoV2 has a higher degree of transmission between humans (Wu et al., 2020). SARS-CoV2 has a multifunctional molecular mechanism associated with an envelope with three structural proteins (glycoproteins): M protein (membrane protein), E protein (envelope protein), S protein (that directly mediates the host's infection) and N protein (viral RNA binding). The structure is presented in Figs. 1 and 2 respectively.

1.1.1. Cell-virus interaction mechanism

The mechanism of disease transmission depends on the spike protein (S-protein as shown in Fig. 1). This protein binds the virus in reaction catalysed by the host cell's surface enzymes. Recent gene
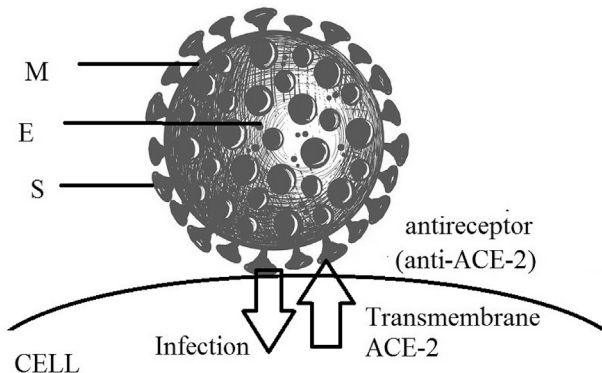
analysis (Coutard et al., 2020) helped to distinguish between the characteristics of the SARS-CoV2 S proteins and those of SARS-CoV, mainly referring to a "cleavage site" peculiar to the furin enzyme (found in CoViD-19 and not in other CoVs). SARS-CoV2 infects the host and gains the protein S, comprised of three regions:

- extensive ectodomain (Secto)
- a transmembrane anchor (single pass)
- a short intracellular tail

The Secto ectodomain, which has two subunits: a subunit-anti-receptor called S1 and a subunit S2 with strong lithic action, (defined as lysis of the host cell membrane to carry out the virus-host fusion and start the infectious process) fuses the virus with the membrane, allowing the viral genome to enter the host. A crosstalk between S1 and S2 then activates the pre-fusion conformation, activating S2 itself. The S1 subunit can be further divided into four core domains: S1a, S1b, S1c and S1d.

Subunit S1 contains RBD (Receptor Binding Domain), which binds to the ACE-2 protein (angiotensin-converting enzyme 2) of the host cell membrane (on the surface of respiratory cells), which, in turn functions as a receptor for the ligand (anti-receptor) constituted by the RBD domain itself. SARS-CoV and SARS-CoV2 share 77.5% similarity in the amino acid sequence of S1. As mentioned above, viral entry in the host requires S protein priming by cellular proteases, which involve S protein cleavage in S1 and S2 subunits where the S2 subunit allows the fusion of viral and cellular membranes; the S protein of SARS-CoV engages ACE-2 as the "entry-receptor" (Li et al., 2003) and employs the cellular serine-protease TMPRSS2 for S-priming (Glowacka et al., 2011). Due to the high percentage identity of amino-acid sequence between SARS-CoV S-protein and SARS-CoV2 S-protein, it is highly possible that both SARS viruses use the same serine-protease. The linking to ACE-2 for virus entry infection shows "high efficiency" in its function and this is a possibly a major factor of SARS-CoV and SARS-CoV2 high transmissibility (Li et al., 2005). Is also possible that a TMPRSS2-inhibitor, approved for clinical use, blocks entry of the virus and constitutes a treatment option (Hoffmann et al., 2020).

The binding with the ACE-2 protein occurs at the core domains S1a and this complex (ACE-2 + S1a) makes the reaction irreversible. During this research, we have used the Hill function formalism to illustrate the ligand-protein binding process. Details are provided in section 3. ACE-2, on the host cell membrane, is usually connected to another protein called B^0AT1. SARS-CoV2 infection begins with the initial formation of a hetero trimer consisting of the RBD domain of the S1 subunit, from ACE-2 protein (in its PD domain) and B^0AT1 protein. For future research, this trimeric complex can be explored and manipulated. A study conducted by (Wrapp et al., 2020) suggests that ACE-2 must dimerise to become activate, resulting in two PD homo-dimer domains capable of binding the SARS-CoV2 S protein and forming a trimer. Furthermore, they reported that SARS-CoV2 protein S forms trimers with two of the RBD domains: one facing upwards and one facing downwards. This structure is presented in Fig. 2.

Researchers around the world are making every effort to develop a vaccine against the virus, for example a recent study reported the dependence of "SARS-CoV-2 cell entry" on ACE2 (Hoffmann et al., 2020). The molecular bond between SARS-CoV2 spike protein and ACE2 looks similar to that of the SARS-CoV, but there are some differences in the precise amino acids used to bind SARS-CoV2 to that ACE2 receptor compared to SARS-CoV. Most likely the high affinity, probably derived from such small variations of 2019-nCoV S protein for human ACE2, may contribute to the apparent ease with which 2019-nCoV can spread from human to human (Wrapp et al., 2020). Recently (Wang et al., 2020), have
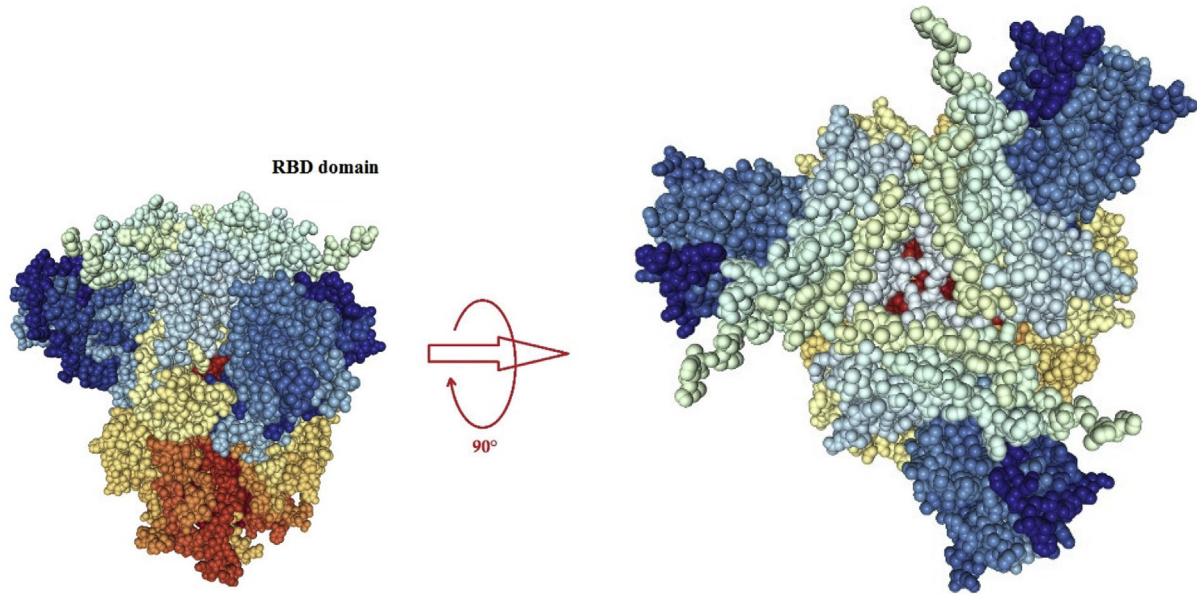

Fig. 1. Illustration of SARS-CoV2 protein-cell interaction. E, S and M are the proteins respectively.

Fig. 2. Schematic of 2019-nCoV S primary structure colored by domain. Spike protein with RBD domain; on right show the top projection after a 90° rotation. Also shown the RBD domain in up-configuration.

discussed the possibility of using a monoclonal antibody as an antiviral therapy: i.e. "47D11" antibody that targets the trimeric complex involving glycoprotein S. In this research, tests performed by the ELISA cross-reactivity system with S1-SARS-S hybridomas in supernatant indicated four types of hybridomas showing a cross reactivity with S1 subunits. Among these hybridomas, the one called 47D11 showed neutralization activity for both SARS-CoV and SARS-CoV2 and targeted the RBD domain of S1b. Furthermore, Wang et al. (2020) showed that 47D11 has a greater affinity for the ectodomain of the S (Secto) protein of SARS-CoV than SARS-CoV2, despite the antigen being present in equimolar amounts. This bond, however, does not compete with the binding of S1b with ACE-2.47D11 is included among the monoclonal antibodies that can inactivate the "S-protein" of SARS-CoV2 by preventing the forma- tion of the pre-fusion structure (Walls et al., 2019). A further structural analysis has shown that "S1b" core domain can be divided into a core subdomain of about 338−506 amino acids and a receptor-binding subdomain of about 438−498 amino acids; the

latter consists of a "loop" protruding from the β-sheet structure of the core subdomain. Currently, the inhibition mechanism operated by 47D11 is not understood. Thus in this review we investigate through a quantitative mathematical model, the possible existence of a "cooperative binding" that CoViD-19 shows in the infection process as described above. Our hypothesis is supported by the need for activation of the infection system by the virus, given by the particular molecular kinetics that leads to the formation of the "infection trimer" given by the viral S1 protein and the ACE-2 re- ceptor complex. Some studies have shown the possibility of an analysis in this sense conducted through the "Hill function" and have shown a "positive cooperativity" in the case of infection of para-flu viruses that act by binding with a Hemagglutinin- Neuraminidase Receptor (Tappert et al., 2013). The interaction, demonstrated by 47D11 monoclonal antibody, does not affect the bond between the protein S and ACE-2 but has affinity with the RBD domain of the same protein S, suggesting that there is a "facilitation" implying an amplification of the process that, in turn,
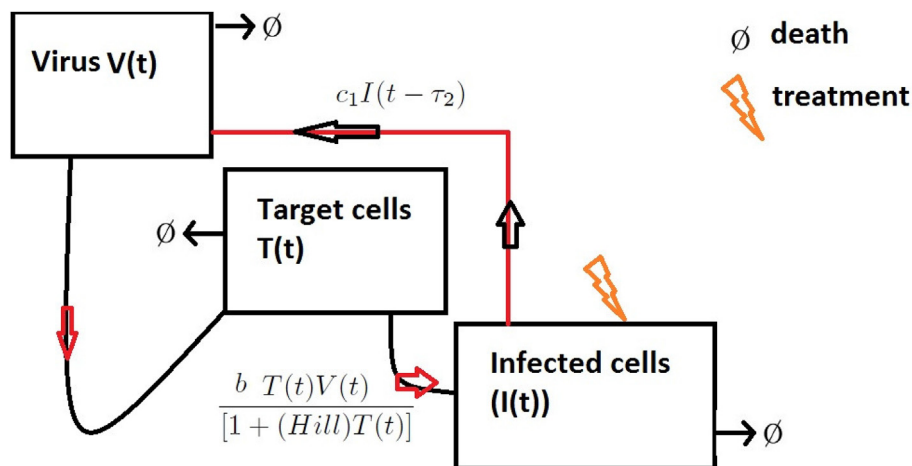

Fig. 3. Schematic description of the disease dynamics.

Table 1
Parametric approximation.

Parameters	Mean	Standard Deviation	MC-err
A	0.014818	0.00040725	3.988×10^{-5}
c_1	4.9561	7.731e-14	1.895e-14
Treatment	−0.42597	8.9973e-15	2.0631e-15

causes an increase in the infection capacity. In Section 2 the background and the development of the model are described, whereas Section 3 discusses the facts linked with the novel virus onset. The conclusions and future work are provided after the discussion.

The rest of the manuscript is organised as follows: Section 2 presents the background and the development of the model. Section 3 provides the discussion based on the facts linked with the novel virus onset. The conclusions and future work are provided after the discussion.

2. Computational framework

Based on the structural dynamics and infectious mechanism of SARS-CoV2, discussed in the Introduction, the step by step development of the model is as follows:

2.1. Hill function formalism

While developing the computational framework, the virus-target cell interaction was studied in depth. For this purpose, Hill function formalism, which helps to depict the dynamical behaviour and the biochemical interactions, is utilized. The model development can be understood with the aid of the schematic presented in Fig. 3. A standard mathematical description of this function is given as:

$$f(x(t)) = \frac{x^k}{a^k + x^k}$$

where k is kept equal to 1 in this study. Further details of the biochemical reactions and hill function formalism can be found in (Iftikhar et al., 2020; Sherin et al., 2018). The normalized version $\frac{T(t)}{1+(Hill)T(t)}$ of this standard hill function is considered in this study. Here '$Hill$' is used as a constant.

2.2. Delay in virus-target cell interaction

The term $\frac{V(t)T(t)}{1+(Hill)T(t)}$ is then used in equation (1) to show the interaction between virus and target cells. After a delay of τ_1, in light of delayed dynamics (Reddy and Sansom, 2016), the term
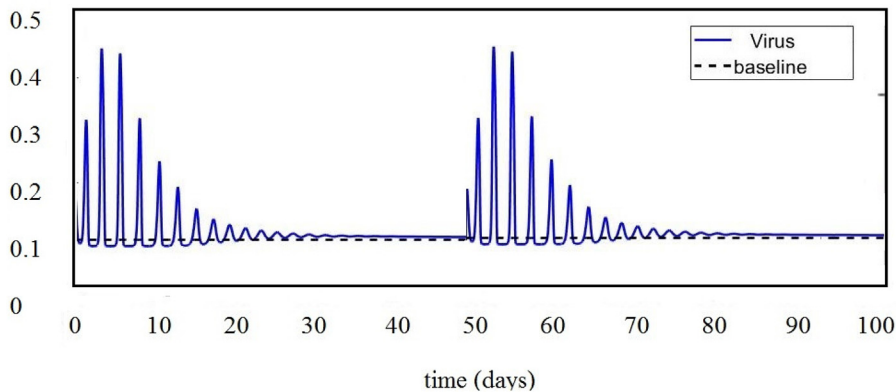

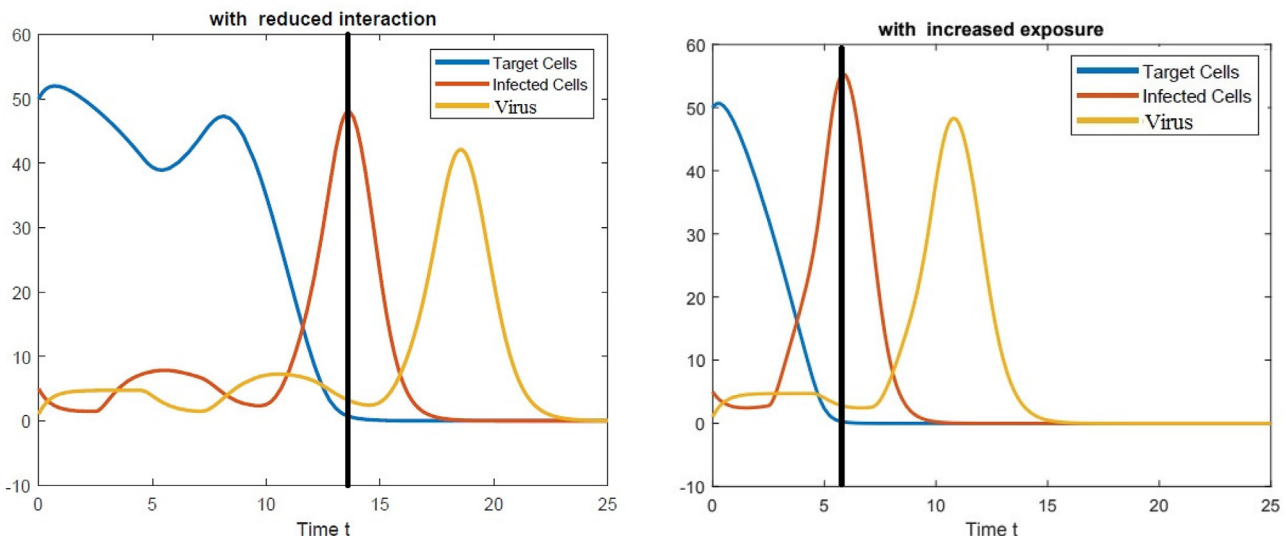**Fig. 4.** Visualization of repeated virus attacks.

Fig. 5. For different interaction rates, different dynamics in disease onset are depicted. We can see the infected cell count increased after a short period of time.
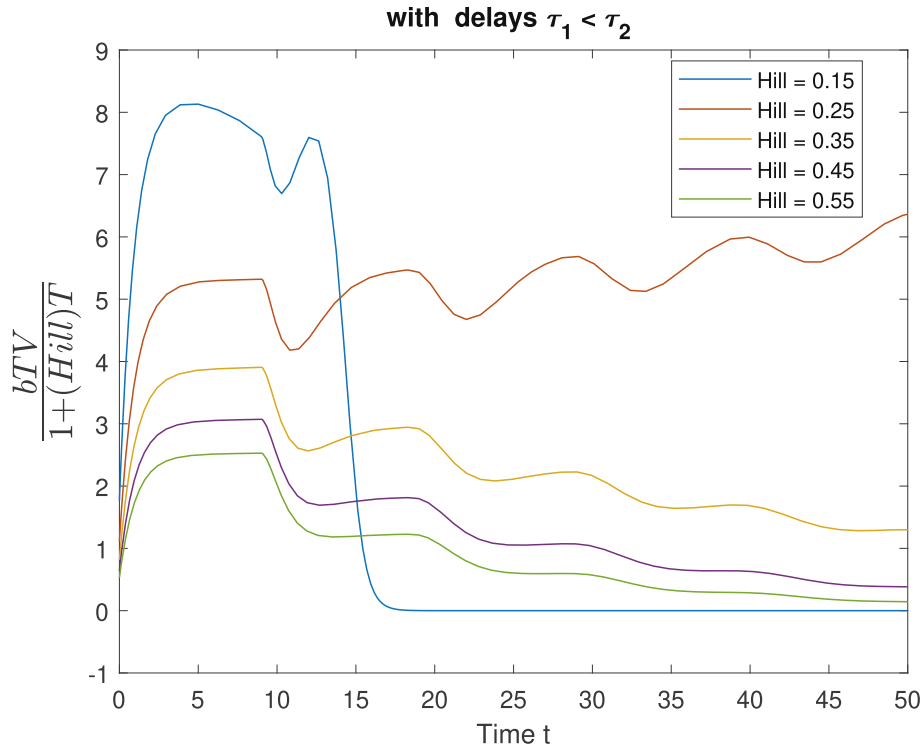
Fig. 6. For different values of the coefficient "Hill", helpful dynamics in disease onset are depicted.

$\frac{V(t-\tau_1)T(t-\tau_1)}{1+(Hill)T(t-\tau_1)}$ is responsible for the onset of infection and thus giving rise to the density of infected cells as shown in equation (2). The model is then analyzed for further analysis and is discussed in the following section.

2.3. Model development

The following model is designed to explore the virus-target cell interactions. A schematic is presented in Fig. 3.

$$\frac{d}{dt}T(t) = aT(t) - b(1 - Treatment)\frac{T(t)V(t)}{[1 + (Hill)T(t)]} \qquad (1)$$

$$\frac{d}{dt}I(t) = b(1 - Treatment)\frac{T(t-\tau_1)V(t-\tau_1)}{[1 + (Hill)T(t-\tau_1)]} - cI(t) \qquad (2)$$

$$\frac{d}{dt}V = c_1I(t-\tau_2) - dV(t) \qquad (3)$$

3. Analysis of SARS-CoV2 & target cell interaction

A computational framework to interpret the dynamics of viral infection is presented in this manuscript. The motivation behind the model development was to address the challenges linked with the fact that SARS-CoV2 lack self-sufficiency, i.e. it is a dormant particle unless it enters a host cell, where it hijacks the metabolism of host cell and produces copies of itself. Therefore for scientific research, such as for vaccine development, it is difficult to culture it and a computational framework can provide a better opportunity to explore the dynamics of the infection.

Two important questions about the disease onset are:

1. Infectious Dose: amount of SARS-CoV2 virus, the host is exposed to.
2. Incubation Period: time taken by the virus to hijack the target cells.

To address the questions one by one, we have run some numerical experiments based on the set of parametric values (source files provided as supplementary material). First, we focused on the infectious dose (Shen and Woo, 2020), i.e. the amount of virus, the host was exposed to. From the mathematical model presented in section 2, we can see that the density V of SARS-CoV2 depends on the term:

$c_1I(t-\tau_2)$

Since S are obligate intracellular pathogens (they can replicate depending on the metabolism of a host cell), c_1 is the parameter that measures the rate at which the infected cells are hijacked by the virus and provide a place to replicate the virus. The density V also depends on the delay caused by the infected cell and virus interaction, represented as τ_2. The concept of delay in cell interactions and its growth is not new and dates back in 2002, when "Interaction of the Coronavirus Nucleoprotein with Nucleolar Antigens and the Host Cell" was reported by (Chen et al., 2002).

In general, developing the mathematical models for biological problems is really challenging since it is really important to.

- determine the structure of the biological problem through system of differential equations,
- determine/forecast parametric values that will govern the dynamics of the mathematical model.

The two points are the central issues of computational systems biology. This type of approach ia also considered as a "reverse engineering process" (Zhan and Yeung, 2011). The parameter
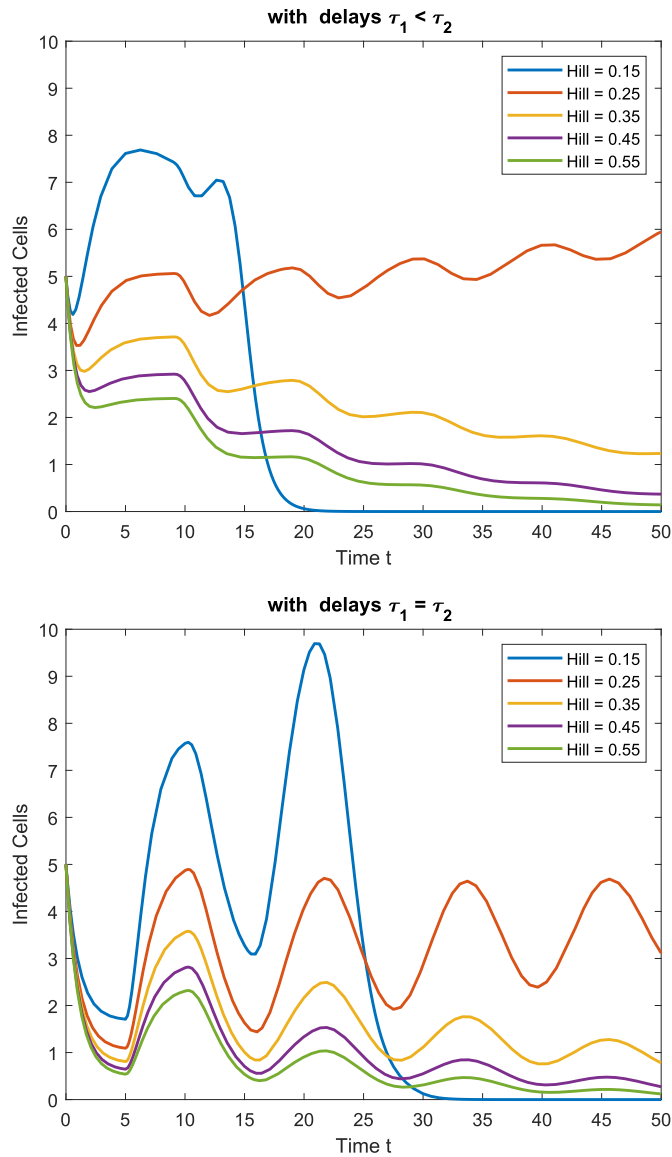
Fig. 7. For different values of the coefficient "Hill", helpful dynamics in disease onset are depicted.

estimation problem is generally formulated as an optimization problem that minimizes an objective function which represents the fitness of the model with respect to a set of experimental data. For the given problem, since the experimental data is not available yet, we have used MCMC to forecast the parametric values based on the research hypothesis and evidence available in the recent literature.

In this manuscript we have obtained the most appropriate value for the parameter c_1 with the aid of MCMC (Markov chain Monte Carlo) reverse engineering tool of Matlab^{TM}. The statistical metrics are provided in Table 1. Thus the mathematical model presented here can depict the infectious dose.

Next, we discussed the incubation period, this parameter is reported differently in different studies on Corona virus (Assiri et al., 2013; Backer et al., 2019).

The density of the virus, also depends on the body's immune response, we can consider it as a treatment against the viral attack. The higher the immune response, the fewer are the chances of the virus to survive. Unfortunately, in case of CoViD-19 outbreak, it has been reported that the virus replicates very quickly, prior the host gets a chance to prevent its rapid action with an immune response and thus it can not control the virus and starts going berserk (as documented in a report by (Coronavirus, 2020) Anthony Fehr, a virologist at the University of Kansas).

It has been reported in the literature that patients are being re-admitted after having already been infected, we have provided a graphical visualization of such situation, where virus density is presented in Fig. 4.

This *"cytokine storm,"* causes the immune system to stat sending cells ready to do battle into the lung. It is reported in the literature ((Reddy and Sansom, 2016)) that the immune response and the virus then work together to damage the body of the infected host. This effect is presented in Table 1, where we can see the negative sign with the treatment term, when the model was analyzed using reverse engineering based on the reported studies.

To explore the resulting density of virus after interaction with the human cells, we have varied the parametric values (for increased and reduced interaction scenarios) and have showed in Fig. 5 that the time taken by the virus to hijack target cells and to convert these into infected cells was really short when the exposure to virus was increased (right panel), whereas it took almost double the time in disease onset when the interaction was limited.

Next, the Hill function coefficient "Hill" was varied for below and above 1/2 to check the dynamics of the disease. We can see
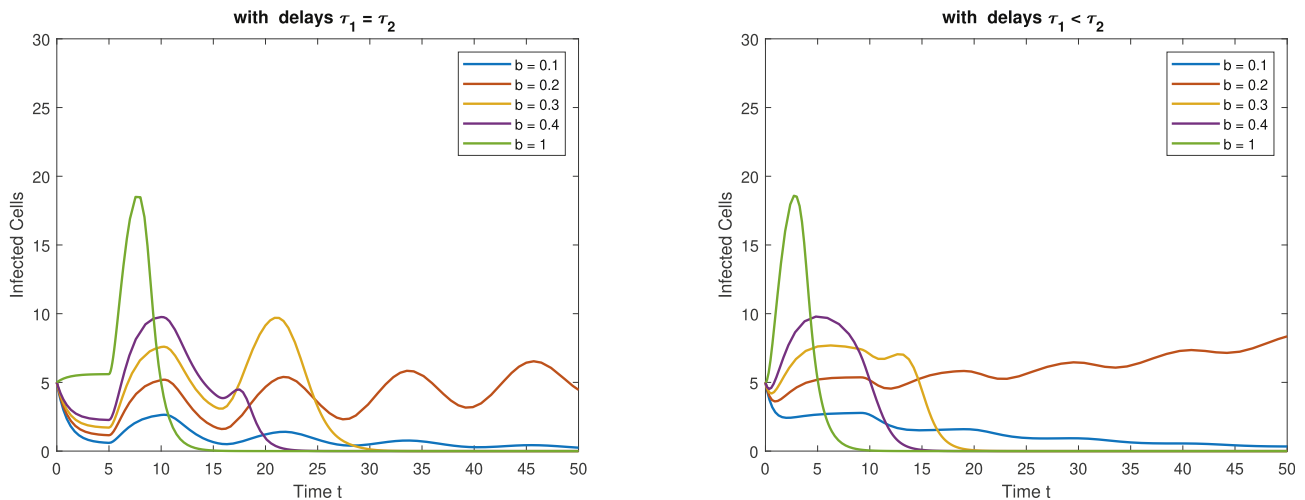

Fig. 8. For different values of the coefficient "b", helpful dynamics in disease onset are depicted.

from Fig. 6 out of phase dynamics with different periods of inter-action. The left panel shows the dynamics in interaction and the right panel provides the phase space portraits for virus and infected cells densities. Fig. 7 depicts the density of infected cells relative to time for equal and different values of delays for target cells, virus (τ_1) and infected cells (τ_2). Fig. 8 provides the clear understanding of the spread of infection relative to different values of parameter b (1).

With the aid of mathematical model we have concluded the following:

1. Delay from one compartment to the other plays an important role in the timeframe of the disease onset.
2. The Hill function is important in kinetic modelling and the Hill coefficient is important parameter to forecast a complete cycle of infection.
3. The reverse engineering approach proves to be efficient in parametric approximation for such complex dynamics.

4. Conclusions and future work

The manuscript presents a computational framework to explore the dynamics of novel SARS-CoV2 onset. The model helps to analyze the time period required by the virus to hijack the cells and spread through the body by infecting the target cells. Emphasis is made in this manuscript on the accurate understanding of the delay in "target cell-SARS-CoV2" interaction, which can help to forecast the expected duration of incubation period. It is concluded that the density of virus varies from the exposure time (of virus to the target cells) to interaction time. This variation can be monitored through the parameters responsible for the interaction rates. Currently, the pandemic is ongoing. The model described here can be implemented to future laboratory generated cell data.

Declaration of competing interest

The authors declare that there is no conflict of interest.

CRediT authorship contribution statement

Ayesha Sohail: Conceptualization, Methodology, Software, Data curation, Writing - original draft, Writing - review & editing. **Alessandro Nutini:** Conceptualization, Methodology, Software, Visualization, Investigation, Software, Validation.

Acknowledgments

References

Assiri, A., McGeer, A., Perl, T.M., Price, C.S., Al Rabeeah, A.A., Cummings, D.A., Alabdullatif, Z.N., Assad, M., Almulhim, A., Makhdoom, H., et al., 2013. Hospital outbreak of middle east respiratory syndrome coronavirus. N. Engl. J. Med. 369 (5), 407–416.

J. A. Backer, D. Klinkenberg, and J. Wallinga. Incubation period of 2019 novel coronavirus (2019-ncov) infections among travellers from wuhan, China, 20–28 january 2020. Euro Surveill., 25(5), 2020.

Chen, H., Wurm, T., Britton, P., Brooks, G., Hiscox, J.A., 2002. Interaction of the coronavirus nucleoprotein with nucleolar antigens and the host cell. J. Virol. 76 (10), 5233–5250.

Coronavirus, C. Y. Johnson, 2020. What You Need to Read. Technical report, the Washington Post.

Coutard, B., Valle, C., de Lamballerie, X., Canard, B., Seidah, N., Decroly, E., 2020. The spike glycoprotein of the new coronavirus 2019-ncov contains a furin-like cleavage site absent in cov of the same clade. Antivir. Res. 176, 104742.

Glowacka, I., Bertram, S., Müller, M.A., Allen, P., Soilleux, E., Pfefferle, S., Steffen, I., Tsegaye, T.S., He, Y., Gnirss, K., et al., 2011. Evidence that tmprss2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. J. Virol. 85 (9), 4122–4134.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.-H., Nitsche, A., et al., 2020. Sars-cov-2 cell entry depends on ace2 and tmprss2 and is blocked by a clinically proven pro-tease inhibitor. Cell.

Iftikhar, M., Iftikhar, S., Sohail, A., Javed, S., 2020. Ai-modelling of molecular iden-tification and feminization of wolbachia infected aedes aegypti. Prog. Biophys. Mol. Biol. 150, 104–111.

Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J., Hsueh, P.-R., 2020. Severe acute respi-ratory syndrome coronavirus 2 (sars-cov-2) and corona virus disease-2019 (covid-19): the epidemic and the challenges. Int. J. Antimicrob. Agents, 105924.

Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., et al., 2003. Angiotensin-converting enzyme 2 is a functional receptor for the sars coronavirus. Nature 426 (6965), 450–454.

Li, W., Zhang, C., Sui, J., Kuhn, J.H., Moore, M.J., Luo, S., Wong, S.-K., Huang, I.-C., Xu, K., Vasilieva, N., et al., 2005. Receptor and viral determinants of sars-coronavirus adaptation to human ace2. EMBO J. 24 (8), 1634–1643.

Reddy, T., Sansom, M.S., 2016. Computational virology: from the inside out. Biochim. Biophys. Acta Biomembr. 1858 (7), 1610–1618.

Shen, S., Woo, R., 2020. Coronavirus Incubation Could Be as Long as 27 Days, Chi-nese Provincial Government Says. Technical Report. The Thomson Reuters Trust Principles.

Sherin, L., Farwa, S., Sohail, A., Li, Z., Bég, O.A., 2018. Cancer drug therapy and sto-chastic modeling of "nano-motors". Int. J. Nanomed. 13, 6429.

Tappert, M.M., Porterfield, J.Z., Mehta-D'Souza, P., Gulati, S., Air, G.M., 2013. Quan-titative comparison of human parainfluenza virus hemagglutinin-neuraminidase receptor binding and receptor cleavage. J. Virol. 87 (16), 8962–8970.

Walls, A.C., Xiong, X., Park, Y.-J., Tortorici, M.A., Snijder, J., Quispe, J., Cameroni, E., Gopal, R., Dai, M., Lanzavecchia, A., et al., 2019. Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. Cell 176 (5), 1026–1039.

Wang, C., Li, W., Drabek, D., Okba, N.M., van Haperen, R., Osterhaus, A.D., van Kuppeveld, F.J., Haagmans, B.L., Grosveld, F., Bosch, B.-J., 2020. A Human Monoclonal 1 Antibody Blocking Sars-Cov-2 Infection. bioRxiv.

Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.-L., Abiona, O., Graham, B.S., McLellan, J.S., 2020. Cryo-em structure of the 2019-ncov spike in the prefusion conformation. Science 367 (6483), 1260–1263.

Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., et al., 2020. A new coronavirus associated with human respiratory disease in China. Nature 579 (7798), 265–269.

Zhan, C., Yeung, L.F., 2011. Parameter estimation in systems biology models using spline approximation. BMC Syst. Biol. 5 (1), 14.