Modeling COVID-19 as a National Dynamics with a SARS-CoV-2 Prevalent Variant: Brazil - A Study Case

S. Celaschi
CTI Renato Archer, Rod. D Pedro I (SP-65), Km 143,6 - CEP 13069-901, Campinas – SP, 13069, Brazil
Corresponding Author: sergio.celaschi@cti.gov.br

September, 2020

This preliminary report has not been certified by peer review. It should not be relied on to guide health-related behavior, and should not be reported in news media as established information.

----------ABSTRACT----------

COVID-19 global dynamics is modeled by an adaptation of the deterministic SEIR Model, which takes into account two dominant lineages of the SARS-CoV-2, and a time-varying reproduction number to estimate the disease transmission behavior. Such a methodology can be applied worldwide to predict forecasts of the outbreak in any infected country. The pandemic in Brazil was selected as a first study case. Brazilian official published data from February 25th to August 30th, 2020 was used to adjust a few epidemiologic parameters. The estimated time-dependence mean value to the infected individuals (confirmed cases) presents - in logarithmic scale - standard deviation SD = 0.08 for over six orders of magnitude. Data points for additional three weeks were added after the model was complete, granting confidence on the outcomes. By the end of 2020, the predicted numbers of confirmed cases in Brazil, within 95% credible intervals, may reach 6 Million (5 -7), and fatalities would accounts for 180 (130 – 220) thousands. The total number of infected individuals is estimated to reach 13 ± 1 Million, 6.2% of the Brazilian population. Regarding the original SARS-CoV-2 form and its variant, the only model assumption is their distinct incubation rates. The variant form reaches a maximum of 96% of exposed individuals as previously reported for South America.

KEYWORDS; *COVID-19; SARS-Cov-2; mutation; global dynamics; Brazil; study case.*

Introduction

The SARS-Cov-2 has revealed important facts about governments, and health organizations world wise. This powerful new virus demonstrates that national health organizations are incapable of prompt responding to such an emergency. Reasons why a central government cannot cope to this crisis are many[1]. Infectious disease outbreaks show how countries are all tied. To deal with Covid-19 countries must work together on a collectively and coherent response strategy to drive the infection to a low level. Developing and emerging countries are particularly vulnerable by the pandemic. In such countries it is difficult to sustain social distancing measures. Some reasons are: their weak defenses, health systems are often poor, governments have fewer resources to intervene, and a significant fraction of the population has no savings and depend on daily sources of income[1]. COVID-19 was first reported in Brazil in February 2020 – EPI week # 9. One semester latter, the country became one of the worst affected globally. Brazil comprises many states with vulnerable communities, and a relatively weak social protection system. After six months from the first reported case, the number of confirmed cases and deaths crossed 3.9 million and 120 thousand, respectively. Facts that rise doubts on the availability of public health care for the sections of society that cannot afford private care[2]. Since the start of the epidemic in Brazil, several types of Non-pharmaceutical Interventions - NPI have been adopted with varied success by the country's 27 federal units and 5,596 municipalities. Virus transmission has dropped substantially in most units. However, by September 2020, the estimated reproduction number remains above the unit[3]. Thus, only mitigation (and not suppression) of the epidemic has been achieved so far. Current interventions remain insufficient to extinguish virus transmission in Brazil. Closer surveillance of viral transmission quantifying the impact of different control measures on transmission may help to minimize future infections. Moreover, continued monitoring of the genetic diversity of the virus lineages circulating globally became necessary. As recently suggested, virus diversity are of prime importance in the SARS-CoV-2 transmission[4,5]. While the COVID-19 pandemic continues in the Americas, new highlights into the Brazilian are: the high SARS-CoV-2 transmission through the country; the role of its large urban centers; the lack of simultaneous lockdown regions; the non-equitable access to tests and reports. Those factors potentially contributed to sustain pandemic spread. By the end of February 2020, before the implementation of NPIs and travel bans, different SARS-Cov-2 lineages had emerged in Brazil from Europe[6]. In the beginning, the epidemic had spread mostly locally and within-state borders[6].

The present study aims to derive a consistent and quantified model on the COVID-19 outbreak as a national dynamics, presenting Brazil as a study case. The major results are obtained applying an adaptation of a well known epidemiological model on official published data. Epidemiological models are commonly stochastic, network-based, spatially diffusive, using meta-population approaches[7,8,9]. However, the parameters of dynamic deterministic models are directly related to interpretable physical processes[10]. As a drawback, deterministic models impose restrictive analysis, once the dynamics of the host population and the

virus are not deterministic. The population has free will, and the virus undergoes favorable mutations upon natural selection[5]. Host dynamics was account for into the model by spreading the use of time-varying reproduction number[11]. Regarding the pathogen, the variant SARS-CoV-2 D614G mutation[5], which became globally proposed as dominant, was incorporated in our model. In short, the methodology employed is the application of an adaptation to the deterministic SEIR Model, which takes into account two lineages of the SARS-CoV-2, and a time-varying reproduction number to account for changes in the dynamics of the COVID-19 transmission behavior. Such a methodology can be applied worldwide to predict forecasts of the outbreak in any infected country.

The SARS-CoV-2 Increasing Frequency on Global Distribution

Even though SARS-CoV-2 sequence diversity is reported to be low, natural selection acts upon favorable mutations[12]. The persistence of the pandemic might enable accumulation of immunologically relevant mutations. Recent identification of spike amino acid variant has been reported which increasing frequency globally[5] but not necessarily brings evidence of positive selection for 614G[13].The spike protein determines the infectivity of the virus, and its transmissibility in the host. Mutations in the gene encoding Spike (S) protein are being reported since SARS-CoV-2 was identified in humans. Those mutations of S proteins affect viral life cycle and its interaction with the host. Very recent studies suggest that the observed increased transmission reported in Europe and Americas may be associated with the most dominant variant[5]. Those authors present experimental evidence that the so called D614G spike variant is associated with greater infectivity as well as higher viral loads. Global tracking data show that the G614 variant in Spike has spread faster than D614 comparing two time periods apart by a 2-week gap. The fraction for South America prior to March 1st are D = 38%, G = 62%, and D = 4 %, G = 96% in the period March 2020 21-30. They also show running weekly average counts of sampled sequences exhibiting the D614 and G614 variants for South America between January 12 and May 12. The results are explained assuming the virus being more infectious. On the other hand, the authors did not find evidence of G614 effects on disease severity. As recently reported elsewhere, the Spike protein 614 polymorphism has little variance in growth rates among clusters and presented no significant difference in initial growth rates[13]. Those facts are incorporated in our model and evidences are demonstrated, assuming the same time-varying R_t number for both Spike protein 614 morphisms[13].

Prevalence of Asymptomatic Infection

The deadly face of COVID-19, which by the end of September 2020 had claimed more than 866 000 lives worldwide, is not the only face. SARS-CoV-2 has a dual nature: tragically lethal to many persons and surprisingly harmless to others. Infected persons who remain asymptomatic and pre-symptomatic (here on named non-symptomatic) play a significant role in the ongoing COVID-19 pandemic. The absence symptoms in persons infected with SARS-CoV-2 does not mean an absence of harm. The procedure to classify non-symptomatic hosts is to carry repeated observations of the individual over time in a group of individuals sharing a statistical factor. Therefore, the prevalence of non-symptomatic SARS-CoV-2 infection has to be included in modeling any COVID-19 pandemic outbreak. Applying statistical analysis on recent published and reliable data[14], the percentage value of non-symptomatic and symptomatic hosts were estimated to be respectively (54 +/- 9) % and (46 +/- 9) %.

Serial Intervals and Inference of time-varying reproduction number

Serial interval (SI) is an essential metric for estimating other epidemiological parameters, which in turn are used to predict disease trends, interventions and health care demands. SI depends on the pathogen incubation period which quantifies the biological process of relevant virus mutation and disease progression and tends to follow distributions resulting from genetic differences[15]. Variations in SI can occur and may have significant implications for the transmission dynamics affecting the estimation of epidemic parameters. The real-time transmissibility of an infectious disease is often characterized by the instantaneous reproduction number (R_t), which is defined as the expected number of secondary infections caused by an infector within a short time window. Equivalently, R_t can be expressed as the transmission rate $\beta(t)$ divided by the rate γ_o at which infected people recover or die. The aim of control interventions is typically to reduce the R_t value. A number of methods are available to estimate effective reproduction numbers during epidemics[16]. which provide a superior dynamics of the epidemic. With this aim, a method for estimating the instantaneous reproduction number using branching processes was developed[17]. It relies on two inputs: a disease incidence time series (the numbers of new observed cases at successive times) and an estimate of the

distribution of serial intervals. More recently, this statistical framework was extended to allow adding data on known pairs of index and secondary cases from which the serial interval is directly estimated[11]. During the COVID-19 outbreak in Brazil, the time-varying reproduction number was estimated following the procedure available on-line[18].

Estimative of Infected Individuals and Fatal Outcomes

There are two measures used to assess the proportion of infected individuals with fatal outcomes, the infection fatality ratio (IFR), which estimates this proportion of deaths among all infected individuals, and case fatality ratio (CFR), which estimates this proportion of deaths among identified confirmed cases. The level of transmission is generally underestimated for infectious diseases, which difficulty the IFR estimative[19]. A significant proportion of infected people are undetected because they are non-symptomatic, and fail to present at healthcare facilities. Furthermore, testing capacity have been limited in Brazil, and restricted to people with severe cases and priority risk groups. To accurately measure IFR, the complete number of infections, and deaths caused by the disease must be known. An important method to estimate the true number of infected individuals is the serological testing of a representative random sample of the population to detect evidence of exposure to a pathogen. Very recently, a careful evaluation of the IFR was reported to the COVID-19 outbreak in Brazil[20]. A country-wide average IFR of 1.05% (95% CI: 0.96 – 1.17%) was reported by the authors. On the other hand, reliable CFRs are generally obtained at the end of an outbreak, after most infected individuals either died or recovered. CFR calculated using the above relation during ongoing epidemics is influenced by lags in report dates for cases and deaths. Obtaining both IFR and CFR from available data would allow estimating the fraction of symptomatic hosts to apply in the model.

Modeling COVID-19 Brazilian Pandemics

Epidemiological models are useful in estimating main parameters of a pathogen spread like SARS-CoV-2. Mathematically, such models may evaluate the efficacy of specific interventions, identify efficient strategies, and forecast future scenarios[21]. Models provide estimative on the basic and effective reproduction numbers, i.e., the number of new infections caused by each infectious individual, the case fatality ratio (CFR), which estimates this proportion of deaths among identified confirmed cases. They also may quantify the infection fatality ratio (IFR), clarifying ambiguities create by asymptomatic infected hosts and delays between incubation, infection and heal/death. Important to note that official reported confirmed COVID-19 cases may miss asymptomatic infections, which can bias estimates of disease severity and IFR, so epidemiological models could help reduce this uncertainty. The Susceptible-Exposed-Infected-Removed - SEIR model was adapted in this study to take into account the original SARS-Cov-2 D-form and its dominant G-variant with their own incubation rates, a pre-estimated fraction of symptomatic hosts, and a pre-inferred time-varying reproduction number. COVID-19 has a latent or incubation period, during which the individual is said to be infected but not infectious. Members of this population in latent stage are labeled as Exposed (but not infectious). Taken into consideration the original SARS-Cov-2 D-form and its dominant G-variant labeled as D and G, the deterministic model with the groups: Susceptible, Exposed (D and G), Infected (D and G), and Removed (recovered and deaths/fatalities) is labeled as the $SE_DE_GI_DI_GR$ Model. The number of fatalities is assumed dependent on the confirmed COVID-19 cases (Infected D and G). This dependence is not linear neither monotonic, and is obtained from official reported cumulative values. COVID-19 cases in Brazil are reported by public health and private services, and interrelated on a website which summarizes daily the aggregated counts.[22] Worth to mention that by the time this study was conducted, reliable estimative on the number of sub notifications and information on the number of non-symptomatic were unreliable. Under those assumptions, the set of ordinary differential equations – ODE governing our $SE_DE_GI_DI_GR$ model follows;

$$\frac{dS(t)}{dt} = -\beta(t).[I_D(t) + I_G(t)].\frac{S(t)}{N} \quad (1)$$

$$\frac{dE_D(t)}{dt} = \beta(t).I_D(t).\frac{S(t)}{N} - \sigma_D.E_D(t) \quad (2)$$

$$\frac{dE_G(t)}{dt} = \beta(t).I_G(t).\frac{S(t)}{N} - \sigma_G.E_G(t) \quad (3)$$

$$\frac{dI_D(t)}{dt} = -\gamma_o.I_D(t) + \sigma_D.E_D(t) \quad (4)$$

$$\frac{dI_G(t)}{dt} = -\gamma_o.I_G(t) + \sigma_G.E_G(t) \quad (5)$$

$$\frac{dRe(t)}{dt} = \gamma_o.I_D(t) + \gamma_o.I_G(t) \quad (6)$$

where *S(t), $E_D(t)$, $E_G(t)$, $I_D(t)$, $I_G(t)$, and Re(t),* are respectively daily numbers of Susceptible, Exposed (D and G), Infected (D and G), Removed (recovered and deaths) individuals. $S(t) + E_D(t) + E_G(t) + I_D(t) + I_G(t) + Re(t) = N =$ Constant. $\beta(t)= R(t)/\gamma_o$, *where R(t)* is the time-varying reproduction number. $R(t=0)=R_o$. γ_o is the removed rate that infected individuals (symptomatic and non-symptomatic) recover, leaving the infected groups $I_D(t)$ *and* $I_G(t)$. The accumulated SARS-Cov-2 confirmed cases are obtained from:

$$C(t) = \alpha_s.\int_0^t [E_D(\tau) + E_G(\tau)].d\tau \quad (7)$$

where $\alpha_s = 0.47 \pm 0.09$ is the estimated fraction of symptomatic individuals. Major assumptions for this model are: $\sigma_D = 0.5$ and $\sigma_G = 0.25$ are respectively the incubation rates for D and G SARS-CoV-2 exposed hosts. It is assumed that both D and G groups share the same *R(t)* time dependence. In short, *N*, σ_D and σ_G are the only fitting parameters to data. The susceptible number $N=60 \times 10^6$ (28.7% of the Brazilian population) is chosen as the minimum number of susceptible to account for the accumulated confirmed cases in the study period. Fatalities were modeled as a function of the confirmed cases *C(t),* and this dependence, obtained from official reported cases and deaths, is not linear neither monotonic as discussed latter.

Results and Discussion

Official data, from February 25th to August 19th, 2020 EPI week #9, provided by the Ministry of Health of Brazil[22], was considered to estimate part of the parameters that govern the dynamics established by Eqns. (1) to (7). All model parameters were estimated by minimizing the mean squared quadratic errors. A key parameter in deterministic transmission models is the reproductive number R, which is quantified by both, the pathogen and the susceptible population in which it circulates. Thus, a single pathogen, like the SARS-CoV-2, will have different R values depending on lineage and transmission dynamics of the population experiencing the outbreak. When infection is spreading through a population that may be partially non-symptomatic, an accurate estimation of the instantaneous R value is crucial to plan and control the infection.[11] During the COVID-19 outbreak in Brazil, the methodology to estimate the basic reproduction number was presented in a previous publication.[16] The time-varying reproduction number was estimated from a Gamma Distribution Time Generation (Mean Value = 3.6, Variance = 4.8) on confirmed cases from February 25th to August 19th. Table 1 presents some estimated values obtained for a 7 days running window.

Table 1. Selected time-varying reproduction values estimated from a Gamma Distribution Time Generation (Mean Value = 3.6, Variance = 4.8) applied on a 7 days running window of confirmed cases.

t_start	t_avr	t_end	Mean(R)	Std(R)	Quantile 0.025(R)	Quantile 0.05(R)	Quantile 0.25(R)	Median (R)	Quantile 0.75(R)	Quantile 0.95(R)	Quantile 0.975(R)
2	5	8	2.24	0.19	1.89	1.94	2.11	2.24	2.37	2.56	2.63
9	12	15	2.28	0.14	1.9	1.94	2.07	2.17	2.26	2.41	2.46
16	19	22	1.97	0.06	1.87	1.89	1.95	2	2.04	2.1	2.12
23	26	29	1.55	0.02	1.87	1.88	1.9	1.92	1.93	1.95	1.96
30	33	36	1.45	0.01	1.45	1.45	1.46	1.47	1.47	1.48	1.49
37	40	43	1.34	0.01	1.42	1.42	1.43	1.43	1.43	1.44	1.44
44	47	50	1.27	0	1.32	1.32	1.32	1.33	1.33	1.33	1.33
51	54	57	1.23	0	1.26	1.26	1.27	1.27	1.27	1.27	1.27
58	61	64	1.24	0	1.23	1.23	1.23	1.24	1.24	1.24	1.24
65	68	71	1.23	0	1.23	1.23	1.23	1.23	1.24	1.24	1.24
72	75	78	1.2	0	1.22	1.22	1.22	1.22	1.22	1.22	1.22
79	82	85	1.2	0	1.2	1.2	1.2	1.2	1.2	1.2	1.2
86	89	92	1.17	0	1.19	1.19	1.19	1.19	1.19	1.19	1.19
93	96	99	1.16	0	1.17	1.17	1.17	1.17	1.17	1.17	1.17
100	103	106	1.13	0	1.15	1.15	1.15	1.15	1.15	1.15	1.15
107	110	113	1.11	0	1.12	1.12	1.12	1.12	1.12	1.12	1.12
114	117	120	1.11	0	1.11	1.11	1.11	1.11	1.11	1.11	1.11
121	124	127	1.09	0	1.11	1.11	1.11	1.11	1.11	1.11	1.11
128	131	134	1.08	0	1.09	1.09	1.09	1.09	1.09	1.09	1.09
135	138	141	1.07	0	1.08	1.08	1.08	1.08	1.08	1.08	1.08
142	145	148	1.07	0	1.07	1.07	1.07	1.07	1.07	1.07	1.07
149	152	155	1.06	0	1.07	1.07	1.07	1.07	1.07	1.07	1.07
156	159	162	1.06	0	1.06	1.06	1.06	1.06	1.06	1.06	1.06
163	166	169	1.05	0	1.06	1.06	1.06	1.06	1.06	1.06	1.06
170	173	176	1.04	0	1.05	1.05	1.05	1.05	1.05	1.05	1.05
177	180	183	1.04	0	1.04	1.04	1.04	1.04	1.04	1.04	1.04

The complete estimated sequence of *R(t)* values as well as its Credible Intervals (5 – 95%) in Fig. 1 where time is shown in units of EPI weeks. An epidemiological week, commonly referred to as an EPI week is simply a standardized method of counting weeks to allow for the comparison of data year after year. Time-varying R(t) describes COVID-19 time shifts incidence. Rapid epidemic growth phases are associated with R(t) values higher than 2. In contrast, epidemic decline phases may be characterized by 2 > R > 1. The beginning of the pandemic spread in Brazil presented, for a one-week time window, R-values larger than 2. However, after two weeks, R(t) shows a monotonic decline, reaching values 4% above the unit, as a consequence of collective acquired immunity and NPI regional interventions across de country (Fig.1). NPIs in Brazil was implemented between EPI week #10 and #12 across the countries' 27 federal states, and consisted of social distancing, mask wearing, school, and stores closures. At the start of the epidemics, value of $R_o > 2$ was reported in São Paulo concurrent with the timing of state mandatory NPIs[6, 16].

$SE_DE_GI_DI_GR$ model fittings to data on accumulated confirmed cases (symptomatic infected individuals), and fatalities are show in Fig. 1. The only fitting values are: $N = 6x10^7$ (28.7% of Brazilian population), $\beta(t=0) = 1.265$, $\sigma_D = 0.25$, $\sigma_G = 2\sigma_D$. Initial conditions: $S(0) + E(0) = N$, $E(0) = E_D(0) + E_G(0) = 8$, $I_D(0) = I_G(0) = R_e(0) = 0$. The fitting to the infected individuals (logarithmic scale) presents standard deviation SD = 0.08 over six orders of magnitude. The SARS-Cov-2 spike variant G associated with greater infectivity and higher viral loads is modeled by a higher (two fold) incubation rate σ_G as compared to σ_D. The fraction of symptomatic hosts $\alpha_s = 0.47 \pm 0.09$ within 95% credible interval was estimated comparing IFR to CFR reported values[20]. The estimated mean value and deviation are consistent with previous analysis.[14] Fatalities were modeled as a function of the confirmed cases *C(t)*. By the end of 2020, the predicted numbers of confirmed cases in Brazil, within 95% credible intervals, may reach 6 Million (5 - 7) and fatalities would accounts for 184 thousand (134 – 218). The total number of infected individuals is estimated to reach (13 ± 1) Million, 6.2% of the population.

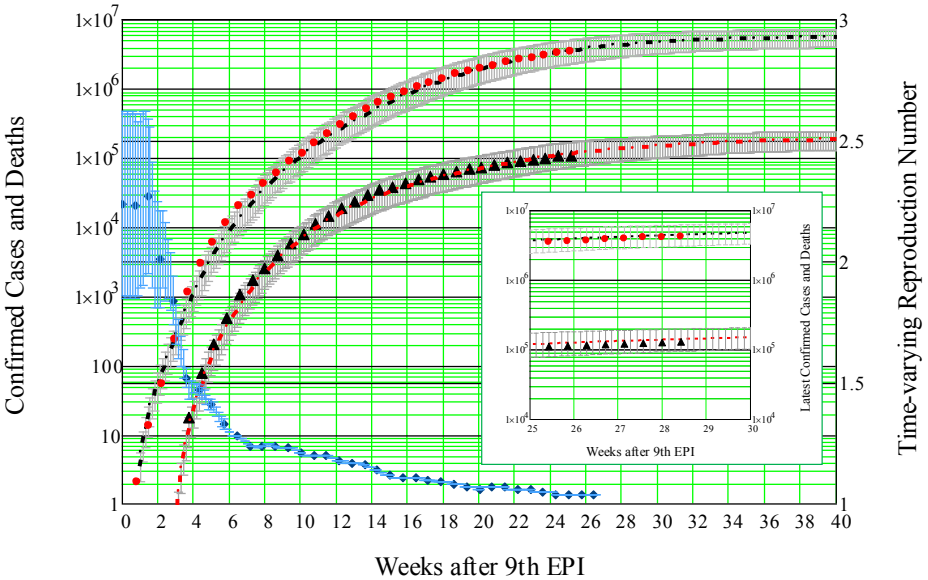

Figure 1. Data on confirmed cases and fatalities are shown by solid red circles, and solid black triangles respectively. Error bars account for the RMS values on the SEIR parameters. The fitting values to SEIR model are: $N = 6.10^7$, $\beta o = 1.265$, $\sigma_D = 0.25$, $\sigma_G = 2\sigma_D$ $\alpha_s = 0.47 \pm 0.09$. Average values and error bars (vertical in black) for infected individuals, and fatalities are shown respectively in solid red circles and solid black triangles. The mean fitting to the infected individuals (logarithmic scale) presents standard deviation SD = 0.08. Right vertical axis shows the time-varying R(t) (solid blue diamond) as the result of control interventions, and collective immunity[11]. Vertical bars represent 5 to 95% CRI. Data points for the last three weeks (Figure Inset - solid circles and triangles) were added after the model was complete, granting confidence on the outcomes.

Fatalities as a function of the confirmed cases *C(t),* present a time-dependence that is not linear neither monotonic. As shown in Fig. 2, the non-zero ratio of deaths/(confirmed cases) by COVID-19 (open red circles), begins three weeks after the first reported case, rises monotonically, after roughly seven weeks reaches a maximum of 7%, and drops to a ratio around 2.6%. Possible explanations are: NPIs as social distancing; the increased number of tests and reports available after the first weeks of the epidemic outbreak; and ICU improved medical procedures. COVID-19 was detected in Brazil in the 9th epidemiological week of 2020, and the test for the SARS-CoV-2 virus was effectively included in the surveillance four weeks later.

Furthermore, ventilators were scarce at the first weeks. The fitting equation (Fig. 2 - dashed blue curve) to the experimental data multiplied by Eq.7 allows estimating the accumulated fatalities as presented in Fig. 1.

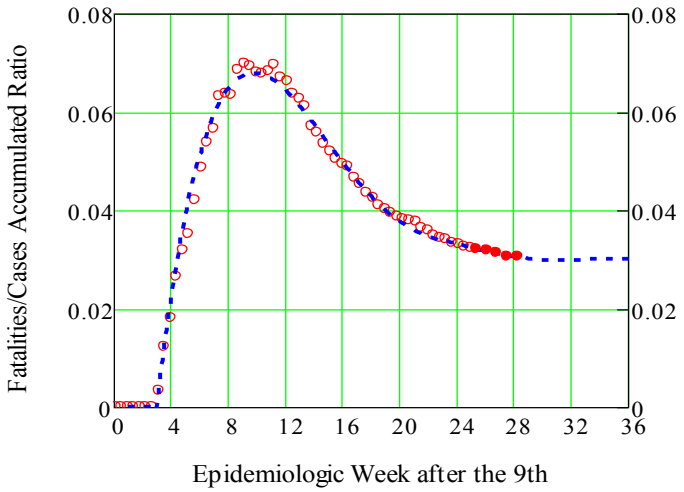

Figure 2. Time dependence of the ratio Fatalities/(Confirmed Cases). This ratio presents a time-dependence that is not linear neither monotonic. The non-zero ratio of deaths/(confirmed cases) by COVID-19 (open red circles), begins three weeks after the first reported case, rises monotonically, after roughly seven weeks reaches a maximum of 7%, and drops to a ratio around 2.6 %. The fitting equation (dashed blue curve) to the experimental data multiplied by Eq.7 allows estimating the accumulated fatalities as shown in Fig. 1. Data points for the last three weeks (solid red circles) were added after the model was complete, granting confidence on the outcome.

Regarding the original SARS-CoV-2 D-form and its G-variant, the only distinction in modeling is their own incubation rates. Additionally, it is assumed both forms produce the same disease severity/fatalities, and share the same instantaneous reproductive number, which is an oversimplified assumption. The initial percentage values set in the ODE for individuals exposed to the original D-form and its G-variant were respectively 66% and 34%.[4] Figure 3 presents the evolution of exposed and infected host to the original D-form and its G-variant according to model. These results confirm the prevalence of the G-variant SARS-CoV-2 form in COVID-19 pandemic as globally predicted. The G-variant form reaches a maximum of 96% of exposed individuals as previously reported for South America.[4] The resulted D- to G-shift in the epidemiologic spread suggests that the G-variant form may have a fitness advantage. These findings support continuing surveillance of SARS-CoV-2 mutations to support development of immunological interventions.

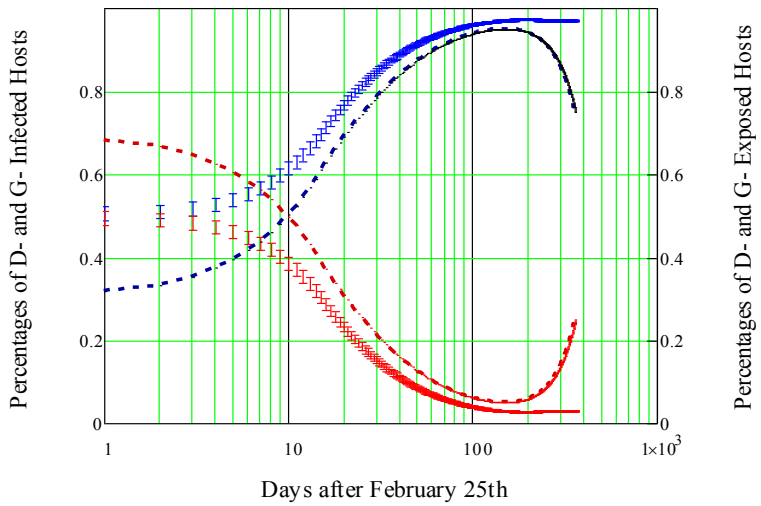

Figure 3. Estimated evolution of exposed and infected host to SARS-CoV-2 original D-form and G-variant. Broken red and blue lines present respectively the evolution of D and G exposed hosts to SARS-CoV-2. Vertical red and blue bars are respectively the evolution $\alpha_s = 0.47 \pm 0.09$ and $E(0) = 6 \pm 2$) of D and G infected hosts. According to model, the percentage of exposed hosts to prevalent G-form of SARS-CoV-2 quickly rises and reaches a maximum of 96% as previously reported for South America.

Brazil is the largest country in South America. The time interval comprising the first confirmed case in S. Paulo, and the first confirmed death in the Tocantins (in the Amazon forest) , the last Brazilian state to report a COVID-19 fatality encompass 8 EPI weeks.[6] Applying a single time frame for a close surveillance of viral transmission across the country limits an overall analysis. An alternative to overall visualize the pandemic stage is to follow the dependence of new weekly cases and fatalities on their cumulative values, as shown in Fug. 4. Solid black circles are official reported values, and red solid circles are average predicted values. This alternative view clearly estimates the predicted final numbers of confirmed cases in the vicinity of 6 Million and fatalities approaching under 200 thousand.

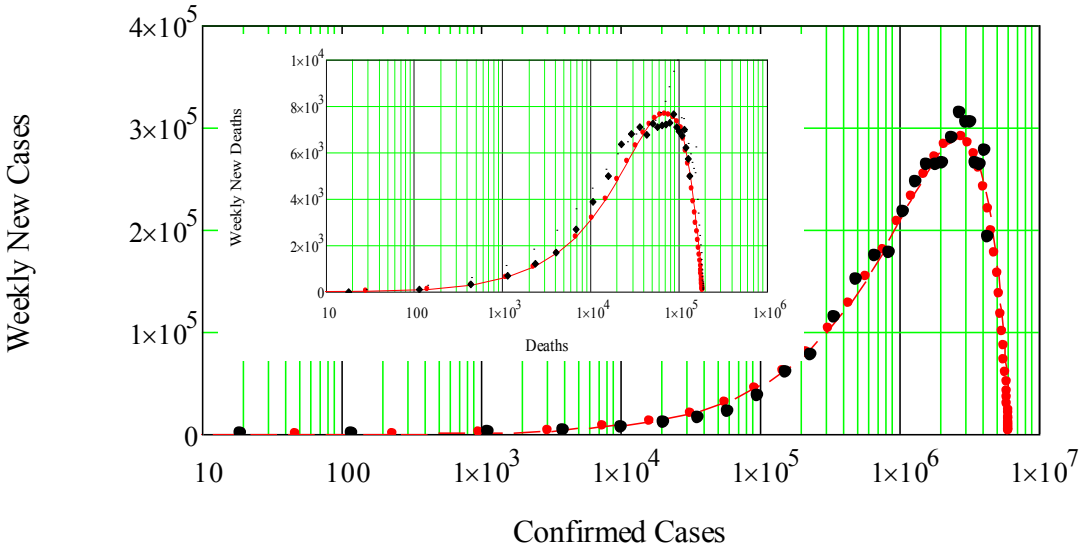

Figure 4. Weekly new cases and deaths (figure inset) versus confirmed cases and deaths respectively. Solid black circles are official reported values, and red solid circles are average predicted values. This alternative view clearly estimates the end numbers of confirmed cases in the vicinity of 6 Million and fatalities approaching under 200 thousand.

Conclusions

COVID-19 global dynamics was modeled by an adaptation of the deterministic SEIR epidemiological model, taking into consideration two lineages of the SARS-CoV-2, and a time-varying reproduction number to estimate the disease transmission behavior. Such a methodology can be applied worldwide to predict forecasts of the outbreak in any infected country. The pandemic in Brazil was selected as a first study case. Brazilian official published data was used to adjust a few model parameters. The estimated time-dependence ranges (5 – 95% CRI) to the infected individuals (confirmed cases), and fatalities agrees with reported values for over six and five orders of magnitude respectively. Data points for additional three weeks were added after the model was complete, granting confidence on the outcomes. By the end of 2020, the predicted numbers of confirmed cases in Brazil, within 95% credible intervals, may reach 6 Million (5 - 7) and fatalities would accounts for 184 thousand (134 – 218). The total number of infected individuals is estimated to reach (13 ± 1) Million, 6.2% of the population. Regarding the original SARS-CoV-2 form and its variant, the only model assumption is their distinct incubation rates. The variant form reaches a maximum of 96% of exposed individuals as previously reported for South America. An important conclusion worth to point out follows. Modeling any infectious disease reveals that the spread of infection depletes susceptible individuals. Despite the social and mental commotion that COVID-19 has imposed so far, most of the Brazilian and the world populations are still susceptible to SARS-CoV-2 infection. By the end of 2020, the depletion of susceptible Brazilian individuals was estimated in 10%. Sufficient depletion of susceptibility (by NPIs or not) has to be achieved to reach herd immunity threshold, and weak the global dynamics spread.

Acknowledgments

The author acknowledges the financial support from Project Plat IoT, granted by FUNTTEL *("Fundo para o Desenvolvimento Tecnológico das Telecomunicações"*). Financial grant # 01.16.0053.01 FINEP/MCTI from Brazilian Ministry of Science, Communications, and Innovation. This paper reflects only the author's views and the Agencies are not responsible for any use that may be made of the information contained therein.

References

1. Arab-Zozani, M., Amery, H., Lessons learned on how to manage a crisis such as: how the COVID-19 revealed our weaknesses?, International Journal of Epidemiologic Research, (), pp. Article in press (2020). Science: https://wellcome.org/news/more-health-crisis-covid-19-impacts-are-far-reaching-and-long-term.
2. de Souza, W.M., Buss, L.F., Candido, D.d.S. et al. "Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil". Nature Hum. Behav. 4, 856–865 (2020). https://doi.org/10.1038/s41562-020-0928-4
3. Felix, H. C. Francisco, Fontenele J., Instantaneous R calculation for COVID-19 epidemic in Brazil. medRxiv preprint doi: https://doi.org/10.1101/2020.04.23.20077172
4. Qianqian Li, et al., "The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity", Cell, July 2020, https://doi.org/10.1016/j.cell.2020.07.012
5. Korber et al., Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus, Cell (2020), https://doi.org/10.1016/j.cell.2020.06.043
6. Darlan S. C., "Evolution and epidemic spread of SARS-CoV-2 in Brazil", medRxiv preprint doi: https://doi.org/10.1101/2020.06.11.20128249 .
7. Zlojutro, A., Rey, D. & Gardner, L., "A decision-support framework to optimize border control for global outbreak mitigation", Nature Sci Rep 9, 2216 (2019). doi.org/10.1038/s41598-019-38665-w.
8. Wang, L., Wu, J.T., "Characterizing the dynamics underlying global spread of epidemics". Nature Commun 9, 218 (2018). doi: 10.1038/s41467-017-02344-z
9. Ruiyun Li, Sen Pei et. Al., "Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2)", Science 01: Vol. 368, Issue 6490, pp. 489-493, May 2020. doi: 10.1126/science.abb3221
10. Kissler, S.M., Tedijanto, C., Goldstein, E., ET. Al., "Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period", Science 14 Apr 2020. doi: 10.1126/science.abb5793
11. Thompson RN, Stockwin ., van Gaalen R.D, Polonsky JA, et al. "Improved inference of time-varying reproduction numbers during infectious disease outbreaks", Epidemics (2019) https://doi.org/10.1016/j.epidem.2019.100356
12. Fauver, J.R., et al. (2020), "Coast-to-Coast Spread of SARS-CoV-2 during the Early Epidemic in the United States". Cell 181, 990–996.e5 (2020)
13. Volz, E., Hill, Verity, et.al., Evaluating the effects of SARS-CoV-2 Spike mutation D614G on transmissibility and Pathogenicity medRxiv preprint doi: https://doi.org/10.1101/2020.07.31.20166082
14. Obadia et al. "The Ro package: a toolbox to estimate reproduction numbers for epidemic outbreaks" BMC Medical Informatics and Decision Making, 12: 147. Published online 2012 Dec 18. doi: 10.1186/1472-6947-12-147
15. Ali, S.T, Lin Wang E., Lau.Y, Xu, Xiao-Ke, Du, Z., Wu, Y., Leungn, G.M., Cowling, B.J., Serial interval of SARS-CoV-2 was shortened over time by NPI interventions, Science 28 Aug 2020: Vol. 369, Issue 6507, pp. 1106-1109 DOI: 10.1126/science.abc9004
16. Celaschi, S., "Quantifying Effects, Forecasting Releases, and Herd Immunity of the Covid-19 Epidemic in S. Paulo – Brazil", The Intern. J. of Engineering and Science (IJES), Vol. 9, Issue 07, pp 33-42, (2020), ISSN (e): 2319-1813 ISSN (p): 20-24-1805, doi: 10.9790/1813-0907023342
17. Cori, A., Ferguson, N.M., Fraser, C., Cauchemez, S., "A new framework and software to estimate time-varying reproduction numbers during epidemics". Am. J. Epidemiol., 178, 1505–1512, (2013.
18. Felix, H.C., Fontenele J.B., "Instantaneous R calculation for COVID-19 epidemic in Brazil" MedRxiv preprint doi: https://doi.org/10.1101/2020.04.23.20077172
19. Estimating mortality from COVID-19, WHO reference number: WHO-2019-nCoV-Sci_Brief-Mortality-2020.1 Scientific Brief 4 August 2020
20. Marr, V., and Quartin, M., "An estimate of the COVID-19 infection fatality rate in Brazil base on a seroprevalence survey", medRxiv preprint doi: https://doi.org/10.1101/2020.08.18.20177626.
21. Grassly, N., Fraser, C., "Mathematical models of infectious disease transmission". Nature Rev Microbiol 6, 477–487 (2008). doi: 10.1038/nrmicro1845.
22. https://covid.saude.gov.br/ Visited September 10th, 2020.