

Investigating the first stage of the COVID-19 pandemic in Ukraine using epidemiological and genomic data

Yuriy Gankin ^{a,1}, Alina Nemira ^{b,1}, Vladimir Koniukhovskii ^{c,1}, Gerardo Chowell ^d, Thomas A. Weppelmann ^e, Pavel Skums ^{b,2}, Alexander Kirpich ^{d,2,*}

^a Quantori, Cambridge, Massachusetts, United States of America

^b Department of Computer Science, Georgia State University, Atlanta, Georgia, United States of America

^c EPAM Systems, Saint Petersburg, Russian Federation

^d Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, Georgia, United States of America

^e Department of Internal Medicine, University of South Florida, Tampa, Florida, United States

ARTICLE INFO

Keywords:

SARS-CoV-2

Agent-based

COVID-19

Ukraine

Phylogenetics

Phylodynamics

ABSTRACT

The novel coronavirus SARS-CoV-2 was first detected in China in December 2019 and has rapidly spread around the globe. The World Health Organization declared COVID-19 a pandemic in March 2020 just three months after the introduction of the virus. Individual nations have implemented and enforced a variety of social distancing interventions to slow the virus spread, that had different degrees of success. Understanding the role of non-pharmaceutical interventions (NPIs) on COVID-19 transmission in different settings is highly important. While most such studies have focused on China, neighboring Asian countries, Western Europe, and North America, there is a scarcity of studies for Eastern Europe. The aim of this epidemiological study is to fill this gap by analyzing the characteristics of the *first months of the epidemic* in Ukraine using agent-based modelling and phylodynamics. Specifically, first we studied the dynamics of COVID-19 incidence and mortality and explored the impact of epidemic NPIs. Our stochastic model suggests, that even a small delay of weeks could have increased the number of cases by up to 50%, with the potential to overwhelm hospital systems. Second, the genomic data analysis suggests that there have been multiple introductions of SARS-CoV-2 into Ukraine during the early stages of the epidemic. Our findings support the conclusion that the implemented travel restrictions may have had limited impact on the epidemic spread. Third, the basic reproduction number for the epidemic that has been estimated independently from case counts data and from genomic data suggest sustained intra-country transmissions.

1. Introduction

SARS-CoV-2 virus causing COVID-19 was first detected in December 2019 in the Chinese city of Wuhan (Zhu et al., 2020; Lu et al., 2020; Huang et al., 2020; Chen et al., 2020; Ciotti et al., 2020), and has rapidly spread around the globe, prompting the World Health Organization (WHO) to declare a pandemic in March, 2020 (WHO, 2020), just three month after the first reported case. Despite having much lower case-fatality rate than other recent coronavirus pandemics such as the severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS), the novel coronavirus has claimed more lives just within a few months of introduction than both of those epidemics combined (Mahase, 2020). As of June 29,

2021 there were more than 182 million infections with over 3.9 million deaths (Johns Hopkins University, 2021). In the absence of vaccines during the early pandemic period, non-pharmaceutical interventions (specifically, non-pharmaceutical epidemic mitigation interventions) were the only tools at the disposal of public health authorities to prevent and to mitigate the virus spread (Chen et al., 2020; Caddy, 2020; Mullard, 2020). The strategies implemented and enforced by governments around the world were highly variable and included frequent sanitation of public spaces, enforced social distancing, wearing of masks, closure of schools, churches, and ban of mass gatherings (Islam et al., 2020; WHO, 2020; CDC, 2020).

Even well after a year since the epidemic started, fundamental questions regarding the effects of non-pharmaceutical interventions

* Corresponding author.

E-mail address: akirpich@gsu.edu (A. Kirpich).

¹ Equal Contribution.

² Equal Contribution.

(NPIs) (Di Gennaro et al., 2020; Iyengar et al., 2020) and the genomic evolution of SARS-CoV-2 (Tizaoui et al., 2020; Yi, 2020; Li et al., 2020; Benvenuto et al., 2020) during the introductory period remain. Additionally, recent modeling efforts aimed at shedding light on those questions have mostly focused on China (Zhu et al., 2020; Lu et al., 2020; Huang et al., 2020; Chen et al., 2020), the rest of Asia (Wei et al., 2020; Yamamoto and Bauer, 2020; Roy et al., 2020), Western and Central Europe (Salje et al., 2020; Yamamoto and Bauer, 2020; Fokas et al., 2020), and North America (Ghinai et al., 2020; Ellington et al., 2020; Silverman et al., 2020; Fauver et al., 2020; Ciotti et al., 2020), largely neglecting Africa, the Middle East, and Eastern Europe.

In Eastern Europe, post-socialist economics and healthcare systems are inherently different from Western Europe. The available SARS-CoV-2 transmission models for Eastern Europe are based on relatively simple SIR or similar compartmental models (Issanov et al., 2020; Nesteruk, 2020; Kyrychko et al., 2020) where individuals are assigned to groups and all individuals within a given group are expected to have the same characteristics. To the best of our knowledge no agent-based modeling studies have been conducted for Ukraine to evaluate the impact of spatial heterogeneities in key transmission drivers such as density of infected individuals and their geographic locations. Furthermore, the number of genomic epidemiology studies on the COVID-19 pandemic in Eastern Europe has been limited. In this paper, we sought to fill the knowledge gap for the Ukrainian epidemic Åslund (2020), which provides a unique setting for studying the COVID-19 spread under the ex-USSR healthcare system, and with the epidemic mitigation policies similar to the rest of Europe.

The first confirmed case in Ukraine was reported on March 3, 2020 and was an individual who has recently traveled from Italy. The first death was reported on March 13, 2020 (Kyiv, 2020; Åslund, 2020). The Ukrainian government started to implement quarantine measure on March 12, 2020 (Åslund, 2020; Kyrychko et al., 2020) while the cases continued to rise possibly because of the delayed detections of existing infections and returns of infected Ukrainians from abroad (Kyrychko et al., 2020) (Fig. 3B). As a result, more strict measures have been implemented on April 6, 2020 which included the closure of schools, universities, shopping malls, and mandatory mask regiment in public places (Kyrychko et al., 2020). Those measures were slightly softened on April 24, 2020 and many services resumed even though some restrictions lasted till the end of June 2020 (Åslund, 2020). As a summary, Ukrainian officials took the epidemic very seriously from the beginning and started to implement the mitigation efforts and corresponding regulations almost immediately after multiple cases in the country have been detected. At the same time implementation of the proposed mitigation efforts did vary from region to region, and so did the compliance with those regulations (UN report, 2020; Atlantic Council, 2021).

2. Methods

2.1. Agent-based stochastic model

To investigate the COVID-19 epidemic in Ukraine and to assess its dynamics under different mitigation scenarios, we utilized our general stochastic agent-based modeling framework (Kirpich et al., 2021). The model was adjusted to the Ukrainian settings and fit into the observed Ukrainian data. The summary of the framework together with the adaptation details are outlined below.

In brief, the model simulates the epidemics evolving over the discrete time interval $(1, \dots, T)$ with time points $1 \leq t \leq T$ corresponding to calendar days and over the certain geographical area projected on a plain. Infected individuals are represented as agents with multiple characteristics that include geographic coordinates; age; infection time, severity and current status; disease stage; infectivity rate and infectivity radius which determines how frequently and where it produces secondary infections. The summaries of empirical reproduction numbers of individual agents which are generated by model simulations are used for the estimation of the population basic reproduction number R_0 (Milligan and Barrett, 2015). The geographical part of the model includes circular local epidemic spread areas $E = \{E_1, E_2, \dots, E_T\}$ characterized by their centers and radii. The centers of these areas represent hotspots of the infection introduction into the local population (e.g., transport hubs or administrative centers). The model incorporates NPI measures via a reduced infection transmission parameters which are effective starting from a certain calendar date customizable within the model.

In this study, we used epidemic spread areas and the corresponding incidence and mortality data reported by the National Security and Defence Council of Ukraine (National Security And Defence Council of Ukraine, 2021). It includes daily reports for individual administrative regions ("oblast") under the control of the Ukrainian government starting from March, 2020. The reported data was separated into three parts. The initially reported cases from March 12, 2020 to April 12, 2020 were retrospectively incorporated into the model as the initial conditions (Kirpich et al., 2021). The reported and model-produced data from April 22, 2020 to July 12, 2020 were used for model calibration, and from July 13, 2020 to August 1, 2020 – for model validation. The data before April 22, 2020 were used solely for the initial conditions to increase the model fit robustness, since the initial number of cases was relatively small in comparison to subsequent periods. The August 1, 2020 has been selected as the end date of our simulations to agree with the dates of genomic analysis based on available analyzed SARS-CoV-2 sequences collection times (GISAID, 2020).

Optimization of model parameters has been performed by minimizing the sum of squared differences between the model-produced outputs (across multiple runs) and the calibration data using the Nelder-Mead

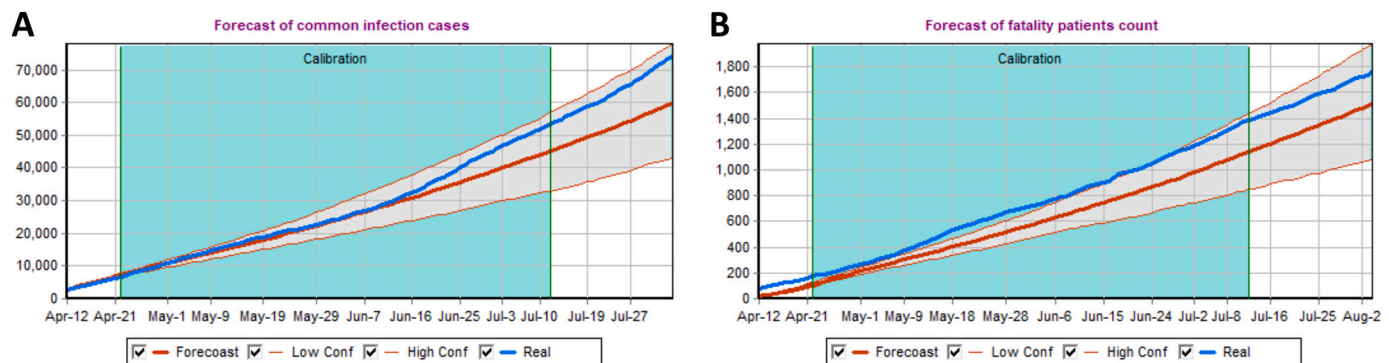


Fig. 1. The model predictions together with the reported cases (Panel A) and reported death (Panel B) are presented. The model calibration time interval is highlighted in green (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.). Red lines correspond to the median of the five hundred model-produced runs together with the corresponding 90% prediction bands to quantify the uncertainty. The actual observed case and death counts are displayed in blue for visual comparison.

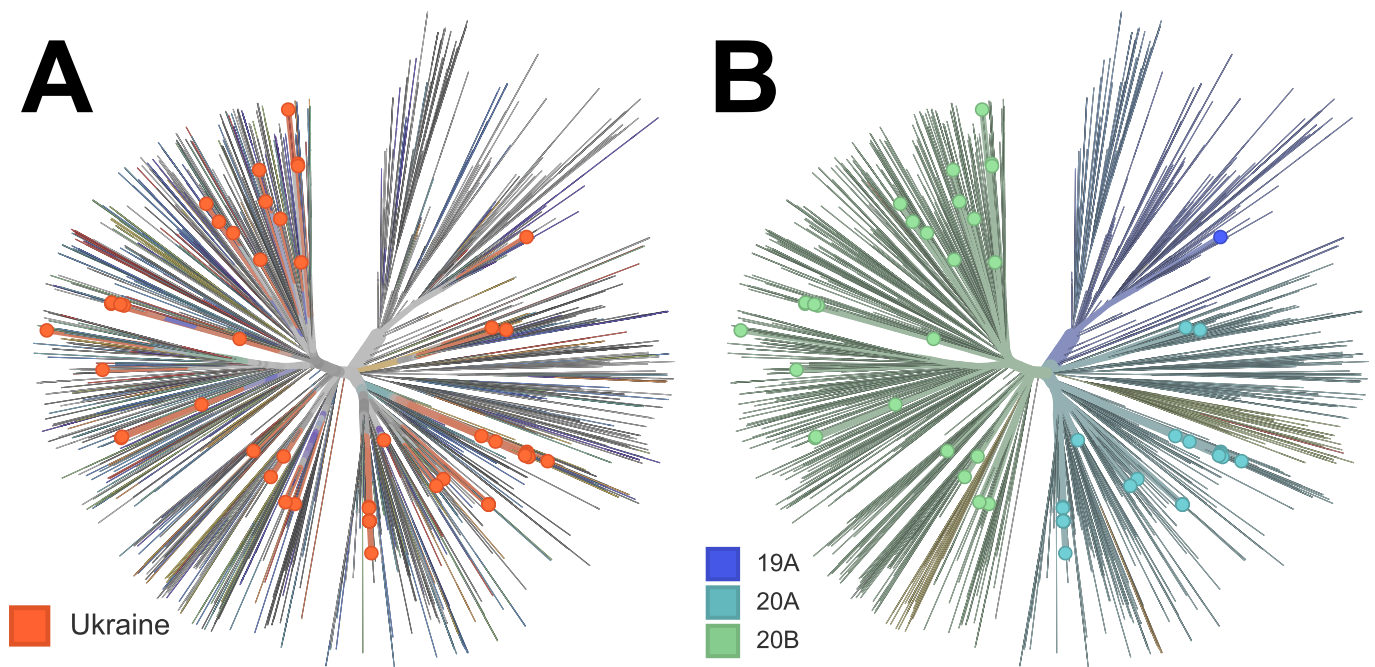


Fig. 2. The global phylogenetic tree of SARS-CoV-2 genomes: A) distribution of Ukrainian SARS-CoV-2 genomes inside global SARS-CoV-2 population, B) the same tree with global SARS-CoV-2 lineages highlighted.

numerical minimization method (Nelder and Mead, 1965). The population basic reproduction number \mathcal{R}_0 Milligan and Barrett (2015) has been estimated from the model-produced distribution quantiles (5%, median, 95%) of the reproduction numbers of individuals and summarized across multiple stochastic runs (Kirpich et al., 2021). The estimates for \mathcal{R}_0 were produced from the model fit to real data with the assumption that interventions have started almost immediately after the virus introduction.

In addition to the simulations based on the model fit to the actual case count, mortality and NPI data, two alternative simulation scenarios were considered under the hypothetical assumptions that NPIs that caused reduced transmissibility were implemented one (on April 19, 2020) and two (on April 26, 2020) weeks after the simulation start time. The results of simulations under these three scenarios were compared to assess the effect of timely NPI implementations.

The additional details about the model can be found in our earlier study (Kirpich et al., 2021), and the model implementation tailored to Ukrainian data is available at <https://github.com/quantori/COVID19-Ukraine-Transmission>.

2.2. Genomic epidemiology analysis

For the first stage of the analysis, we utilized a SARS-CoV-2-specific pipeline implemented by Nextstrain ([github.com](https://github.com/nextstrain), 2021). Sixty high-quality SARS-CoV-2 genomes from Ukraine sampled between April 24, 2020 and August 7, 2020 were extracted from GISAID (GISAID, 2020). These genomes were complemented by a representative subsample of 6479 sequences from other geographic regions obtained using a country-specific subsampling algorithm of Nextstrain. Briefly, it first selects sequences from the focal country, and then complements them with samples from other geographical areas, prioritizing sequences from its geographic neighbors and the sequences that are genetically close to the “focal” sequences. This reduces the size of the genomic dataset to that allowing for feasible phylogenetic inference. The maximum likelihood phylogenetic tree for the obtained dataset was constructed by IQ-TREE (Nguyen et al., 2015) using Hasegawa-Kishino-Yano (HKY)+ Γ nucleotide substitution model (Hasegawa et al., 1985). For the obtained timed phylogeny, ancestral geographic traits of internal nodes were inferred using the “mugration model” of nextstrain (Hadfield et al.,

2018). Using these traits, intra-country transmission clusters were identified as clades with the most recent common ancestors (MRCA) estimated as originating from Ukraine. For each cluster, confidence intervals for emergence times for MRCA and its parent were also obtained.

Next, a phylodynamic analysis of the three largest clusters and the entire Ukrainian SARS-CoV-2 population was performed using BEAST v1.10.4 (Suchard et al., 2018). We used a strict molecular clock, HKY+ Γ nucleotide substitution model, a tree prior with exponential growth coalescent. Priors for the parameters were defined in BEAUti v 1.10.4 and were the following: a) normal $\mathcal{N}(\text{mean} = 8.0\text{e-}4, \text{st. dev} = 2.0\text{e-}5)$ for the clock rate, b) log-normal $\mathcal{LN}(\text{mean} = 1.0, \text{st. dev} = 1.25)$ for the population size, c) double exponential (Laplace) distribution $\text{DEXP}(\mu = 0, b = 100)$ for the growth rate, d) normal $\mathcal{N}(\text{mean} = 0, \text{st. dev} = 1)$ for the freqParameter, e) exponential $\text{EXP}(\text{mean} = 0.5, \text{offset} = 0)$ for the gammaShape parameter, and f) log-normal $\mathcal{LN}(\text{mean} = 1.0, \text{st. dev} = 1.25)$ for the kappa parameter. The detailed parameters file is available in XML format at <https://github.com/alanira/COVID19-Ukraine-phylodynamics>. The parameters were estimated after 30,000, 000 iterations of Markov Chain Monte Carlo (MCMC) sampling, with the initial 10% values discarded as burn-in. The results were accepted if the effective sample sizes were above 200 for all parameters. The estimated exponential growth rates were used to calculate the basic reproduction numbers \mathcal{R}_0 under the assumption that SARS-CoV-2 generation intervals (i.e., times between infection onset and onward transmission) were gamma-distributed (Li et al., 2020). We used the formula

$$\hat{\mathcal{R}}_0 = \left[1 + \frac{\hat{f}\hat{\sigma}^2}{\hat{\mu}} \right]^{\frac{2}{\hat{\sigma}^2}}, \quad (1)$$

where $\hat{\mu}$ and $\hat{\sigma}$ are the mean and standard deviation of the aforementioned gamma distribution (BEAST, 2020a,b; Wallinga and Lipsitch, 2007; Grassly and Fraser, 2008). For these values, we used the estimates $\hat{\mu} = 5.20$ and $\hat{\sigma} = 1.72$ from Ganyani et al. (2020) and $\hat{\mu} = 3.95$ and $\hat{\sigma} = 1.51$ from Ganyani et al. (2020). The formula Eq. (1) defines a strictly monotone transformation of \hat{f} , and, therefore, it also straightforwardly transforms the 95% highest posterior density (HPD) intervals for \hat{f} into those for \mathcal{R}_0 .

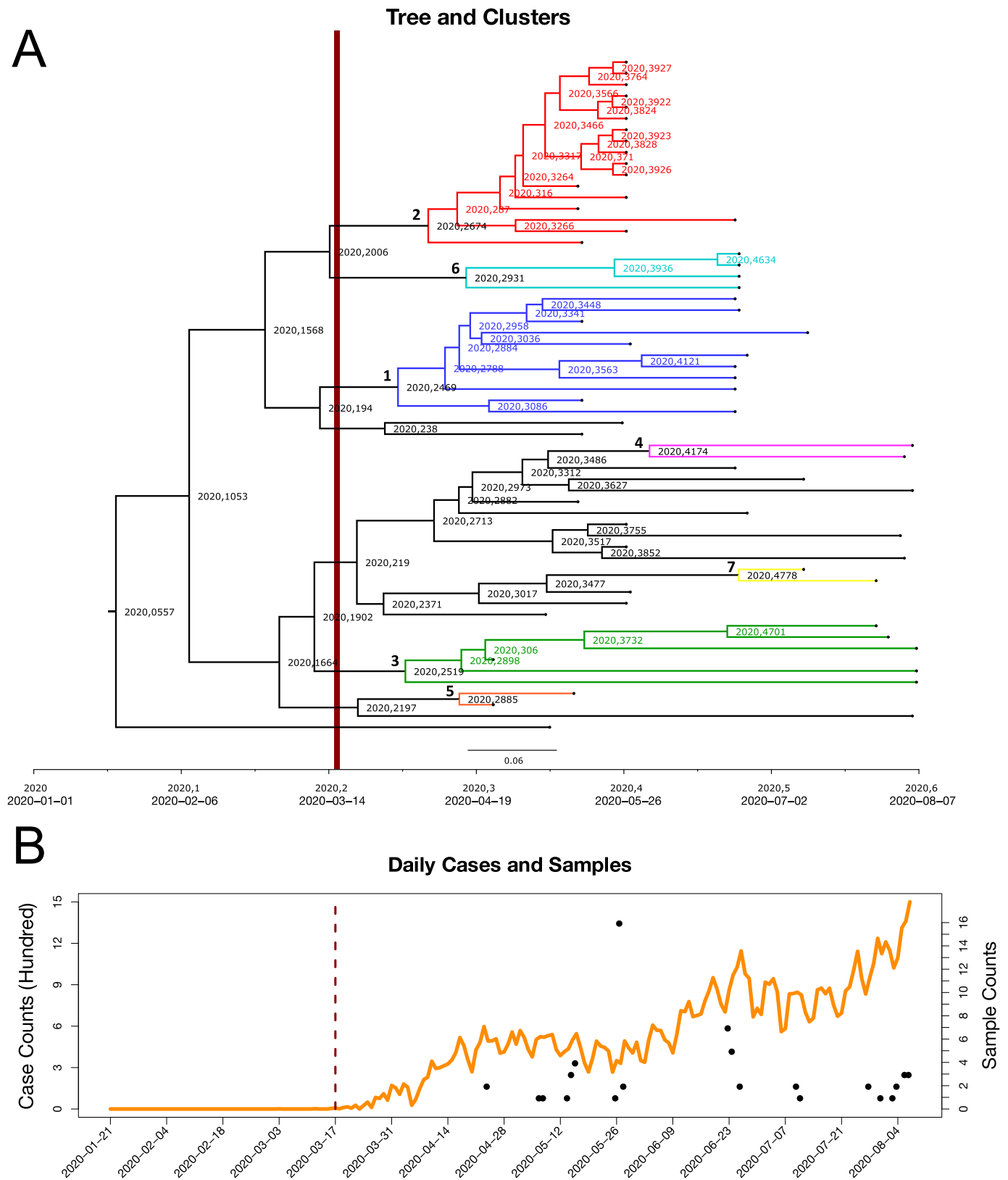


Fig. 3. Panel A: The SARS-CoV-2 clusters are presented in the Ukrainian phylogenetic tree. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) Clusters colored by blue, red, green, pink, orange, azure, yellow, and numbered from one to seven, respectively. On the tree the length of the entire calendar year is equal to one while the decimal fractions correspond to dates within the year. Dates in standard format are provided as on the axis for the reference. Panel B: Daily incidence of reported cases for Ukraine (orange) together with the sample counts and collection dates for sequenced samples (black). The travel restriction has happened on March 17, 2020, which is indicated by a vertical dark red bar time separator in both panels.

3. Results

3.1. Agent-based stochastic model

The visual results of the first scenario (model fit) and the corresponding outputs are summarized in Fig. 1. Blue curves in Fig. 1 correspond to the reported data. They are captured by the model fits which is also indicated by the corresponding median and 90% pointwise model prediction bands across five hundred runs. The calibration interval is highlighted by cyan background.

For each of the three considered scenarios the median value across five hundred simulations were computed at each time point and presented together with the corresponding 5-th and 95-th percentiles across five hundred stochastic realizations to form the 90% prediction intervals (PI-s). The corresponding results are summarized in Tables 1 and 2 for the model-predicted cases and deaths, respectively. The three scenario summaries from Table 1 can be directly compared. For comparison the actual number of reported cases by August 1, 2020 was 71,056 (National Security And Defence Council of Ukraine, 2021) which validates the model fit since August 1, 2020 was outside of the calibration interval. The hypothetical April 19th and April 26th intervention start dates produce larger number of cases in comparison to the original fitted scenario. The median estimates can be compared directly. The hypothetical April 19th scenario results in 16% predicted increase in cumulative number of cases on June 1, 2020 and in 20% predicted increase in cumulative number of cases on August 1, 2020 in comparison to the fitted scenario. The hypothetical April 26th scenario results in 36% predicted increase in cumulative number of cases on June 1, 2020 and in 46% predicted increase in cumulative number of cases on August 1, 2020 when compared to the fitted scenario.

The hypothetical April 19th scenario results in 14% increase in cumulative number of deaths predicted on June 1, 2020 and in 20% increase in cumulative number of deaths predicted on August 1, 2020. The hypothetical April 26th scenario results in 32% increase in cumulative number of deaths predicted on June 1, 2020 and in 46% increase in cumulative number of deaths predicted on August 1, 2020.

Interestingly, the median results for the hypothetical April 19th scenario displayed better alignment with the actual data. This suggests the delayed impact of NPIs in transmission mitigation caused by the time needed to put the prescribed measures into effect. Furthermore, the obtained results demonstrate the importance of the early epidemic mitigation measures which cause the reduction in transmission probability parameters and, therefore, a reduction in the number of cases, and (more importantly) deaths. At the same time the results for later mitigation efforts implementation dates should only be interpreted as sensitivity analysis, since the Ukrainian government has implemented quarantine measures from the beginning of the epidemic and there were no data to properly estimate the corresponding non-intervention transmission probability parameters (github.com, 2021). Therefore, the corresponding non-quarantine probability parameters have been adopted from the previous analysis (Kirpich et al., 2021).

The population basic reproduction number R_0 estimate during the

Table 1

The model outputs are presented together with the reported data. The predicted number of cumulative cases produced by the model over time for three different epidemic mitigation scenarios for three initiation dates together with the corresponding 90% prediction intervals.

Scenario	June 1,2020	July 1,2020	August 1,2020
Real Data	24,012	44,998	71,056
April 12, 2020	23,724 (19,093;28,250)	38,932 (29,181;48,782)	57,810 (41,864;75,029)
April 19, 2020	27,511 (22,681;32,872)	46,193 (36,238;57,100)	69,358 (50,642; 88,407)
April 26, 2020	32,220 (26,016;38,645)	55,141 (41,974;69,992)	84,227 (63,992;109,069)

Table 2

The model outputs are presented together with the reported data. The predicted number of cumulative death produced by the model over time for three different epidemic mitigation scenarios for three initiation dates together with the corresponding 90% prediction intervals.

Scenario	June 1,2020	July 1,2020	August 1,2020
Real Data	718	1,173	1,709
April 12, 2020	568 (464;667)	967 (738;1,203)	1,461 (1,050;1,904)
April 19, 2020	653 (541;783)	1,143 (889;1,412)	1,762 (1,308; 2,241)
April 26, 2020	752 (612;906)	1,364 (1,049;1,705)	2,141 (1,604;2,719)

intervention was estimated to be 1.10 (median) with the corresponding 90% confidence interval from quantiles equal to (0.24;1.88).

3.2. Genomic epidemiology of SARS-CoV-2

Despite a sparse sampling, the observed genomic diversity of SARS-CoV-2 in Ukraine is substantial, indicating both multiple introductions of the virus and sustained intra-country evolution (Fig. 2). This agrees well with the patterns observed in other countries (Geoghegan et al., 2020), and emphasizes the contribution of global movement of people to the rapid spread of SARS-CoV-2. Specifically, Ukrainian sequences are distributed among eight lineages by the classification of (Rambaut et al., 2020) as follows: $B.1$ - 50.0% of genomes, $B.1.1$ - 28.3%, $B.1.1.243$ - 8.3%, $B.1.527$ - 5.0%, $B.1.1.325$ - 3.3%, and 1.7% for each $B.1.131$, $B.1.1.194$, B . Similarly, by Nextstrain classification the distribution of lineages is: 19A - 1.7%, 20A - 51.7% and 20B - 46.7% (Fig. 2).

Seven Ukrainian clusters contain multiple sequences and jointly constitute 73.3% of all sampled genomes. Presence of these clusters and the corresponding intra-country lineages indicate sustained internal transmissions (Fig. 3A and Figure A1–A7 in Appendix). For each such lineage, a confidence interval of its introduction time can be assessed by the union of the confidence intervals for inferred dates of its Ukrainian MRCA v and the non-Ukrainian parent of v (Table A2 in Appendix).

We analyzed these introduction times relatively to the implementation time of the travel ban, that was established on March 16, 2020 (baltsi.mfa.gov.ua, 2021) for foreign citizens and on March 17, 2020 (Closed borders, 2021) for all travelers with the exception of Ukrainian citizens returning from abroad. It turned out that three out of seven lineages were most likely introduced into the country after the travel ban date, as indicated by their introduction confidence intervals (Fig. 4). Similarly, a single lineage was likely imported before that date; for three remaining lineages the travel ban date falls into their confidence intervals, preventing us from the decisive conclusion, even though the date lies closer to the left ends of all intervals. Thus, the analysis support the hypothesis that the travel restrictions had limited effect on the virus importation control.

The estimates of the basic reproduction number R_0 for three largest lineages are summarized in Table 3. All estimates are significantly above one, indicating sustained local transmission of SARS-CoV-2 during the first months of the epidemic in Ukraine.

4. Discussion

In this study, we have detailed the epidemic characteristics of the first months of the COVID-19 pandemic in Ukraine and studied the effects of NPIs. We considered two complementary approaches based on the stochastic modeling applied to incidence data and genomic epidemiology methods applied to sequencing data. Different types of data reflect various aspects of the epidemics, and are prone to different biases. Therefore, such synthetic approach facilitates ubiquitous understanding of the early stages of the epidemic in Ukraine.

COVID-19 pandemic is characterized by a richness of available data, that allow to utilize agent-based modelling and genomic analysis at the finest possible resolution. In Ukraine, we have an access to public health

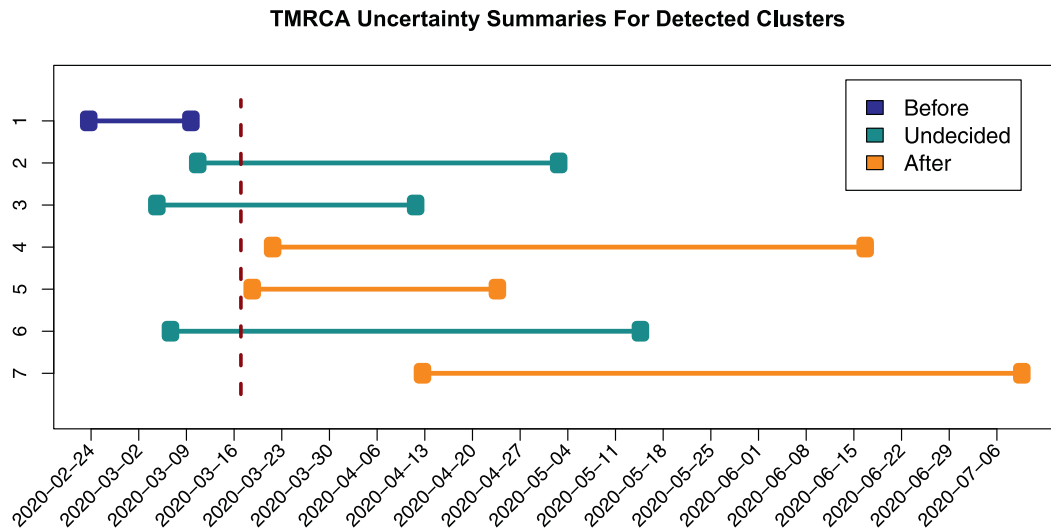


Fig. 4. Introduction times for seven largest Ukrainian transmission lineages visualized from Table A2 from the supplement. The estimated intervals for introduction times are depicted as horizontal lines, the border closure date (March 17, 2020) is indicated by a vertical line.

Table 3

The estimates of the basic reproduction number \mathcal{R}_0 for three largest clusters together with the corresponding 95% confidence intervals (CI-s). The results are reported for two pairs of generation interval distribution parameters $\hat{\mu}$ and $\hat{\sigma}$ reported by two studies.

Cluster	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mathcal{R}}_0$ & 95% CI	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mathcal{R}}_0$ & 95% CI
First	5.20	1.72	1.31 (1.12; 1.52)	3.95	1.51	1.23 (1.09; 1.37)
Second	5.20	1.72	1.47 (1.1; 1.98)	3.95	1.51	1.34 (1.07; 1.68)
Third	5.20	1.72	1.48 (1.16; 1.94)	3.95	1.51	1.35 (1.12; 1.65)

data on the level of individual regions, which makes agent-based model predictions more comprehensive. Similarly, the advances and cost reduction of next-generation sequencing (NGS) methods allowed rapid genomic data acquisition at early stages of the epidemic (GISAID, 2020). These data processed by advanced phylogenetic and phylodynamic models allow to assess the virus importation and intra-country transmission dynamics from a “different angle” (Armstrong et al., 2019). Furthermore, in cases when two methods produced independent estimations of the basic reproduction number \mathcal{R}_0 , the obtained results are comparable, thus highlighting their consistency. The uncertainty estimates for the stochastic estimates are wider, which may be due to the fact that stochastic model has more parameters and higher variability in the outputs while phylodynamic models has pretty strong priors.

The study has limitations since the available surveillance incidence and genomic molecular data are limited. Ukraine is one of the poorest countries in Europe (based on GDP per capita), and, therefore, the health care infrastructure in Ukraine lacks in some parts the resources of developed countries (emerging-europe.com, 2021). In particular, the shortage of screening tests and the limited reporting of incidence data during the initial epidemic likely substantially underestimated the burden of disease in terms of incidence counts. Likewise, the scarcity of sequencing capacities limited the number of genomic data for the phylogenetic analysis. As such, the actual number of viral clusters of local transmission remains unknown and should be interpreted as a lower bound which only form the “tip of the iceberg” of all transmission clusters. Moreover, the local population compliance with the NPI regulations implemented by officials is always a question, which might have reduced the effectiveness of such measures (UN report, 2020; Atlantic Council, 2021).

In summary, this study was among the first to explore the characteristics of the initial pandemic as it spread to Ukraine and provided

additional genomic analysis not previously published.

5. Conclusions

Ukraine is a country of 44 million inhabitants (7th largest in Europe by population and second by territory), but our knowledge about COVID-19 epidemic there is highly limited, especially, in comparison to most other European countries. The presented work aimed at filling this knowledge gap. Specifically, we have conducted a comprehensive study of multiple aspects of the COVID-19 epidemic in Ukraine by applying and combining agent-based, and phylodynamics methodologies. This approach allowed us to produce a holistic picture of the epidemics, even given the relative scarcity of specific types of data.

In particular, the stochastic model suggested that even a small delay of weeks could have increased the number of cases by up to 50%, with the potential to overwhelm hospital systems. The genomic data analysis suggested that there have been multiple introductions of SARS-CoV-2 into Ukraine during the early stages of the epidemic. Our findings also support the conclusion that the implemented travel restrictions may have had limited impact on the epidemic spread. The estimates of the basic reproduction number \mathcal{R}_0 for the epidemic have been obtained independently from case counts data and from genomic data and both have suggested sustained intra-country transmissions.

Declaration of Interests

The authors declare no conflict of interests.

Funding

AN was supported by the GSU Brains and Behavior fellowship. PS was supported by the National Institutes of Health grant 1R01EB025022 and by the National Science Foundation grant 2047828.

Acknowledgements

The authors would like to acknowledge the authors, originating and submitting laboratories of the sequences from GISAID. The list of laboratories is provided in Appendix. The authors would also like to acknowledge Sergey Bratsky for his help with Ukraine incidence data collection and acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.meegid.2021.105087>.

References

- Armstrong, G.L., MacCannell, D.R., Taylor, J., Carleton, H.A., Neuhaus, E.B., Bradbury, R.S., Posey, J.E., Gwinn, M., 2019. Pathogen genomics in public health. *N. Engl. J. Med.* 381, 2569–2580.
- Åslund, A., 2020. Responses to the covid-19 crisis in russia, ukraine, and belarus. *Eurasian Geogr. Econ.* 1–14.
- Atlantic Council, 2021. Ukraine's local authorities and the Covid-19 pandemic <https://www.atlanticcouncil.org/blogs/ukrainealert/ukraines-local-authorities-and-the-covid-19-pandemic/>, Accessed: 2021-06-29.
- baltis.mfa.gov.ua, 2021. No Entry for Foreigners - May 26, 2020 <https://baltis.mfa.gov.ua/ru/news/timchasova-zaborona-do-24-kvitnya-2020-r-na-peretinannya-derzhavnogo-kordonu-na-vyizd-v-ukrayinu-dlya-inozemciv-ta-osib-bez-gromadyanstva>, Accessed: 2021-02-20.
- BEAST, 2020a. Software https://beast.community/estimating_R0.html, Accessed: 2021-02-20.
- BEAST, 2020b. Software https://beast.community/phylogenetics_of_epidemic_influenza, Accessed: 2021-02-20.
- Benvenuto, D., Giovanetti, M., Ciccozzi, A., Spoto, S., Angeletti, S., Ciccozzi, M., 2020. The 2019-new coronavirus epidemic: evidence for virus evolution. *J. Med. Virol.* 92, 455–459.
- Caddy, S., 2020. Russian SARS-CoV-2 vaccine. *BMJ(Online)* 370.
- CDC, 2021. <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html>, 2020. Accessed: 2021-02-20.
- Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., et al., 2020. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in wuhan, china: a descriptive study. *Lancet* 395, 507–513.
- Chen, W.-H., Strych, U., Hotez, P.J., Bottazzi, M.E., 2020. The SARS-CoV-2 vaccine pipeline: an overview. *Curr. Trop. Med. Rep.* 1–4.
- Ciotti, M., Ciccozzi, M., Terrinoni, A., Jiang, W.-C., Wang, C.-B., Bernardini, S., 2020. The covid-19 pandemic. *Crit. Rev. Clin. Lab. Sci.* 57, 365–388.
- Closed borders, 2021. <https://www.kmu.gov.ua/news/uryad-timchasovo-zaboroniv-vyizd-v-ukrayinu-inozemciv-ta-zakriv-mizhnarodne-pasazhirske-aviaspoluchennya>, 2021b. Accessed: 2021-02-20.
- Di Gennaro, F., Pizzol, D., Marotta, C., Antunes, M., Racalbutto, V., Veronese, N., Smith, L., 2020. Coronavirus diseases (covid-19) current status and future perspectives: a narrative review. *Int. J. Environ. Res. Public Health* 17, 2690.
- Ellington, S., Strid, P., Tong, V.T., Woodworth, K., Galang, R.R., Zambrano, L.D., Nahabedian, J., Anderson, K., Gilboa, S.M., 2020. Characteristics of women of reproductive age with laboratory-confirmed SARS-CoV-2 infection by pregnancy status-united states, january 22-june 7, 2020. *Morb. Mort. Wkly Rep.* 69, 769.
- Emerging-europe.com, 2021. <https://emerging-europe.com/news/imf-ukraine-is-emerging-europes-poorest-country/>, 2021. Accessed: 2021-03-01.
- Fauver, J.R., Petrone, M.E., Hodcroft, E.B., Shioda, K., Ehrlich, H.Y., Watts, A.G., Vogels, C.B., Brito, A.F., Alpert, T., Muyombwe, A., et al., 2020. Coast-to-coast spread of SARS-CoV-2 during the early epidemic in the united states. *Cell*.
- Fokas, A., Dikaos, N., Kastis, G., 2020. Mathematical models and deep learning for predicting the number of individuals reported to be infected with SARS-CoV-2. *J. R. Soc. Interf.* 17, 20200494.
- Ganyani, T., Kremer, C., Chen, D., Torneri, A., Faes, C., Wallinga, J., Hens, N., 2020. Estimating the generation interval for coronavirus disease (covid-19) based on symptom onset data, march 2020. *Eurosurveillance* 25, 2000257.
- Geoghegan, J.L., Ren, X., Storey, M., Hadfield, J., Jelley, L., Jefferies, S., Sherwood, J., Paine, S., Huang, S., Douglas, J., et al., 2020. Genomic epidemiology reveals transmission patterns and dynamics of SARS-CoV-2 in aotearoa new zealand. *Nat. Commun.* 11, 1–7.
- Ghinai, I., McPherson, T.D., Hunter, J.C., Kirking, H.L., Christiansen, D., Joshi, K., Rubin, R., Morales-Estrada, S., Black, S.R., Pacilli, M., et al., 2020. First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in the usa. *Lancet*.
- GISAID, 2020. A global initiative on sharing avian flu data <https://www.gisaid.org/>, Accessed: 2021-02-20.
- github.com, 2021. <https://github.com/nextstrain/ncov>, 2021. Accessed: 2020-11-01.
- github.com, 2021. <https://github.com/quantori/COVID19-Ukraine-Transmission>, 2021. Accessed: 2020-11-01.
- Grassly, N.C., Fraser, C., 2008. Mathematical models of infectious disease transmission. *Nat. Rev. Microbiol.* 6, 477–487.
- Hadfield, J., Megill, C., Bell, S.M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford, T., Neher, R.A., 2018. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 34, 4121–4123.
- Hasegawa, M., Kishino, H., Yano, T.-a., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial dna. *J. Mol. Evol.* 22, 160–174.
- Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al., 2020. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *Lancet* 395, 497–506.
- Islam, N., Sharp, S.J., Chowell, G., Shabnam, S., Kawachi, I., Lacey, B., Massaro, J.M., D'Agostino, R.B., White, M., 2020. Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. *BMJ* 370.
- Issanov, A., Amanbek, Y., Abbay, A., Adambekov, S., Aljofan, M., Kashkynbayev, A., Gaipov, A., 2020. Covid-19 Outbreak in Post-Soviet States: Modeling the Best and Worst Possible Scenarios.
- Iyengar, K., Mabrouk, A., Jain, V.K., Venkatesan, A., Vaishya, R., 2020. Learning opportunities from covid-19 and future effects on health care system. *Diabetes Metab. Syndr.* 14, 943–946.
- Johns Hopkins University, 2021. Johns Hopkins University - Coronavirus Resource Center, <https://coronavirus.jhu.edu/map.html>, 2021. Accessed: 2021-03-01.
- Kirpich, A., Koniukhovskii, V., Shvartc, V., Skums, P., Weppelmann, T.A., Imanyitov, E., Semyonov, S., Barsukov, K., Gankin, Y., 2021. Development of an interactive, agent-based local stochastic model of COVID-19 transmission and evaluation of mitigation strategies illustrated for the state of Massachusetts, USA. *PLoS ONE* 16, e0247182.
- Kyiv Post, 2021. <https://www.kyivpost.com/ukraine-politics/developing-first-coronavirus-case-identified-in-ukraine.html>, 2020. Accessed: 2021-02-20.
- Kyrychko, Y.N., Blyuss, K.B., Brovchenko, I., 2020. Mathematical modelling of the dynamics and containment of covid-19 in ukraine. *Sci. Rep.* 10, 1–11.
- Li, C., Yang, Y., Ren, L., 2020. Genetic evolution analysis of 2019 novel coronavirus and coronavirus from other species. *Infect. Genet. Evol.* 104285.
- Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K.S., Lau, E.H., Wong, J.Y., et al., 2020. Early transmission dynamics in wuhan, china, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.*
- Lu, H., Stratton, C.W., Tang, Y.-W., 2020. Outbreak of pneumonia of unknown etiology in wuhan china: the mystery and the miracle. *J. Med. Virol.*
- Mahase, E., 2020. Coronavirus: covid-19 has killed more people than sars and mers combined, despite lower case fatality rate. *BMJ*.
- Milligan, G.N., Barrett, A.D., 2015. An Essential Guide. Wiley Online Library.
- Mullard, A., 2020. How COVID vaccines are being divvied up around the world. *Nature*. <https://doi.org/10.1038/d41586-020-03370-6>. URL: <https://doi.org/10.1038/d41586-020-03370-6>.
- National Security And Defence Council of Ukraine, <https://covid19.mbo.gov.ua/>, 2021. Accessed: 2021-02-20.
- Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. *Comput. J.* 7, 308–313.
- Nesteruk, I., 2020. Long-Term Predictions for Covid-19 Pandemic Dynamics in Ukraine, Austria and Italy. Europe PMC.
- Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. Iq-tree: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274.
- Rambaut, A., Holmes, E.C., O'Toole, Á., Hill, V., McCrone, J.T., Ruis, C., du Plessis, L., Pybus, O.G., 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* 5, 1403–1407.
- Roy, D., Tripathy, S., Kar, S.K., Sharma, N., Verma, S.K., Kaushal, V., 2020. Study of knowledge, attitude, anxiety & perceived mental healthcare need in indian population during covid-19 pandemic. *Asian J. Psychiatr.* 102083.
- Salje, H., Kiem, C.T., Lefrancq, N., Courtejoie, N., Bosetti, P., Paireau, J., Andronico, A., Hozé, N., Richet, J., Dubost, C.-L., et al., 2020. Estimating the burden of SARS-CoV-2 in france. *Science*.
- Silverman, J.D., Hupert, N., Washburne, A.D., 2020. Using influenza surveillance networks to estimate state-specific prevalence of SARS-CoV-2 in the united states. *Sci. Transl. Med.* 12.
- Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using beast 1.10. *Virus Evol.* 4, vey016.
- Tizaoui, K., Zidi, I., Lee, K.H., Abou Ghayda, R., Hong, S.H., Li, H., Smith, L., Koyanagi, A., Jacob, L., Kronbichler, A., et al., 2020. Update of the current knowledge on genetics, evolution, immunopathogenesis, and transmission for coronavirus disease 19 (covid-19). *Int. J. Biol. Sci.* 16, 2906.
- UN report, 2021. UN report: Impact of COVID-19 on human rights in Ukraine, https://www.ohchr.org/Documents/Countries/UA/Ukraine_COVID-19_HR_impact_EN.pdf, 2020. Accessed: 2021-06-29.
- Wallinga, J., Lipsitch, M., 2007. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. R. Soc. B: Biol. Sci.* 274, 599–604.
- Wei, W.E., Li, Z., Chiew, C.J., Yong, S.E., Toh, M.P., Lee, V.J., 2020. Presymptomatic transmission of SARS-CoV-2-singapore, january 23-march 16, 2020. *Morb. Mort. Wkly Rep.* 69, 411.
- WHO, 2021. <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-COVID-19-11-march-2020>, 2020. Accessed: 2021-02-20.
- WHO, 2021. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public>, 2020. Accessed: 2021-02-20.
- Yamamoto, N., Bauer, G., 2020. Apparent difference in fatalities between central europe and east asia due to SARS-CoV-2 and covid-19: four hypotheses for possible explanation. *Med. Hypotheses* 110160.
- Yi, H., 2020. 2019 novel coronavirus is undergoing active recombination. *Clin. Infect. Dis.*
- Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., et al., 2020. A novel coronavirus from patients with pneumonia in china, 2019. *N. Engl. J. Med.*