The COVID-19 Pandemic Dynamics and Incomes

Igor Nesteruk^{1,2}, Oleksii Rodionov³

¹ Institute of Hydromechanics, National Academy of Sciences of Ukraine, Kyiv, Ukraine

2 SBIDER (Systems Biology & Infectious Disease Epidemiology Research) Centre at University of Warwick, UK

³ Private consulting office, Kyiv, Ukraine

Abstract

Objectives. Large differences in the number of registered SARS-CoV-2 cases per capita in different countries encourage research into the causes of this phenomenon. In particular, the accumulated numbers of cases per million (CC) demonstrated strong linear correlations with the gross domestic product per capita (GDP) and the median age of populations. In this paper the possible correlations between GDP and numbers of cases CC and deaths (DC) per million, case fatality risks CFR=DC/CC, vaccinations and testing levels will be investigated. As well non-linear correlations of CC, DC and CFR values versus vaccinations and testing levels will be considered.

Methods. A non-linear correlation and John Hopkins University (JHU) datasets for African and European countries corresponding to August 1, 2022 were used.

Results. The numbers of CC, DC and CFR increase for richer countries, the same trends were revealed for DC and CFR values in Africa, but opposite ones in Europe. As expected, the testing and vaccination levels increase with the growth of GDP. Higher levels of testing probably allowed revealing more cases and COVID-19 related deaths in rich countries. CC values showed a very strong increasing trend with the increase of numbers of tests per capita (TC). Unexpectedly, the same increasing trend was revealed for CC and DC values versus percentage of fully vaccinated people (VC). Nevertheless, the decrease of CFR with the increase of VC demonstrates a positive effect of vaccinations.

Conclusions. In some countries, the number of undetected COVID-19 cases may be tens or even hundreds of times higher than the number of registered ones due to the differences in testing levels and age structure. This fact increases the probability of the appearance of new dangerous SARS-CoV-2 strains and has to be taken into account in further investigations of impact of different factors on the pandemic dynamics.

Keywords

COVID-19 pandemic, epidemic dynamics in Africa, epidemic dynamics in Europe, gross domestic product per capita, non-linear correlation, statistical methods.1

Introduction

The general characteristics of the COVID-19 pandemic dynamics require further research despite the vast number of available publications, including studies comparing the COVID-19 pandemic dynamics in different regions and the impact of various factors [1-18]. In particular, a strong linear correlation was revealed in [18] between the gross domestic product per capita (GDP) [19] and the numbers of cases per million (CC) registered in African countries as of February 1, 2022, [20]. In this study, a non-linear correlation between incomes and values of CC, accumulated numbers of deaths per million (DC) and the case fatality risk (CFR=DC/CC) will be investigated with the use of datasets for African and European countries corresponding to August 1, 2022, [20]. We will also discuss the possible influence of the accumulated numbers of tests per capita (TC) and the percentage of fully vaccinated people (VC) on the CC, DC and CFR values.

inesteruk@yahoo.com (I. Nesteruk); aleksei.rdnv@gmail.com (O. Rodionov) 0000-0001-7250-2729 (I. Nesteruk); 0000-0002-4466-2183 (O. Rodionov)

 \bigcirc © 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). CEUR Workshop Proceedings (CEUR-WS.org)

ProfIT AI 2024: 4th International Workshop of IT-professionals on Artificial Intelligence (ProfIT AI 2024), September 25–27, 2024, Cambridge, MA, USA

Data, the non-linear regression and Fisher test

We will use the data sets regarding the gross domestic product per capita (GDP) based on purchasing power parity (PPP) available in [19] and some COVID-19 characteristics reported by John Hopkins University (JHU) as of August 1, 2022, [20]. The figures corresponding to the versions of files available on September 4, 2022 are presented in supplementary Tables S1 and S2 and shown in the Figure for African (black markers) and European countries (blue markers).

The following non-linear correlation will be applied:

$$
y = a + b(x + c)^{\gamma}; \ a \ge 0, \ b > 0, \ x > -c \tag{1}
$$

to find links between GDP, TC, and VC (variable x) and CC, DC, DC/CC, TC, and VC (variable y). At γ =1 relationship (1) reduces to the linear one. It can be also reduced to the linear correlation by using new random variables *z* and *w*≡log(*x*+*c*), [11, 14]:

$$
z \equiv \log(y-a) = \log(b) + \gamma \log(x+c)
$$
 (2)

The constant parameters γ , $log(b)$ and corresponding best fitting lines can be found with the use of standard linear regression formulas [21] for different values of constant parameters *a* and *c*. Their optimal values correspond to the maximum of the correlation coefficient magnitude |*r*| or the ratio of the Fisher functions $F/F_c(k_1,k_2)$, $(k_1=1, k_2=n-2, n$ is the number of observations, i.e., the number of countries in datasets), [6, 11, 14]. The corresponding experimental values *F* can be calculated with the use of formula (S1), [21], the critical values $F_c(k_1,k_2)$ of the Fisher function at a desired significance or confidence level can be found in [22]. If $F/F_c(k_1,k_2)$ < 1, the hypothesis about the relationship (1) is not supported by the results of observations. The highest values of $F/F_c(k_1,k_2)$ correspond to the most reliable hypotheses.

Results

The optimal values of parameters for different non-linear correlations (1) are listed in Table 1. Corresponding best fitting lines are shown in the Figure by the black color for African datasets, blue - for Europe and red - for complete datasets (Africa + Europe). Rows 1-3 of Table 1 and solid lines in Figure illustrate that the accumulated numbers of cases per million CC always increase with the increase of the incomes. Nevertheless, the number of deaths DC and case fatality risk CFR decrease in richer European countries (see rows 5, 8 and blue dashed and dotted lines; the correlation DC versus GDP is supported at confidence level 0.025; F_c (1, 40) =5.47). What is surprising is the increase in DC and CFR values with increasing income in Africa and for the full data sets (see rows 4, 6, 7, 9 and corresponding dashed and dotted lines; the correlation CFR versus GDP for complete dataset is supported at confidence level 0.005; $F_c(1,94) = 8.33$).

 As expected, the vaccination and testing levels (VC and TC) always increase with rising incomes (see rows 10-15 and corresponding lines). Rows 16 and 18 represent the correlation of CC values versus TC and VC, respectively. The strongest link between the number of cases and the testing level (r=0.9496, and the highest $F/F_c(k_1,k_2)$ ratio, see row 16) and the strong link between TC and GDP values (see row 12) allows us to conclude that high CC values in rich countries are probably connected with the higher testing level. Numbers of cases and deaths per capita for complete dataset (Africa + Europe) increase with the increase of percentage of vaccinated people VC (see rows 18 and 20). Opposite trend was revealed only for CFR values (row 22). Thus, the positive effect of vaccinations is visible only in decreasing the probability to die for persons tested positive. To eliminate the influence of the testing level the same correlations were investigated for 15 European countries with TC>3. Corresponding CC, DC and CFR values demonstrate decreasing trend with the increase of VC, but it was not supported even at the significance level 0.05 (F_c (1,13)= 4.67; see rows 19, 21 and 23).

Table 1.

Optimal values of parameters in eq. (1), correlation coefficients and the results of Fisher test applications for Africa, Europe and complete datasets (Africa + Europe).

Figure 1: Characteristics of the COVID-19 pandemic in Africa (black) and Europe (blue) versus gross domestic product based on purchasing power parity GDP (PPP) per capita in international US dollars

The characteristics accumulated as of August 1, 2022 are: numbers of cases per million (CC, "circles"), numbers of deaths per million (DC, "crosses"), percentage of fully vaccinated people (VC, "dots"), numbers of test per capita (TC, "squares"). The case fatality risk was calculated with the use of formula CFR=DC/CC and shown by "triangles. Lines represent the best fitting relationships (1) with the optimal values of parameters listed Table 1: the black color corresponds to African countries, the blue one – to European, the red one – to complete datasets (Africa + Europe).

Discussion

The large difference between the number of registered and real COVID-19 cases [23-33] has to be taken into account to investigate the effects of different factors on the pandemic dynamics. In particular, different healthcare infrastructures, public health policies, and social behaviors could significantly change the pandemic dynamics and the analysis of these factors needs further investigations. Here we will focus on some specific influence of the testing rate. In particular, the TC values could approach some critical level, which allows revealing almost all COVID-19 cases. To check this hypothesis, let us consider the countries with TC>3. Their relatively large number - 16 (all countries are located in Europe) - allows drawing some statistical conclusions (see row 17). First, there is no correlation between CC and TC values even for significance level 0.1 (*Fc* (1, 14) =3.14). It means that 3 or more tests per person were enough to reveal the majority of cases before August 1, 2022. Its average value CC_a is approximately 460,834 and can be used to calculate the visibility coefficient

$$
\beta = \frac{CC_a}{CC} \tag{3}
$$

as the ratio of real to registered number of cases.

There are some theoretical and experimental estimations of the visibility coefficient for different periods of COVID-19 pandemic [10, 23-25]. For example, a total testing in Slovakia (89.5% of population was tested on October 31- November 7, 2020) revealed a number of previously undetected cases, equal to about 1.63% of the population [23, 24]. Taking into account that the number of detected cases in Slovakia was approximately 1% of population [20], we can estimate $\beta \approx 2.63$ for that period in Slovakia. As of August 1, 2022 the corresponding value CC=473,844 for Slovakia (see Table S2) is slightly larger than CC_a showing the good detection level in this country with TC=9.41.

A random testing in two kindergartens and two schools in the Ukrainian city of Chmelnytskii [25] revealed the value 3.9 in December 2020. Theoretical estimations based on the generalized SIR model [6, 10] yielded values from 3.7 to 20.4 for Ukraine and 5.4 for Qatar in different periods of the COVID-19 pandemic. As of August 1, 2022 formula (3) yields the *β* values 3.8 for Ukraine and 3.0 for Qatar (CC=152,375.8, [20]). Corresponding visibility coefficients are: 4.4 in Japan; 1.7 in US and 14.7 in India.

The value $CC_a = 460,834$ and formula (3) is probably not applicable for China and other Zero-COVID countries [34], where the total control and maximum suppression of the pandemic were applied. For example, mainland China has achieved the testing level TC=6.46 already on April 11, 2022, [20]. The value CC=636 registered on August 1, 2022 is much lower than the CC_a figure. Nevertheless, CC values in Australia, New Zealand and South Korea (where the zero tolerance policy was not as severe as in China) CC values vary from 317,619 to 384,572 (see [20]). The testing levels in Hong Kong (TC= 6.59 as of May 24, 2022, [20]) and in mainland China are very close. The huge difference in the registered numbers of cases per million (CC= 181,231 in Hong Kong as of August 1, 2022, [20]) probably is connected with much higher values of the tests per case ratio in mainland China, [14].

The lack of appropriate testing makes it especially difficult to detect the first cases of a new disease, which for SARS-CoV-2 probably appeared long before December 2019 [26]. In particular, theoretical estimates give the date of the appearance of the first case at the beginning of August 2019, [6].

The insufficient testing and high values of visibility coefficients can lead to controversial conclusions about the influence of vaccinations. For example, for complete datasets, unexpected upward trends for CC and DC values with the increasing VC were revealed at the very high significance level (see rows 18 and 20). Similar correlations were also found in [11] for average daily numbers of COVID-19 cases and deaths. In some countries (e.g., Israel, Japan, New Zealand), high vaccination levels did not prevent new severe pandemic waves [10, 14] with record numbers of cases and deaths, [10, 14]. Statistical studies support the fact that vaccinations diminished CFR, but their ability to reduce infections should be questioned [10, 11, 13, 14] and needs further investigation.

It would also be interesting to investigate the reasons for the increase in CC and DC values with the increase in incomes (see rows 3 and 6 in Table 1). One of them could be a lower mobility and less number of contacts in poor countries [18]. The age of population is another important factor in the visible COVID-19 pandemic dynamics [11, 35], since the percentage of asymptomatic (and unregistered) patients is much higher in children [27-30]. In particular, a one-year increment in the median year of population yields a 12-18 thousand increase in *CC* values and 52-83 increase in *DC* values (both figures correspond December 31, 2022), [35]. Taking into account the 24 year difference in the median age (18 in Africa and 42 in Europe, [36]) we can expect 288- 432 thousand higher CC figures and 1.2-2 thousand higher DC figures in Europe. The huge number of undetected COVID-19 cases increases the probability of the appearance of new dangerous SARS-CoV-2 variants.

Conclusions

Non-linear correlation analysis (using JHU datasets for Europe and Africa corresponding to August 1, 2022) demonstrated that the numbers of COVID-19 cases CC and deaths DC per capita and case fatality risks CFR=DC/CC increase for richer countries. The same trends were revealed for DC and CFR values in Africa, but opposite ones in Europe. As expected, the testing and vaccination levels increase with the growth of GDP. Higher levels of testing probably allowed revealing more cases and COVID-19 related deaths in rich countries. CC values showed a very strong increasing trend with the increase of numbers of tests per capita (TC). Unexpectedly, the same increasing trend was revealed for CC and DC values versus percentage of fully vaccinated people (VC). Nevertheless, the decrease of CFR with the increase of VC demonstrates a positive effect of vaccinations.

In some countries, the number of undetected COVID-19 cases may be tens or even hundreds of times higher than the number of registered ones due to the differences in testing levels and age structure. This fact increases the probability of the appearance of new dangerous SARS-CoV-2 strains and has to be taken into account in further investigations of impact of different factors on the pandemic dynamics.

Conflict of Interest

The authors declare no conflict of interests

Ethical Approval statement

 The study does not use any experiments with humans or animals. The data sources are available on the Internet.

Acknowledgments

The authors are grateful to Robin Thompson, Matt Keeling, and Paul Brown for their support and providing very useful information. Igor Nesteruk was supported by INI-LMS Solidarity Programme at the University of Warwick, UK.

References

- [1] Byass, P. Eco-epidemiological assessment of the COVID-19 epidemic in China, January-February 2020. medRxiv 2020, doi:10.1101/2020.03.29.20046565.
- [2] Hamzah, F.; Binti, A.; Lau, C.; Nazri, H.; Ligot, D.V.; Lee, G.; Tan, C.L. CoronaTracker: Worldwide COVID-19 Outbreak Data Analysis and Prediction. Bull. World Health Organ. 2020, 1, 32.
- [3] Fanelli, D.; Piazza, F. Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos Solitons Fractals 2020, 134, 109761.
- [4] Distante, C.; Piscitelli, P.; Miani, A. Covid-19 Outbreak Progression in Italian Regions: Approaching the Peak by the End of March in Northern Italy and First Week of April in Southern Italy. Int. J. Environ. Res. Public Health 2020, 17, 3025.
- [5] KY Ng, MM Gui. COVID-19: Development of a robust mathematical model and simulation package with consideration for ageing population and time delay for control action and resusceptibility. Physica D: Nonlinear Phenomena.Volume 411, October 2020, 132599. https://doi.org/10.1016/j.physd.2020.132599.
- [6] Nesteruk I. COVID19 pandemic dynamics. Springer Nature, 2021, https://link.springer.com/book/10.1007/978-981-33-6416-5.
- [7] Pardhan S,. Drydakis N. Associating the Change in New COVID-19 Cases to GDP per Capita in 38 European Countries in the First Wave of the Pandemic. Front Public Health. 2021 Jan 20;8:582140. doi: 10.3389/fpubh.2020.582140.
- [8] Ryan M Barber, Reed J D Sorensen, David M Pigott, Catherine Bisignano et al. Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov

14, 2021: a statistical analysis, The Lancet, Volume 399, Issue 10344, 2022, pp. 2351-2380, https://doi.org/10.1016/S0140-6736(22)00484-6.

- [9] Rossman H., Shilo S., Meir T., Gorfine M., Shalit U., Segal E. COVID-19 dynamics after a national immunization program in Israel. NATURE MEDICINE.VOL 27, June 2021, pp. 1055– 1061. https://doi.org/10.1038/s41591-021-01337-2.
- [10] Nesteruk I. Influence of Possible Natural and Artificial Collective Immunity on New COVID-19 Pandemic Waves in Ukraine and Israel. Explor Res Hypothesis Med 2021; Published online: Nov 11, 2021. doi: 10.14218/ERHM.2021.00044.
- [11] Nesteruk I. Trends of the COVID-19 dynamics in 2022 and 2023 vs. the population age, testing and vaccination levels. *Frontiers in Big Data*; 2024, 6:1355080. doi: 10.3389/fdata.2023.1355080
- [12] Alireza Mohammadi, Ievgen Meniailov, Kseniia Bazilevych, Sergey Yakovlev, Dmytro Chumachenko. Comparative study of linear regression and SIR models of COVID-19 propagation in Ukraine before vaccination. Radioelectronic and Computer Systems, 2021, Vol. 0, no. 3,pp. 5 – 18 . DOI https://doi.org/10.32620/reks.2021.3.01.
- [13] Nesteruk I. Should we ignore SARS-CoV-2 disease? *Epidemiology and Infection*. 2024;152:e57. doi:10.1017/S0950268824000487
- [14] Nesteruk I. Impact of vaccination and testing levels on the COVID-19 waves. Research Gate, Uploaded June 13, 2024. DOI: 10.13140/RG.2.2.12941.35047
- [15] Yamauchi T, Takeuchi S, Uchida M, Saito M, Kokaze A. The association between the dynamics of COVID-19, related measures, and daytime population in Tokyo. Scientific Reports, 23 Feb 2022, 12(1):3063. DOI: 10.1038/s41598-022-06716-4.
- [16] Chumachenko, D., Dudkina, T., Yakovlev, S., Chumachenko, T. Effective Utilization of Data for Predicting COVID-19 Dynamics: An Exploration through Machine Learning Models, *International Journal of Telemedicine and Applications*, 2023, 9962100, 24 pages, 2023. https://doi.org/10.1155/2023/9962100
- [17] Alali Y., Harrou F., and Sun Y., A proficient approach to forecast COVID-19 spread via optimized dynamic machine learning models, *Scientific Reports*. (2022) 12, no. 1, https://doi.org/10.1038/s41598-022-06218-3, 35165290.
- [18] Nesteruk I, Rodionov O. Coronavirus and incomes: the COVID-19 pandemic dynamics in Africa in February 2022. MedRxiv. Posted April 26, 2022. doi: https://doi.org/10.1101/2022.04.22.22274058.
- [19] International Monetary Fund. https://en.wikipedia.org/wiki/List_of_countries_by_GDP_(PPP)_per_capita.
- [20] COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). https://github.com/owid/covid-19 data/tree/master/public/data.
- [21] Draper NR, Smith H. Applied regression analysis. 3rd ed. John Wiley; 1998.
- [22] https://onlinepubs.trb.org/onlinepubs/nchrp/cd-22/manual/v2appendixc.pdf
- [23] https://edition.cnn.com/2020/11/02/europe/slovakia-mass-coronavirus-test-intl/index.html
- [24] https://www.voanews.com/covid-19-pandemic/slovakias-second-round-coronavirus-testsdraws-large-crowds.
- [25] An experiment with mass testing for COVID-19 was conducted in Khmelnytsky| Podillya News [Internet]. Podillya News | https://podillyanews.com/2020/12/17/u-shkolahhmelnytskogo-provely-eksperyment-z-testuvannyam-na-covid-19/
- [26] Weinberger DM et al. Estimating the early death toll of COVID-19 in the United States. [Preprint.] MEDRXIV, 2020 Apr. https://doi.org/10.1101/2020.04.15.2006643.
- [27] Statsenko Y. et al. Impact of Age and Sex on COVID-19 Severity Assessed From Radiologic and Clinical Findings. Front. Cell. Infect. Microbiol., 25 February 2022. Sec. Clinical Microbiology Volume 11 – 2021.| https://doi.org/10.3389/fcimb.2021.777070
- [28] Davies NG, Klepac P, Liu Y, Prem K, Jit M; CMMID COVID-19 working group; Eggo RM. Agedependent effects in the transmission and control of COVID-19 epidemics. Nat Med. 2020 Aug;26(8):1205-1211. doi: 10.1038/s41591-020-0962-9. Epub 2020 Jun 16. PMID: 32546824.
- [29] Fowlkes A.L., Yoon S.K., Lutrick K., Gwynn L., Burns J., Grant L., Phillips A., Ellingson K., Ferraris M., LeClair M., et al. Effectiveness of 2-Dose BNT162b2 (Pfizer BioNTech) mRNA Vaccine in Preventing SARS-CoV-2 Infection Among Children Aged 5-11 Years and Adolescents Aged 12–15 Years -PROTECT Cohort, July 2021–February 2022. MMWR Morb. Mortal. Wkly. Rep. 2022;71:422–428. doi: 10.15585/mmwr.mm7111e1.
- [30] Shang W, Kang L, Cao G, Wang Y, Gao P, Liu J, Liu M. Percentage of Asymptomatic Infections among SARS-CoV-2 Omicron Variant-Positive Individuals: A Systematic Review and Meta-Analysis. Vaccines (Basel). 2022 Jun 30;10(7):1049. doi: 10.3390/vaccines10071049. PMID: 35891214; PMCID: PMC9321237.
- [31] Schreiber, P.W., Scheier, T., Wolfensberger, A. *et al.* Parallel dynamics in the yield of universal SARS-CoV-2 admission screening and population incidence. *Sci Rep* **13**, 7296 (2023). https://doi.org/10.1038/s41598-023-33824-6
- [32] You Y, Yang X, Hung D, et al. Asymptomatic COVID-19 infection: diagnosis, transmission, population characteristics. BMJ Supportive & Palliative Care 2024;14:e220-e227.
- [33] Kronbichler A, Kresse D, Yoon S, Lee KH, Effenberger M, Shin JI. Asymptomatic patients as a source of COVID-19 infections: a systematic review and meta-analysis. *Int J Infect Dis*. 2020;98:180-186. doi:10.1016/j.ijid.2020.06.052
- [34] https://en.wikipedia.org/wiki/Zero-COVID
- [35] Igor Nesteruk, Matt Keeling. Population age as a key factor in the COVID-19 pandemic dynamics. Preprint. Research Square. Posted November 30, 2023. https://doi.org/10.21203/rs.3.rs-3682693/v1
- [36] https://www.visualcapitalist.com/mapped-the-median-age-of-every-continent/. Retrieved September 30, 2023.

Supplementary tables

Table S1. Gross domestic product per capita (GDP) based on purchasing power parity (PPP), accumulated and daily characteristics of the COVID-19 pandemic dynamics in African countries as of August 1, 2022 (figures corresponding to other days are specified in notes).

Figures corresponding to different days in 2022:
February: $1 - 13$; $2 - 18$; March: 3 -7; 4 -10; 5 -12; 6 -24;
May: 7 -1; 8 -5; 9 -18; 10 -19; 11 -20; 12 -20; 19 -22; 20 -23;
June: 13 -2; 14 -12; 15 -24; 24 -27; 25 -31;
July: 21 -3; August: $26 - 2$; $27 - 7$; $28 - 14$; $29 - 21$.

Table S2. Gross domestic product per capita (GDP) based on purchasing power parity (PPP), accumulated and daily characteristics of the COVID-19 pandemic dynamics in European countries as of August 1, 2022 (figures corresponding to other days are specified in notes)

Figures corresponding to different days: ¹ -December, 21, 2021; in 2022: January: ² -5; 3 -10; 4 -29; 5 -31; February: ⁶ -18; 7 -23; 8 -27; March: 9 -10; ¹⁰ -29; April: ¹¹ -14: May: $12 - 10$; $13 - 11$; $14 - 19$; $15 - 22$; June: ¹⁶ -1; 17 -2; 18 -11; 19 -12; 20 -16; 21 -17; 22 -18; 23 -19; 24 -20; 25 -21; 26 -22; 27 -23; July: 28 -5; 29 -11; 30 -26; 31 -28; 32 -29; 33 -31; August: 34 -2; 35 -3; 36 -11; 37 -21.

Fisher function

The experimental values of the Fisher function can be calculated with the use of the formula:

$$
F = \frac{r^2(n-m)}{(1-r^2)(m-1)}
$$
(S1)

where n is the number of observations (number of countries and regions taken for statistical analysis); m=2 is the number of parameters in the linear regression equation (2), [22].