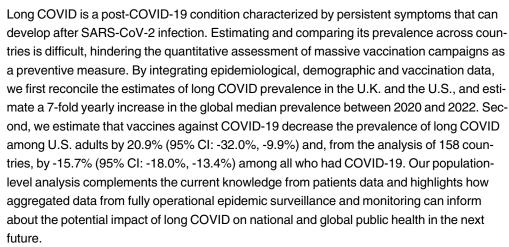


Prevalence of long COVID decreases for increasing COVID-19 vaccine uptake

Manlio De Domenico 1,2*

- 1 Department of Physics and Astronomy "Galileo Galilei", University of Padova, Padova, Italy, 2 Padua Center for Network Medicine, Padova, Italy
- * manlio.dedomenico@unipd.it







OPEN ACCESS

Citation: De Domenico M (2023) Prevalence of long COVID decreases for increasing COVID-19 vaccine uptake. PLOS Glob Public Health 3(6): e0001917. https://doi.org/10.1371/journal.pgph.0001917

Editor: Srinivasa Rao Mutheneni, CSIR-Indian Institute of Chemcial Technology, INDIA

Received: November 29, 2022 Accepted: April 24, 2023 Published: June 21, 2023

Copyright: © 2023 Manlio De Domenico. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All the raw data used in this study are publicly available from online repositories and published research papers, as detailed and fully referenced in Materials and methods. The data produced in this study are publicly available at this URL: https://osf.io/g69vp.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Introduction

There is increasing evidence that post COVID-19 sequelae might occur in individuals following SARS-CoV-2 infection. The definition of post COVID-19 condition by a Delphi consensus requires that symptoms should not find an alternative diagnosis, should persist 3 months after the onset of COVID-19 and should last for at least 2 months [1]. Long COVID, also known as Post-acute sequelae of COVID-19 (PASC), is characterized by a constellation of symptoms ranging from fatigue to shortness of breath and cardiac abnormalities [2–6], which are not limited to the respiratory system [7–9] but might include other systems too, reflecting the onset of metabolic, gastrointestinal, cardiovascular and neurocognitive disorders [10], with risk increasing with the severity of the acute COVID-19 infection [11].

The pathogenesis of long COVID is still unknown, making it difficult to fully characterize, anticipate, diagnose and treat this multisystem disease [12] and risk factors [13]. Consequently, it is difficult to measure its incidence and its prevalence at national and regional scales. Among possible causes, it has been hypothesized that fragments of the virus can persist for months in tissues driving chronic inflammation [14, 15]. Other mechanisms include the triggering of autoimmunity, the dysbiosis of microbiome or virome, and unrepaired tissue damage [16]. It

is possible that multiple mechanisms are at work, leading distinct long COVID outcomes: understanding attributes, predictors and risk factors of this disease [17–19] and developing biomarkers [20] might help to stratify patients for a more precise diagnosis and potential treatment (see [21] for the latest review on this topic).

While long COVID is becoming a disease of increasing concern [22–29], here we focus on at least two important questions relevant for policy making that remain to be answered. First, estimates of long COVID prevalence should be reconciled to allow for a better characterization of the underlying epidemiology: in fact, the heterogeneity of the patient cohorts does not allow for a direct comparison of distinct results, which depend on the presence (or absence) of matched control groups, the stratification by immunological status (due to vaccination protocols and natural infections), pre-existing conditions and comorbidities, age and severity of the acute phase of in the infection. The number of clusters corresponding to all combinations of the aforementioned features is so large that it is difficult, in practice, to study all of them. We refer to recent systematic reviews of existing studies for further details [30, 31].

Second, once the prevalence is measured, it is worth quantifying the impact of massive vaccination campaigns on long COVID. While a biological mechanism to support a direct relationship between COVID-19 vaccines and long COVID is still matter of investigation [21, 32], it is plausible to assume that vaccination might induce viral reservoir clearance, according to one of the mechanisms behind the pathogenesis. Recent studies from individual patients data suggest that long COVID outcomes are less likely in vaccinated groups [33, 34].

Understanding the post-viral symptoms of COVID-19 and quantifying the prevalence of long COVID will provide invaluable insights to better understand other existing post-acute infection syndromes [35] and will provide guidelines for policy in future pandemics [36].

Materials and methods

Overview of the data sets

The data that support the findings of this study are available from public repositories and/or published papers, without restrictions. A dedicated repository, with the data integrated and processed in this study, is publicly available.

Subgroups used for stratification. In the following, we will consider specific subgroups of the population to perform our stratified analysis in a geographic region (e.g., a U.S. state or a whole country). Specifically, we use the following labels:

- 18+: to indicate the subset of the whole population consisting of adults only
- 18+ who had COVID-19: to indicate the subset of the whole population consisting only of adults who had COVID-19 in the reference period
- Same age group: to indicate the subset of the population referring to a specific age group. The age brackets considered to compare results from the U.K. and the U.S. are 18–29, 30–39, 40–49,50–59, 60–69, 70–79 and 80–99. Another set of age brackets, resampled into narrower bins than the previous ones, consists of 0–5, 6–12, 13–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80–89 and 90–99.
- Same age group who had COVID-19: to indicate the subset of the population referring to a specific age group, who had COVID-19 in the reference period
- All who had COVID-19: to indicate the subset of the whole population who had COVID-19
 in the reference period
- All: to indicate the whole population

Long COVID prevalence in the U.K. The prevalence at national level and for distinct age groups is measured through the Coronavirus (COVID-19) Infection Survey, with respect to the whole population, by the Office for National Statistics (ONS) and made publicly available at this URL [37]. Data refers to self-reported long COVID after infection with the Omicron variant in the UK: we used the 04 August 2022 release and, specifically, the estimated percentage of people living in private households with self-reported long COVID of any duration (four week period ending 02 July 2022). The sample size is n = 221, 164 [38]. The data provides also lower and upper 95% confidence limits.

Long COVID prevalence in the U.S. The prevalence at state/national level and for distinct age groups is measured through the Household Pulse Survey, with respect to the distinct adult populations, by the Centers for Disease Control and Prevention (CDC) and made publicly available [39]. Data refers to responses to an internet questionnaire, and estimates are weighted to adjust for nonresponse and to match Census Bureau estimates of the population by age, sex, race and ethnicity. We used the reference period corresponding to 29 June—11 July 2022 (n = 57, 534). The data provides also lower and upper 95% confidence limits. The source reports that confidence intervals only reflect the potential for sampling error, whereas nonsampling errors (e.g., measurement, coverage, nonresponse and processing errors) can also occur and are more likely for this type of surveys.

Cumulative number of infections. The usually reported number of cases refers to confirmed cases. However, it is well known that not all infections are detected, e.g., because it is more difficult to find and test asymptomatic or paucisymptomatic individuals or because contact tracing capacity might be limited. Therefore, we relied on model estimates of the true number of cases over time and, consequently, its cumulative number. Specifically, we used data routinely published by the Institute for Health Metrics and Evaluation (IHME) and the "COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University" [40, 41]. For all countries considered in this study, we have used the IHME data hosted by Our World In Data (OWID), referring to the period 04 February 2020—06 June 2022 [42]. The mean estimate that we consider has been already used in several global public health contexts, e.g. to discuss in detail the weakness of the current global preparedness system [36], discuss policies for the next future [43] and estimate excess mortality [44]. Nevertheless, as any other model, it relies on assumptions and its estimates should be used with caution.

Since we use the IHME estimates, we cannot model counterfactual scenarios, for instance to show what would have happened in the absence of vaccination.

Long COVID prevalence estimates at global level. Estimates by country of the number of patients affected by long COVID in 2020 and 2021, regardless of the severity of their symptoms, are obtained from [45]. The data is are based on information for 1,906 community infections and 10,526 hospitalized patients from ten collaborating cohorts, three of which included children, as well as ICD-coded medical record data concerning 1.3 million infections.

Vaccination uptake. For the correlation analysis in the U.S. states we have used the data from CDC [46] and hosted by OWID [47]. Specifically we have considered the share of people in each state who received at least one dose of a COVID-19 as of 15 Aug 2022, and the share of people in the state who completed the initial COVID-19 vaccination protocol as of 15 Aug 2022. Note that there can be differences from country to country about the definition of a fully vaccinated individual, depending on the fact that an individual has completed the initial vaccination protocol followed by boosters or not.

For consistency with the data about the long COVID prevalence at global level, and the corresponding analysis, we have used the share of people who completed the initial COVID-19 vaccination protocol by 29 May 2022, provided by OWID, which should reduce any potential bias.

Nevertheless, it should be noted that vaccination uptake, alone, can only be considered as a proxy for population immunity, since natural infection contributes to develop some level of immunity over a limited temporal window.

Resampling age-stratified data

To overcome some of the limitations of individual, microscopic, analysis due to heterogeneity of cohorts, we use a population-based approach typical of descriptive epidemiology, with a large-scale, macroscopic, analysis to reconcile the estimates of long COVID prevalence across distinct countries and quantify the effectiveness of vaccination to reduce it. Our approach has the main advantage to integrate over all the intervening and characterizing factors of long COVID mentioned above, thus providing a population-level picture that can be used for comparisons across countries.

To map the share of a population or the count (e.g., of confirmed cases) provided for a given age group into different age group we need information about the original reference population, the 1-year age pyramid of the country. For the U.K. and the U.S. we used age pyramids provided for 2022 by the World Population Prospects of the United Nations [48].

For the U.S. total population at the state level, we have used the 2019 Census estimates taken by the United States Census Bureau [49]. Here, we have rescaled the population of each state by the ratio between the U.S. total population in 2022 and the corresponding U.S. state total population in 2019, to have a consistent total population for the subsequent analysis.

For the age-stratified population at the U.S. state level, we have combined the above data with the share of population per distinct age group provided, for 2020, by Kaiser Family Foundation estimates based on the Census Bureau's March Current Population Survey (CPS: Annual Social and Economic Supplements), 2017-2021.

Indicating by $r_{A,B|P}$ the share of a reference population P for ages in the group [A, B], we estimate the population count n_k for age $k \in [A, B]$ by

$$n_k = r_{A,B|P} P \frac{P_k}{\frac{P_k}{B}},$$

$$\sum_{\ell=A} P_{\ell}$$
(1)

and the corresponding share of population by $r_k = n_k/P_k$, where P_k indicates the population count for age k. Given a distinct age group, e.g., indicated by [A', B'], and another reference population P', the corresponding rate is obtained by

$$r_{A',B'|P'} = \frac{\sum_{k=A'}^{B'} n_k}{P'} = r_{A,B|P} \frac{P}{P'} \frac{\sum_{k=A'}^{B'} P_k}{\sum_{\ell=A}^{B} P_\ell}.$$
 (2)

Regression models and odds ratio

Since the errors on the estimate are not available, in our linear regression models we account for the presence of confidence intervals on the estimates via weights. Being intervals asymmetric, for each point estimate x_i , with bounds $[l_i, u_i]$, we calculate $[(l_i - x_i)^2 + (u_i - x_i)^2]^{-\frac{1}{2}}$, thus giving more importance to more precise estimates.

Given the prevalence of long COVID estimated, with the aforementioned linear model, for the extremal cases of 0% and 100% vaccine uptake, and indicating with $b \pm \epsilon_b$ the estimate on

the intercept and with $a \pm \epsilon_a$ the estimate on the slope, all measured in percentages, the odds ratio is given by

$$OR = \frac{b/(100 - b)}{(a+b)/(100 - a - b)}. (3)$$

Visualizing results with density-equalizing maps

Prevalence data are usually visualized via maps which, however, are known to lead to potential misinterpretation of the burden of a disease when sparsely populated large areas dominate densely populated smaller areas. A solution to this problem is provided by density-equalizing maps, also known as cartograms, which reduce the impact of known limitations by applying nonlinear transformations to the political areas (e.g., states in the case of the U.S. or countries in world maps) in order to preserve the proportionality to the variabile of interest. Here, we use the method, based on diffusion, proposed by Gastner and Newman [50].

Results

We computed the age-stratified prevalence with respect to the same reference populations—the adults, the adults who had confirmed COVID-19 and the whole population of the two countries (Methods)—using data from two national surveys in the U.K. and the U.S., provided by the Office for National Statistics (ONS) and the Centers for Disease Control and Prevention (CDC), respectively.

In Fig 1A we show the comparison between the long COVID prevalence in the two countries by using age groups according to CDC, showing that it is systematically higher in the U. S., regardless of the reference population and the age group. For a meaningful comparison, it is of particular interest the population of adults who had COVID-19, where the difference in prevalence is less marked in those aged 18-29 and 30-39. Focusing on the UK only, Fig 1B shows the prevalence for the age groups as provided by the ONS, that we have estimated with respect to distinct reference populations, including the same age group. The original data is provided with respect to the whole population: for the populations related to COVID-19 cases we had to reconstruct the cumulative number of confirmed cases in each age group, and correct it for the estimate of the total number of cases included undetected ones (Methods). People aged 12-16, 17-24, 25-34 and 35-49, show comparable prevalence, stable and close to 1 in 5, with respect to their age group and to who had COVID-19 within their age group. The prevalence is much higher in older age groups. To gain more insights, we have resampled the estimates into narrower age groups: we find that the prevalence reaches 1 in 3 for people aged 60-69 and 1 in 2 for people aged 70-79. For children aged 0-5 and 6-12 the estimated prevalence is about 1 in 10.

The CDC data also provides information for each U.S. state, although it is not stratified by age. Considering each state as an independent geographic area, we have compared their distribution of the estimated prevalence with respect to distinct reference populations and against the value measured for the U.K. and the Netherlands [51], the latter providing only a mean estimate. Fig 2A shows that the distribution is rather homogeneous, with well identified mean and dispersion, and in good agreement with the estimates from the two European countries. We argue that the signal might be even stronger if measured from the vaccinated population of adults only, although the lack of publicly available data prevented us from further investigations. Fig 2B shows the geographic prevalence of long COVID stratified by age group.

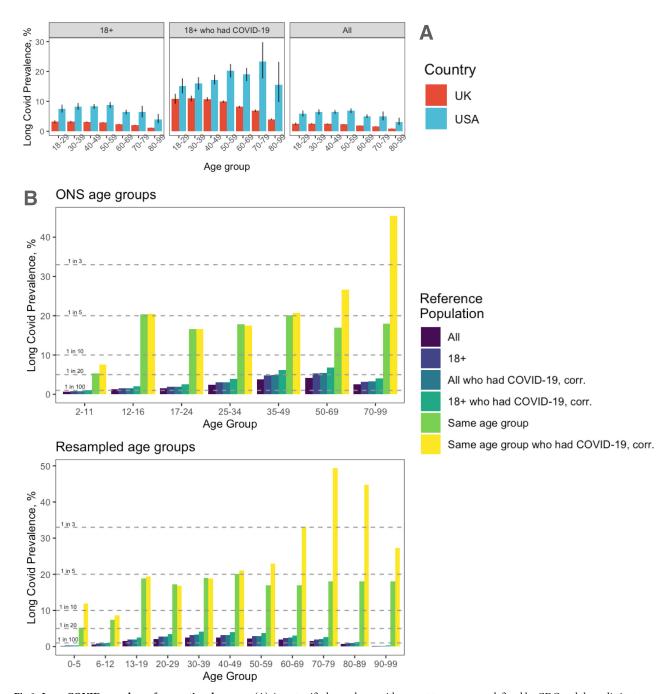


Fig 1. Long COVID prevalence from national surveys. (A) Age-stratified prevalence with respect to age groups defined by CDC and three distinct reference populations (adults, adults who had COVID-19 and the whole population), for the U.K. and the U.S. (B) Age-stratified prevalence in the U.K. with respect to age groups defined by ONS (top) and resampled into narrower bins (bottom). More reference populations are considered and, for the ones involving people who had COVID-19, we corrected the confirmed number of cases by a factor accounting for undetected ones (Methods).

Stable and consistent estimates of the prevalence allowed us to analyze the impact of COVID-19 vaccination campaigns on long COVID. Fig 3A shows the prevalence against the percentage of the population participating to the vaccination campaign, in each state separately. Fig 3B and 3C shows density-equalizing maps highlighting the geographic distribution of the vaccine uptake and the long COVID prevalence.

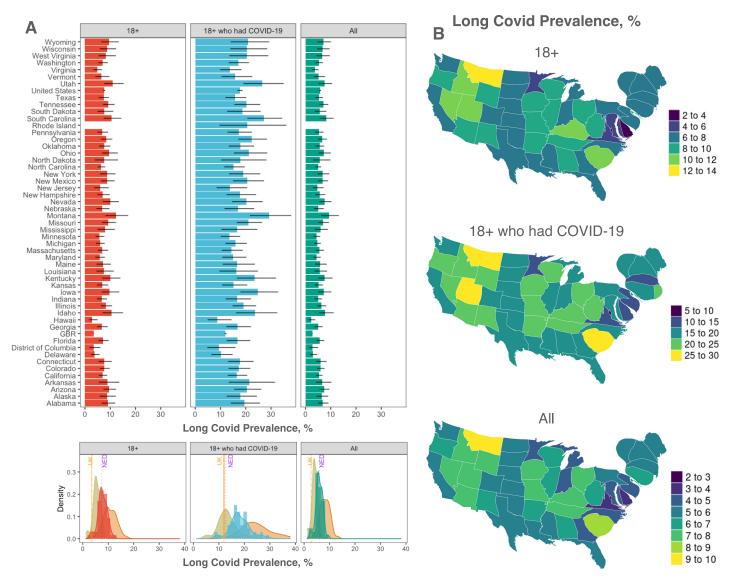


Fig 2. Long COVID prevalence from national surveys. (A) Long COVID prevalence in each U.S. state, the U.S. and the U.K. (labeled as GBR) for the same reference populations in panel A (top) and the distribution of their mean, color-coded consistently with the top panels, and the distribution of the corresponding extremal values—i.e., the 95% lower and upper bounds—(middle). For comparison, we have reported the estimates from the U.K. and the Netherlands [51] as vertical lines. (B) Geographic distribution of the prevalence in the corresponding reference populations by means of density-equalizing maps (Methods). Maps are produced with the software R, package maps, based on Natural Earth free vector and raster map data.

We considered two indicators: the population vaccinated with at least one dose and the population who completed the initial vaccination protocol consisting of two doses. Let us focus on the adult population who had COVID-19, which is the one where the effects of long COVID are less smeared out. First, we measure a decreasing trend, highlighting that also from correlation analyses at population level it is possible to measure the signal of the beneficial effects of vaccines. Second, the result is stronger for the population who received two doses: the slope is -20.9% \pm 5.5%, suggesting that the difference between a population covered at 100% by the initial vaccination protocol and a population not covered at all leads to a difference of about 21% in developing long covid after an infection. Specifically, the rate for the

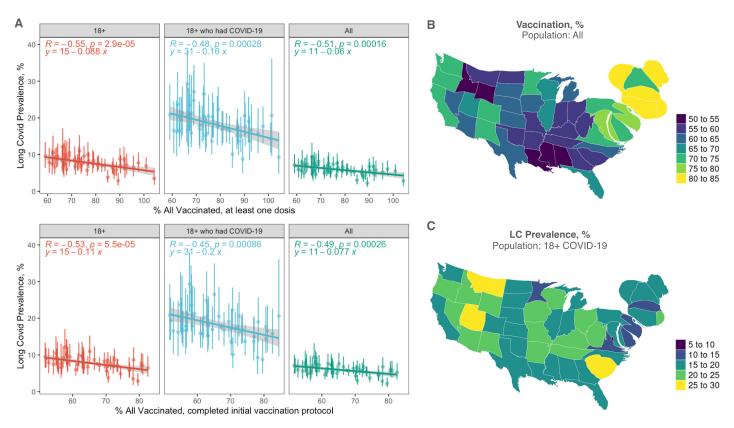


Fig 3. Long COVID prevalence decreases with vaccine uptake in the U.S. (A) Prevalence in U.S. states and the U.S. exhibits a decreasing trend with respect to vaccine uptake, both in the population vaccinated with at least one dose (top) and two doses (bottom), with the largest gap between 100% vaccinated and 100% unvaccinated scenarios observed in the reference population of adults who had COVID-19. The measured correlations are significant ($p < 10^{-3}$) in all cases and highlights that the larger the vaccine uptake the lower the prevalence of long covid, with the effect further increasing with more robust vaccination protocol. (B) Global density-equalizing maps highlight the geographic distribution of the vaccine uptake. (C) As in B, for the long COVID prevalence. Maps are produced with the software R, package maps, based on Natural Earth free vector and raster map data.

former is 10.7% (95% CI: 0%-24.0%) and the rate for the latter is 31.6% (95% CI: 24.2%-39.1%). The odds ratio between the unvaccinated and vaccinated scenarios is 3.9 (95% CI: 2.7-5.4).

We have performed the same analysis at a global level with data from 158 countries, using recent estimates of the prevalence in 2021 [45] and global vaccine uptake rates with respect to the whole population. Fig 4A, where information is further stratified by WHO region and income group, shows patterns similar to the ones observed for the U.S. states, with remarkable differences. Fig 4B and 4C shows density-equalizing maps highlighting the geographic distribution of the vaccine uptake and the long COVID prevalence.

In Africa and in the Eastern Mediterranean regions the correlation signal is not significant due to the low vaccine uptake and, most likely, to the vaccine equity crisis [52–54]. At a global level, the reduction of prevalence is -15.7% (95% CI: -18.0%, -13.4%), which is compatible with the estimate obtained from US for the prevalence among adults only -20.9% (95% CI: -32.0%, -9.9%).

Finally, we use national and global estimates to gain additional insights about the yearly prevalence of long COVID with respect to the whole population. The global median in 2020 was 0.06% (95% CI: 0.02%, 0.13%), while the global median in 2021 was 0.4% (95% CI: 0.1%, 1.0%), corresponding to a 7 fold increase. The estimate for the U.S. in the first half of 2022 is

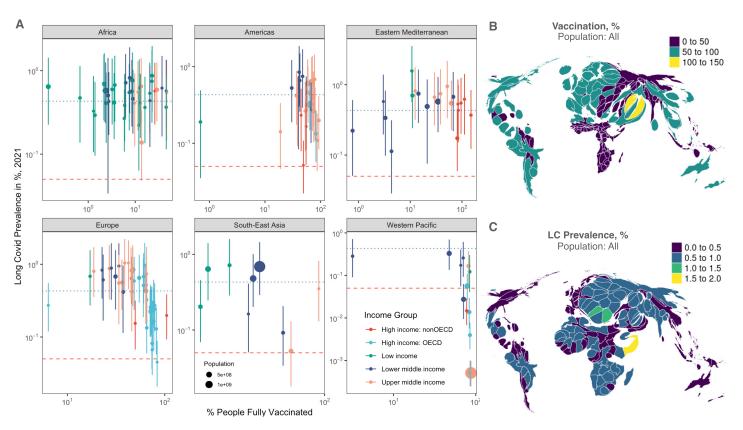


Fig 4. Long COVID prevalence decreases with vaccine uptake, globally. (A) Analysis similar to the one in Fig 4 performed at a global level, using long COVID prevalence estimates for 2021 in the whole population, stratified by WHO region and income group. Dotted blue lines indicate the global median, whereas the dashed red lines indicate the 1 in 2000 limit, used to define a rare disease. (B) Global density-equalizing maps highlight the geographic distribution of the vaccine uptake. (C) As in B, for the long COVID prevalence. Maps are produced with the software R, package maps, based on Natural Earth free vector and raster map data.

5.8% (95% CI: 5.5%, 6.1%), a 13 fold increase from the 2021 global median. The estimate for the U.K. is 2.8% (95% CI: 2.7%, 2.8%), a 6 fold increase from the 2021 global median.

Discussion

Estimating the prevalence of long COVID, the pathological condition that might follow a COVID-19 infection, from ecological data is a difficult task because of the (i) lack of age-stratified data at country level and, overall, (ii) the lack of data for the vast majority of countries. The task is made further difficult by the different reference populations (e.g., all adults above 18 years, adults who had COVID-19, whole population, so forth and so on), which often make difficult to compare the available estimates even across countries releasing the corresponding data. Furthermore, another pressing question highly relevant for global public health concerns the impact of COVID-19 vaccination to reduce the risks of developing long COVID and, consequently to decrease its prevalence at population level.

Concerning the first goal of this study, we used publicly available population data and statistical techniques, we have derived consistent estimates at national and global scales that allowed us to compare, in a meaningful way, the long COVID prevalence across distinct reference populations and age groups. We identified consistent results, with the prevalence in the U.S. being larger than the prevalence in the U.K. Across age groups, long COVID is currently affecting

about 1 in 10 children aged 0-12 and 1 in 5 people aged 13-59, with at least a 2 fold prevalence among those ones aged 60-89.

Concerning the second goal of this study, we used our estimates of the prevalence to report a decreasing trend with respect to increasing vaccine uptake, with a reduction of long COVID between 15% and 20%, compatible with recent reports from individual patients data. These results, in the U.S. and global, are in good agreement with estimates from cohorts of individual patients. A recent study, using the US Department of Veterans Affairs national healthcare databases to build a cohort of 33,940 individuals, has estimated a reduction of the risk of long COVID due to vaccinations of about 15%, by analyzing information from 13 million people including contemporary (n = 4, 983, 491), historical (n = 5, 785, 273) and vaccinated (n = 2, 566, 369) controls [33]. Another study, focusing on 9 Italian health care facilities and 2,560 participants, reported that the number of vaccine doses was associated with lower long COVID prevalence: 41.8% (95% CI, 37.0%–46.7%) in unvaccinated patients, 30.0% (95% CI, 6.7%–65.2%) with 1 dose, 17.4% (95% CI, 7.8%–31.4%) with 2 doses, and 16.0% (95% CI, 11.8%–21.0%) with 3 doses [34]. Note that the difference between the case with 3 doses and the unvaccinated rates is about -25.8%, again in good agreement with our estimate from U.S. data.

While results are encouraging, since they provide some evidence that vaccinations are associated with lower prevalence of long COVID, it would be desirable to understand which biological mechanism(s) could support causal relationships. According to the viral remnant hypothesis of long COVID, one of the mechanisms behind the pathogenesis, vaccines could be able to accelerate clearance of SARS-CoV-2 in the human body. Instead, according to another mechanism for pathogenesis, namely the immune/inflammatory hypothesis, vaccines could reduce the amplified immune/inflammatory response to the virus. It is worth remarking that it is possible that multiple mechanisms are at work, although the effective ones are still unknown [31] and require further investigation.

On the one hand, the main advantage of our population-based study is that it naturally integrates over all potential factors (e.g., immunological status, pre-existing conditions, so forth and so on, see [31]) affecting existing estimates, thus providing an effective way to evaluate the effects of vaccines at national level, from an ecological perspective. On the other hand, a limitation of our approach is that it does not allow to quantify the impact of each specific factor contributing to long COVID, as it can be done from individual, microscopic, information. Consequently, there are specific advantages in both ecological and cohort-based analysis, which should be employed together to: (i) characterize the mechanisms leading to post-COVID-19 conditions and develop tailored pharmaceutical therapies; (ii) monitor the prevalence of those conditions at scale, across distinct reference populations, in order to quantify the impact on national health systems and, accordingly, devise adequate public health policies. There are some limitations worth discussing in this context. For instance, the vaccination data does not perfectly reflect the immunity of the population, since infections and breakthrough infections might play a role that cannot be disentangled at an ecological level. Another limitation concerns the use of the IHME estimates, preventing us from modeling counterfactual scenarios—e.g., to show what would have happened in the absence of vaccination—or to perform a sensitivity analysis on its parameters. Therefore, although it is encouraging that our population-level analysis provides estimates that are in good agreement with individual-based results, where it is easier to control for intervening factors, our results should be considered with the adequate caution and cannot prescind from further validation with different models and nonecological analyses.

Together, our findings suggest that long COVID prevalence and its positive, but not exceptional, response to vaccination have the potential to stress national health systems. The

persistent emergence of new SARS-CoV-2 variants such as the sub-lineage BQ.1 in Nigeria (July 2022) and XBB.1.5 in the U.S. (October 2022), followed by localized surges in COVID-19 cases that spread globally within a few months [55, 56], poses important challenges to global public health in the next future. In fact, to characterize the impact on long COVID prevalence of a new SARS-CoV-2 variant it is necessary to wait several months: an amount of time that could have dramatic public health effects in case of sustained community transmission. Consequently, decreasing the risk of infection by reducing community transmission might be an effective approach to control the long COVID prevalence, since it allows to buy the time required to (i) collect new data about the risk of developing post-COVID-19 conditions due to the emerging variant and (ii) further develop pharmaceutical therapies, from antivirals to more effective vaccines. To this aim, data-driven models [55, 57-66] can be used to monitor and control the transmission at national level, as well as to devise effective pro-active measures [67]. For instance, monitoring and controlling air quality [68] can reduce the risk of airborne transmission for any indoor environment [69], but requires clear targets on air changes per hour and air filtration to become an effective public health intervention [70, 71]. Keeping a low infodemic level—i.e., maintaining the circulation of reliable information about the multifaceted aspects of the ongoing pandemic [72-74]—with periodic and adequate communication campaigns, might enhance compliance with public health guidelines [75]. Additionally, coordination among countries is needed to enhance global preparedness [43, 76, 77] to emerging SARS-CoV-2 variants, as well as to novel pathogens with pandemic potential, as the ones due to zoonotic spillover induced by climate change, that according to a recent study, are not going to be unlikely [78].

Complementing vaccination campaigns with potential therapies, as well as further investing on prevention, might reduce the prevalence of long COVID and, consequently, should be considered for public health policies in the next future.

Acknowledgments

M.D.D. thanks V. d'Andrea and S. Campioni for interesting discussions. The author is also grateful to the organizations, institutions and researchers providing publicly available or sharing data during the COVID-19 pandemic.

Author Contributions

Conceptualization: Manlio De Domenico.

Data curation: Manlio De Domenico.

Formal analysis: Manlio De Domenico.

Investigation: Manlio De Domenico.

Methodology: Manlio De Domenico.

Supervision: Manlio De Domenico. **Validation:** Manlio De Domenico.

Visualization: Manlio De Domenico.

Violatization Manno De Domenico.

Writing - original draft: Manlio De Domenico.

Writing - review & editing: Manlio De Domenico.

References

- World Health Organization et al. A clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021. In: A clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021; 2021.
- Crook H, Raza S, Nowell J, Young M, Edison P. Long covid-mechanisms, risk factors, and management. BMJ (Clinical Research ed). 2021; 374:n1648–n1648. PMID: 34312178
- Michelen M, Manoharan L, Elkheir N, Cheng V, Dagens A, Hastie C, et al. Characterising long COVID: a living systematic review. BMJ Global Health. 2021; 6(9):e005427. https://doi.org/10.1136/bmjgh-2021-005427 PMID: 34580069
- Aiyegbusi OL, Hughes SE, Turner G, Rivera SC, McMullan C, Chandan JS, et al. Symptoms, complications and management of long COVID: a review. Journal of the Royal Society of Medicine. 2021; 114 (9):428–442. https://doi.org/10.1177/01410768211032850 PMID: 34265229
- Davis HE, Assaf GS, McCorkell L, Wei H, Low RJ, Re'em Y, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. EClinicalMedicine. 2021; 38:101019. https://doi.org/10.1016/j.eclinm.2021.101019 PMID: 34308300
- Hampshire A, Trender W, Chamberlain SR, Jolly AE, Grant JE, Patrick F, et al. Cognitive deficits in people who have recovered from COVID-19. EClinicalMedicine. 2021; 39:101044. https://doi.org/10.1016/j.eclinm.2021.101044 PMID: 34316551
- Tran VT, Porcher R, Pane I, Ravaud P. Course of post COVID-19 disease symptoms over time in the ComPaRe long COVID prospective e-cohort. Nature Communications. 2022; 13(1):1–6. https://doi.org/ 10.1038/s41467-022-29513-z PMID: 35383197
- Kedor C, Freitag H, Meyer-Arndt L, Wittke K, Hanitsch L, Zoller T, et al. A prospective observational study of post-COVID-19 chronic fatigue syndrome following the first pandemic wave in Germany and biomarkers associated with symptom severity. Nature Communications. 2022; 13(1):1–13. https://doi. org/10.1038/s41467-022-32507-6
- Whitaker M, Elliott J, Chadeau-Hyam M, Riley S, Darzi A, Cooke G, et al. Persistent COVID-19 symptoms in a community study of 606,434 people in England. Nature communications. 2022; 13(1):1–10. https://doi.org/10.1038/s41467-022-29521-z
- of Disease Long COVID Collaborators GB. Estimated Global Proportions of Individuals With Persistent Fatigue, Cognitive, and Respiratory Symptom Clusters Following Symptomatic COVID-19 in 2020 and 2021. JAMA. 2022; 328(16):1604–1615. https://doi.org/10.1001/jama.2022.18931
- Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19.
 Nature. 2021; 594(7862):259–264. https://doi.org/10.1038/s41586-021-03553-9 PMID: 33887749
- Ayoubkhani D, Khunti K, Nafilyan V, Maddox T, Humberstone B, Diamond I, et al. Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study. BMJ (Clinical research ed). 2021; 372:n693–n693. https://doi.org/10.1136/bmj.n693 PMID: 33789877
- Thompson EJ, Williams DM, Walker AJ, Mitchell RE, Niedzwiedz CL, Yang TC, et al. Long COVID burden and risk factors in 10 UK longitudinal studies and electronic health records. Nature communications. 2022; 13(1):1–11. https://doi.org/10.1038/s41467-022-30836-0 PMID: 35764621
- Gaebler C, Wang Z, Lorenzi JC, Muecksch F, Finkin S, Tokuyama M, et al. Evolution of antibody immunity to SARS-CoV-2. Nature. 2021; 591(7851):639–644. https://doi.org/10.1038/s41586-021-03207-w PMID: 33461210
- Natarajan A, Zlitni S, Brooks EF, Vance SE, Dahlen A, Hedlin H, et al. Gastrointestinal symptoms and fecal shedding of SARS-CoV-2 RNA suggest prolonged gastrointestinal infection. Med (N Y). 2022; 3 (6):371–387.e9. https://doi.org/10.1016/j.medj.2022.04.001 PMID: 35434682
- Merad M, Blish CA, Sallusto F, Iwasaki A. The immunology and immunopathology of COVID-19. Science. 2022; 375(6585):1122–1127. https://doi.org/10.1126/science.abm8108 PMID: 35271343
- Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer RC, et al. Attributes and predictors of long COVID. Nature medicine. 2021; 27(4):626–631. https://doi.org/10.1038/s41591-021-01292-y PMID: 33692530
- Su Y, Yuan D, Chen DG, Ng RH, Wang K, Choi J, et al. Multiple early factors anticipate post-acute COVID-19 sequelae. Cell. 2022; 185(5):881–895. https://doi.org/10.1016/j.cell.2022.01.014 PMID: 35216672
- Subramanian A, Nirantharakumar K, Hughes S, Myles P, Williams T, Gokhale KM, et al. Symptoms and risk factors for long COVID in non-hospitalized adults. Nature Medicine. 2022; p. 1–9. https://doi.org/10. 1038/s41591-022-01909-w PMID: 35879616
- 20. Klein J, Wood J, Jaycox J, Lu P, Dhodapkar RM, Gehlhausen JR, et al. Distinguishing features of Long COVID identified through immune profiling. medRxiv. 2022. https://doi.org/10.1101/2022.08.09. 22278592 PMID: 35982667

- Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. Nature Reviews Microbiology. 2023; p. 1–14.
- Venkatesan P. NICE guideline on long COVID. The Lancet Respiratory Medicine. 2021; 9(2):129. https://doi.org/10.1016/S2213-2600(21)00031-X PMID: 33453162
- 23. Thompson RC, Simons NW, Wilkins L, Cheng E, Del Valle DM, Hoffman GE, et al. Molecular states during acute COVID-19 reveal distinct etiologies of long-term sequelae. Nature Medicine. 2022; p. 1–11. https://doi.org/10.1038/s41591-022-02107-4 PMID: 36482101
- Zhang H, Zang C, Xu Z, Zhang Y, Xu J, Bian J, et al. Data-driven identification of post-acute SARS-CoV-2 infection subphenotypes. Nature Medicine. 2022; p. 1–10. https://doi.org/10.1038/s41591-022-02116-3 PMID: 36456834
- Bowe B, Xie Y, Al-Aly Z. Acute and postacute sequelae associated with SARS-CoV-2 reinfection. Nature medicine. 2022; 28(11):2398–2405. https://doi.org/10.1038/s41591-022-02051-3 PMID: 36357676
- Xu E, Xie Y, Al-Aly Z. Long-term neurologic outcomes of COVID-19. Nature medicine. 2022; 28 (11):2406–2415. https://doi.org/10.1038/s41591-022-02001-z PMID: 36138154
- Puntmann VO, Martin S, Shchendrygina A, Hoffmann J, Ka MM, Giokoglu E, et al. Long-term cardiac pathology in individuals with mild initial COVID-19 illness. Nature Medicine. 2022; 28(10):2117–2123. https://doi.org/10.1038/s41591-022-02000-0 PMID: 36064600
- 28. Montani D, Savale L, Noel N, Meyrignac O, Colle R, Gasnier M, et al. Post-acute COVID-19 syndrome. European Respiratory Review. 2022; 31(163). https://doi.org/10.1183/16000617.0185-2021 PMID: 35264409
- 29. Routen A, O'Mahoney L, Ayoubkhani D, Banerjee A, Brightling C, Calvert M, et al. Understanding and tracking the impact of long COVID in the United Kingdom. Nature medicine. 2022; 28(1):11–15. https://doi.org/10.1038/s41591-021-01591-4 PMID: 34811549
- Iqbal FM, Lam K, Sounderajah V, Clarke JM, Ashrafian H, Darzi A. Characteristics and predictors of acute and chronic post-COVID syndrome: A systematic review and meta-analysis. EClinicalMedicine. 2021; 36:100899. https://doi.org/10.1016/j.eclinm.2021.100899 PMID: 34036253
- Notarte KI, Catahay JA, Velasco JV, Pastrana A, Ver AT, Pangilinan FC, et al. Impact of COVID-19 vaccination on the risk of developing long-COVID and on existing long-COVID symptoms: A systematic review. EClinicalMedicine. 2022; 53(101624):101624. https://doi.org/10.1016/j.eclinm.2022.101624
 PMID: 36051247
- Ayoubkhani D, Bermingham C, Pouwels KB, Glickman M, Nafilyan V, Zaccardi F, et al. Trajectory of long covid symptoms after covid-19 vaccination: community based cohort study. bmj. 2022; 377. https://doi.org/10.1136/bmj-2021-069676 PMID: 35584816
- Al-Aly Z, Bowe B, Xie Y. Long COVID after breakthrough SARS-CoV-2 infection. Nat Med. 2022; 28 (7):1461–1467. https://doi.org/10.1038/s41591-022-01840-0 PMID: 35614233
- Azzolini E, Levi R, Sarti R, Pozzi C, Mollura M, Mantovani A, et al. Association between BNT162b2 vaccination and long COVID after infections not requiring hospitalization in health care workers. JAMA. 2022; 328(7):676–678. https://doi.org/10.1001/jama.2022.11691 PMID: 35796131
- Choutka J, Jansari V, Hornig M, Iwasaki A. Unexplained post-acute infection syndromes. Nature Medicine. 2022; 28(5):911–923. https://doi.org/10.1038/s41591-022-01810-6 PMID: 35585196
- Sachs JD, Karim SSA, Aknin L, Allen J, Brosbøl K, Colombo F, et al. The Lancet Commission on lessons for the future from the COVID-19 pandemic. The Lancet. 2022; 400(10359):1224–1280. https://doi.org/10.1016/S0140-6736(22)01585-9 PMID: 36115368
- 37. Office for National Statistics. Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK; 2022. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/health andsocialcare/conditionsanddiseases/datasets/alldatarelatingtoprevalenceofongoingsymptoms followingcoronaviruscovid19infectionintheuk.
- Office for National Statistics. Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK: 4 August 2022; 2022. Available from: https://www.ons.gov.uk/peoplepopulationand community/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptoms followingcoronaviruscovid19infectionintheuk/4august2022.
- Long COVID—Household Pulse Survey—COVID-19—cdc.gov;. https://www.cdc.gov/nchs/covid19/ pulse/long-covid.htm.
- GitHub—CSSEGISandData/COVID-19: Novel Coronavirus (COVID-19) Cases, provided by JHU CSSE—github.com;. https://github.com/CSSEGISandData/COVID-19.
- Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020; 20(5):533–534. https://doi.org/10.1016/S1473-3099(20)30120-1 PMID: 32087114

- Daily new estimated COVID-19 infections from the IHME model—ourworldindata.org;. https://ourworldindata.org/grapher/daily-new-estimated-covid-19-infections-ihme-model.
- Lazarus JV, Romero D, Kopka CJ, Karim SA, Abu-Raddad LJ, Almeida G, et al. A multinational Delphi consensus to end the COVID-19 public health threat. Nature. 2022; 611(7935):332–345. https://doi.org/10.1038/s41586-022-05398-2 PMID: 36329272
- 44. Wang H, Paulson KR, Pease SA, Watson S, Comfort H, Zheng P, et al. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. The Lancet. 2022; 399(10334):1513–1536. https://doi.org/10.1016/S0140-6736(21)02796-3
- 45. Hanson SW, Abbafati C, Aerts JG, Al-Aly Z, Ashbaugh C, Ballouz T, et al. A global systematic analysis of the occurrence, severity, and recovery pattern of long COVID in 2020 and 2021. medRxiv. 2022;. https://doi.org/10.1101/2022.05.26.22275532 PMID: 35664995
- COVID-19 Vaccinations in the United States—cdc.gov;. https://covid.cdc.gov/covid-data-tracker/ #vaccinations.
- Coronavirus (COVID-19) Vaccinations—ourworldindata.org;. https://ourworldindata.org/covid-vaccinations.
- **48.** World Population Prospects—Population Division—United Nations—population.un.org;. https://population.un.org/wpp/.
- 49. Bureau UC. Census.gov—census.gov;. https://www.census.gov/.
- Gastner MT, Newman ME. Diffusion-based method for producing density-equalizing maps. Proceedings of the National Academy of Sciences. 2004; 101(20):7499–7504. https://doi.org/10.1073/pnas.0400280101 PMID: 15136719
- 51. Ballering AV, van Zon SKR, Olde Hartman TC, Rosmalen JGM, Lifelines Corona Research Initiative. Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. Lancet. 2022; 400(10350):452–461. https://doi.org/10.1016/S0140-6736(22)01214-4 PMID: 35934007
- The Lancet Infectious Diseases. COVID-19 vaccine equity and booster doses. Lancet Infect Dis. 2021; 21(9):1193. https://doi.org/10.1016/S1473-3099(21)00486-2
- McIntyre PB, Aggarwal R, Jani I, Jawad J, Kochhar S, MacDonald N, et al. COVID-19 vaccine strategies must focus on severe disease and global equity. The Lancet. 2022; 399(10322):406–410. https://doi.org/10.1016/S0140-6736(21)02835-X PMID: 34922639
- Sacco PL. The vaccine equity crisis is a stress test for all future major environmental challenges. Science of The Total Environment. 2022; 825:154073. https://doi.org/10.1016/j.scitotenv.2022.154073 PMID: 35202684
- 55. Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. The Lancet Infectious Diseases. 2020; 20(7):793–802. https://doi.org/10.1016/S1473-3099(20) 30230-9 PMID: 32247326
- Davis JT, Chinazzi M, Perra N, Mu K, y Piontti AP, Ajelli M, et al. Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave. Nature. 2021. https://doi.org/10.1038/s41586-021-04130-w
- 57. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science. 2020; 368(6489):395–400. https://doi.org/10.1126/science.aba9757 PMID: 32144116
- Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. Science. 2020; 369 (6500). https:// doi.org/10.1126/science.abb9789 PMID: 32414780
- 59. Maier BF, Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. Science. 2020; 368(6492):742–746. https://doi.org/10.1126/science. abb4557 PMID: 32269067
- 60. Arenas A, Cota W, Gómez-Gardeñes J, Gómez S, Granell C, Matamalas JT, et al. Modeling the spatio-temporal epidemic spreading of COVID-19 and the impact of mobility and social distancing interventions. Physical Review X. 2020; 10(4):041055. https://doi.org/10.1103/PhysRevX.10.041055
- Kraemer MUG, Hill V, Ruis C, Dellicour S, Bajaj S, McCrone JT, et al. Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence. Science. 2021; 373(6557):889–895. https://doi.org/10.1126/science.abj0113 PMID: 34301854
- 62. Sun K, Wang W, Gao L, Wang Y, Luo K, Ren L, et al. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. Science. 2021; 371 (6526). https://doi.org/10.1126/science.abe2424 PMID: 33234698
- Domenico LD, Sabbatini CE, Boëlle PY, Poletto C, Crépey P, Paireau J, et al. Adherence and sustainability of interventions informing optimal control against the COVID-19 pandemic. Communications Medicine. 2021; 1(1). https://doi.org/10.1038/s43856-021-00057-5

- 64. Marziano V, Guzzetta G, Mammone A, Riccardo F, Poletti P, Trentini F, et al. The effect of COVID-19 vaccination in Italy and perspectives for living with the virus. Nature Communications. 2021; 12(1):1–8. https://doi.org/10.1038/s41467-021-27532-w PMID: 34907206
- 65. Nouvellet P, Bhatia S, Cori A, Ainslie KE, Baguelin M, Bhatt S, et al. Reduction in mobility and COVID-19 transmission. Nature communications. 2021; 12(1):1–9. https://doi.org/10.1038/s41467-021-21358-2 PMID: 33597546
- 66. Cramer EY, Ray EL, Lopez VK, Bracher J, Brennen A, Castro Rivadeneira AJ, et al. Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the United States. Proceedings of the National Academy of Sciences. 2022; 119(15):e2113561119. https://doi.org/10.1073/pnas. 2113561119 PMID: 35394862
- Sacco PL, Valle F, De Domenico M. Proactive vs. reactive country responses to the COVID19 pandemic shock. PLOS Global Public Health, In press. 2021.
- Dowell D, Lindsley WG, Brooks JT. Reducing SARS-CoV-2 in Shared Indoor Air. JAMA. 2022; 328
 (2):141–142. https://doi.org/10.1001/jama.2022.9970 PMID: 35671318
- Buonanno G, Ricolfi L, Morawska L, Stabile L. Increasing ventilation reduces SARS-CoV-2 airborne transmission in schools: A retrospective cohort study in Italy's Marche region. Frontiers in Public Health. 2022; 10:4922. https://doi.org/10.3389/fpubh.2022.1087087
- Allen JG, Ibrahim AM. Indoor air changes and potential implications for SARS-CoV-2 transmission. Jama. 2021; 325(20):2112–2113. https://doi.org/10.1001/jama.2021.5053 PMID: 33861316
- Srikrishna D, Karan A, Dhillon RS. SARS-CoV-2 Reduction in Shared Indoor Air. JAMA. 2022; 328 (21):2162–2163. https://doi.org/10.1001/jama.2022.18028 PMID: 36472600
- **72.** Zarocostas J. How to fight an infodemic. The lancet. 2020; 395(10225):676. https://doi.org/10.1016/ S0140-6736(20)30461-X PMID: 32113495
- Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. Journal of medical Internet research. 2020; 22(6):e19659. https://doi.org/10.2196/19659 PMID: 32558655
- 74. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. Nature human behaviour. 2020; 4(12):1285–1293. https://doi.org/10.1038/s41562-020-00994-6 PMID: 33122812
- 75. van der Linden S. Misinformation: susceptibility, spread, and interventions to immunize the public. Nature Medicine. 2022; 28(3):460–467. https://doi.org/10.1038/s41591-022-01713-6 PMID: 35273402
- Lal A, Abdalla SM, Chattu VK, Erondu NA, Lee TL, Singh S, et al. Pandemic preparedness and response: exploring the role of universal health coverage within the global health security architecture. Lancet Glob Health. 2022; p. e1675–e1683. https://doi.org/10.1016/S2214-109X(22)00341-2 PMID: 36170734
- 77. Klamser P, d'Andrea V, Di Lauro F, Zachariae A, Bontorin S, di Nardo A, et al. Enhancing global preparedness during an ongoing pandemic from partial and noisy data. medRxiv. 2022.
- Carlson CJ, Albery GF, Merow C, Trisos CH, Zipfel CM, Eskew EA, et al. Climate change increases cross-species viral transmission risk. Nature. 2022; p. 1–1. PMID: 35483403