

RESEARCH

Open Access



# Impact of vaccine hesitancy on secondary COVID-19 outbreaks in the US: an age-structured SIR model

Alfonso de Miguel-Arribas<sup>1,2</sup>, Alberto Aleta<sup>3\*</sup> and Yamir Moreno<sup>1,2,3</sup>

## Abstract

**Background:** The COVID-19 outbreak has become the worst pandemic in at least a century. To fight this disease, a global effort led to the development of several vaccines at an unprecedented rate. There have been, however, several logistic issues with its deployment, from their production and transport, to the hesitancy of the population to be vaccinated. For different reasons, an important amount of individuals is reluctant to get the vaccine, something that hinders our ability to control and—eventually—eradicate the disease.

**Materials and methods:** Our aim is to explore the impact of vaccine hesitancy when highly transmissible SARS-CoV-2 variants of concern spread through a partially vaccinated population. To do so, we use age-stratified data from surveys on vaccination acceptance, together with age-contact matrices to inform an age-structured SIR model set in the US.

**Results:** Our results show that per every one percent decrease in vaccine hesitancy up to 45 deaths per million inhabitants could be averted. A closer inspection of the stratified infection rates also reveals the important role played by the youngest groups. The model captures the general trends of the Delta wave spreading in the US (July–October 2021) with a correlation coefficient of  $\rho = 0.79$ .

**Conclusions:** Our results shed light on the role that hesitancy plays on COVID-19 mortality and highlight the importance of increasing vaccine uptake in the population, specially among the eldest age groups.

**Keywords:** COVID-19, Mathematical modeling, Age-structured SIR, Vaccination, Hesitancy

## Background

After more than two years since the onset of the COVID-19 outbreak, firstly reported by the Chinese authorities on December 31, 2019, it is clear that this pandemic has become the worst one in at least a century.

Multiple aspects of our life have been severely affected at various scales: psychological [1, 2] and social [3]; human-related systems and infra-structures [4, 5]; supply chains [6, 7]; and the economy in general [8, 9]. To

manage the disease, it was mandatory to adopt a plethora of measures aimed at reducing the mixing and interaction among individuals in order to mitigate SARS-CoV-2 transmission and propagation. Lockdowns [10–14], curfews and mobility restrictions [15–18], social distancing [19–21], personal protection [22, 23] are now part of the new normalcy across the world. This “new normal” [24–26], being its impact as critical as the virus itself, was conceived and promised as something that should be ephemeral, a toll to pay, just until the ultimate solution arrives: the vaccines.

A rapid and massive scientific effort [27, 28] to develop a vaccine against SARS-CoV-2 was deployed and

\*Correspondence: [albertoaleta@gmail.com](mailto:albertoaleta@gmail.com)

<sup>3</sup> ISI Foundation, Via Chisola 5, 10126 Turin, Italy

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

successfully achieved in less than a year; another unprecedented fact. From December 2020, just a year after the onset of the pandemic, several nations started their vaccination campaigns in the pursue of herd immunity to control the pandemic. But again, further problems proliferated. From lack of confidence due to the relative short time for vaccine development and approval, to the fear of suffering serious side effects, or due to outlandish conspiracy theories, some people hesitate or are reluctant to vaccination [29–32]. For good or bad, this phenomenon is neither exclusive nor new [29, 33]. To be vaccinated (or not) is, in most countries, a choice of the individual, even though the consequences of such a choice go beyond the self and affect the social sphere. Hesitancy poses an ethical problem since if a critical fraction of individuals declines vaccine uptake for any disease, resurgence is to be expected [34–36]. This already happened in the UK, which was declared measles-free in 2017 but lost this status just 2 years later due to sub-optimal vaccination uptake [37].

Our aim is to quantify the effects of vaccine hesitancy in the US during the COVID-19 pandemic. In particular, given the large heterogeneity found across the US population, we perform our analysis on each state. It is worth noting that we do not intend to replicate the real trajectory of the COVID-19 pandemic in the US, neither we aim to accurately forecast the unfolding of future outbreaks and epidemic sizes. Instead, we focus on looking at correlations between variables related to hesitancy and the disease impact on the population, i.e., attack rates and deaths. We also explore the role that age structure plays in conditioning the outcomes and estimate potentially averted deaths if a 1% point more of the hesitant fraction of the population would change its attitude. Finally, we compare the model output from our hypothetical scenario with the epidemic impact caused during the COVID-19 Delta wave during July–November 2021.

## Materials and methods

We make use of an age-structured SIR model to simulate the spreading dynamics, which is fed with real and up-to-date data of the US age-distributed population and contact matrices, as well as with survey-based seroprevalence estimations [38]. We propose a hypothetical scenario in which COVID-19 outbreaks emerge in each state, independently, with a mitigated propagation due to the presence of some restrictions, while there is an ongoing vaccination campaign designed following the information obtained from public surveys [39]. Once the vaccination and this first outbreak end, we assume a *back to normal* situation, where all restrictions are lifted, disease awareness vanishes, and a new outbreak sets in. We assume that these successive outbreaks happen for a more transmissible variant of the virus,

mimicking in this way the observed evolutionary path of the SARS-CoV-2 variants of concern.

### Epidemic model

Given the utmost relevance of age in the effects of COVID-19, it is compulsory to introduce the age distribution of the population and the specific interaction between age groups to adequately model the dynamics of the disease [40–42]. We use the estimated age-contact matrices provided by Mistry et al. [43] updated to the population structure of 2019 [44, 45]. Then, we build an age-structured SIR model defined by this set of equations [43]:

$$\begin{aligned} \frac{dS_a}{dt} &= -\lambda_a S_a, \\ \frac{dI_a}{dt} &= \lambda_a S_a - \gamma I_a, \\ \frac{dR_a}{dt} &= \gamma I_a, \end{aligned} \quad (1)$$

where  $S_a$  is the number of susceptible individuals of age  $a$ ,  $I_a$  is the number of infected individuals of age  $a$ ,  $R_a$  is the number of removed individuals of age  $a$ , and  $\gamma^{-1}$  is the infectious period, which is assumed to be the same for all age classes and equal to  $\gamma^{-1} = 4.5$  days. COVID-19 is a disease with a more complex natural history than a SIR model can account for, being required to add some pre-symptomatic or asymptomatic compartments, as well as a latency period, for certain applications. Nonetheless, it has been shown that SIR models can correctly describe the overall evolution of the disease [46], which is enough for the scope of this paper. Lastly,  $\lambda_a$  is the force of infection for individuals of age  $a$  and it is expressed as

$$\lambda_a = \beta \chi_a \sum_{a'} M_{aa'} \frac{I_{a'}}{N_{a'}}, \quad (2)$$

where  $\beta$  is the transmissibility of the virus,  $N_a$  is the total number of individuals of age  $a$ , and  $M_{aa'}$  measures the average number of contacts of an individual of age  $a$  with individuals of age  $a'$ . Finally,  $\chi_a$  is an age-dependent susceptibility factor accounting for the lower susceptibility of children to the disease, i.e.  $\chi_a = 0.56$  if  $a \leq 19$  and 1 otherwise [47].

The basic reproductive number  $R_0$  is defined in this model as

$$R_0 = \frac{\beta}{\gamma} \rho(\chi M), \quad (3)$$

where  $\rho(\chi M)$  is the spectral radius, or largest eigenvalue, of the age-contact matrix (in this case also incorporating the susceptibility factor) [48].

## Scenario

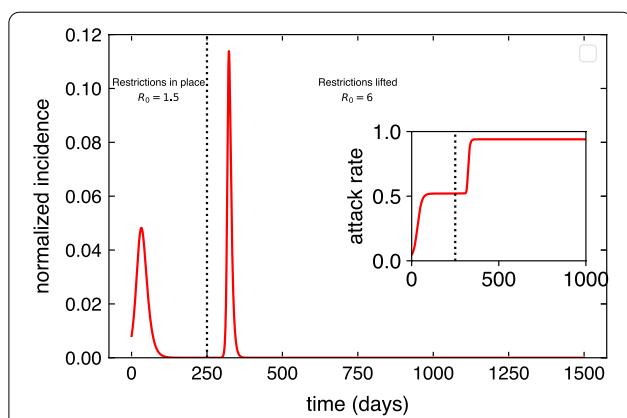
First, we collect the seroprevalence data measured in September 2020 for each age group and US state. We set the corresponding fraction of the population in each state into the removed compartment. Second, we simulate an initial outbreak with an  $R_0 = 1.5$ . This basic reproductive number is below the estimated  $R_0$  value for unmitigated transmission of the original variant of SARS-CoV-2 that is around 2.5–3 [49]. With this choice, we mimic a scenario in which there are some restrictions, social distancing and other protective measures in place, yielding a smaller effective  $R_0$ . During this outbreak, we implement a vaccination campaign (described below). By the end of the campaign, all individuals that have not refused vaccine uptake will have been vaccinated.

Once the vaccination campaign is completed, we assume that societies have returned to normalcy, i.e., any kind of restrictions and precautionary measures are lifted. Then, a new outbreak is seeded in each state, emulating a spill over from other states in the country or importation from other countries. In this second outbreak, we set  $R_0 = 6$  which is closer to the dominant variant of concern (delta variant) [50], predominant in the US since mid 2021 until the emergence of the Omicron strain by the end of 2021 [51, 52], being this last variant of concern more highly transmissible with respect to Delta even though with lower pathogenicity [53]. The reason to propose these secondary outbreaks in each US state with an  $R_0$  more in line with Delta is the availability of data to compare our hypothetical scenario with reality. Note that we assume that no awareness or other non-pharmaceutical interventions are in place during this outbreak. Thus, it can be thought as the worst case scenario of resurgence after a vaccination campaign.

As a visual example of the proposed scenario, in Fig. 1 we show how the incidence would evolve at the level of state for the full epidemic if no vaccination campaigns were deployed during the first outbreak. When an aggressive variant sets in, secondary outbreaks may still cause havoc. The inset depicts the evolution of the prevalence, which can reach almost the whole population for large enough  $R_0$ . In Sect. 3, we explore how the vaccination efforts modify this baseline scenario.

## Vaccination

We use data from The COVID States Project (<https://covidstates.org>), in particular the surveys in Report #43: COVID-19 vaccine rates and attitudes among Americans [39]. These surveys provide information on vaccination acceptance/hesitancy by age at the state level. Therein, several degrees of predisposition toward vaccines are reported. The following categories are



**Fig. 1** Proposed baseline scenario. Following the first wave of the epidemic, part of the population acquires natural immunity. Then, we simulate the propagation of a mitigated outbreak due to the presence of some restrictions, social distancing and prophylaxis measures, leading to a slower propagation of the original variant of the disease ( $R_0 = 1.5$ ). After the outbreak extinguishes a back-to-normal situation is assumed and all prevention measures are lifted. Then, an outbreak is seeded again with a higher basic reproductive number,  $R_0 = 6$ . On top of this baseline scenario, we will introduce a vaccination campaign during the first outbreak and explore the impact of vaccination hesitancy on the second outbreak

distinguished: individuals who are “already vaccinated”, individuals who are inclined to be vaccinated “as soon as possible”, “after at least some people I know”, “after most people I know”, and finally people who “would not get the COVID-19 vaccine”. The shares of people in each category is given at a national level for different age groups. The data shows an important amount of heterogeneity in each of those categories by age group. However, at the level of state, the data is not disaggregated by age groups, only the share of people in each vaccine acceptance category is shown. More specifically, we are looking for the coefficients  $g_{c,a}^{\text{state}}$ , which represent the share of people for every acceptance category,  $c$ , and age-class,  $a$  in every US state. These coefficients satisfy  $N_c^{\text{state}} = \sum_a g_{c,a}^{\text{state}} N_a^{\text{state}}$ , where  $N_a^{\text{state}}$  is the population of the state in the age class  $a$  and  $N_c^{\text{state}}$  is the population of the state in the acceptance category  $c$ . These  $N_c^{\text{state}}$  values are provided in appendix A within the report, but the information is not disaggregated by age at the level of state [39].

The report offers information at a national level about how people are distributed within acceptance categories by age groups. We refer to the shares shown in the report as  $h_{c,a}^{\text{national}}$ , which are normalized by age-class, that is,  $1 = \sum_a h_{c,a}^{\text{national}}$  for a particular category  $c$ . The quantity  $\sum_a h_{c,a}^{\text{national}} N_a^{\text{state}}$  would be the number of people if national coefficients apply for a certain state and vaccine acceptance category  $c$ . We relate these

coefficients  $h_{c,a}^{\text{national}}$  to coefficients  $g_{c,a}^{\text{state}}$  through a linear transformation:

$$g_{c,a}^{\text{state}} = h_{c,a}^{\text{national}} \frac{N_c^{\text{state}}}{\sum_a h_{c,a}^{\text{national}} N_a^{\text{state}}} \quad (4)$$

This transformation preserves the shares of people in a certain vaccine acceptance category  $c$  in every state and also allows for the introduction of age heterogeneity adapted from the national-level data.

Vaccination campaigns are complex and depend highly on several properties of the population: age, risk groups, professions, supplies, infrastructure, etc. Since we are mostly interested in the aftermath after the vaccination campaign, we adopt a simple scheme. From the aforementioned surveys, we extract the fraction of the population within each age group and state that is willing to be vaccinated,  $V_a^{\text{state}}$ . We set the length of the vaccination campaign to be  $\Delta t_v = 150$  days and assume that the fraction of population vaccinated per unit time is constant and equals to  $V_a^{\text{state}} / \Delta t_v$ . Both susceptible and recovered individuals can be vaccinated. For simplicity, the vaccine is assumed to be 100% effective in preventing the infection.

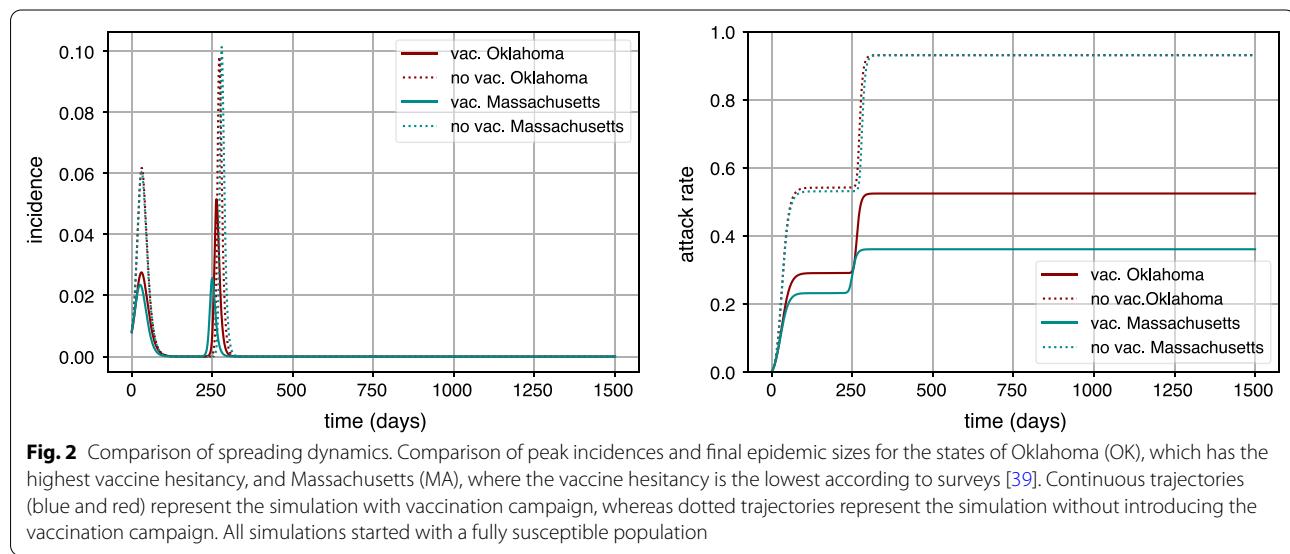
## Results

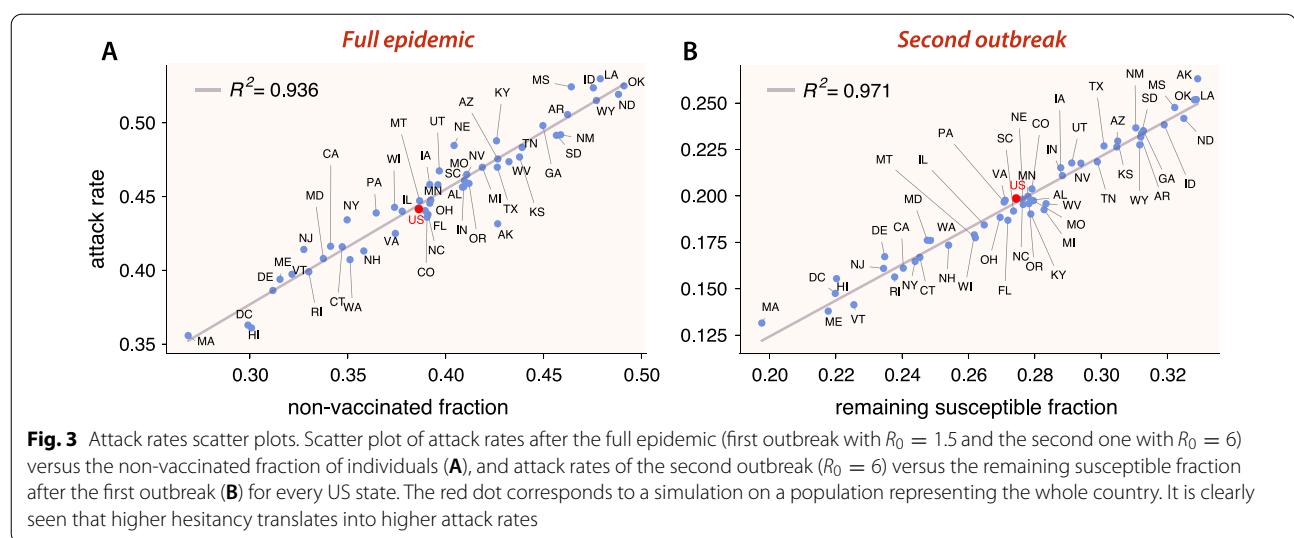
In Fig. 2, we show how the incidence and prevalence of the disease changes from the baseline scenario depicted in Fig. 1 when vaccination is in place. In particular, we consider the state with the highest vaccine hesitancy, Oklahoma (OK), and the state with the lowest one, Massachusetts (MA). Additionally, for a fairer comparison, the simulations were started with a null initial condition

for prevalence ( $R_a(t=0)=0$ ) (i.e. considering that the whole population is in the susceptible state). Dotted lines in the figure show the case without vaccination. We can see that the impact, in each isolated outbreak and for the full epidemic, is more or less the same for both states, differences owing to population internal structure. When vaccination is introduced in the model (continuous lines), we can appreciate the reduction of peak incidence and epidemic final sizes for both states during the first outbreak. However, when we simulate the second outbreak, the state with the lowest vaccine hesitancy shows a remarkably lower impact, while the other state experiences a sizable second outbreak. The peak of the outbreaks is similar in both outbreaks for Massachusetts, while in Oklahoma, the secondary outbreak is around twice as large as the first outbreak.

Next, we focus on the overall effect of vaccination on the spreading. We explore the relationship between the final attack rate (total fraction of the population that was infected) and the fraction of non-vaccinated individuals in each state. At the time of the surveys vaccine uptake on underage people was not being considered and there was no data regarding the attitudes of this age group. Thus, this set of individuals is composed by both underage people and adults who manifested vaccine hesitancy in the aforementioned surveys [39].

In order to look for a possible correlation between state-level attack rates and the fraction of non-vaccinated individuals, we performed a linear regression. Figure 3A depicts a scatter plot of the attack rates versus the fraction of non-vaccinated individuals for the simulated full epidemic unfolding in every state. The coefficient of determination,  $R^2 = 0.936$ , shows a clear relationship

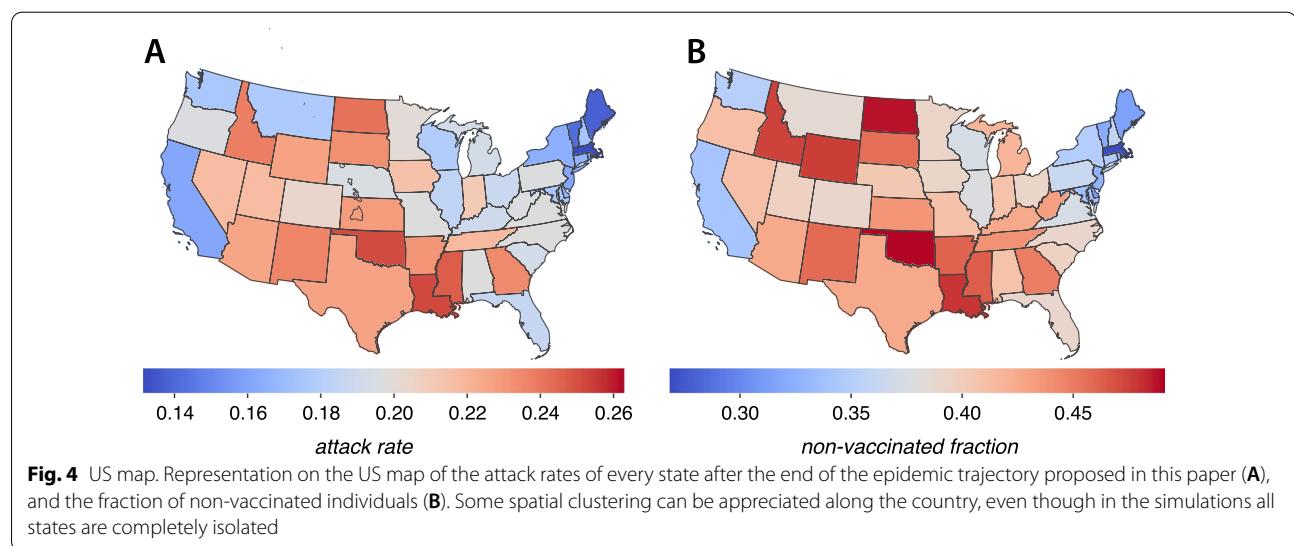




between attack rates and vaccine hesitancy for the full period. Note that we have added a simulation for the whole country (the red dot in the scatter plots) with the age-structure from the whole population. Figure 3B shows a scatter plot of the attack rate of the second outbreak, versus the fraction of remaining susceptible individuals at the end of the first outbreak. Here, the coefficient of determination is also very high,  $R^2 = 0.971$ . Note that the use of the remaining susceptible fraction rather than directly the fraction of non-vaccinated individuals owes to the fact that once the first outbreak and the vaccination campaign have ended, the demographic structure of the pool of susceptible individuals has changed dramatically. This pool is all conformed by individuals that either declined vaccination or are underage.

Since, according to data, hesitancy rates are low in older people, there is a predominance now of younger susceptible individuals. Additionally, Fig. 4 represents the very same data of Fig. 3A on the map of the United States. We observe some geographical clustering, even though we are treating each state as a completely isolated population. The states with higher attack rates or, similarly, the states with a higher fraction of vaccine hesitancy, are concentrated mainly in the interior of the country (inner Pacific west, Intermountain), ranging from north (Midwest) to south (inner Southeast).

Let us next try to get a deeper understanding of what is happening during our simulated second outbreak. Looking at some particular extreme examples, we can appreciate that the state of Massachusetts (MA), with the lowest



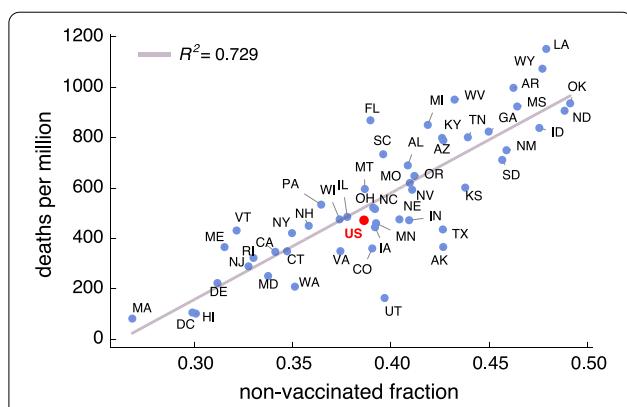
vaccine hesitancy (9% of adult population), has the lowest epidemic size during the first outbreak and also during the second outbreak. On the other hand, Alaska (AK) shows one of the lowest attack rates in the first outbreak, but the highest one in the second outbreak, together with the highest fraction of remaining susceptible at the end of the first outbreak, whereas its hesitancy amounts to 23% of the adult population, way behind the most reluctant states. Interestingly, there are other states with a relative low hesitancy rate that also show a sizable second outbreak. This is the case of the state of Utah (UT), with a hesitancy of about 15% among the adult population but nevertheless ranking high in the size of the second outbreak. One could hypothesize that these two states should have a similar number of deaths during the second outbreak. But, remarkably, as we show below, there is indeed more than a simple extrapolation of the correlation between the outbreak size and the number of non-vaccinated/susceptible individuals when it comes to forecasting mortality. The reason is that the age of non-vaccinated and/or remaining susceptible matters, not only because it usually determines behavior (and risk of infection) but also because the infection fatality rate heavily depends on it. As [40] emphasizes, considering transmission through the lens of (age-based) contact patterns is fundamental to understanding which population groups are driving disease transmission. Several reports, at least for the US, point to the fact that transmission dynamics shifted from older adults in the first stages of the pandemic to younger groups later [40, 54, 55]. This is understandable since once the harshest lockdowns were lifted, naturally younger groups are more socially active. In contrast, elders are less active and due to epidemic awareness, one should expect that they mix more carefully. Regarding mortality, it has been well documented the increasing risk of suffering severe disease and death for the oldest age groups, specially beyond 65 years old [41, 56–60].

In Fig. 5, we show a scatter plot of deaths per million individuals in the second outbreak versus the fraction of non-vaccinated individuals at the end of the first outbreak. We estimate the number of deaths in each age group by applying the corresponding infection fatality rate (IFR) [56], so that:

$$D_a(\infty) = \text{IFR}_a \times R_a(\infty), \quad (5)$$

where  $R_a(\infty)$  and  $D_a(\infty)$  are, respectively, the prevalence and the number of deceased individuals at the end of a particular outbreak.

Even though a high coefficient of determination is obtained, its explanatory power is smaller than for the attack rate, which suggests that there are other factors playing a relevant role. Certainly, we can appreciate that



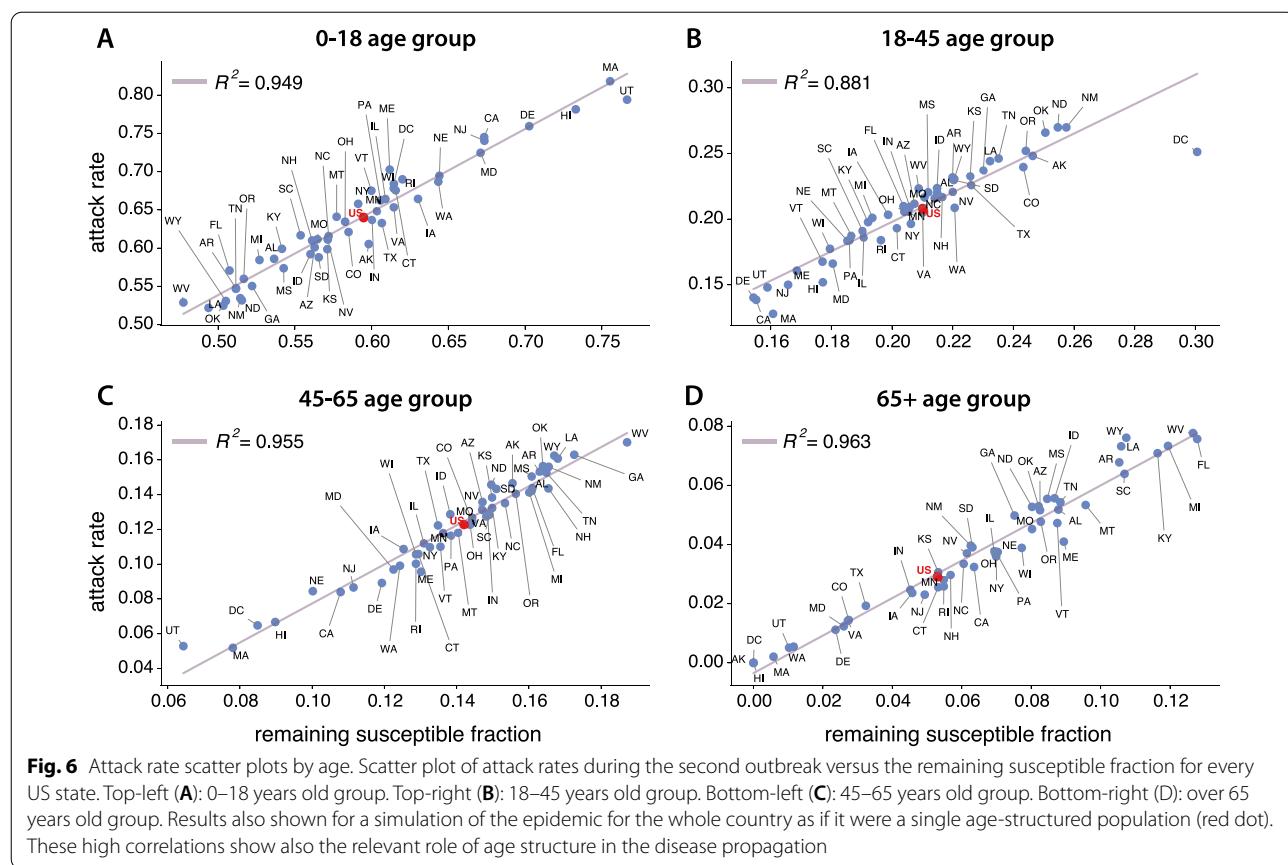
**Fig. 5** Death scatter plots. Scatter plot of deaths per million during the second outbreak versus the non-vaccinated fraction at the end of the first outbreak for every US state. Results also shown for a simulation of the epidemic for the whole country as if it were a single age-structured population (red dot). The model does not include deaths as part of the dynamics, but they can be estimated by applying the infection fatality rate to the final fraction of individuals in the removed compartment for each age class (Eq. (5))

higher proportions of deaths tend to occur in those states with higher hesitancy. Bringing back the case of Alaska (AK), and contrary to what could be naively expected, we see that it has been overtaken by several states. Even more striking is the case of Utah (UT), being in the lower part of the ranking. This clearly reveals that apart from vaccine hesitancy, the age structure is playing a key role in the disease dynamics and COVID-related fatalities [56, 58, 59].

To understand better these interdependencies, we next look at the attack rates during the second outbreak by coarse-graining the 85 age groups resolved in our model into four main relevant categories for the sake of the analysis. In Fig. 6, we show results for 0–18 (A), 18–45 (B), 45–65 (C), and more than 65 year old age groups (D). For each one, the attack rates during the second outbreak are computed as  $R_a(\infty)/R(\infty)$ , while the fractions of remaining susceptible individuals at the end of the first outbreak are computed as  $S_{1a}/S_1$ , where  $R(\infty)$  is the final attack rate, and  $S_1$  is the total fraction of remaining susceptible subjects. Thus, these figures tell us the share of people in each group  $a$  within the susceptible and removed pools.

## Discussion

The results by age groups exhibit a very high correlation for the linear fittings, which indicate the relevance of age structure in the transmission of the disease. For every age group, states with higher hesitancy tend to experience larger epidemic sizes. Regarding the cases mentioned before, namely, Utah (UT) and Alaska (AK), one can see



that their fractions of remaining susceptible individuals are large in the youngest age groups and rather small (null for Alaska) in the 65+ age strata. This ultimately explains why these two states undergo large secondary outbreaks that are not translated into a higher number of deaths. Finally, we also note that the fraction of remaining susceptible individuals is the highest for roughly every state in the two youngest age brackets (around or higher than 50%), which means that the younger age groups will be the driving group of the second outbreak.

It is thus clear the relevance of both vaccine hesitancy and of age heterogeneity in order to project the impact of the epidemic spreading on a territory. To showcase this, we look for an estimation of how many deaths could potentially be averted just by reducing the fraction of individuals in the “would not get the COVID-19 vaccine” category in one percentage point. It may occur that for states with an important share of younger population and not very high hesitancy, an extra effort does not pay off. Conversely, in states with an older population and for those with high hesitancy, such additional increase in the percentage of vaccinated may represent an important benefit. We believe these are important considerations for public health policy making. Table 1 shows the

number of averted deaths per million people if vaccine hesitancy is reduced by one per cent during the vaccination campaign in every state. A first look at the table would take us to believe that, overall, states with higher hesitancy will tend to avert more deaths by improving their vaccine roll out. But if we look for correlations between the total number of averted deaths in both outbreaks and the fraction of non vaccinated individuals, we obtain a not so high Pearson coefficient of  $\rho = 0.61$ , signalling correlations but not quite strongly. We have learnt throughout the discussion and related literature review the severe impact that disease has on the eldest groups within a population and specially in this work, the importance of having a low pool of remaining susceptible individuals in the oldest age groups (65+) for having lower death rates in secondary outbreaks. If then we correlate this quantity with the total number of averted deaths, we obtain a Pearson coefficient of  $\rho = 0.92$  and thus very high correlation and greater explanatory power.

The results shown in this work are based on a standard and sound epidemiological approach based on compartmental ODE modeling with a heavily-based data-driven input for several aspects: population age structure, mixing patterns by age, and vaccine uptake

**Table 1** Average number of averted deaths per million (95% CI in brackets), separately in the 1st and 2nd outbreak, due to reducing vaccine hesitation in one percent point

State	Averted deaths (1st outbreak)	Averted deaths (2nd outbreak)	State	Averted deaths (1st outbreak)	Averted deaths (2nd outbreak)
AK	5.12 [2.83–10.57]	16.78 [9.37–35.14]	MT	12.03 [6.62–23.86]	34.49 [18.93–67.57]
AL	8.62 [4.73–17.13]	39.18 [21.49–76.94]	NC	8.08 [4.44–16.11]	34.3 [18.83–67.7]
AR	10.23 [5.61–20.29]	42.38 [23.17–82.64]	ND	10.83 [5.92–21.52]	35.56 [19.34–69.23]
AZ	9.66 [5.29–19.12]	39.74 [21.73–77.58]	NE	9.75 [5.34–19.35]	27.7 [15.11–54.23]
CA	8.79 [4.81–17.45]	24.96 [13.62–49.08]	NH	12.1 [6.67–24.09]	33.04 [18.19–65.35]
CO	7.66 [4.21–15.43]	26.57 [14.62–53.06]	NJ	4.49 [2.47–8.91]	27.55 [15.08–54.42]
CT	10.43 [5.72–20.71]	28.18 [15.43–55.57]	NM	10.73 [5.9–21.38]	37.28 [20.45–73.19]
DC	4.09 [2.22–8.28]	9.43 [5.16–20.01]	NV	7.57 [4.16–15.1]	34.65 [19.04–68.39]
DE	8.34 [4.59–16.6]	26.71 [14.71–52.98]	NY	4.18 [2.29–8.29]	31.78 [17.35–62.54]
FL	10.3 [5.62–20.2]	48.74 [26.54–94.48]	OH	10.55 [5.79–20.98]	32.53 [17.83–63.96]
GA	5.89 [3.24–11.8]	38.9 [21.38–76.83]	OK	9.74 [5.34–19.41]	38.24 [20.93–74.83]
HI	9.31 [5.09–18.44]	10.87 [6.06–22.76]	OR	11.05 [6.07–21.97]	35.75 [19.61–70.23]
IA	7.31 [4.0–14.56]	31.54 [17.25–62.05]	PA	7.84 [4.3–15.2]	36.33 [19.88–71.2]
ID	10.18 [5.59–20.29]	35.19 [19.29–68.9]	RI	10.26 [5.63–20.38]	26.66 [14.59–52.55]
IL	9.28 [5.08–18.44]	30.61 [16.73–60.12]	SC	9.66 [5.32–19.19]	39.87 [21.91–78.17]
IN	9.00 [4.95–18.04]	30.06 [16.51–59.43]	SD	10.53 [5.79–21.03]	35.13 [19.24–68.87]
KS	9.08 [4.99–18.16]	33.38 [18.29–65.69]	TN	9.54 [5.24–19.0]	39.14 [21.47–76.84]
KY	11.88 [6.52–23.52]	37.12 [20.32–72.48]	TX	5.99 [3.29–12.05]	27.89 [15.3–55.49]
LA	7.68 [4.22–15.25]	45.27 [24.79–88.51]	UT	5.68 [7.38–26.53]	17.02 [9.31–34.05]
MA	8.37 [4.59–16.67]	16.94 [9.29–33.83]	VA	8.31 [4.57–16.66]	27.93 [15.37–55.55]
MD	5.64 [3.1–11.27]	25.95 [14.27–51.71]	VT	13.40 [7.38–26.53]	31.36 [17.21–61.56]
ME	14.13 [7.77–27.89]	29.7 [16.29–58.21]	WA	8.14 [4.48–16.35]	22.54 [12.4–45.11]
MI	12.03 [6.6–23.82]	40.84 [22.34–79.66]	WI	10.70 [5.88–21.28]	30.5 [16.71–59.9]
MN	7.42 [4.08–14.82]	31.34 [17.18–61.9]	WV	13.25 [7.28–26.2]	44.72 [24.5–87.12]
MO	10.60 [5.81–21.04]	34.9 [19.09–68.37]	WY	12.39 [6.82–24.64]	39.78 [21.81–77.73]
MS	8.78 [4.82–17.48]	38.48 [21.09–75.36]	US	7.95 [4.36–15.85]	32.15 [17.62–63.43]

Results shown for every state and for the whole country (US)

attitudes. The proposed scenario on which the model is run, however, does not map exactly any real situation experienced in the US and thus our aim was not to reproduce neither forecast accurately realistic COVID-19 trajectories. The particular outcomes brought should be regarded as what-if scenarios or hypothetical outcomes of what to expect overall given the premises hold. If, for instance, vaccine acceptance fractions should differ, as well as the transmissibility of a given virus strain, the specific figures could change dramatically, but not so the underlying conclusions. Additionally, from the beginning the model confection and the devised scenario was not conceived to realistically simulate or reproduce the myriad of complexities and heterogeneities involved in COVID-19 spreading country-wise. However, we can still check to a certain point how the model performed with respect to reality and thus gain confidence on the conclusions derived from it or rather discard the approach as insufficient or unsatisfactory.

As explained above, we proposed a hypothetical situation in which the epidemic spreading was ongoing but under rather mild transmission conditions due to disease awareness and general restrictions, while at the same time a mass vaccination campaign was deployed. After vaccination was completed and the epidemic wave was rather halted, then we simulated that societies turned back to normal from very low daily incidence, but the virus was still there and new secondary outbreaks emerged.

In reality, vaccine roll out in the US took off in the beginning of 2021 among concerns about vaccine hesitancy, a few months before we posed our research questions and designed the aforementioned hypothetical scenario. By mid February the highest epidemic wave experienced by the country until that moment was ending and reaching a plateau of rather low incidence. Overall, the situation stayed under control except for a very slight peak around mid April and then a relaxed decrease

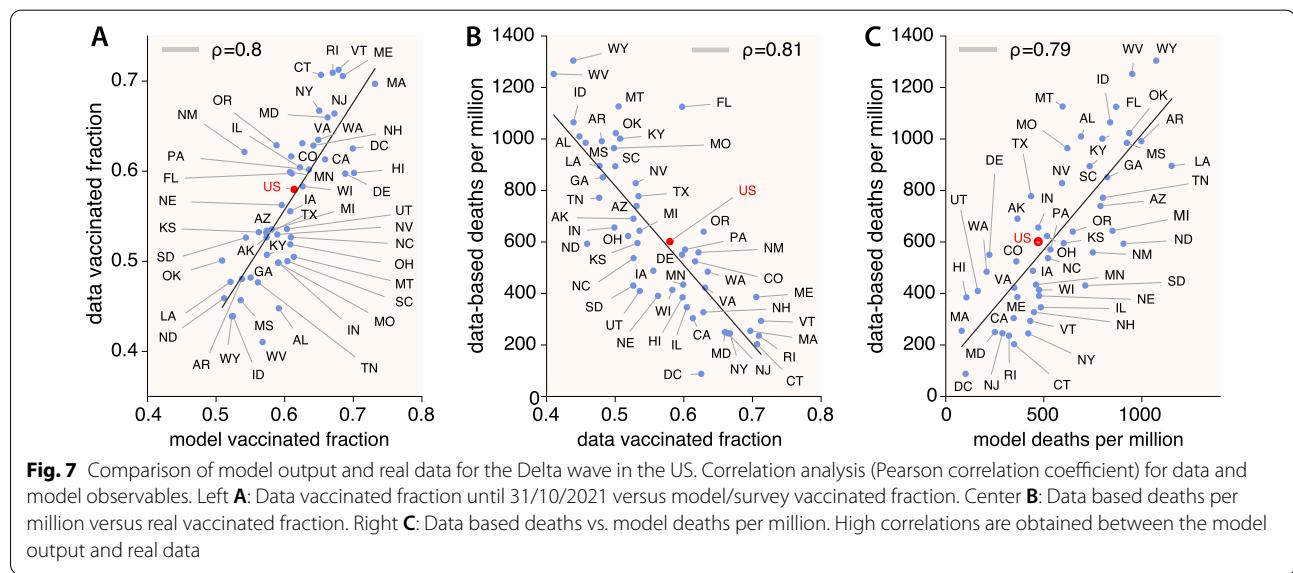
until reaching the lowest national incidence levels by June 2021 since the beginning of the pandemic. Thus, we could draw some parallelisms here with our first wave in our experiment: a rather controlled and decaying progression with vaccination going on. Then, restrictions were overall lifted up and a new wave started to build up by the beginning of July 2021. This wave was mainly driven by the Delta variant of concern, peaking on the first days of September, reaching a higher plateau than the previous wave by early November 2021, and immediately followed by the huge Omicron wave. Given the history of the spreading dynamics in the US, we consider that our hypothesized secondary outbreaks in every state, describing a rather unmitigated scenario with a more aggressive strain, could match reasonably well the Delta wave that took place across the described period. Therefore, it is informative to check how the model output relates to the real impact of the epidemic during the aforementioned period. In Fig. 7, we looked at the following pairs of observables to check for correlations: real data [61] vs. model/survey vaccination fractions (Fig. 7A), data-based deaths [62] vs. real vaccination fractions (Fig. 7B), and data-based deaths vs. model deaths (Fig. 7C). We can see that all confronted observables show a high correlation. First, in panel A, comparing data and model vaccination fractions we obtain a Pearson correlation coefficient of  $\rho = 0.8$ . Overall, we can say that surveys on attitude towards vaccination were good, strengthening subsequent model results. Second, in panel B we obtain a value of  $\rho = 0.81$  when correlating what happened in reality regarding deaths and vaccination. As expected, the higher the vaccination fraction in a state, the lower the deaths that took place due to the

COVID-19. When comparing in the paper the model deaths per million with the non-vaccinated fraction (for the matter, this is equivalent to comparing it to the vaccinated fraction), we obtained an  $R^2 = 0.729$  and therefore a Pearson coefficient of  $\rho = 0.85$ , a very high signal of correlation. We can then see that our model projections captured the real trend quite satisfactorily and this emphasizes the dominant role of the vaccine of mitigating the impact on the population. Finally, in panel C we obtained a Pearson coefficient  $\rho = 0.79$  when comparing this time the real data-based estimation of deaths and the model-based estimation of deaths. Thus we find again a high correlation between our hypothesized scenario of secondary outbreaks and the Delta wave.

## Conclusions

In this work, we have explored SARS-CoV-2 transmission dynamics on a population that is partially vaccinated and is seeded again with the virus when restrictions are fully lifted. We explored, in particular, to what extent vaccine hesitancy may still drive sizable outbreaks in a context where a more transmissible SARS-CoV-2 variant of concern is dominant. We used data from vaccination acceptance surveys, together with up-to-date age distributed populations and contact matrices in the US to inform an age-structured SIR model.

Our results show a clear correlation between the size of experienced outbreaks, once all kinds of measures are lifted, and the fraction of vaccine hesitancy or, similarly, the fraction of remaining susceptible individuals at the onset of a second outbreak. Higher vaccine hesitancy ratios expose the population to larger outbreaks and,



inversely, higher vaccine acceptance ratios can mitigate the impact to the point of negligible secondary waves due to immunity of the population.

We have also inspected in detail the role of the age structure of the population in both the attack rate and the mortality of secondary outbreaks. Our findings reveal that the prevalence is highly correlated with the fraction of remaining susceptible individuals by age classes, with the youngest contributing the most to the attack rate. It is however not immediate to project such a correlation to the expected number of deaths, as here too age plays a role, though in the opposite direction, e.g., the younger the population, the lower the mortality.

Lastly, we estimated the number of potentially averted deaths during the course of the simulated epidemic if the number of people reluctant to vaccine uptake were reduced in one percentage point. Results show again the relevance of regarding age structure in transmission since not all the states with higher hesitancy rank highest in averting deaths. It is very relevant the fraction of hesitant individuals in the older groups.

To round up the analysis, we investigated how the model fared when comparing with real data. Even though the devised experiment here was not intended to accurately replicate or forecast real COVID-19 trajectories, the data-driven approach and sound modeling offered very high correlations when comparing survey/model vaccination against real vaccination, and model death estimation against real data-based deaths during the Delta wave in the US.

We acknowledge that our model has several limitations. One is at the core of its compartmental structure, not including a more detailed progression of the natural history of the disease, which might affect our estimation of deaths, and does not consider hospitalizations of any kind. The vaccination campaign could be implemented in a more realistic way and owing to each state idiosyncrasy but, more importantly, vaccines are revealing to be not sterilizing and thus not fully preventing transmission and, on top of that, immunity decays over time. These facts do not affect the overall dynamics explored in this paper, but should be incorporated to provide reliable estimations on the exact amount of expected infections or deaths. Additionally, the behavioral responses are not completely accounted for. All these factors open important challenges for future works.

To conclude, the most important implications of the results reported here include: (i) data on vaccination by age is important to accurately capture the evolution of mortality in secondary waves; (ii) surveys on vaccination attitude are a valuable proxy to estimate the hesitancy of the population; (iii) allocation of additional resources is more important in states with relative high hesitancy rates but specially in states where the remaining

susceptible population is older; (iv) reintroduction of restrictions could be needed in states with very high attack rates to reduce pressure over healthcare systems; and (v) incentives to vaccination will reduce the number of deaths if they focus on the older generations.

#### Acknowledgements

Not applicable.

#### Author contributions

AA and YM conceived the study. AdMA curated the data, implemented the computational model and obtained results. All authors contributed to result analysis and discussion. AdMA wrote a first draft of the manuscript. YM wrote the final version and all authors revised it. All authors read and approved the final manuscript.

#### Funding

A. d.M.A. is funded by an FPI Predoctoral Fellowship of MINECO. We acknowledge partial support from the Government of Aragon, Spain and "ERDF A way of making Europe" through grant E36-20R (FENOL) to A. d.M. A. A. and Y. M., from Ministerio de Ciencia e Innovación, Agencia Española de Investigación (MCIN/AEI/10.13039/501100011033) Grant No. PID2020-115800GB-I00 to A. d.M. A. A. and Y.M., from Banco Santander (Santander-UZ 2020/0274) to A. A. and Y.M., and by Soremartec S.A. and Soremartec Italia, Ferrero Group to A. A. and Y.M. The funders had no role in the study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

#### Availability of data and materials

All the data used in this paper is publicly available. The age-contact matrices can be directly downloaded from <https://zenodo.org/record/4287574> and the information on vaccination attitudes is contained in the tables of the Appendix A of <https://osf.io/rnw8z/>.

#### Declarations

##### Ethics approval and consent to participate

The Declaration of Helsinki refers to "research on identifiable human material and data". This study does not use any individual human data. The public, aggregated statistics used in the study cannot be de-identified. Therefore, no ethics approval was needed.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

##### Author details

<sup>1</sup>Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, 50018 Zaragoza, Spain. <sup>2</sup>Department of Theoretical Physics, University of Zaragoza, 50009 Zaragoza, Spain. <sup>3</sup>ISI Foundation, Via Chisola 5, 10126 Turin, Italy.

Received: 17 November 2021 Accepted: 23 May 2022

Published online: 01 June 2022

#### References

1. Saladino V, Algeri D, Auriemma V. The psychological and social impact of Covid-19: new perspectives of well-being. *Front Psychol*. 2020;11:2550.
2. Tan BY, Chew NW, Lee GK, Jing M, Goh Y, Yeo LL, et al. Psychological impact of the COVID-19 pandemic on health care workers in Singapore. *Ann Intern Med*. 2020;173(4):317–20.
3. Romero CS, Delgado C, Catalá J, Ferrer C, Errando C, Iftimi A, et al. COVID-19 psychological impact in 3109 healthcare workers in Spain: the PSIMCOV group. *Psychol Med*. 2020;52:1–7.

4. Verelest F, Kuylen E, Beutels P. Indications for healthcare surge capacity in European countries facing an exponential increase in coronavirus disease (COVID-19) cases, March 2020. *Eurosurveillance*. 2020;25(13):2000323.
5. Armocida B, Formenti B, Ussai S, Palestro F, Missoni E. The Italian health system and the COVID-19 challenge. *Lancet Public Health*. 2020;5(5):e253.
6. Hobbs JE. Food supply chains during the COVID-19 pandemic. *Can J Agric Econ/Revue canadienne d'agroéconomie*. 2020;68(2):171–6.
7. Guan D, Wang D, Hallegatte S, Davis SJ, Huo J, Li S, et al. Global supply-chain effects of COVID-19 control measures. *Nat Hum Behav*. 2020;4:1–11.
8. McKibbin W, Fernando R. The economic impact of COVID-19. *Economics in the Time of COVID-19*. 2020;45.
9. Martin A, Markhvida M, Hallegatte S, Walsh B. Socio-economic impacts of COVID-19 on household consumption and poverty. *Econ Disasters Climate Change*. 2020;4(3):453–79.
10. Lau H, Khosrawipour V, Kocabach P, Mikolajczyk A, Schubert J, Bania J, et al. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *J Travel Med*. 2020;27(3):taaa037.
11. Koh D. COVID-19 lockdowns throughout the world. *Occup Med*. 2020;70(5):322–32.
12. Ren X. Pandemic and lockdown: a territorial approach to COVID-19 in China, Italy and the United States. *Eurasian Geogr Econ*. 2020;61(4–5):423–34.
13. Lancet T. India under COVID-19 lockdown. *Lancet (London, England)*. 2020;395(10233):1315.
14. Meunier TA. Full lockdown policies in Western Europe countries have no evident impacts on the COVID-19 epidemic. *MedRxiv*. 2020
15. Kraemer MU, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*. 2020;368(6490):493–7.
16. 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.
17. Aleta A, Martin-Corral D, y Piontti AP, Ajelli M, Litvinova M, Chinazzi M, et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nat Hum Behav*. 2020;4(9):964–71.
18. Bonaccorsi G, Pierri F, Cinelli M, Flori A, Galeazzi A, Porcelli F, et al. Economic and social consequences of human mobility restrictions under COVID-19. *Proc Natl Acad Sci USA*. 2020;117(27):15530–5.
19. Briscese G, Lacetera N, Macis M, Tonin M. Compliance with covid-19 social-distancing measures in Italy: the role of expectations and duration. *Natl Bureau Econ Res*; 2020.
20. Thunström L, Newbold SC, Finnoff D, Ashworth M, Shogren JF. The benefits and costs of using social distancing to flatten the curve for COVID-19. *J Benefit-Cost Anal*. 2020;11(2):179–95.
21. Lewnard JA, Lo NC. Scientific and ethical basis for social-distancing interventions against COVID-19. *Lancet Infect Dis*. 2020;20(6):631–3.
22. World Health Organization. Rational use of personal protective equipment for coronavirus disease (COVID-19): interim guidance. Geneva: World Health Organization; 2020.
23. Schünemann HJ, Akl EA, Chou R, Chu DK, Loeb M, Lotfi T, et al. Use of facemasks during the COVID-19 pandemic. *Lancet Respir Med*. 2020;8(10):954–5.
24. Rab S, Javaid M, Haleem A, Vaishya R. Face masks are new normal after COVID-19 pandemic. *Diabetes Metab Syndr*. 2020;14(6):1617–9.
25. Lee TH. Creating the new normal: the clinician response to COVID-19. *NEJM Catal Innov Care Deliv*. 2020;1(2):2.
26. Zinn JO. A monstrous threat: how a state of exception turns into a 'new normal'. *J Risk Res*. 2020;23(7–8):1083–91.
27. Le TT, Andreadakis Z, Kumar A, Román RG, Tollesfesen S, Saville M, et al. The COVID-19 vaccine development landscape. *Nat Rev Drug Discov*. 2020;19(5):305–6.
28. Wang J, Peng Y, Xu H, Cui Z, Williams RO. The COVID-19 vaccine race: challenges and opportunities in vaccine formulation. *AAPS PharmSciTech*. 2020;21(6):1–12.
29. Schuster M, Eskola J, Duclos P, et al. Review of vaccine hesitancy: rationale, remit and methods. *Vaccine*. 2015;33(34):4157–60.
30. Dror AA, Eisenbach N, Taiber S, Morozov NG, Mizrahi M, Zigrion A, et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. *Eur J Epidemiol*. 2020;35(8):775–9.
31. Khubchandani J, Sharma S, Price JH, Wiblishauser MJ, Sharma M, Webb FJ. COVID-19 vaccination hesitancy in the United States: a rapid national assessment. *J Commun Health*. 2021;46(2):270–7.
32. Romer D, Jamieson KH. Conspiracy theories as barriers to controlling the spread of COVID-19 in the US. *Soc Sci Med*. 2020;263: 113356.
33. Larson HJ, Jarrett C, Eckersberger E, Smith DM, Paterson P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007–2012. *Vaccine*. 2014;32(19):2150–9.
34. Wong L, Wong P, AbuBakar S. Vaccine hesitancy and the resurgence of vaccine preventable diseases: the way forward for Malaysia, a South-east Asian country. *Hum Vaccines Immunotherap*. 2020;16(7):1511–20.
35. Porteous GH, Hanson NA, Sueda LAA, Hoaglan CD, Dahl AB, Ohlson BB, et al. Resurgence of vaccine-preventable diseases in the United States: anesthetic and critical care implications. *Anesthesia Analgesia*. 2016;122(5):1450–73.
36. Poland GA, Jacobson RM, Ovsyannikova IG. Trends affecting the future of vaccine development and delivery: the role of demographics, regulatory science, the anti-vaccine movement, and vaccinomics. *Vaccine*. 2009;27(25–26):3240–4.
37. Measles in England - Public health matters; 2021. <https://publichealthmatters.blog.gov.uk/2019/08/19/measles-in-england> accessed 11 Sep 2021.
38. Bajema KL, Wiegand RE, Cuffe K, Patel SV, Iachan R, Lim T, et al. Estimated SARS-CoV-2 Seroprevalence in the US as of September 2020. *JAMA Internal Med*. 2020;181(4):450–60.
39. Lazer D, Ognyanova K, Baum M, Druckman J, Green J, Gitomer A, et al. The COVID States Project# 43: COVID-19 vaccine rates and attitudes among Americans. 2021.
40. Monod M, Blenkinsop A, Xi X, Hebert D, Bershan S, Tietze S, et al. Age groups that sustain resurging COVID-19 epidemics in the United States. *Science*. 2021;371(6536):eabe8372.
41. Levin AT, Hanage WP, Owusu-Boaitey N, Cochran KB, Walsh SP, Meyerowitz-Katz G. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *Eur J Epidemiol*. 2020;35(12):1123–38.
42. Davies NG, Klepac P, Liu Y, Prem K, Jit M, Eggo RM. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat Med*. 2020;26(8):1205–11.
43. Mistry D, Litvinova M, y Piontti AP, Chinazzi M, Fumanelli L, Gomes MF, et al. Inferring high-resolution human mixing patterns for disease modeling. *Nat Commun*. 2021;12(1):1–12.
44. <https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-detail.html>.
45. Arregui S, Aleta A, Sanz J, Moreno Y. Projecting social contact matrices to different demographic structures. *PLoS Comput Biol*. 2018;14(12): e1006638.
46. 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. <https://doi.org/10.1126/science.abb9789>.
47. Russell M, Viner P. Susceptibility to SARS-CoV-2 infection among children and adolescents compared with adults: a systematic. *JAMA Pediatr*. 2021;175(2):143–56.
48. Diekmann O, Heesterbeek JAP, Metz JA. On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *J Math Biol*. 1990;28(4):365–82.
49. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J Travel Med*. 2020;27(2):8.
50. Liu Y, Rocklöv J. The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus. *J Travel Med*. 2021;28(7):taab124.
51. Lin L, Zhao Y, Chen B, He D. Multiple COVID-19 waves and vaccination effectiveness in the United States. *Int J Environ Res Public Health*. 2022;19(4):2282.
52. Lambrou AS, Shirk P, Steele MK, Paul P, Paden CR, Cadwell B, et al. Genomic surveillance for SARS-CoV-2 variants: predominance of the Delta (B.1.617.2) and omicron (B.1.1.529) variants—United States, June 2021–January 2022. *Morb Mortal Wkly Rep*. 2022;71(6):206.

53. Suzuki R, Yamasoba D, Kimura I, Wang L, Kishimoto M, Ito J, et al. Attenuated fusogenicity and pathogenicity of SARS-CoV-2 Omicron variant. *Nature*. 2022;603:1–6.
54. Boehmer TK, DeVies J, Caruso E, van Santen KL, Tang S, Black CL, et al. Changing age distribution of the COVID-19 pandemic—United States, May–August 2020. *Morb Mortal Wkly Rep*. 2020;69(39):1404.
55. Harris JE. Data from the COVID-19 epidemic in Florida suggest that younger cohorts have been transmitting their infections to less socially mobile older adults. *Rev Econ Household*. 2020;18(4):1019–37.
56. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis*. 2020;20(6):669–77.
57. Ghisolfi S, Almås I, Sandefur JC, von Carnap T, Heitner J, Bold T. Predicted COVID-19 fatality rates based on age, sex, comorbidities and health system capacity. *BMJ Glob Health*. 2020;5(9): e003094.
58. Kang SJ, Jung SI. Age-related morbidity and mortality among patients with COVID-19. *Infect Chemotherapy*. 2020;52(2):154.
59. Bonanad C, García-Blas S, Tarazona-Santabalbina F, Sanchis J, Bertomeu-González V, Fáclla L, et al. The effect of age on mortality in patients with COVID-19: a meta-analysis with 611,583 subjects. *J Am Med Dir Assoc*. 2020;21(7):915–8.
60. Esteve A, Permanyer I, Boertien D, Vaupel JW. National age and coresidence patterns shape COVID-19 vulnerability. *Proc Natl Acad Sci USA*. 2020;117(28):16118–20.
61. Our World in Data. <https://ourworldindata.org/us-states-vaccinations> accessed 08 Apr 2020
62. The New York Times GitHub COVID-19 repository. <https://github.com/nytimes/covid-19-data> accessed 08 Apr 2020

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](http://biomedcentral.com/submissions)

