Response to the Open Letter by Haber et al.

Haber et al. wrote an open letter, claiming that our recently published paper (Zhang et al., Identifying airborne transmission as the dominant route for the spread of COVID-19, *Proc. Natl. Acad. Sci. USA*, 117, 14857, 2020) is flawed and requesting retraction of the paper. Scientific research is to discover the truth, and a fundamental principle in scientific research is that it shall be evidence-based. Similarly, an assessment of a scientific research paper shall also be evidence-based. What is the scientific evidence in the paper by Zhang et al. and in the open letter by Haber et al.?

The paper by Zhang et al. (2020) presented and analyzed the COVID-19 pandemic trends in the three epicenters, Wuhan, Italy, and New York City (NYC). Our scientific results are illustrated in five figures of the paper. Figs. 1 to 3 show publicly available data for COVID-19 together with a statistical analysis. Fig. 4 represents our synopsis of the virus transmission routes, and Fig. 5 presents our assessment of the impacts of mitigation measures and explanations for the observed pandemic trends. Our main conclusions are that airborne transmission represents the dominant route to spread the disease and face covering represents the determinant in shaping the pandemic trends. Figs. 1 to 3 in the paper are straightforward presentations and analysis of the data, and projection is made to estimate the number of infections prevented by face-covering in Fig. 2.

SARS-CoV-2 is a novel virus with unprecedented transmission efficiency and interventions undertaken. While the appropriate methodology to model the transmission and intervention for a chaotic system such as COVID-19 has yet to be established (Siegenfeld et al., Opinion: What models can and cannot tell us about COVID-19. *Proc. Natl. Acad. Sci. USA*, doi.org/10.1073/pnas.2011542117, 2020), any empirical modeling framework for infectious disease dynamics, which Haber et el. advocated in their criticisms to our approach, cannot be trusted without scientific validation.

Indeed, the curve for total confirmed infections exhibits two distinct growth stages in the world, U.S., Italy, and NYC (**attached Fig. 1**), in contrast to conventional infectious disease dynamics that typically treats the entire process as a single episode. The initial growth is characterized by a sub-exponential curve in the number of total confirmed infections (middle column in **attached Fig. 1**), which is followed by a remarkably linear growth (right column in **attached Fig. 1**). This initial sub-exponential growth was not the topic addressed in our paper, since it has been well discussed in several previous studies (e.g., Maier et al., Effective containment explains sub-exponential growth in recent confirmed COVID-19 cases in China. Science 2020; 368, 742). In our paper, we discussed that "*It is possible, however, that these measures ('social distancing, quarantine, and isolation') likely alter the slope of the infection curve, that is, by reducing the rate of infections during the early stage of the pandemic*". It is also possible that the combined measures of social distancing and lockdown/stay-at-home orders converted the exponential growth to the linear growth.

A critical finding in our paper lies in that the total infection curve is reflected by a remarkable linearity after implementing social distancing/stay-at-home measures, which follows immediately the exponential growth. This linearity extends from April 1 to May 9 in the world and U.S. without implementing mandated face covering and to the onset of mandated face covering in Italy and

NYC. Note that we did not assume the linearity; it is the property from the publicly available COVID-19 data. This linearity can be understood by a dynamic equilibrium between first-order transmission (mainly airborne), intervention (social distancing/stay-at-home), and the interaction between transmission and intervention as well as combined (addition or canceling) second-order effects (referred to as remnants depicted by Paths A and B in Fig. 5 by Zhang et al., 2020). Such an equilibrium is only significantly disrupted by other first-order processes, such as implementing mandated face-covering (Path B in Fig. 5 by Zhang et al., 2020), which causes significant departure from the linearity and curve flattening in Italy and NYC.

We further estimated the number of infections prevented from face covering by projecting the infection curve in Italy and NYC using the linear regression obtained prior to mandated face covering. We examined two different time intervals for the linear regression, i.e., 15 vs 26 days prior to the mandated face covering to evaluate the time lag effects, which can include possible delay in data reporting and incubation period, among other secondary processes. We found little difference in the linear regression using 15 and 26 days (Fig. S1 in Zhang et al., 2020). The onset and significance of the deviation from linearity in Italy and NYC can also be impacted by other secondary processes, including remnants defined in our work and face covering prior to the mandated measure, but the latter effect does not impact the outcome of our projection.

The validity of our approach is further supported by additional scientific evidence. An analysis of the pandemic trends for the 15 top-infected states in the U.S. (**attached Fig. 2**) shows that the total infections grow linearly for all states without mandated face covering but start significant flattening only for states after implementing mandated face covering. The results in **attached Fig. 2** are adopted from a preprint, which is accessible at the following link: http://centromariomolina.org/english2/wp-content/uploads/2020/06/COVID19-Zhang-MM-Paper2-6-26-2020_Understanding-Transmission-and-intervention-for-the-COVID-19-pandemic.pdf.

We summarize below the key aspects in Zhang et al. (PNAS, 2020):

(i) Our approach captured and represented the essential first-order processes, i.e., the transmission routes of SARS-CoV-2 (contact vs airborne transmission), the interventions (social distancing/stay-at-home and mandated face-covering), and the interaction between transmission routes and interventions. Specifically, our work addressed the first-order effects of intervention measures, "*social distancing and stay-at-home measures, in conjunction with hand sanitizing, minimized short-range contact transmission but did not prevent long-range airborne transmission. Mandated face covering effectively prevented airborne transmission by blocking atomization and inhalation of virus-bearing aerosols and contact transmission by blocking viral shedding of droplets. The combined face-covering and social distancing measures offered dual protection*".

(ii) Our work explained the pandemic trend at multiple locations (**attached Fig. 1 and Fig. 3**).

(iii) Our work provided guidance to development of intervention policies to constrain the spread of COVID-19 pandemic (the recommended paths B and C in Fig. 5 of our paper), including the following evidence-based statement: "*We conclude that wearing of face masks in public corresponds to the most effective means to prevent interhuman* *transmission, and this inexpensive practice, in conjunction with extensive testing, quarantine, and contact tracing, poses the most probable fighting opportunity to stop the COVID-19 pandemic, prior to the development of a vaccine*".

(iv) Our paper is scientifically sound and has produced major policy impacts.

In contrast, Haber et al. offered no scientific evidence to invalidate our results and conclusions and to substantiate their own assertions. They made several statements that were largely taken out of the context of our paper. A critical flaw by Haber et al. lies in that they focused exclusively on detailed secondary processes but ignored the physical insight and essential first-order processes, which is characteristic of mismodeling for COVID-19 (Siegenfeld et al., 2020). In addition, they provided several non-scientific and inconsistent statements. We elaborate below our objections to the open letter by Haber et al.

The following statements by Haber et al. are untrue: "*The main conclusions of this paper are based in comparison of linear case count slopes within and between regions, with mask mandates as the observed variable of interest. It ignores other clear differences in disease control policy between these areas, including broader heterogeneity in face mask policy*". Our work captured and represented the essential first-order processes, i.e., the transmission routes of SARS-CoV-2 (contact transmission vs airborne transmission), the interventions (social distancing/stay-at-home and mandated face-covering), and the interaction between transmission routes and interventions. We acknowledged the uncertainty and alternative secondary processes and concluded "*However, it is implausible that the limitations of mitigation measures alone contributed dominantly to the global pandemic trend, as exemplified by the success in China*". On the contrary, the approach advocated by Haber et al. is characteristic by mismodeling. According to Siegenfeld et al. (2020), "*But if a model's assumptions do not yield the same general large-scale behaviors of the system being modeled, adding additional details to the model will serve only to create a false sense of confidence.*"

The following statements by Haber et al. are false: "*In one critical example, the paper asserts that "after April 3, the only difference in regulatory measures between NYC and the United States lies in face coverings on April 17 in NYC." This is verifiably false, based on widely available (e.g., HIT-COVID2) sources. It is flatly untrue that there were no other regulatory differences between NYC and the rest of the US on those dates; it is also untrue that NYC was the only region in the US mandating use of face coverings*". The fact is that after April 3 both NYC and the United States as a nation have comparable social distancing and stay-at-home measures (see **attached Table 1**), but the U.S. as a nation did not implement mandated face covering. The above statements by Haber et al. also provide evidence for their failure to differentiate between first-order and second-order phenomena.

The following statements by Haber et al. are untrue: "*In another example, the paper asserts that airborne transmission is the dominant route for COVID-19 spread. To justify this headline conclusion, ... In fact, in April, many regions (e.g., Sweden, parts of the United States) were not in lockdown, and quarantine and isolation were not in place in most parts of the world*". The fact is that guidelines for social distancing, quarantine, isolations were issued by U.S. CDC on March 16, and guideline for physical distancing was recommended by WHO on March 18. By April 3, many countries with more than 82% of the total COVID-19 cases globally were under lockdown/stay-at-home orders, and only states with less than 4% of the total confirmed cases in the U.S did not implement stay-at-home order (see **attached Table 1**). This is also another evidence that Haber et al. did not differentiate between first-order and second-order phenomena. Further, Haber et al. failed to comprehend that our conclusion "*airborne transmission of COVID-19 represents the dominant route for infection*" is based on scientifically-sound reasoning.

The following statements by Haber et al. are erroneous: "*Beyond this, the study is afflicted with serious methodological errors that undermine any confidence in its findings…. Any one of the above issues in isolation would be cause for serious concern, but in combination, they are alarming*". The criticisms made by Haber et al. focused exclusively on detailed secondary processes but ignored the physical insight and critical first-order processes relevant to virus transmission, the effects of mitigation measures, and the interaction between transmission and intervention. While detailed modeling for the COVID-19 pandemic trend was beyond the scope of our paper, the methodology suggested by Haber et al. is intangible and physically-flawed. Specifically, its main deficiencies include:

(i) Secondary processes are overly emphasized and cannot be quantified.
(ii) First-order processes are completely ignored.
(iii) Their framework for infectious disease dynamics is characteristic of mismodeling (Siegenfeld et al., 2020).
(iv) Their approach cannot explain the pandemic trend.
(v) Their approach cannot provide insight to guide development of regulatory policies.

It is important to note that Haber et al. suggested that "*Case counts were modeled with a simple linear regressions, which is not consistent with infectious disease dynamics*". They failed to recognize that this linearity is not an assumption to "model" the pandemic trend in our work, but is the property from the publicly available data of the pandemic trend. The linearity is an important scientific finding from our analysis, reflecting a dynamic equilibrium between first-order transmission mechanisms and social distancing/stay-at-home measures as well as the combined second-order effects.

Throughout their open letter, Haber et al. provided no scientific evidence to invalidate our results (the five figures in our paper) and conclusions (airborne transmission represents the dominant route to spread the disease and face covering represents the determinant in shaping the pandemic trends) and to support their own assertions. At a minimum, they need to provide some evidence to illustrate what figures in our paper are incorrect and by what level of inaccuracy, by presenting their own results using their perceived reliable approach – 'case counts modeled using infectious disease dynamics'. The data to reproduce our work are readily obtainable on public domain. Haber et al. need to perform at least a single case study relevant to our paper to illustrate that their perceived approach is feasible and valid to model the COVID-19 pandemic. Also, they did not provide a single reference to point out the correct modeling method relevant to COVID-19, particularly that capable of dealing with the multi-growth stages for the total infections, as illustrated in **Attached Fig. 1**.

Haber et al. stated: "*While masks are almost certainly an effective public health measure for preventing and slowing the spread of SARS-CoV-2*", but provided no scientific evidence to justify their own claim.

Haber et al. asserted "*For this reason, we would argue that it may be time for PNAS to reconsider its policies on the Contributed Submissions track under which this paper was published, as this mechanism effectively bypasses editorial decisions and undermines peer review*", but rather selected a non-peer-reviewed path in the form of an open letter to engage in their own scientific debate.

In summary, Haber et al. failed to properly comprehend our work, denied our scientific evidence, proposed an infeasible and physically flawed framework to deal with COVID-19, and erroneously criticized our work without a scientific basis. Haber et al.'s claims referring to our paper to be "highly flawed" were made merely on the basis of their perceptions on what and how research should be conducted in modeling the COVID-19 pandemic trend, in accordance with empirical modeling. While the subject on the methodological approaches for understanding the COVID-19 pandemic as well as our method, results, and conclusions can certainly be debated in a legitimate and scientific fashion, their request for retraction of our paper is unjustified and unwarranted.

Mario J. Molina
Department of Chemistry and Biochemistry,
University of California - San Diego

Renyi Zhang
Departments of Atmospheric Sciences and Chemistry
Texas A&M University, College Station

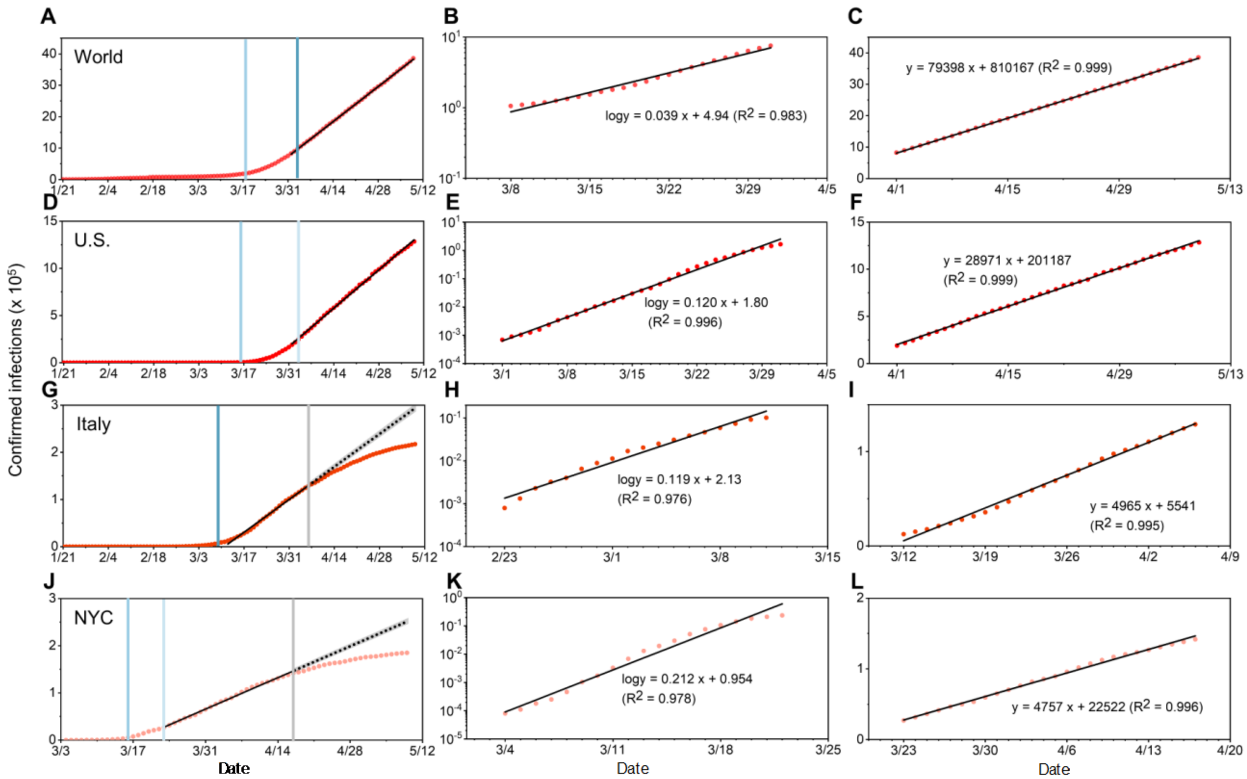

Fig. 1. Total confirmed infections in the World (A-C), U.S. (D-F), Italy (G-I), and NYC (J-L). The left column corresponds to the trends of total confirmed infections, and the middle and right columns correspond to the sub-exponential and linear growth portions, respectively. The vertical lines mark the dates for implementing mitigation measures: March 18 for physical distancing recommended by WHO and April 3 for lockdown/stay-at-home order for countries with 82% of the global total infections; March 16 for social distancing and April 3 for stay-at-home orders implemented in most states with 96% of total infections in the U.S.; March 9 for lockdown and April 6 for mandated face-covering in Italy; March 16 for social distancing, March 22 for stay-at-home order, and April 17 for mandated face-covering in NYC. The solid black line denotes linear regression through the data, and the dotted black line (G and L) denotes projection of infections without face covering based on linear regression for the data prior to the onset of mandated face covering. The gray shade (G and J) represents 95% confidence interval for the projection.

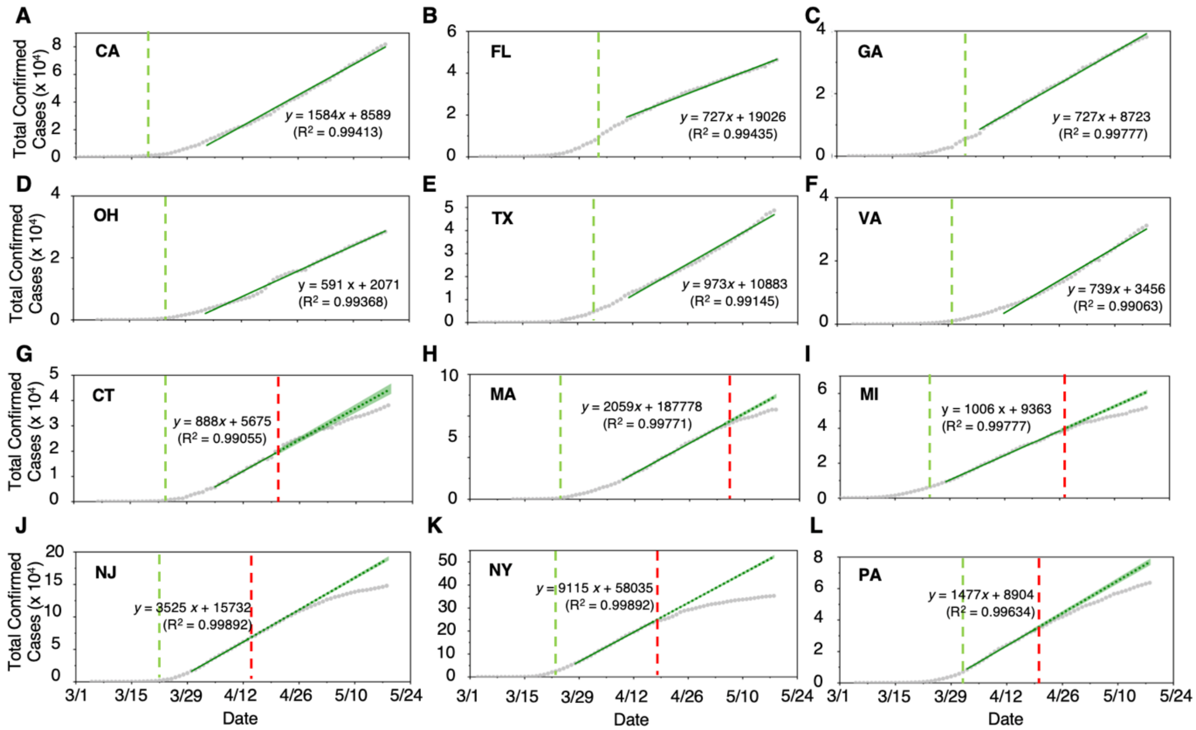

Fig. 2. Additional evidence for linearity and curve flattening in the total confirmed infections among 12 selected states of the 15-top infected states in the U.S. Among the 15 states, 9 and 6 states did and did not implement mandated face covering, respectively, which collectively accounted for about 78% of the total confirmed infections in the nation. Panels **A** to **F** correspond to all six states without mandated face covering, and panels **G** to **L** correspond to six states with mandated face covering. The vertical green and red dashed lines label the onsets for stay-at-home orders and mandated face covering, respectively. For comparison, guidelines for social distancing, quarantine, and isolation were issued by the federal government on March 16, 2020. The solid green line denotes linear regression through the data, and the dotted green line denotes projection of infections without face covering based on linear regression for the data prior to the onset of mandated face covering. The green shade (**G** to **L**) represents 95% confidence interval for the projection. Data taken from Li et al. (*Understanding transmission and intervention for the COVID-19 pandemic in the United States*, under submission).

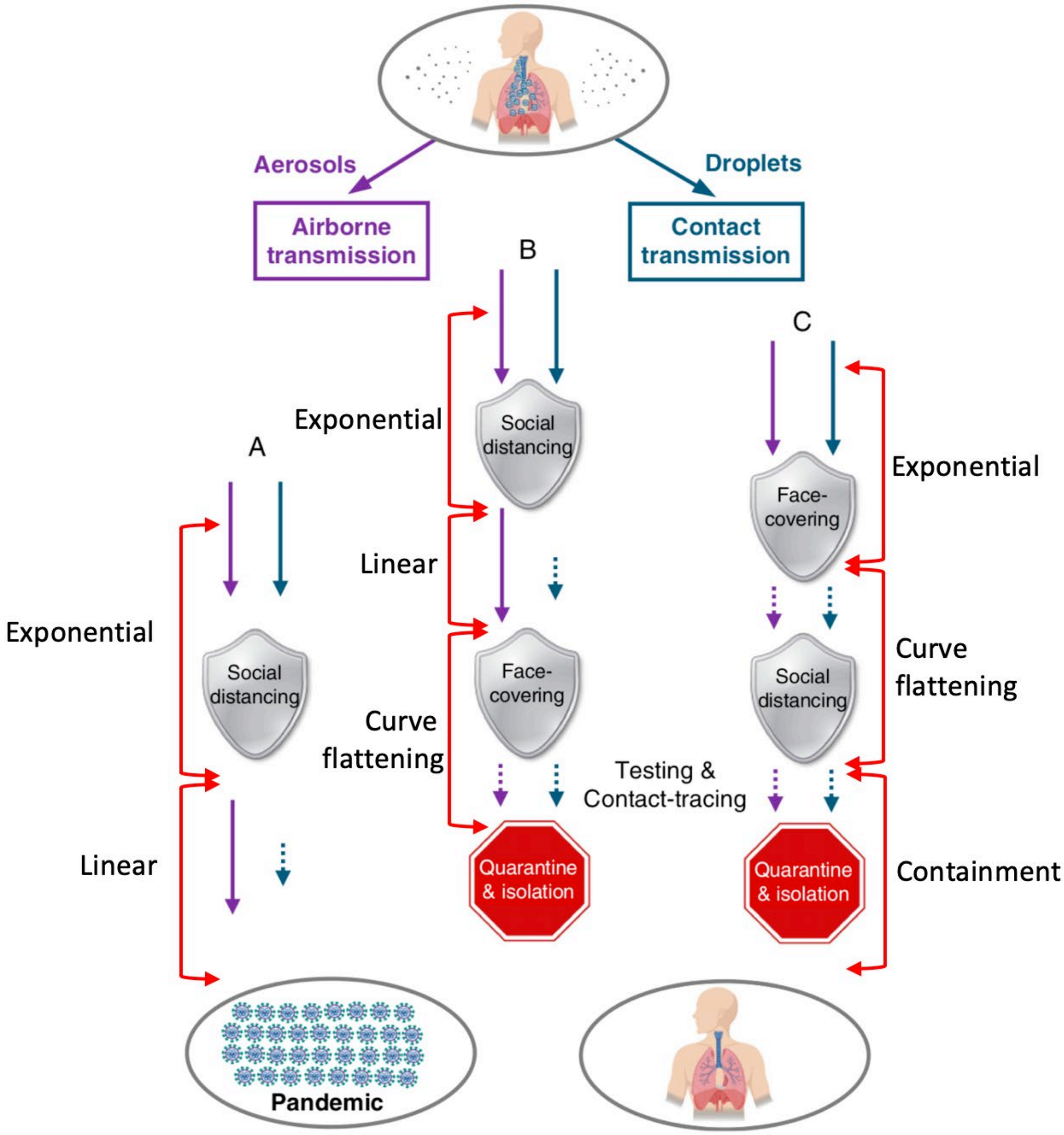

Fig. 3. Understanding the COVID-19 pandemic trends. Path A explains the pandemic trends in the world and the U.S. (Fig. 1C in Zhang et al., 2020) as well as in Panels A to F in attached Fig. 2. Path B explains the pandemic trends in Italy and NYC (Fig. 2B,C in Zhang et al., 2020) as well as in Panels G to L in attached Fig. 3. Path C explains the pandemic trends in China (Fig. 1B in Zhang et al., 2020) and Wuhan (Fig.2A in Zhang et al., 2020).

Table 1. Percentages of total infections of COVID-19 for countries with lockdown/stay-at-home orders globally and for U.S. states without stay-at-home orders by April 3. The data were from COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (https://coronavirus.jhu.edu/map.html, accessed on June 18).

* US denotes the sum of states in the US with stay-at-home orders.

	Total cases on April 3	Dates of lockdown/stay-at-home orders		Total cases on April 3
Major countries with lockdown/stay-at-home measures by April 3			**States in the U.S. without stay-at-home order by April 3**	
US*	266774	Mar 19~April 7	Alabama	1495
Italy	119827	Mar 9	Missouri	2124
Spain	119119	Mar 14	Oklahoma	990
Germany	91159	Mar 23	South Carolina	1700
China	82511	Jan 23	North Dakota	173
France	65202	Mar17	South Dakota	187
Iran	53183	Mar 14	Nebraska	279
UK	38689	Mar 23	Iowa	699
Belgium	16770	Mar 18	Arkansas	704
Austria	11524	Mar 16	Wyoming	162
Australia	5330	Mar 23	Utah	1248
Ireland	4273	Mar 12		
Russia	4149	Mar 28		
Malaysia	3333	Mar 18		
Pakistan	2818	Mar 24		
India	2567	Mar 25		
Thailand	1978	Mar 25		
Mexico	1688	Mar 23		
Greece	1613	Mar 23		
Peru	1595	Mar 16		
Percentage of total infections of COVID-19 for countries with lockdown/stay-at-home orders globally by April 3			**Percentage of total infections of COVID-19 for U.S. states without stay-at-home by April 3**	
Global infection			U.S. Infection	
Total:	1097000		Total:	276535
Percentage	82%		Percentage	4%