Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Information Sciences 561 (2021) 211–229

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

An investigation of testing capacity for evaluating and modeling the spread of coronavirus disease

Choujun Zhan [a], Jiaqi Chen [b], Haijun Zhang [b,*]

[a] School of Computing, South China Normal University, Guangzhou 510641, China
[b] Department of Computer Science, Harbin Institute of Technology, Shenzhen 518055, China

ARTICLE INFO

Article history:
Received 14 October 2020
Received in revised form 26 January 2021
Accepted 31 January 2021
Available online 16 February 2021

Keywords:
Testing capacity
COVID-19
Pandemic evaluation
Epidemiological model
Time series prediction

ABSTRACT

Despite the consistent recommendation to scale-up the testing of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), comprehensive analysis on determining the desirable testing capacity (*TC*) is limited. This study aims to investigate the daily *TC* and the percentage of positive cases over the tested population (*PPCTP*) to evaluate the novel coronavirus disease 2019 (COVID-19) trajectory phase and generate benchmarks on desirable *TC*. Data were retrieved from government facilities, including 101 countries and 55 areas in the USA. We have divided the pandemic situations of investigated areas into four phases, i.e., low-level, suppressing, widespread, or uncertain transmission phase. Findings indicate each country should increase TC to roughly two tests per thousand people each day. Additionally, based on *TC*, a susceptible-unconfirmed-confirmed-recovered (SUCR) model, which can capture the dynamic growth of confirmed cases and estimate the group size of unconfirmed cases in a country or area, is proposed. We examined our proposed SUCR model for 55 areas in the USA. Results show that the SUCR model can accurately capture the dynamic growth of confirmed cases in each area. By increasing *TC* by five times and applying strict control measures, the total number of COVID-19 patients would reduce to 33%.

1. Introduction

The novel coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a new infectious disease [16,41,44] with high transmissibility and a reproduction number ranging from 1.4 to 6.49 [40,28]. According to world health organization (WHO) reports, approximately 40% of COVID-19 patients will have mild symptoms; 40% will have moderate symptoms; 15% of patients will have severe symptoms; and 5% of patients will experience critical disease [6]. A large proportion of patients infected with COVID-19 are asymptomatic [33,14]. Some studies report that asymptomatic infections exceed 50–75% of the total number of infected people [11], while other studies indicate that the proportion of asymptomatic infections is approximately 20–30% [31,11,32]. The large group of asymptomatic patients play an important role in the spread of COVID-19 [4,19,36,1]. The average incubation period of COVID-19 is estimated to be approximately five days, with a maximum reported incubation period longer than two weeks [23,24,3,26]. Prior to symptom onset, pre-symptomatic (or asymptomatic) patients also have person-to-person transmissivity [16,44,35,47] and are acutely contagious during this asymptomatic incubation period [40]. Confirmed cases are normally isolated, and contact tracing is a common intervention for

* Corresponding author.
E-mail addresses: zchoujun2@gmail.com (C. Zhan), hjzhang@hit.edu.cn (H. Zhang).

https://doi.org/10.1016/j.ins.2021.01.084
0020-0255/

controlling infectious disease, which results in a low transmission ability of confirmed cases [18]. However, unconfirmed patients may disobey community-level measures to reduce contact between individuals and/or travel outside of their homes for the purpose of errands, entertainment, or exercise [6]. Modeling studies estimated that approximately 86% of infections were undetected and undocumented [25]. Unconfirmed patients may migrate from one city to another [21,45,49], and even from one country to another through the aviation network [7], leading to COVID-19 importations [15]. If no strict control measures are implemented and enforced, the spread of the epidemic will increase dramatically and transit to the phase of community transmission [27], which brings much greater challenges for epidemic prevention and control [9]. In addition, when there is a bona fide epidemic outbreak, a large number of patients experience severe or critical disease and require intensive medical care, which overwhelms medical systems and finally results in a high mortality rate [6,9]. Antibody tests suggest that there exist a large group of unconfirmed cases (including patients in the incubation period, mild patients, and asymptomatic patients) [30], which is the main reason for the rapid spread of the virus. In just a few months, the number of confirmed cases and deaths worldwide exceeded 7,000,000 and 110,000, respectively. COVID-19 has resulted in a global pandemic, as determined officially by the WHO, and presents a markedly severe health crisis [10].

A comprehensive empirical investigation to generate benchmarks on the scale of SARS-CoV-2 tests for risk assessment and trajectory evaluation is imperative for three primary reasons. First, countries or areas differ in pandemic progression, and thus sufficient testing capacity (*TC*) is essential to elucidate the progress of the pandemic [34]. For example, if the number of diagnoses in a certain area is low or stable, this may not necessarily mean that the epidemic is under control, i.e., most patients may remain unidentified due to inadequate testing power. Monitoring and documentation of daily *TC* and infected numbers over time enable an accurate trend analysis to determine COVID-19 trajectory [9]. Second, sufficient COVID-19 tests can provide a clear picture of how the epidemic has spread in a country or area. Then, policy-makers can initiate and implement effective public health responses according to specific local conditions, and further establish isolation and strategic triage functions for patients with COVID-19. Indeed, previous studies have demonstrated that isolation and contact tracing are the core of early prevention of the outbreak [18,25]. Inadequacy in *TC* impedes contact tracing and implementation of protective measures, as well as disease dissemination; whereas, sufficient *TC* can significantly reduce the speed of disease transmission and decrease the number of infected groups [39,34]. For example, 'blanket' testing was applied in an isolated Italian village of approximately 3000 people, and all confirmed cases were isolated. Remarkably, the number of daily confirmed cases decreased by more than 90% in less than ten days [11]. The average time of confirmation of COVID-19 patients also markedly influences control of the epidemic, in which even a short delay of confirmation of infected cases can significantly decrease epidemiological risks [18]. Third, it is important to estimate the number of unconfirmed cases in an area to control the epidemic [25]. Any disease transmission model should consider *TC*, and it should be able to estimate the size of the undiagnosed population and predict the direction of disease transmission [48].

Epidemiological models alert that 40–70% of the population could become infected unless protective measures are taken, highlighting the importance of early testing to identify infected populations [2,37]. The WHO and healthcare professionals around the world advocate for improving *TC* for SARS-CoV-2 to ascertain infection numbers, evaluate the impact of surveillance strategies, examine transmission dynamics, estimate reproductive numbers, evaluate epidemiological risks, and inform the development of practical and contextually-appropriate guidance to prevent onward transmission [34,38]. Some researchers suggested conducting population-based testing in a short period due to the highly contagious attribute of COVID-19 [34,25]. However, population-based testing is challenging to achieve because it is necessary to manufacture many reliable test kits, disseminate them to the entire population, and correctly collect, process, and evaluate the testing results. On the other hand, some guidelines recommended testing prioritization for individuals with symptoms, contact exposure (i.e., healthcare workers, suspected cases/clusters), vulnerable populations, and others according to the discretion of clinicians [13]. A large proportion of the infected population with mild, limited, or a lack of symptoms, and without access to healthcare, were overlooked, making the scope of the pandemic undeterminable [25]. Although large-scale testing is mostly advocated [11], the problem remains of deciding on the optimal number of tests performed by countries or areas.

At present, two kinds of tests exist for testing COVID-19. One type of test checks for the presence of the COVID-19 virus, aiming to establish whether an individual is currently infected. The other kind of test searches for the presence of antibodies, which can indicate whether an individual has been infected in the past or not, even if this individual has recovered and is not currently infected with the COVID-19 virus. A summary of the current state of testing technologies and their associated implementations can be found in [20]. Currently, a quantitative real-time reverse transcriptase-polymerase chain reaction (RT-PCR) assay is utilized for detecting SARS-CoV-2 from respiratory secretions as a definitive diagnosis of COVID-19 [42]. The RT-PCR test, however, cannot obtain 100% accuracy. For this reason, it is typical for the same individual to be tested for COVID-19 more than once. In addition, countries or areas are reporting testing data in different ways. For example, some areas report the number of tests performed, while other areas report the number of individuals tested or even combined data [43]. This difference has a significant consequence, in which people may be tested multiple times, and the number of tests that an individual receives is likely to vary across countries. As a result, no country or area can determine the number of people infected with COVID-19 with 100% accuracy. Indeed, it is only possible to know the infection status of those who have been tested. This is a serious constraint, of course, because the number of confirmed cases, as well as the number of individuals who have tested positive, only reflect part of the total number of infected individuals. We collected official up-to-date data released by the centers for disease control and prevention (CDC) of different countries and areas globally. We used testing data to investigate and evaluate the spread of COVID-19 in countries and areas. Our data comprise 101 countries (total population of 5.26 billion, 69.2% of the global population) and 55 areas in the USA (including 50 states, 4 unincorporated

territories and District of Columbia). By September 30, 2020, which was the end of our study period, 33,785,178 infection cases and 1,010,147 deaths were verified globally. There were 405,140,000 occurrences of testing performed from these 101 countries, which constitute an average of 72.70 tests per thousand people. The USA, specifically, has performed 107,300,299 tests, which are an average of 327.9 tests per thousand people.

Evaluation of COVID-19 trajectory and *TC* must be evidence-based. Real-time, accurate data on the testing of suspected cases and confirmed cases will inform decision-making [17]. This study aims to generate a recommendation on the desirable *TC* for a country or area to assess pandemic trajectory and inform health response. We employed data-driven approaches based on dynamic profiles of *TC* and the percentage of positive cases over the tested population (*PPCTP*). The *TC* was operationalized as a proportional concept, regarding tested numbers over the population, considering diversified people worldwide. This study also summarizes the *TC* and *PPCTP* results of areas with different COVID-19 transmission phases, ranging from low-level, suppression, widespread, to uncertain. Experimental results indicate that *TC* in different countries is uneven. The pandemic situation of 101 countries in the six continents is also different. By September 30, 2020, Europe and Oceania controlled the spread of COVID-19 well and Asia slowed down the spread, while North and South America attempted to suppress transmission. However, most of the African and Asian countries should scale-up *TC*. Then, we take advantage of detailed data about the testing number of countries or areas to develop a *TC*-based model to estimate unconfirmed cases and evaluate the *TC*. A susceptible-unconfirmed-confirmed-recovered (SUCR) model is proposed to predict the spread and distribution of COVID-19 infections across the world and 55 areas in the USA. Furthermore, it can also evaluate whether or not the *TC* in a country or area is sufficient. Results indicate that the total number of COVID-19 patients in the USA would reduce to 33% by increasing *TC* by five times and applying strict control measures.

The rest of this paper is organized as follows. In Section 2, we propose a data-driven method to evaluate the pandemic situation based on the *TC* and *PPCTP*. In Section 3, a new SUCR model considering *TC* is proposed and applied to estimate the number of confirmed and unconfirmed cases in 55 prefectures in the USA. Finally, conclusion and future work propositions are presented in Section 4.

2. Pandemic situation evaluation method

2.1. An illustrative example

At day t, the official pandemic data released by a country or area always include: the number of active confirmed cases $C(t)$; the number of recovered cases $R_c(t)$; and death toll $D_c(t)$. Then, the cumulative confirmed cases are $I_c(t) = C(t) + R_c(t) + D_c(t)$. In addition, $N_T(t)$ represents the number of COVID-19 tests performed in a country or area until day t; N_p is the population of this country or area. Then, $TC(t)$ and $PPCTP(t)$ can be defined as:

$$TC(t) = \frac{N_T(t)}{N_p} \times 1000, \tag{1}$$

$$PPCTP(T) = \frac{\Delta I_c(t)}{N_T} \times 100\%. \tag{2}$$

where $\Delta I_c(t) = I_c(t) - I_c(t-1)$.

From the Middle March to early June 2020, New York and the UK was one of the epicenters in the USA and Europe, respectively. By September 30, 2020, Chile was still experiencing widespread transmission, while Pakistan had a relatively limited testing ability so that the actual pandemic situation was rather vague. Korea successfully contained the spread of COVID-19 since April 2020. The pandemic situation in these areas is representative, so these areas are adopted as illustrative examples. First, we demonstrate $TC(t)$ in different areas (see Fig. 1a), while $PPCTP(t)$ in these areas is shown in Fig. 1b. It is found that New York State, the UK, Pakistan, Chile, and Korea enhanced their $TC(t)$ gradually. In Fig. 1c, the size of circles represents $TC(t)$, while the size of the solid circle stands for the corresponding $PPCTP(t)$. Note that the increase of circle size means the increase of $TC(t)$ or the increase of $PPCTP(t)$, or vice versa. $TC(t)$ and $PPCTP(t)$ variations in these areas reveal the pandemic situation's evolution and management.

Most countries or areas tested individuals who had COVID-19 symptoms, such as fever, cough, shortness of breath, and those who have recently had close contact with COVID-19 infections even though they do not show symptoms [12]. Therefore, a portion of tested individuals should not have been infected by COVID-19 but tested. Here, we assume that the number of individuals who develop COVID-19 symptoms or have close contact with COVID-19 patients is $N_s(t)$, while the number of active unconfirmed COVID-19 infections at time t is $U(t)$. At the beginning of an outbreak, the capacity for testing in an area is always relatively insufficient. Tests may be reserved (or "rationed") for particular high-risk groups. In this case, even though the tested number and the confirmed number are small, the number of confirmed cases almost equals the number of tests, i.e., $PPCTP(T) \approx 100\%$. As a result, a small *TC* with a high $PPCTP(t)$ indicates that the virus has been spreading and is moving towards widespread transmission. For this reason, a small number of confirmed cases is not equal to low-level transmission. Indeed, a small *TC* with a high $PPCTP(t)$ is a danger signal. Moreover, an area with a large $PPCTP(t)$ may indicate that there is not sufficiently broad testing for obtaining a reliable result of the virus's actual spread. While people with highly severe symptoms may have been tested in such countries or areas, there are likely to be many times more people with mild

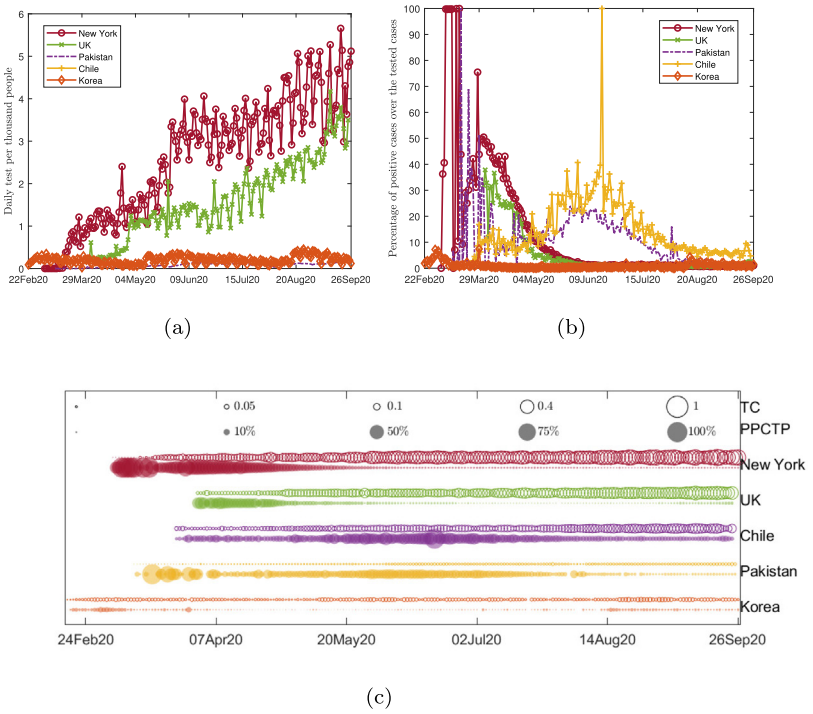


Fig. 1. Pandemic situations of New York State, the UK, Pakistan, Chile, and Korea: (a) the number of daily COVID-19 tests per thousand people in a country or area, $TC(t)$; (b) the percentage of positive cases over the tested population $PPCPT(t)$; and (c) an overview of pandemic situations in terms of $TC(t)$ and $PPCPT(t)$.

or no symptoms that were never tested. Therefore, as the virus infects more people, *TC* should be expanded to provide a reliable estimation of the correct number of infected people. More patients with suspected COVID-19 infection would be tested with an escalation of *TC*. When the widespread transmission has been controlled, *TC* should remain stable. With a decrease in the number of COVID-19 infections, the number of suspected patients with COVID-19 symptoms may not change. As a result, $PPCTP(t)$ will decrease significantly. For instance, New York State and the UK demonstrate an increase in *TC* after mid-April 2020, while the confirmed cases in these two areas decrease correspondingly. This indicates that COVID-19 is suppressed in these two areas by September 30, 2020. However, the situation in Pakistan appears unstable, because it shows inadequate *TC* and high $PPCTP(t)$ values. Korea has less than 0.3 confirmed cases per hundred tests. However, Korea has a relatively low *TC* of 0.2 tests per thousand individuals. Consequently, we suggest that Korea increases its *TC* to avoid another strike of COVID-19. Hence, we can summarize the following observations.

- Initially, a relative lack of awareness of the virus and few countermeasures prevented the spread of COVID-19, which always results in a relatively insufficient $TC(t)$. High positive results would manifest the prominent insufficiency in $TC(t)$ among the tested population in practice since only people with apparent symptoms are tested [9,12]. For this reason, a small $TC(t)$ with a high $PPCTP(t)$ indicates that the virus has been spreading and is moving towards the widespread transmission, which means that $TC(t)$ should be scaled-up.
- If the $TC(t)$ is sufficient, COVID-19 infections with mild or no symptoms, and individuals who have COVID-19 symptoms or who only had contact with COVID-19 patients, would be tested. We can assume that patients with the COVID-19 virus and patients with only COVID-19 symptoms are homogeneously mixed. Then, the increase in cumulative confirmed patients is

$$\Delta I_c \approx N_T(t)\frac{\Delta U(t)}{\Delta U(t)+N_s}, \tag{3}$$

$$PPCTP(T)=\frac{\Delta I_c}{N_T(t)}\approx\frac{U(t)}{U(t)+N_s}. \tag{4}$$

Note that the larger is the number of active unconfirmed cases $U(t)$, the larger is the $PPCTP(t)$. As a result, $PPCTP(t)$ is a good indicator of the group size of unconfirmed cases. If $PPCTP(t)$ decreases, this may indicate that the spreading is slowing down; whereas, a large $PPCTP(t)$ may mean that the spreading is increasing, and the government should scale-up testing to identify unconfirmed cases.

- $\Delta I_c(t) \leqslant U(t)$, which indicates that

$$\frac{\Delta N_T(t)}{\Delta U(t) + N_s \leqslant 1} \Rightarrow \Delta N_T(t) \leqslant \Delta U(t) + N_s. \tag{5}$$

$$PPCTP(t) = \frac{\Delta I_c(t)}{N_T} \times 100\%, \tag{6}$$

Therefore, there exists a threshold $N_{T,thr} = N_s + \Delta U(t)$, and if the TC is larger than $N_{T,thr}$, then most of the unconfirmed cases should be identified.

- If the TC is insufficient or $U(t)$ does not exhibit a clear trend (increasing or decreasing), then, the $PPCTP(t)$ would be not stable, which means that $PPCTP(t)$ also would not show a clear trend of either increasing or decreasing.

2.2. Evaluate the pandemic situation in a prefecture

To further demonstrate the importance of the TC for evaluating the spread of COVID-19 and the epidemiological risks that it poses in different populations in comparison with solely examining the number of confirmed cases in a region, we summarize the results with respect to these two metrics over 101 countries from six continents. Based on the WHO report, the pandemic situation is classified into four phases: ''low-level transmission", ''suppressing transmission", ''widespread transmission", and ''uncertain". The pandemic stage classification is based on the TC and $PPCTP$ of the latest ten days of the study period, which takes the incubation and symptom onset period of COVID-19 into account [3,23,26]. First, a threshold of TC is adopted. If the last ten days' average TC in a country or area is smaller than TC_{thr}, the epidemiological risk of this country or area is deemed unreliable, as the TC is too low to achieve a reliable result. We determine all of the pairs of $\{TC(t), PPCTP(t)\}$, in which all $PPCTP(t) \geqslant 0.5$, which means that the percentage of positive cases over the tested population is larger than 50%. We then derive the 95% confidence interval(CI) of $TC(t)$. The results show that 95% CI is [0.1284, 0.1535]. Therefore, in this study, we adopt $TC_{thr} = 0.1535$. Additionally, we adopt two thresholds (or cut off) referring to the official report of WHO [9]. $PPCTP_c$ is used as cut off for ''low-level transmission", while $PPCTP_s$ is used as cut off for ''suppressing transmission". First, we derive the $PPCTP(t)$ from all the countries and areas to achieve a set of $\{PPCTP(i)|i = 1, 2, \cdots, M\}$. Then, we rank these numbers from smallest to largest, namely, $PPCTP(1) \leqslant PPCTP(2) \leqslant \cdots \leqslant PPCTP(M)$. Next, we calculate $n_c = \lceil \alpha_c M \rceil$ and $n_s = \lceil \alpha_s M \rceil$, where $0 < \alpha_c < \alpha_s < 1$, $\lceil x \rceil$ is a ceiling function mapping x to the least integer greater than or equal to x. Here, we adopted $\alpha_c = 0.2$ and $\alpha_s = 0.4$. In this case, 80% of $PPCTP(t)$ larger than $PPCTP_c = 0.82$, while only 20% of $PPCTP(t)$ is less than 0.80. Similarly, 2.34 is chosen as a threshold because 60% of countries have $PPCTP(t)$ larger than $PPCTP_s$. These thresholds are constant. Additionally, the classification results are robust to slightly variation of $PPCTP_c$, $PPCTP_s$ and time window. In conclusion, the classification standard is as follows:

1. **Low-level transmission:** Suppose a country or area has an average TC larger than TC_{thr} tests per thousand people in the latest ten days. In that case, it is considered as low-level transmission when the latest ten days' $PPCTP(t)$ is smaller than a threshold $PPCTP_c$. Here, 0.80% is adopted as the threshold $PPCTP_c$.
2. **Suppressing transmission:** The pandemic situation is classified as suppressing when at least 80% of the latest ten days exhibited a negative $\Delta PPCTP(t) = PPCTP(t) - PPCTP(t-1)$, or at least 80% of the latest ten days showed a positive rate of less than $PPCTP(t)_s\%$.
3. **Widespread transmission:** When at least 80% of the latest ten days showed a positive $\Delta PPCTP(t)$ or at least 80% of the latest ten days exhibited a positive rate above $PPCTP_s(t)\%$, the epidemiological situation is considered to be a widespread transmission.
4. **Uncertain:** Any other situation.

The classification procedure is summarized in Algorithm 1. To further demonstrate the importance of the TC for evaluating the spread of COVID-19 and the epidemiological risks in different populations, 101 countries from six continents were chosen for the investigation.

Algorithm 1: Epidemiological risk classification

Require:The tested capacity and percentage of positive cases over tested population of an area, namely, $TC(t)$ and $PPCTP(t)$.

ensure

The epidemiological risk in an area.

Initialization:

$$\begin{aligned} \Delta PPCTP(t-9) &= PPCTP(t-9) - PPCTP(t-10) \\ \Delta PPCTP(t-8) &= PPCTP(t-8) - PPCTP(t-9) \\ &\vdots \\ \Delta PPCTP(t) &= PPCTP(t) - PPCTP(t-1) \end{aligned} \tag{6}$$

Classification Process

if $min\{TC(t-10), TC(t-9), \cdots, TC(t)\} < TC_{thr}$

The epidemiological risk is classified as "uncertain".

else $[max\{PPCTP(t-10), PPCTP(t-9), \cdots, PPCTP(t)\} \leqslant PPCTP_s]$

The epidemiological risk is classified as "low-level transmission".

else $[PPCTP_c \leqslant max\{PPCTP(t-10), PPCTP(t-9), \cdots, PPCTP(t)\} \leqslant PPCTP_s]$

The epidemiological risk is classified as "suppressing transmission".

else [The number of negative $\Delta PPCTP(t-i)$ larger than 7]

The epidemiological risk is classified as "suppressing transmission".

else $[PPCTP_s \leqslant min\{PPCTP(t-10), PPCTP(t-9), \cdots, PPCTP(t)\}]$

The epidemiological risk is classified as "widespread transmission".

else [The number of positive $\Delta PPCTP(t-i)$ larger than 7]

The epidemiological risk is classified as "widespread transmission".

else [Otherwise]

The epidemiological risk is classified as "uncertain".

end if

Return The epidemiological risk classification of an area: "low-level transmission", "suppressing transmission", "widespread transmission", or "uncertain".

2.3. The pandemic situation of 101 countries

2.3.1. Uneven TC in different countries

We checked and collected the testing data from national public health institutes in each country, while population data were from Bureau of Statistics of each country. *TC* in different countries is uneven (shown in Fig. 2). The average *TC* of the 101 countries is 0.6941 tests per thousand people. However, 37 countries (36.63%) are above this level, while 64 countries (63.37%) are lower than this level.

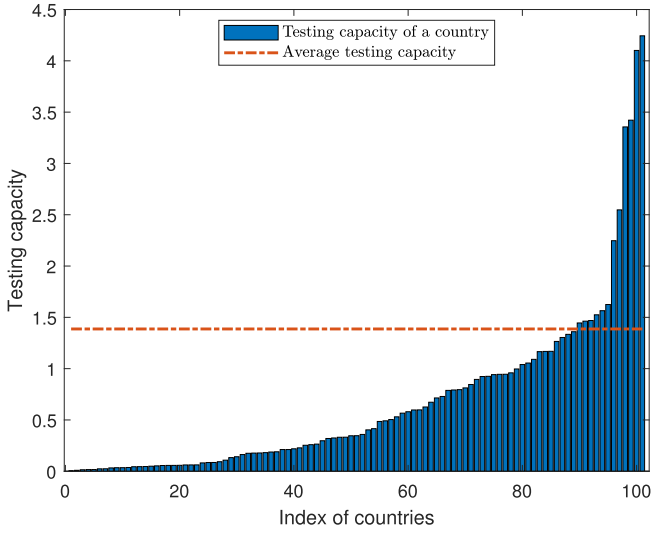


Fig. 2. *TC* in different countries. The red dashed line represents the average *TC* of the 101 countries. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Number of countries in different epidemiological statuses in six continents.

Continent (No. of countries)	Low-level	Suppressing	Widespread	Uncertain
Asia (27)	2	2	12	11
Europe (34)	3	6	9	15
Oceania (3)	2	0	0	1
North America (9)	1	0	5	3
South America (8)	0	0	5	3
Africa (20)	0	0	5	15

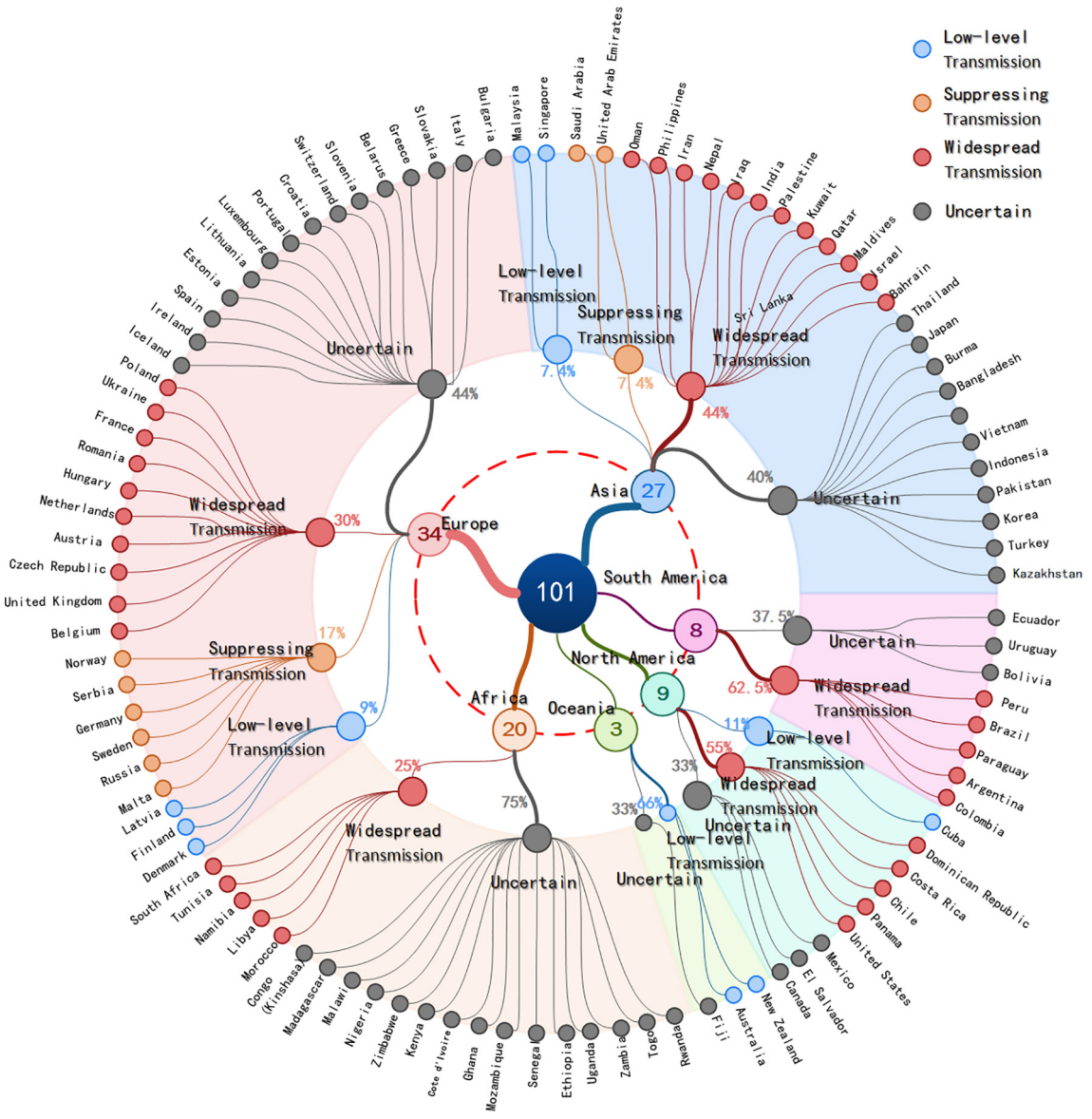


Fig. 3. Classification of the pandemic situation of 101 countries in six continents on September 30, 2020.

2.3.2. Classification of pandemic situation of 101 countries in six continents

Based on the proposed data-driven approach, 101 countries were classified into four categories: low-level transmission, suppressing transmission, widespread transmission, and uncertain situation. We summarize the results with respect to these two metrics over 101 countries from six continents (see Table 1 and Fig. 3). Note that the spread of the pandemic in each country varies moment-to-moment. Our results are derived from data before September 30, 2020. Table 1 shows the number

of countries on six continents in various stages of the pandemic. Based on the aforementioned classification method, the pandemic situation of only 9 countries was found to be the low-level transmission, including 2 Asian countries (Malaysia, Korea), 2 Oceania countries (Australia and New Zealand), 3 European countries (Latvia, Finland, and Denmark), 1 North American country (Cuba), and one African country (Rwanda). These countries have a comparatively higher $TC(> 1.264)$ and a lower $PPCTP(< 3.027\%)$ in the latest ten days by September 30, 2020. A total of 8 countries were classified as suppressing transmission, including 2 Asian countries (Saudi Arabia and the United Arab Emirates), 6 European countries (Norway, Serbia, Germany, Sweden, Russia, and Malta). These countries demonstrate a moderate *TC*, and a decreasing or stable trend of COVID-19 morbidity rate. 37 countries were in the widespread transmission stage, including 12 Asian countries (Oman, Philippines, Iran, Nepal, Iraq, India, Palestine, Kuwait, Qatar, Maldives, Israel, and Bahrain), 9 European countries (Poland, Ukraine, France, Romania, Hungary, Netherlands, Austria, Czech Republic, United Kingdom, and Belgium), 5 North American countries (Dominican Republic, Costa Rica, Chile, Panama, and the USA), 5 South American countries (Peru, Brazil, Paraguay, Argentina, and Colombia), and 5 African countries (South Africa, Tunisia, Namibia, Libya, and Morocco). These countries have a comparatively lower *TC* and higher *PPCTP*. Moreover, the increasing trend of COVID-19 morbidity demonstrated by countries in this category indicates a pandemic outbreak. Finally, 48 countries demonstrated an uncertain status, including 11 Asian counties (Thailand, Japan, Burma, Bangladesh, Sri Lanka, Vietnam, Indonesia, Pakistan, Korea, Turkey, and Kazakhstan), 15 European countries (Bulgaria, Italy, Slovakia, Greece, Belarus, Slovenia, Switzerland, Croatia, Portugal, Luxembourg, Lithuania, Estonia, Spain, Ireland, and Iceland), 1 Oceania country (Fiji), 3 North American countries (Mexico, El Salvador, Canada), 3 South American countries (Ecuador, Uruguay, and Bolivia), and 15 African countries (Madagascar, Malawi, Nigeria, Zimbabwe, Kenya, Cote d'Ivoire, Ghana, Mozambique, Senegal, Ethiopia, Uganda, Zambia, Togo, and Rwanda). The disease outbreak of these countries was unclear, which is why they were categorized as uncertain situation.

In conclusion, the results show that many European countries (9 of 33) have controlled the spread of the epidemic, while most North and South American countries (16 of 17) were experiencing widespread transmission or the uncertain stage, and only one country was in a low-level and suppressing transmission state. African countries require improving their *TC*, as 15 of 20 countries were in an uncertain stage. In Asia, 2 of 27 countries contained the spread of the pandemic and had a low-level transmission, while 12 countries were in the widespread transmission stage. More than one-third of Asian countries (11 of 27) were in an uncertain stage. In total, there were 48 countries in an uncertain stage, i.e., 47.5% of the 101 countries

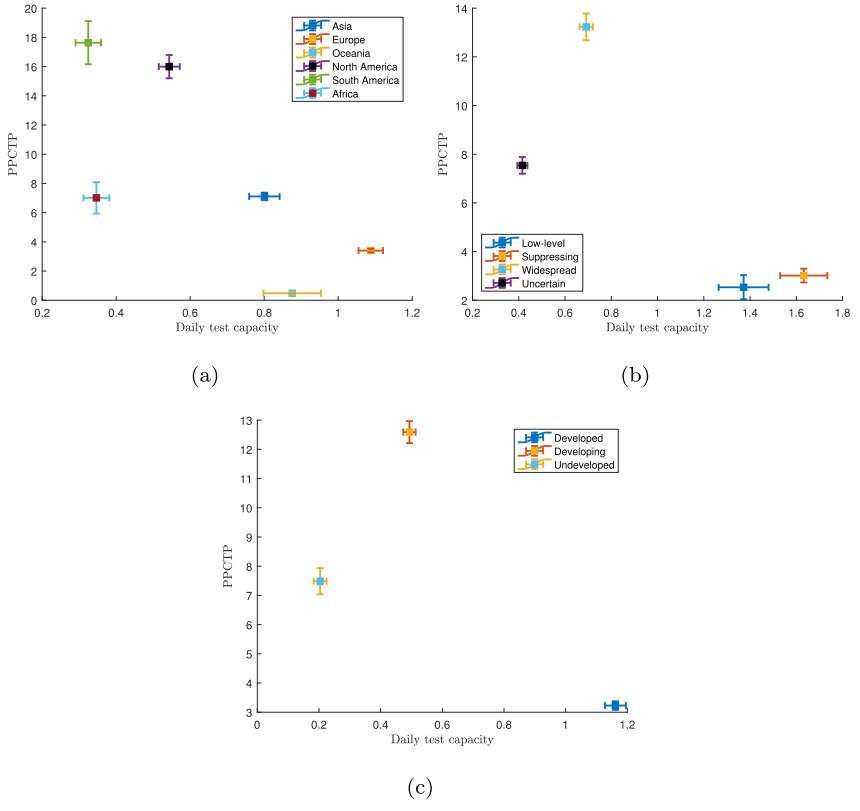


Fig. 4. Results of daily *TC* against *PPCTP* (Squares represent the mean values, while the horizontal and vertical bars indicate the lower and upper bound of the 95% CI, respectively): (a) The mean values and 95% CI of *TC* and *PPCTP* of countries in six different continents; (b) The mean values and 95% CI of *TC* and *PPCTP* of countries in four different situations; (c) The mean values and 95% CI of *TC* and *PPCTP* of countries in developed, developing, and undeveloped countries.

had insufficient *TC*, or the pandemic situation of these countries was unclear. Countries with a higher *TC* and lower *PPCTP* than expected might indicate more successful public health responses. In conclusion, the *TC* and *PPCTP* changes of 101 countries from Feb. 4 to Sep. 30, 2020, allow us to accurately monitor, evaluate, and compare pandemic development and containment nationally and internationally.

2.3.3. Statistical results of TC and PPCTP in six continents

Fig. 4a shows the mean values and 95% CIs of *TC* and *PPCTP* of countries in different continents, including Asia, Europe, Oceania, North America, South America and Africa, while the values are shown in Table 2. The results demonstrate that countries in Oceania have best controlled the COVID-19 pandemic. The countries in Oceania have the largest mean *TC*, with a mean value of 0.8557 tests per thousand people each day, while the *PPCTP* is only 0.4799%. Most of the European countries have also suppressed the spread of COVID-19. The mean *TC* in Europe is 1.0881, while the mean *PPCTP* is 3.4010%. Asia has slowed down the spread of COVID-19. However, a larger proportion of Asian countries should increase their *TC*, since the *TC* is relatively low, with an average of 0.8006 tests per thousand people each day. North and South America are currently suffering from the pandemic. The *TC* of North America is 0.5334 tests per thousand people, while the *TC* of South America is only 0.3245 tests per thousand people, which is the lowest *TC* in the six continents. Moreover, the *TC* of Africa is only 0.3245. In conclusion, the pandemic situation in the six continents is different. Europe and Oceania have controlled the spread of COVID-19 the best, while North and South America are attempting to suppress transmission. Asia has slowed down the spread of COVID-19. Many Asian and African countries should scale-up their *TC*.

2.3.4. Statistical results of TC and PPCTP of four phase countries

Fig. 4b shows the mean values and 95% CIs of *TC* and *PPCTP* of countries in four different statuses, including low-level transmission, suppressing transmission, widespread transmission, and uncertain stage. Table 3 presents the *TC* and the *PPCTP* values for countries with different COVID-19 containment statuses. Countries with good pandemic containment demonstrate an average of 2.5328% positive results among the tested population, with a 95% CI of 2.0386% to 3.0270%. The average *TC* among these countries with desirable pandemic control was 1.3721 [95% CI: 1.2642 to 1.4500]. Countries that show a suppressing trend of pandemic management have an average of 3.0127% positive results, with a 95% CI of 27293% to 3.2961%. The average *TC* among these countries was 1.6325 [95% CI: 1.5305 to 1.7345]. Countries that remain in the stage of the pandemic outbreak have an average of 13.2316% positive results, with a 95% CI of 12.6836% to 14.7796%. The average *TC* among countries classified as being in the outbreak stage was 0.6916 [95% CI: 0.6629 to 0.7202]. Countries categorized as uncertain situation in terms of the pandemic situation have the lowest *TC* (mean = 0.4157 [95% CI: 0.3938 to 0.4375]). The limited number of tests made the estimation of the pandemic situation in these countries challenging. From these results, however, we can empirically make a suggestion that the *TC* should be 2 (rounded the results from of the average *TC* 1.6325 of countries suppressing the spread) test per thousand people each day. Only 7 countries pass this level, while other countries do not have sufficient *TC*. If each country has a *TC* larger than two tests per thousand people each day, a total of 22,290,000 should be tested each day. As of September 30, 2020, only 8,361,389 tests are conducted each day, reaching 37% of the recommended testing scale. However, since the *TC* is uneven, the actual situation is much worse than this.

If one test costs 100 USD and the *TC* of the world is two people per thousand people a day, the total cost of testing each day reaches 1.113 Billion USD and 122.04 billion USD annually. The world GDP was 87.698 trillion USD in 2019. Promoted *TC* only costs 0.39% of the world's GDP. However, if we do not increase *TC*, it may be necessary to lock down the whole world. Indeed, the economic loss of a global lockdown is much higher than the cost of worldwide COVID-19 testing.

Table 2
TC and *PPCTP* in six continents.

Continent	Mean (*TC*)	95%CI (*TC*)	Mean (*PPCTP*)	95%CI (*PPCTP*)
Asia	0.8006	[0.7597,0.8416]	7.1082%	[6.8517,7.3646]%
Europe	1.0881	[1.0550,1.1213]	3.4010%	[3.2796,3.5224]%
Oceania	0.8557	[0.7981,0.9534]	0.4799%	[0.3728,0.5870]%
North America	0.5334	[0.5150,0.5717]	15.9961%	[15.1956,16.7966]%
South America	0.3245	[0.2901,0.3590]	17.6341%	[16.1571,19.1112]%
Africa	0.3465	[0.3116,0.3815]	7.0090%	[5.9387,8.0794]%

Table 3
TC and *PPCTP* among countries with different pandemic control statuses.

Pandemic situation	Mean (*TC*)	95%CI (*TC*)	Mean (*PPCTP*)	95%CI (*PPCTP*)
Low-level	1.3721	[1.2642,1.4500]	2.5328%	[2.0386,3.0270]%
Suppressing	1.6325	[1.5305,1.7345]	3.0127%	[2.7293,3.2961]%
Widespread	0.6916	[0.6629,0.7202]	13.2316%	[12.6836,13.7796]%
Uncertain	0.4157	[0.3938,0.4375]	7.5393%	[7.1978,7.8809]%

Table 4
Statistics on the pandemic situation in developed, developing, and undeveloped countries.

Continent (No. of countries)	Low-level	Suppressing	Widespread	Uncertain
Developed (33)	**6**	**4**	8	15
Developing (57)	2	4	**27**	**24**
Undeveloped (11)	0	0	2	9

2.4. Pandemic situation in developed, developing, and undeveloped countries

According to the World Bank, the 101 countries are classified into developed, developing, and undeveloped countries [8]. Table 4 shows the classification results. Many developed countries have controlled the spread of COVID-19. Based on the proposed method, the pandemic situation of 6 developed countries was considered low-level transmission. A total of 4 developed countries have suppressed the spread of COVID-19, while 8 countries are in the widespread transmission phase, and 15 countries are in the uncertain stage. Furthermore, only 2 developing countries are in the low-level transmission stage, and 4 developed countries have suppressed the spreading of COVID-19. However, 27 developing countries are in the widespread transmission stage. 42% of developing countries (24 out of 57 countries) have an insufficient *TC*, which means that most developed countries should increase their *TC*. The pandemic situation is uncertain concerning the undeveloped countries, with 9 out of 11 undeveloped countries being in the uncertain phase. Therefore, 48% of developing and undeveloped countries (33 out of 68) are in the uncertain stage. Developed counties have the largest mean *TC*, with a mean value of 1.1612 tests per thousand people each day, while the *PPCTP* is only 3.2288% (shown in Table 5). Most of developing countries should increase their *TC*, since the *TC* is relatively low, with an average of 0.4933 tests per thousand people each day. The mean *TC* of undeveloped countries is only 0.2037 tests per thousand people, while the *PPCTP* is 7.4842%. Undeveloped countries have a low *TC*.

2.5. The Pandemic situation of 55 areas in the USA

2.5.1. TC in different areas is uneven

Similar to the above analysis method, we further summarize the pandemic results of *TC* and *PPCTP* over the 55 areas of the USA. Obviously, *TC* is markedly uneven among states (see Fig. 5). The average *TC* is 1.5045 tests per thousand people. Moreover, 21 areas are above this level, while 34 areas are lower than this level.

2.5.2. Classification of the pandemic situation of 55 areas

We divide the 55 states into two categories based on the governor's party: the governor belongs to the Republican Party, and the governor belongs to the Democratic Party. Note that the pandemic spreading in each area varies moment-to-moment, and these results are only derived from data before September 30, 2020. Table 6 shows the number of areas in various pandemic stages. Results show that 8 and 37 areas were in suppressing and widespread stages, respectively; whereas, only 3 areas were in the low-level stage. Fig. 6 summarizes the *TC* and *PPCTP* of the 55 areas from Feb 28 to September 30,

Table 5
TC and *PPCTP* in developed, developing and undeveloped countries.

Continent	Mean (*TC*)	95%CI (*TC*)	Mean (*PPCTP*)	95%CI (*PPCTP*)
Developed	1.1612	[1.1271,1.1953]	3.2288%	[3.0833,3.3743]%
Developing	0.4933	[0.4733,0.5133]	12.5899%	[12.2143,12.9655]%
Undeveloped	0.2037	[0.1827,0.2247]	7.4842%	[7.0355,7.9329]%

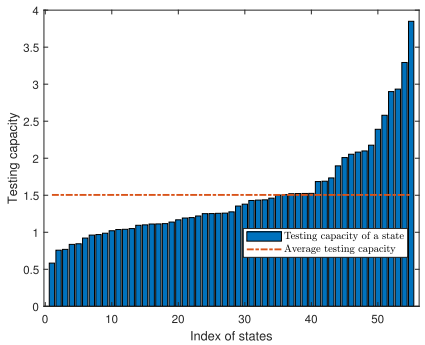


Fig. 5. The average *TC* of 55 areas in the USA from Feb 27, 2020 to September 30, 2020.

Table 6
Pandemic status of 55 areas in the USA.

Party (No. of areas)	Low-level	Suppressing	Widespread	Uncertain
Republican (28)	2	2	19	5
Democratic (27)	1	6	18	2

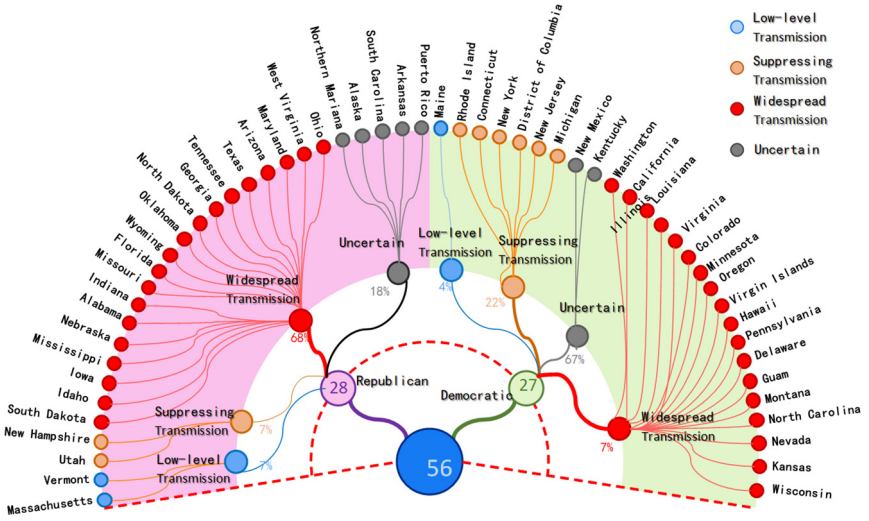


Fig. 6. Classification of the pandemic situation of 55 areas in the USA on September 30, 2020.

2020. The size of the bubbles is proportional to the value of *PPCTP*, while the size of the circles is proportional to the value of *TC*. Only 3 areas were considered as low-level transmission. These states have a comparatively higher $TC(> 0.064)$ and a lower $PPCTP(< 0.43\%)$ in the latest 10 days. A total of 8 areas were classified as suppressing transmission, including 2 Republican areas (Vermont and Alaska) and 6 Democratic areas (New York, Montana, Connecticut, Hawaii, New Jersey, and Rhode Island). These areas demonstrate a moderate *TC* and a decreasing or stable trend of COVID-19 morbidity rate. 37 areas were in the widespread transmission stage, including 19 Republican states (Ohio, Indiana, Maryland, Oklahoma, Nebraska, Missouri, Iowa, South Dakota, Tennessee, Arkansas, Utah, Idaho, Georgia, South Carolina, Texas, Florida, Alabama, and Arizona), and 18 Democratic areas (Minnesota, Oregon, Kentucky, Delaware, Pennsylvania, Colorado, Virginia, California, Wisconsin, Louisiana, North Carolina, Kansas, and Nevada). These areas have a comparatively lower *TC* and a higher *PPCTP*. In addition, the increasing trend of COVID-19 morbidity demonstrated by areas in this category indicates a pandemic outbreak. Finally, 7 areas demonstrated an uncertain pandemic evolution status, including 5 Republican areas (Northern Mariana Islands, Mississippi, Wyoming, and Puerto Rico), and 2 Democratic areas (Virgin Islands, Guam, New Mexico, and Washington). The disease outbreak trend was unclear, which caused these countries to be categorized as uncertain in terms of the pandemic situation.

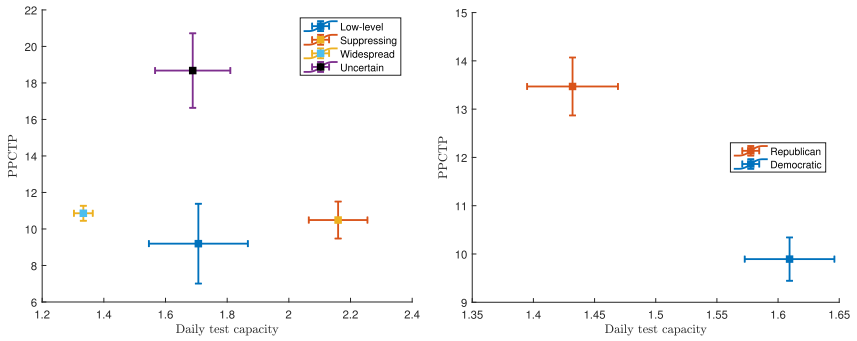
In conclusion, the results show that 4 of 28 Democratic areas have suppressed the spread of the epidemic, while 7 of 27 Democratic areas have suppressed the spread of the epidemic. More than half of the USA areas were experiencing widespread transmission, including 19 Republican areas and 18 Democratic areas. There were 7 areas in the uncertain stage, including 5 Republican areas and 2 Democratic areas. In conclusion, the pandemic situation in Democratic states was slightly better than in Republican areas.

2.5.3. Statistical results of TC and PPCTP of countries in different phases

Fig. 7a and Table 7 show the mean values and 95% CIs of *TC* and *PPCTP* of 55 states in four different statuses, including low-level, suppressing, widespread, and uncertain. Areas that exhibit a suppressing trend of pandemic management have an average of 10.4892% positive results, with a 95% CI of 9.4765% to 11.5018%. The average *TC* among these areas is 2.1598 [95% CI: 2.0646 to 2.2550]. Areas still in the stage of the pandemic outbreak have an average of 10.8603% positive results, with a 95% CI of 10.4497% to 11.2708%. The average *TC* among countries classified as being in the outbreak stage was 1.3335 [95% CI: 1.3030 to 1.3640]. Areas categorized as uncertain in terms of the pandemic situation have a *TC* ($Mean = 0.7472$ [95% CI: 0.6754 to 0.8190]).

2.5.4. Statistical results of TC and PPCTP of 55 areas according to governors

Table 8 and Fig. 7b show the mean values and 95% CIs of *TC* and *PPCTP* of 55 states with different governors. The average *TC* and *PPCTP* of states with Republican governors were 1.4320 and 13.4693%, while the average *TC* and *PPCTP* of states with



(a) (b)

Fig. 7. Results of daily *TC* against *PPCTP* (Squares represent the mean values, while the horizontal and vertical bars indicate the lower and upper bound of the 95% CI, respectively): (a) The mean values and 95% CIs of *TC* and *PPCTP* of 55 states in the four different situations; (b) The mean values and 95% CIs of *TC* and *PPCTP* of states, of which the governor belongs to the Democratic Party or the Republican Party.

Table 7
TC and *PPCTP* among areas with different pandemic control statuses.

Pandemic situation	Mean (*TC*)	95%CI (*TC*)	Mean (*PPCTP*)	95%CI (*PPCTP*)
Low-level	1.7065	[1.5461, 1.8670]	9.1955%	[7.0102,11.3808]%
Suppressing	2.1598	[2.0646,2.2550]	10.4892%	[9.4765,11.5018]%
Widespread	1.3335	[1.3030,1.3640]	10.8603%	[10.4497,11.2708]%
Uncertain	1.6882	[1.5662,1.8103]	18.6768%	[16.6360,20.7175]%

Table 8
TC and *PPCTP* in the 55 states according to governors.

Governor	Mean (*TC*)	95% CI (*TC*)	Mean (*PPCTP*)	95%CI (*PPCTP*)
Republican	1.4320	[1.3949,1.4691]	13.4693%	[12.8692,14.0694]%
Democratic	1.6094	[1.5727,1.6461]	9.8938%	[9.4453,10.3424]%

Democratic governors were 1.6094 and 9.8938, respectively. There was no obvious difference between them, while the pandemic situation in states with Democratic governors was slightly better than states with Republican governors.

3. Susceptible-unconfirmed-confirmed-removed (SUCR) model based on *TC*

3.1. The SUCR model

The above data-driven analysis strongly suggests that the *TC* plays a major role in elucidating the spreading status of the virus. Based on this observation, we develop an epidemic spreading model by leveraging *TC* to investigate the impact of this factor of COVID-19. Here, we design a new SUCR model by considering *TC*. Susceptible individuals can acquire the virus through contact with infectious individuals and become unconfirmed cases, meaning that they are infected, but not confirmed and quarantined, and thus could transmit the virus freely. Unconfirmed cases progress to the confirmed stage with a rate that is proportional to the *TC*. In this situation, confirmed cases would be isolated and hospitalized, and progress into the removed state with a rate proportional to the number of confirmed cases. Removed individuals refer to those who can no longer infect others, meaning that they have either recovered or died. Note that different countries and areas have various non-pharmaceutical interventions, and hence in some areas, confirmed cases also could transmit the virus to people who are susceptible to it [5,22].

Then, at time t, the official pandemic data released by an area include: the number of confirmed active cases $C(t)$; the number of confirmed recovered cases $R_c(t)$; and death toll $D_c(t)$. At the same time, however, there would be active unconfirmed infected cases $U(t)$, unconfirmed recovered cases $R_u(t)$, and unconfirmed death cases $D_u(t)$. As a consequence, individuals can be classified into five different groups: susceptible $S(t)$; unconfirmed active cases $U(t)$; confirmed active cases $C(t)$; officially recorded recovery/death cases $R_{cm}(t) = R_c(t) + D_c(t)$; and unrecorded recovery/death cases $R_{um}(t) = R_u(t) + D_u(t)$. Then, the test-based SUCR model is shown as:

$$\begin{cases} \Delta S(t) = -\alpha C(t-1)S(t-1) - \beta U(t)S(t-1), \\ \Delta U(t) = \alpha C(t-1)S(t-1) + \beta U(t)S(t-1) - k_t N_t(t)\frac{U(t)}{U(t)+N_s} - \gamma_u U(t), \\ \Delta C(t) = k_t N_T(t)\frac{U(t)}{U(t)+N_p} - \gamma_c C(t-1), \\ \Delta R_{cm}(t) = \gamma_c C(t-1), \\ \Delta R_{um}(t) = \gamma_u U(t). \end{cases} \tag{7}$$

The actual number of infected cases, recovered cases, and death cases are $I(t) = C(t) + U(t), R(t) = R_c(t) + R_u(t)$, and $D(t) = D_c(t) + D_u(t)$, respectively. In this model, each parameter has a physical meaning. α represents the exposition rate (infection rate of confirmed cases). For instance, people exposed to the new coronavirus and at risk for contracting COVID-19 might practice quarantine. In some countries/regions, most of the confirmed cases are quarantined in a mobile hospital. Then, we have $\alpha \approx 0$, while others are suggested to self-quarantine at home for 14 days. However, self-quarantine or home quarantine may expose family members to COVID-19, which could cause new infections [5,21], resulting in $\alpha > 0; \beta$ represents the exposition rate (infection rate of unconfirmed cases). Note that unconfirmed cases include asymptomatic patients. Studies have shown that undiagnosed new COVID-19 infection is the main reason for the rapid spread of the virus, and it also constitutes the core of early prevention of the outbreak [25]. Generally, we will have $\beta \geqslant \alpha$, which means that the infection ability of unconfirmed individuals is significantly higher than that of confirmed individuals; γ_c is the removal rate of confirmed cases; γ_u is the removal rate of unconfirmed cases; k_t is a calibrating term; N_s stands for the number of individuals who have COVID-19 symptoms but are not COVID-19 patients.

In conclusion, this model includes five variables

$$X(t) = \{S(t), U(t), C(t), R_{cm}(t), R_{um}(t)\}, \tag{8}$$

and six parameters

$$\theta = \{\alpha, \beta, \gamma_u, \gamma_c, k_t, N_p\}. \tag{9}$$

To the best of our knowledge, almost none of epidemiological models consider the influence from the *TC*. The proposed SUCR model, however, leverages the actual number of *TC* for modeling the spread of COVID-19. Additionally, few models consider unconfirmed cases and cannot estimate the size of unconfirmed cases. In contrast, the SUCR model can estimate the number of unconfirmed cases from officially released pandemic data and COVID-19 testing data. Furthermore, by using the daily predicted cases in our model, we can also calculate a daily risk score for areas based on the difference between their predicted and confirmed cases on any given date.

3.2. Method for parameter estimation

Algorithm 2: Algorithm for calculating time series $\{C(t_i), R_{cm}(t_i)\}$
Input: The set of parameter $\theta = \{\alpha, \beta, \gamma_u, \gamma_c, k_t, N_p\}$ and initial number of susceptible cases S_0, active unconfirmed cases U_0, active confirmed cases C_0, officially recorded/death cases $R_{cm,0}$, and unrecorded recovery/death cases $R_{um,0}$;
Output: Two time series $\{C(t_0), C(t_1), \cdots, C(t_N)\}$ and $\{R_{cm}(t_0), R_{cm}(t_1), \cdots, R_{cm}(t_N)\}$;
Initialization:
1. $S(t_0) = S_0$, $U(t_0) = U_0$, $C(t_0) = C_0$, $R_{cm}(t_0) = R_{cm,0}$, $R_{um}(t_0) = R_{um,0}$.
LOOP Process
2. **for** $i = 0$ to $N - 1$ **do**
3. Calculate $X(t_i + 1)$ from model (11), namely,
4. $\quad X_j(t_i + 1) = X_j(t_i) + f(X_j(t_i), X(t)|\theta)$.
5. Since $C(t_i) \in X_j(t_i)$ and $R_{cm}(t_i) \in X(t_i)$, we can derive $C(t_i)$ and $R_{cm}(t_i)$ from $X(t_i)$.
end for
return $\{C(t_0), C(t_1), \cdots, C(t_N)\}$ and $\{R_{cm}(t_0), R_{cm}(t_1), \cdots, R_{cm}(t_N)\}$.

Algorithm 3: Algorithm for estimating the optimal parameter set θ^*
Input: The initial variable set
$\Theta = \{S(t_0), U(t_0), C(t_0), R_{cm}(t_0), R_{tm}(t_0), \alpha, \beta, \gamma_u, \gamma_c, k_t, N_p\}$.
Output: Optimal parameter set $\bar{\Theta}^*$;
Initialization:
1. Initialize temperature T and set $\Theta_{current} = \Theta_0$. Then,
compute the estimated $\hat{C}_{(t_i)}$ and $R_{cm}(t_i)$ with parameter set $\Theta_{current}$ from Algorithm 2, then derive the objective function (13)
2. $cost_{current} = \sum_{i=1}^{N} w_{C,i}\left(C(t_i) - \widehat{C}(t_i|\Phi)\right)^2 + w_{R,i}\left(R_{cm}(t_i) - \widehat{R}_{cm}(t_i|\Phi)\right)^2$.
3. Initialize temperature T and random starting point
4. $\bar{\Theta}_0 = \Theta_L + k_{rand} * (\Theta_U - \Theta_L)$,
5. where k_{rand} is is a randomly generated real number between 0 and 1.
LOOP Process
6. **for** $i_{iter} = 0$ to i_{max} **do**
7. $i_{iter} = i_{iter} + 1$, $temp_{iter} = 0$, $\Theta_{previous} = \Theta_{current}$, $cost_{previous} = cost_{current}$.
8. **while** $temp_{iter} \leqslant n_{rep}$ **do**
9. $temp_{iter} = temp_{iter} + 1$.
10. Select a new set of parameters ($\Theta_{current}$) from the neighborhood,
11. compute the value of objective function from Algorithm 2. Then, derive $\delta = cost_{current} - cost_{previous}$.
12. **if** $\delta < 0$
13. Accept new parameter set.
14. **else**
15. Accept new parameter set with probability $\exp(-\delta/T)$.
16. **end if**
17. **end while**
18. $T = \alpha * T$, $(0 < \alpha < 1)$.
19. **end for**
20. **return** $\bar{\Theta}^*$

The parameters in the model (7) are unknown, and the historical data can be utilized for parameter estimation, which can be transformed into a nonlinear optimization problem (NOP). The purpose of optimization is to find suitable parameters to make the estimated growth trajectory that matches historical data.

Here, we define $X_0 = \{S(t_0), U(t_0), C(t_0), R_{cm}(t_0), R_{um}(t_0)\}$ as the initial number of susceptible individuals, confirmed cases, unconfirmed cases, official removed cases, and unconfirmed removed cases, respectively. Additionally, there also exist a set of unknown parameters, i.e., $\theta = \{\alpha, \beta, \gamma_u, \gamma_c, k_t, N_p\}$. This set of parameters determine the rate of spreading and recovery in a city, state, or country. Then, the unknown set is as follows:

$$\Theta = \{X_0, \theta\} = \{S(t_0), U(t_0), C(t_0), R_{cm}(t_0), R_{tm}(t_0), \alpha, \beta, \gamma_u, \gamma_c, k_t, N_p\}. \tag{10}$$

and essentially has 11 unknowns, thus requiring great computational effort.

Let $X(t)$ be the extended state vector, i.e., $X(t) = \{S(t), U(t), C(t), R_{cm}(t), R_{um}(t)\}$. Then, the SUCR model can be reformulated as:

$$\Delta X(t_i) = f(X(t_i)|\theta), \tag{11}$$

where $f(x)$ is the right side of the SUCR model (7); and θ is the set of unknown parameters. For computational convenience, we rewrite the SUCR model (7) as follows:

$$X(t_i + 1) = X(t_i) + f(X(t_i)|\theta) \tag{12}$$

Then, by setting X_0, θ, we can derive the estimated variables (see Algorithm 3). Finally, the parameter estimation problem can be formulated as the following constrained nonlinear optimization problem:

$$P_0 : \min_{\Phi} \sum_{i=1}^{N} w_{C,i}\left(C(t_i) - \widehat{C}(t_i|\Phi)\right)^2 + w_{R,i}\left(R_{cm}(t_i) - \widehat{R}_{cm}(t_i|\Phi)\right)^2$$
$$\text{s.t.} \quad \begin{cases} (i) & X(t_i + 1) = X(t_i) + f(X(t_i)|\Phi) \\ (ii) & \Phi_U \geqslant \Phi \geqslant \Phi_L \end{cases}, \tag{13}$$

where $\widehat{C}(t_i|\Phi)$ and $\widehat{R_{cm}}(t_i|\Phi)$ represent the estimated number of confirmed cases and officially removed cases, respectively, at time t_i with parameter set θ and initial condition X_0; and $w_{C,i}$ and $w_{R,i}$ stand for the weighted coefficient. The unknown parameter set is bounded between Φ_U and Φ_L. In this work, an inverse approach is taken to identify the unknown parameters

and states by solving (13) (see Algorithm 3). There are 11 unknown parameters. Furthermore, the computational complexity is approximately $O(11, N_p G_{max})$. Root mean square percentage error (RMSPE) is adopted as a criterion to measure the differences between the propagation breadth generated by the identified model and the official data,

$$e = \sqrt{\frac{1}{N}\sum_{i=1}^{N} w_{C,i}\left(\frac{C(t_i) - \widehat{C}(t_i|\Phi^*)}{C(t_i)}\right)^2 + w_{R,i}\left(\frac{R_{cm}(t_i) - \widehat{R}_{cm}(t_i|\Phi^*)}{R_{cm}(t_i)}\right)^2}, \tag{14}$$

where N is the number of data points.

3.3. Experimental results

We applied this model to 55 areas in the USA. For each area, we ran 10,000 identification procedures to ensure the reliability of the results. This model can estimate the daily number of confirmed cases, unconfirmed cases, and removed cases in

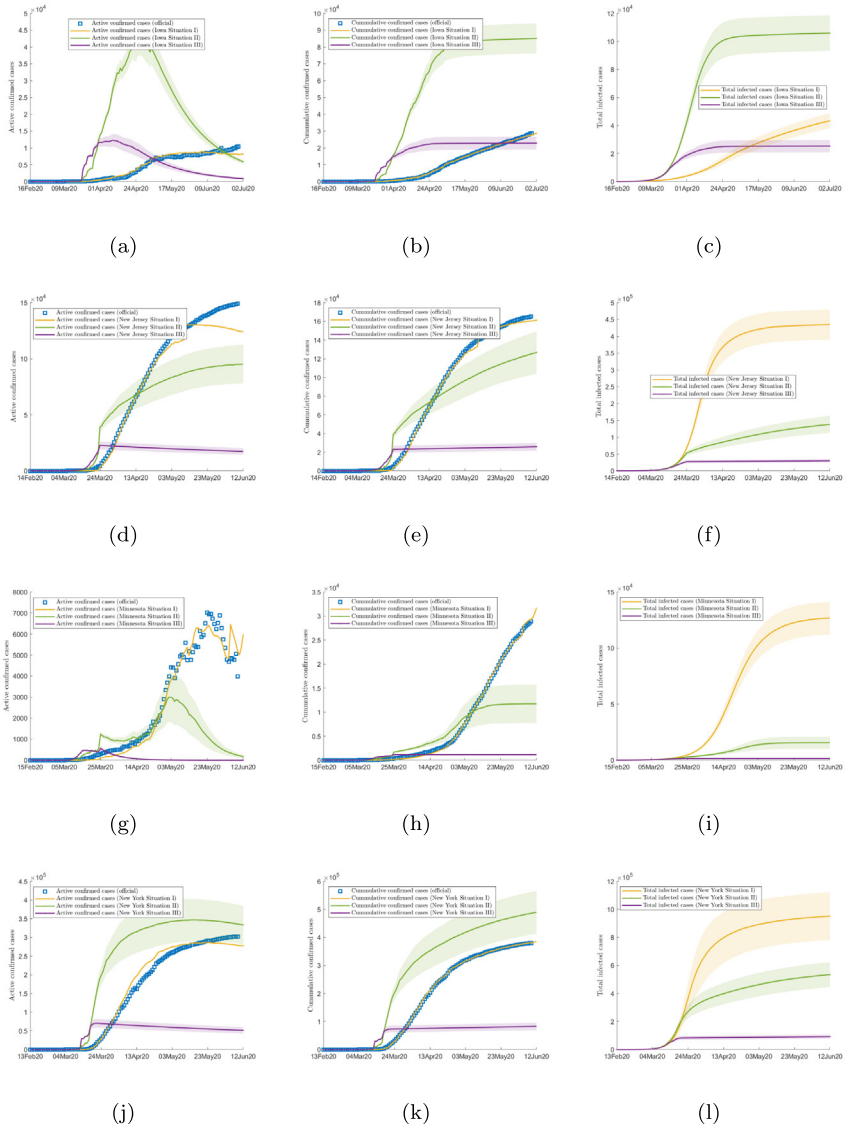


Fig. 8. Estimated historical data of the number of confirmed cases, unconfirmed cases, and total infected cases in four selected states (including Iowa, New Jersey, Minnesota, and New York) in the USA. (a), (d), (g), and (j) show the number of active confirmed cases in each state, in which the squares indicate the real number of confirmed cases, while the three solid lines represent the estimated active number of confirmed cases in three different scenarios: scenario I, scenario II, and scenario III. (b), (e), (h), and (k) show the number of cumulative confirmed cases of the three different scenarios. (c), (f), (i), and (l) show the cumulative of confirmed and unconfirmed cases.

all 55 areas in the USA. Due to space limitations, we only show the results of 4 typical areas in Fig. 8. Assuming that the migration control measures, infection rate, and recovery rate will remain unchanged for a period of time in the future, this model can provide a prediction of the number of infected individuals in each area, as shown in Fig. 8. Based on this model, we estimate all of the confirmed and unconfirmed cases of the 55 areas. In this study, we mainly consider three different scenarios: Fig. 9.

1. In the first scenario, from Feb 7, 2020, to June 15, 2020, the *TC* of an area in the USA is the actual historical *TC*. After June 15, we assume that the *TC* of each area in the USA increases linearly, and the *TC* increases by three times within 120 days, while other conditions remain constant.
2. In the second scenario, from Feb 7, 2020, to June 15, 2020, the *TC* of an area in the USA was five times that of the actual historical *TC*. After June 15, the *TC* of each area in the USA increased linearly, and *TC* increased three times within 120 days, which means that the *TC* after 120 days is 15 times that of the actual *TC* on June 15, 2020. The other conditions do not change.
3. In the third scenario, from Feb 7, 2020, to June 15, 2020, the *TC* of an area in the USA was five times that of the actual historical *TC*. After June 15, the *TC* of each area in the USA increased linearly, and detection increased three times within 120 days. Additionally, each area in the USA applied strict measures to quarantine most of the confirmed cases, which resulted in the reduction of the exposition rate (infection rate of confirmed cases) α. We assume that α only decreased by 50%.

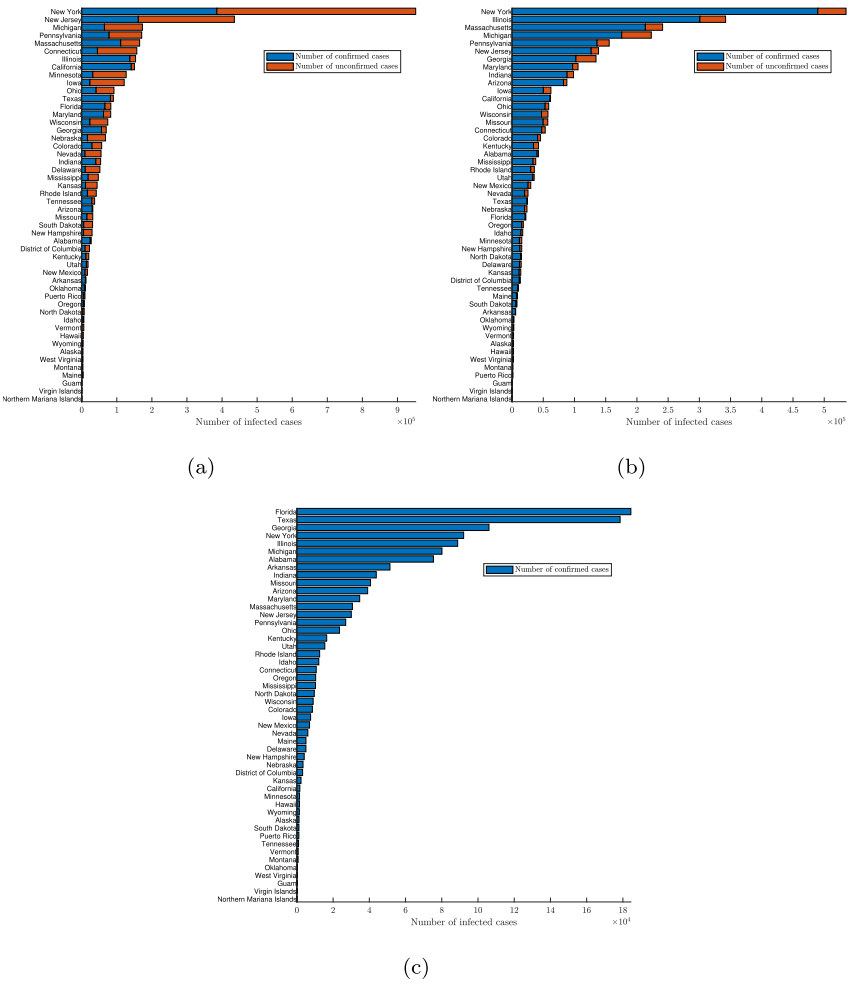


Fig. 9. Number of infected cases in different areas in the USA: (a) The number of confirmed cases (blue bar) and unconfirmed cases in each state in scenario I; (b) the number of confirmed cases (blue bar) and unconfirmed cases in each state under the condition of increasing *TC* by five times (scenario II); (c) the number of confirmed cases (blue bar) and unconfirmed cases in each state under the condition of increasing *TC* by five times and applying strict isolation measures (scenario III). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Experimental results indicate that the proposed SUCR model can accurately capture the spread of COVID-19 in the 55 areas (show in Fig. 8).

The results reveal that the number of infected people has peaked in most areas in the USA in the first wave from Feb 2020 to June 2020. From the estimated propagation profiles of the COVID-19 epidemic for all 55 states in the first wave, the following results are obtained:

- For most areas, the infection numbers will peak between mid-March to early July 2020.
- The peak number of infected individuals will be between 10,000 to 50,000 for most states.
- More than 1% of the population in the USA will get infected before Sep. 2020, as graphically presented.

The results show that the estimated total number of infections would be 3,862,893 until June 30, 2020. If we increased the *TC* by five times, starting on Mar. 15, 2020, the estimated total number of infections would be 2,949,837. Besides, if strict control measures were applied, the total number of infections would be 1,297,414. It can be clearly seen that, just by increasing *TC*, the total number of infected individuals can be greatly reduced. Indeed, by only increasing *TC* by five times, total infections can be reduced to 76.04%. Furthermore, if strict measures are applied, total infections can be reduced to 33.59% (shown in Table 9).

4. Discussion and conclusion

This study is the first to summarize and analyze the *TC* and *PPCTP* worldwide, including 101 countries and 55 areas in the USA, to generate benchmarks using a data-driven approach. Results indicate the necessity and urgency of scaling up *TC* to reach 2 tests per thousand people each day, while less than 10% of the recommended tests have been provided in practice by September 30, 2020. The widespread transmission demonstrated by some of the included countries urge improving *TC* and enhancing containment strategies. This study also shows a substantial uncertainty as some of the included countries have a low *TC* that may yield a large proportion of untested and documented infection. Critical insight from the global COVID-19 trajectory can help in successfully slowing transmission and mitigating health burden. Additionally, it is vital to diagnose accurately and effectively isolate and care for all confirmed cases with no, mild, moderate, and severe illness in health settings or home setting. Although scaling up *TC* is essential to identify infected cases, plan for contextually appropriate health responses, and evaluate the impact of response on pandemic trajectories, it should cooperate with strict public health measures to control the spread of COVID-19. The ultimate goal is to reduce patient death and mitigate the interruption of life and economics caused by COVID-19. The proposed data-driven model can analyze the pandemic situation of an area. However, evaluating current situation is only the first step. Another essential issue is finding how to predict the coming pandemic situations. For instance, predicting the next wave of the COVID-19 epidemic is an emerging topic. State-of-the-art machine learning techniques have shown good performance in many prediction tasks. In line with this, we can develop specific machine learning methods for forecasting the coming pandemic situations in future work.

Apart from testing individuals with symptoms, people with contact exposure, particularly healthcare workers, and suspected cases/clusters must be tested, as infection can be asymptomatic despite high viral loads [29]. The vulnerable population should also be tested, such as residents in the nursing home, patients who visited the hospital, and people who engage in subclinical activities [46]. Investigating the characteristics and travel history of infected cases in a country or city also informs the cause of infection and high-risk areas, whereby to release and implement tailored testing regulations. Countries or areas with a *PPCTP* of 10% should be alerted to improve *TC* as this indicates that the *TC* is insufficient to capture the infected population in a community. The proposed SUCR model can accurately capture the dynamic profiles of the spread of COVID-19. However, this model cannot explain the spreading mechanism in social networks, and cannot evaluate different testing strategies' efficiency. We plan to combine complex network theory to develop a network-based SUCR model to

Table 9
Cumulative number of confirmed, unconfirmed, and total infected cases.

	Mean (SD)	95% CI
Scenarios I		
Confirmed	1,845,571	[1409166,2281976]
Unconfirmed	2,017,322	[1535298, 2499345]
Total	3,862,893	[3219018, 4506768]
Scenarios II		
Confirmed	2,588,951	[1976766, 3201136]
Unconfirmed	360,885	[263151, 458619]
Total	2,949,837	[2429292, 3470382]
Scenarios III		
Confirmed	1,182,605	[902966, 1462245]
Unconfirmed	114,808	[76956, 152660]
Total	1,297,414	[1050804, 1544023]

analyze the spread mechanism on different networks and evaluate different testing strategies for preventing and combating the next wave of COVID-19 epidemic. Moreover, machine learning methods have already shown promising results in forecasting nonlinear dynamic progress [50]. In future work, developing efficient machine learning methods for predicting the whole spread of COVID-19 deserves further investigation.

The study only included countries that reported COVID-19 data publicly, which may reduce its generalizability. It is recommended that the government and health administration should document and share data with the public for combating the COVID-19 collaboratively. Besides, the classification of the transmission phase of COVID-19 was generated by authors using a data-driven approach. More research from designated clinical practice settings is needed to generate practice-based evidence in quantifying the pandemic development trajectory for classification. Scaling up the COVID-19 test is part of the systematic strategies to evaluate and manage the pandemic. The SUCR model reveals that the early detection of infected cases enables multi-level health response to track and treat infected cases to reduce patient death and mitigate health burden. Additionally, public health intervention would be another essential issue for controlling the spread of COVID-19.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by the National Key R&D Program of China under Grant No. 2018YFB1003800, 2018YFB1003805, the National Natural Science Foundation of China under Grant No. 61972112 and No. 61832004.

References

[1] Jaffar A. Al-Tawfiq, Asymptomatic coronavirus infection: Mers-cov and sars-cov-2 (covid-19), Travel medicine and infectious disease, 2020.
[2] Roy M. Anderson, Hans Heesterbeek, Don Klinkenberg, T. Déirdre Hollingsworth, How will country-based mitigation measures influence the course of the covid-19 epidemic? The Lancet 395 (10228) (2020) 931–934.
[3] Jantien A. Backer, Don Klinkenberg, Jacco Wallinga, Incubation period of 2019 novel coronavirus (2019-ncov) infections among travellers from wuhan, China, 20–28 january 2020, Eurosurveillance 25(5) (2020) 2000062.
[4] Yan Bai, Lingsheng Yao, Tao Wei, Fei Tian, Dong-Yan Jin, Lijuan Chen, Meiyun Wang, Presumed asymptomatic carrier transmission of covid-19, Jama 323 (14) (2020) 1406–1407.
[5] Jasper Fuk-Woo Chan, Shuofeng Yuan, Kin-Hang Kok, Kelvin Kai-Wang To, Hin Chu, Jin Yang, Fanfan Xing, Jieling Liu, Cyril Chik-Yan Yip, Rosana Wing-Shan Poon, et al., A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster, The Lancet 395 (10223) (2020) 514–523.
[6] Simiao Chen, Zongjiu Zhang, Juntao Yang, Jian Wang, Xiaohui Zhai, Till Bärnighausen, Chen Wang, Fangcang shelter hospitals: a novel concept for responding to public health emergencies, The Lancet (2020).
[7] Matteo Chinazzi, Jessica T Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, et al., The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak, Science 368 (6489) (2020) 395–400.
[8] World Bank Country and Lending Groups, Archived 11 january 2018 at the wayback machine world bank. Website, 2018. https://web.archive.org/web/20180111190936/https://datahelpdesk.worldbank.org/knowledgebase/articles/906519#High_income.
[9] WHO Covid, Strategy update, 2020, 19.
[10] Domenico Cucinotta, Maurizio Vanelli, Who declares covid-19 a pandemic, Acta Bio-med. Atenei Parmensis 91 (1) (2020) 157–160.
[11] Michael Day, Covid-19: identifying and isolating asymptomatic people helped eliminate virus in italian village. BMJ: Br. Med. J. (Online) 368 (2020).
[12] Centers for Disease Control and Prevention, et al., Priorities for testing patients with suspected covid-19 infection. Available on: https://www.cdc.gov/coronavirus/2019-ncov/downloads/priority-testing-patients.pdf, vol. 25, Oct. 2020.
[13] Centers for Disease Control, Prevention, et al., Evaluating and testing persons for coronavirus disease 2019 (covid-19). National Center for Immunization and Respiratory Diseases (NCIRD), Division of Viral Diseases, 2020.
[14] Monica Gandhi, Deborah S. Yokoe, Diane V. Havlir, Asymptomatic transmission, the Achilles' heel of current strategies to control COVID-19, 2020.
[15] Marius Gilbert, Giulia Pullano, Francesco Pinotti, Eugenio Valdano, Chiara Poletto, Pierre-Yves Boëlle, Eric d'Ortenzio, Yazdan Yazdanpanah, Serge Paul Eholie, Mathias Altmann, et al., Preparedness and vulnerability of african countries against importations of covid-19: a modelling study. The Lancet 395 (10227) (2020) 871–877.
[16] Xi He, Eric H.Y. Lau, Peng Wu, Xilong Deng, Jian Wang, Xinxin Hao, Yiu Chung Lau, Jessica Y. Wong, Yujuan Guan, Xinghua Tan, et al.. Temporal dynamics in viral shedding and transmissibility of covid-19, Nat. Med. 26 (5) (2020) 672–675.
[17] Laura Heavey, Geraldine Casey, Ciara Kelly, David Kelly, Geraldine McDarby, No evidence of secondary transmission of covid-19 from children attending school in ireland, 2020, Eurosurveillance 25 (21) (2020) 2000903.
[18] Joel Hellewell, Sam Abbott, Amy Gimma, Nikos I. Bosse, Christopher I. Jarvis, Timothy W. Russell, James D. Munday, Adam J. Kucharski, W. John Edmunds, Fiona Sun, et al., Feasibility of controlling covid-19 outbreaks by isolation of cases and contacts, Lancet Global Health (2020).
[19] Lei Huang, Xiuwen Zhang, Xinyue Zhang, Zhijian Wei, Lingli Zhang, Xu. Jingjing, Peipei Liang, Xu. Yuanhong, Chengyuan Zhang, Aman Xu, Rapid asymptomatic transmission of covid-19 during the incubation period demonstrating strong infectivity in a cluster of youngsters aged 16–23 years outside wuhan and characteristics of young patients with covid-19: a prospective contact-tracing study, J. Infection (2020).
[20] It is however unclear whether Western. Humanity tested.
[21] Jayson S Jia, Xin Lu, Yun Yuan, Ge Xu, Jianmin Jia, Nicholas A. Christakis, Population flow drives spatio-temporal distribution of covid-19 in china, Nature (2020) 1–5.
[22] Shengjie Lai, Nick W. Ruktanonchai, Liangcai Zhou, Olivia Prosper, Wei Luo, Jessica R. Floyd, Amy Wesolowski, Mauricio Santillana, Chi Zhang, Xiangjun Du, et al., Effect of non-pharmaceutical interventions to contain covid-19 in china, Nature 585 (7825) (2020) 410–413.
[23] Stephen A. Lauer, Kyra H. Grantz, Qifang Bi, Forrest K. Jones, Qulu Zheng, Hannah R. Meredith, Andrew S. Azman, Nicholas G. Reich, Justin Lessler, The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application, Ann. Internal Med. 172 (9) (2020) 577–582.

[24] Qun Li, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren, Kathy S.M. Leung, Eric H.Y. Lau, Jessica Y. Wong, et al., Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia, New England J. Med. (2020).
[25] Ruiyun Li, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, Jeffrey Shaman, Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov-2), Science 368 (6490) (2020) 489–493.
[26] Natalie M. Linton, Tetsuro Kobayashi, Yichi Yang, Katsuma Hayashi, Andrei R. Akhmetzhanov, Sung-mok Jung, Baoyin Yuan, Ryo Kinoshita, Hiroshi Nishiura, Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data, J. Clin. Med. 9 (2) (2020), 538.
[27] Jiaye Liu, Xuejiao Liao, Shen Qian, Jing Yuan, Fuxiang Wang, Yingxia Liu, Zhaoqin Wang, Fu-Sheng Wang, Lei Liu, Zheng Zhang, Community transmission of severe acute respiratory syndrome coronavirus 2, shenzhen, china, 2020, Emerg. Infect. Diseases 26 (6) (Jun. 2020) 1320.
[28] Ying Liu, Albert A. Gayle, Annelies Wilder-Smith, Joacim Rocklöv, The reproductive number of covid-19 is higher compared to sars coronavirus, J. Travel Med. (2020).
[29] Stefan Lohse, Thorsten Pfuhl, Barbara Berkó-Göttel, Jürgen Rissland, Tobias Geißler, Barbara Gärtner, Sören L. Becker, Sophie Schneitler, Sigrun Smola, Pooling of samples for testing for sars-cov-2 in asymptomatic people, Lancet Infect. Diseases (2020).
[30] Smriti Mallapaty, Antibody tests suggest that coronavirus infections vastly exceed official counts, Nature (Lond.) (2020).
[31] Kenji Mizumoto, Katsushi Kagaya, Alexander Zarebski, Gerardo Chowell, Estimating the asymptomatic proportion of coronavirus disease 2019 (covid-19) cases on board the diamond princess cruise ship, yokohama, japan, 2020, Eurosurveillance 25 (10) (2020) 2000180.
[32] Hiroshi Nishiura, Tetsuro Kobayashi, Takeshi Miyama, Ayako Suzuki, Sung-mok Jung, Katsuma Hayashi, Ryo Kinoshita, Yichi Yang, Baoyin Yuan, Andrei R. Akhmetzhanov, et al., Estimation of the asymptomatic ratio of novel coronavirus infections (covid-19), Int. J. Infect. Diseases 94 (2020) 154.
[33] World Health Organization, World Health Organization, et al., Report of the who-china joint mission on coronavirus disease 2019 (covid-19), 2020.
[34] Julian Peto, Covid-19 mass testing facilities could end the epidemic rapidly, Bmj 368 (2020).
[35] Jane Qiu, Covert coronavirus infections could be seeding new outbreaks, Nature (2020).
[36] Camilla Rothe, Mirjam Schunk, Peter Sothmann, Gisela Bretzel, Guenter Froeschl, Claudia Wallrauch, Thorbjörn Zimmer, Verena Thiel, Christian Janke, Wolfgang Guggemos, et al, Transmission of 2019-ncov infection from an asymptomatic contact in Germany, New England J. Med. 382 (10) (2020) 970–971.
[37] Marcel Salathé, Christian L. Althaus, Richard Neher, Silvia Stringhini, Emma Hodcroft, Jacques Fellay, Marcel Zwahlen, Gabriela Senti, Manuel Battegay, Annelies Wilder-Smith, et al., Covid-19 epidemic in switzerland: on the importance of testing, contact tracing and isolation, Swiss Med. Weekly 150 (11–12) (2020) w20225.
[38] Janice Hopkins Tanne, Erika Hayasaki, Mark Zastrow, Priyanka Pulla, Paul Smith, Acer Garcia Rada, Covid-19: how doctors and healthcare systems are tackling coronavirus worldwide, Bmj 368 (2020).
[39] Huaiyu Tian, Yonghong Liu, Yidan Li, Chieh-Hsi Wu, Bin Chen, Moritz UG Kraemer, Bingying Li, Jun Cai, Bo Xu, Qiqi Yang, et al., An investigation of transmission control measures during the first 50 days of the covid-19 epidemic in china, Science 368 (6491) (2020) 638–642.
[40] Kelvin Kai-Wang To, Owen Tak-Yin Tsang, Wai-Shing Leung, Anthony Raymond Tam, Tak-Chiu Wu, David Christopher Lung, Cyril Chik-Yan Yip, Jian-Piao Cai, Jacky Man-Chun Chan, Thomas Shiu-Hong Chik, et al., Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by sars-cov-2: an observational cohort study, Lancet Infect. Diseases (2020).
[41] S. Varela-Santos, P. Melin, A new approach for classifying coronavirus COVID-19 based on its manifestation on chest X-rays using texture features and neural networks, Inf. Sci. 545 (2020) 403–414.
[42] Yishan Wang, Hanyujie Kang, Xuefeng Liu, Zhaohui Tong, Combination of rt-qpcr testing and clinical features for diagnosis of covid-19 facilitates management of sars-cov-2 outbreak, J. Med. Virol. 92 (6) (2020) 538–539.
[43] Benjamin D. Wissel, P.J. Van Camp, Michal Kouril, Chad Weis, Tracy A. Glauser, Peter S. White, Isaac S. Kohane, Judith W. Dexheimer, An interactive online dashboard for tracking covid-19 in us counties, cities, and states in real time, J. Am. Med. Inf. Assoc. (2020).
[44] Roman Wölfel, Victor M. Corman, Wolfgang Guggemos, Michael Seilmaier, Sabine Zange, Marcel A. Müller, Daniela Niemeyer, Terry C. Jones, Patrick Vollmar, Camilla Rothe, et al., Virological assessment of hospitalized patients with covid-2019, Nature 581 (7809) (2020) 465–469.
[45] Joseph T. Wu, Kathy Leung, Gabriel M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study, The Lancet 395 (10225) (2020) 689–697.
[46] Xin Xu, Jian Sun, Sheng Nie, Huiyuan Li, Yaozhong Kong, Min Liang, Jinlin Hou, Xianzhong Huang, Dongfeng Li, Tean Ma, et al., Seroprevalence of immunoglobulin m and g antibodies against sars-cov-2 in china, Nat. Med. (2020) 1–3.
[47] Ping Yu, Jiang Zhu, Zhengdong Zhang, Yingjun Han, A familial cluster of infection associated with the 2019 novel coronavirus indicating possible person-to-person transmission during the incubation period, J. Infect. Diseases 221 (11) (2020) 1757–1761.
[48] K. Choujun Zhan, Tse Chi, Zhikang Lai, Xiaoyun Chen, Mingshen Mo, General model for covid-19 spreading with consideration of intercity migration, insufficient testing and active intervention: application to study of pandemic progression in Japan and USA, JMIR Public Health Surveillance 6 (3) (2020) e18880.
[49] Choujun Zhan, Chi Tse, Yuxia Fu, Zhikang Lai, Haijun Zhang, Modelling and prediction of the 2019 coronavirus disease spreading in china incorporating human migration data, PloS One. doi: 10.1371/journal.pone.0241171, 2020.
[50] M. Zhou, B. Wang, S. Guo, J. Watada, Multi-objective prediction intervals for wind power forecast based on deep neural networks, Inf. Sci. 550 (2021) 207–220.