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.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Data-driven test strategy for COVID-19 using machine learning: A study in Lahore, Pakistan

Chuanli Huang \textsuperscript{a,b,1}, Min Wang \textsuperscript{c,1}, Warda Rafaqat \textsuperscript{a}, Salman Shabbir \textsuperscript{d}, Liping Lian \textsuperscript{e}, Jun Zhang \textsuperscript{a}, Siuling Lo \textsuperscript{b}, Weiguo Song \textsuperscript{a,*}

\textsuperscript{a} State Key Laboratory of Fire Science, University of Science and Technology of China, Hefei, 230027, PR China
\textsuperscript{b} Department of Architectural and Civil Engineering, City University of Hong Kong, Hong Kong, China
\textsuperscript{c} Department of Oncology, The Second Affiliated Hospital of Soochow University, Suzhou, 215004, China
\textsuperscript{d} Program Officer, Punjab Information Technology Board, Arfa Kareem Tower, Ferozpur Road, Lahore, Pakistan
\textsuperscript{e} School of Urban Planning and Design, Shenzhen Graduate School, Peking University, No. 2199, Lishui Road, Nanshan District, Shenzhen, 518055, China

ARTICLE INFO

Keywords:
COVID-19
Test strategy
Policy making
Machine learning
Logistic regression
Time series analysis
Spatial analysis

ABSTRACT

Aims: We aimed at giving a preliminary analysis of the weakness of a current test strategy, and proposing a data-driven strategy that was self-adaptive to the dynamic change of pandemic. The effect of driven-data selection over time and space was also within the deep concern.

Methods: A mathematical definition of the test strategy were given. With the real COVID-19 test data from March to July collected in Lahore, a significance analysis of the possible features was conducted. A machine learning method based on logistic regression and priority ranking were proposed for the data-driven test strategy. With performance assessed by the area under the receiver operating characteristic curve (AUC), time series analysis and spatial cross-test were conducted.

Results: The transition of risk factors accounted for the failure of the current test strategy. The proposed data-driven strategy could enhance the positive detection rate from 2.54\% to 28.18\%, and the recall rate from 8.05\% to 89.35\% under strictly limited test capacity. Much more optimal utilization of test resources could be realized where 89.35\% of total positive cases could be detected with merely 48.17\% of the original test amount. The strategy showed self-adaptability with the development of pandemic, while the strategy driven by local data was proved to be optimal.

Conclusions: We recommended a generalization of such a data-driven test strategy for a better response to the global developing pandemic. Besides, the construction of the COVID-19 data system should be more refined on space for local applications.

1. Introduction

The World Health Organization (WHO) declared COVID-19 a global pandemic on 11 March 2020 [1]. Up to 27 August 2020, there are over 24 million confirmed cases and over 826 thousand deaths around the world. The pandemic of COVID-19 has presented high transmissibility and led to extraordinary socioeconomic disruption due to severe preventative measures by governments, and is likely to be of longer duration and more severe in its economic effects given the greater uncertainty surrounding its nature [2]. The risk mitigation measures facing COVID-19 commonly taken by countries around the world can be collated into six clusters: mobility restriction, Socio-economic restriction, physical distancing, hygiene measures, communication, and international support mechanisms [3]. However, due to the economic downturn, governments are currently looking at designing relaxation efforts by simultaneously considering both public health and economic restart [4–6].

The policy-making is interdisciplinary based on information from risk and exposure assessment [7], which requires a close combination of theory and practice. By applying different plant capacity concepts, researchers seek to measure the use of existing capacities, as well as the evolution and build-up of extra hospital capacity in the Chinese province.
of Hubei during the outbreak of the COVID-19 epidemic in early 2020 [8]. On the other hand, any application of theoretical contribution demands a deep understanding of the development of pandemic. COVID-19 test is our window onto the pandemic and how it is spreading. The importance of the COVID-19 test could be described in four aspects:

a) understanding the spread of the pandemic. No country knows the total number of people infected with COVID-19. All we know is the infection status of those who have been tested. The World Health Organization defines a confirmed case as “a person with laboratory confirmation of COVID-19 infection” [9]. This means that the counts of confirmed cases depend on how many tests a country has conducted [10]. Without testing, there is no data.

b) appropriate response. The test is also the base of many countermeasures. Test combined with contact tracing has been an important measure to control the outbreak [3]. Tests allow us to identify infected individuals, guiding the medical treatment that they receive. It also enables the isolation of those infected and the tracing and quarantining of their contacts [11] and helps allocate medical resources and staff more efficiently [12]. Besides, lockdown relaxation is within consideration in many areas, and the results of the large-scale screening test of COVID-19 is therefore an essential reference [13].

c) pandemic modeling and prediction. Traditional pandemic modeling and epidemiological methods are based on the concept of viral spread due to person-to-person contact, via an empirical transmission parameter, $R_0$, which is adjusted to fit dynamic (time-varying) infection data. Statistical analysis of testing results is the crucial measure for the obtaining of $R_0$ [14–16].

d) Artificial intelligence technology. AI tools are increasingly being used and have demonstrated success in providing insights that can lead to better health policy and management. However, as these technologies are still in their infancy stages, slow progress is being made in their adoption for serious consideration at national and international policy levels [17,18]. Despite that algorithms are gaining accuracy, the database gathering from tests is of the same importance [19].

Although the test results are under close attention, the discrepancy in test strategy of different countries will cause misunderstanding of the development of pandemic in them:

1) The inclusion of test record. The number of tests does not refer to the same in each country—one difference is that some countries report the number of people tested, while others report the number of tests (which can be higher if the same person is tested more than once). And other countries report their testing data in a way that leaves it unclear what the test count refers to exactly [10]. On the other hand, there are many different technologies for COVID-19 testing for different applications of purposes, some currently available, and some still in development [13,20,21]. There are technical differences in how results from these different tests should be interpreted [10]. Current data suggests that other existing testing technologies are subject to very different rates of false positive and false negative results than PCR tests [22].

2) Test policy under limited test capacity. For countries with a high number of (COVID-19) cases per capita, there are likely many people with undetected SARS-CoV-2 infection because testing efforts are not detecting all infected people, including some with clinical disease compatible with COVID-19 [23–26]. The World Health Organization recommends a combination of measures: rapid diagnosis and immediate isolation of cases, rigorous tracking, and precautionary self-isolation of close contacts. Besides, it is also appealed that the identification and testing of potential cases with no or mild disease (e.g., influenza-like illness) needs to be as extensive as is permitted by health care and diagnostic testing capacity. Transmission by people with no or mild symptoms can dampen the power of the isolation strategy because of the reduced likelihood of isolating all cases and tracing all contacts [3,27]. The policy of countries on these two issues, i.e. how to assign testing resources and how to treat the asymptomatic in their test strategy make a great difference to the testing result [16].

To our knowledge, for most countries, there is no unified policy released about how to assign the limited detection chances to all the pending instances including those with travel history, contact history, especially the large amount with wellness issues.

At the start of the outbreak, many countries reserved testing for people who were symptomatic, and in Japan and Europe, testing generally focused on people with severe symptoms. However, testing criteria have evolved with the local and global situation and new scientific evidence. Norway does not recommend widespread testing because of the country’s low infection rate and high probability of false-positive results, thus limiting asymptomatic testing to staff and residents in nursing homes and close contacts of people with confirmed infection. Since August 2020, Norway has introduced new rules to allow everyone who suspects that they might be infected to get tested without an initial assessment by their local community doctor. Meanwhile, South Korea mass tests individuals who have visited public venues or events where people with COVID-19 were present, and thus who might have come into contact with them, regardless of symptoms [28]. Testing strategies should remain flexible and able to adapt rapidly to change depending on the local epidemiological situation, transmission rates, population dynamics, and available resources.

The test strategy was simple in the United States: a) secure as many testing platforms as possible, b) maintain the supply chain of swabs and testing kits, and c) use our precious testing resources wisely [29]. The strategy, however, focused more on “how we will test” rather than “whom we will test.” In addition, the plans being put in place appear to overlook, or ignore a number of fundamental principles guiding the diagnosis and screening for infectious diseases. The “false negative” reality is vital to understand because numerous institutions are promulgating a “test everyone” strategy [30].

According to Commission Recommendation (EU) 2020/1595 of 28 October 2020 on COVID-19 testing strategies, the official journal of the European Union, the Commission underlines again that Member States should define the necessary testing capacities and resources (for taking samples, performing the test and contact tracing) based on testing objectives, demand and supply planning, the latest scientific evidence on the characteristics of the disease. When sufficient capacities are not available Member States should prioritize the testing of individuals showing COVID-19 compatible symptoms, including mild symptoms, and particularly those presenting symptoms of acute respiratory infection. This should be combined, if possible, with testing for influenza and other respiratory infections, e.g. through available multiplex or other relevant assays. Criteria for prioritization of testing should be objective and applied in a non-discriminatory manner [31]. However, the prioritization criteria cannot be quantified according to the Recommendation, thus it is hard to be objective and optimized in practice.

Testing is extremely low in most South Asian countries. For example, the ICMR, India’s apex health research institute, until April 9, 2020, limited testing to individuals with symptoms arriving from abroad, symptomatic medical care workers, facility reported severe acute respiratory infection patients and close contacts of COVID-19 laboratory-confirmed cases. Hence, institutional testing is strategized as an exclusive control measure [32]. There is a need to develop a national testing policy instead of testing guidelines so that all states/provinces are able to prepare their own epidemic preparedness profile to geographically assess the population at risk, and to adhere strictly to the rational testing practice with all laboratory reported or referred cases.

Reflecting on the different testing strategies employed by different countries, it becomes evident that effective, timely, and widespread
testing is key for gaining access to accurate data to employ in combatting COVID-19. When the number of tests was insufficient among multiple nations, prioritizing the individuals in the higher risk groups was essential to ensure that tests were utilized most efficiently [33].

3) Lack of open data for negative cases. WHO calls on that reliable data on testing is necessary to assess the reliability of the data that informs us about the spread of the pandemic: the data on positive cases and deaths [10]. However, for data scientists, the data of negative cases are equally important. The lack of open negative data may cause a problem for researches on statistical analysis or prediction modeling [16]. And for particular regions, Cross-population train/test are used because there are not enough data [19]. However, the influence of the spatial variation, and the time duration of collected data on decision making is rarely studied to our knowledge.

Therefore, for better understanding and studying the pandemic around the world, a reliable standard should be complied with by countries to record and publish their test data. And we expect that not only positive and death data, but more negative data being open for research. We also urge a reliable and unified test strategy being used by countries, especially whose test capacity is deemed insufficient to cover the pending instances. The strategy should be able to adapt to the country’s differences and the development of pandemic. That is what we will propose in this paper, the data-driven strategy.

In the following Section 2 of the paper, we gave a formal definition of the test strategy and introduced the data used in this study. We analyzed how the static policy implemented failed to fit the development of pandemic in Section 3. In Section 4 we proposed a data-driven strategy and demonstrated its superiority over the current strategy prioritizing the travel and contact history. The adaptability of the data-driven strategy on time and space were discussed in Section 5. Conclusions and perspectives were made in Section 6.

2. Test strategy and datasets

2.1. Problem definition

To clarify the focus of this paper, the test strategy optimization problem was defined as the following:

For given instance set \( S \), function \( g(X) = \{ x \in X \mid x \text{ is positive} \} \), and number \( c \in \mathbb{N}^+ \), to find a function \( f(X) \), where \( f(S) = M \), \( M \subseteq S \) and \( \text{card}(M) \leq c \). Maximize: \( n = \text{card}(g(f(S))) \)

In the statement, instance set \( S \) consisted of all instances pending for test. Function \( g(X) \) indicated the test technology, which output the positive instances detected. Number \( c \) was the test capacity, i.e. how many instances could be tested in a given period. To maximize the positive numbers detected under limited test capacity \( c \), an effective selection function \( f(X) \) was demanded to select the instance set \( M \) for the test from the whole set \( S \). In another word, \( f(X) \) was defined as the test strategy.

2.2. Datasets

There were three general approaches to get tested in Lahore: the government regular test, private clinical test, and smart sampling by government. The data used in this study included the information of tested instances from all the three approaches in Lahore from 18 March to 11 July. Totally 35 fields were covered for each individual including some identity information. In this study, The test outcome, travel history, contact history, wellness condition, as well as basic information such as test date, location, age, and gender were considered.

Lahore was the second biggest city in Pakistan, while by July, Pakistan was one of the worst-hit Asian countries in the pandemic [10]. Like many other undeveloped countries, it had been suffering from a serious shortage of test capacity. Less testing meant less infected cases were identified [34]. Sources within hospitals said the patients coming for the tests at the government hospitals were being denied. Secretary Health clarified that the reason behind the artificial shortage of testing kits was an increase in testing need and the budget that had been provided to hospitals for purchasing new kits [35]. The situation Pakistan faced at the time, where test capacity was terribly insufficient, was a common problem for other countries [32]. Thus an intelligent test strategy on Pakistan could also be easily extended to other countries suffering from a shortage of test capacity.

On the other hand, the original report form recording the information of persons being tested, also called “Revised case report form for Confirmed Novel Coronavirus COVID-19”, was recommended by WHO and widely used around the world [36]. Therefore, the coverage of data in the study on Pakistan was also representative for most other countries in the world.

Tests were performed after RNA Extraction using a viral RNA Mini kit. Qualitative dual-target detection of COVID-19 Gene was done on real-time PCR machine. Travel history, contact history, and gender were binary factors tagged with 0 or 1. The wellness scale was an observational assessment assigned by experienced clinicians on a scale ranging from 0 (no symptoms) to 5 (death) before being imported into the data system. The interpretation of the tags of different factors and

| Notation | Factor and Tag | Interpretation |
|----------|----------------|----------------|
| a        | Age            | By years       |
| g        | Gender         | Male           |
|          | 0 Male         |
|          | 1 Female       |
| t        | Travel history | Staying in Pakistan |
|          | 0 Staying in Pakistan |
|          | 1 Back from abroad during the pandemic |
| c        | Contact history| No contact history with confirmed cases |
|          | 0 No contact history with confirmed cases |
|          | 1 Contacted with confirmed cases |
| w        | Wellness scale | No symptoms |
|          | 0 No symptoms |
|          | 1 Mild symptoms: No sense of smell and taste, mild cough and runny nose or diarrhea |
|          | 2 Moderate symptoms: All mild symptoms also high grade fever |
|          | 3 Severe symptoms: Shortness of breath or hypoxemia |
|          | 4 Critical condition: Acute respiratory distress syndrome, septic shock, uncorrectable metabolic acidosis, coagulopathy or multiple organ failure |
|          | 5 Death        |
corresponding notations used in this work were shown in Table 1. A visualized description of the data collected was presented in Fig. 1. The information of 256k + instances was collected totally, including 41k + positive instances and 214k + negative instances. Among the whole instances, 106k + had contact history with confirmed patients, 12k + had travel history abroad, and 27k + had sickness issues, most of which were accompanied with only mild symptoms. The percentage of males reached 73.2%, while the average age was 35, with the maximum age of 99, and the minimum age of neonates.

2.3. Data source

We were thankful to the District Health Authority, Lahore for giving access to the daily data of COVID-19 Patients. District Health Authority, Lahore was among the members of the district health emergency response committee formed by the Punjab Government Lahore Pakistan. With the collaboration of Punjab Information Technology board it is working on COVID-19 decision making and Area selection for locked down. Traceability of suspected and positive cases and their contacts, Collection of samples of suspected cases, and dispatching them to concerned laboratories, and follow up of the results report fell under its responsibility [37].

To maintain all the information records, the Interim case reporting form for COVID-19 Contact Listing Form of confirmed and probable cases, the WHO Minimum Data Set Report Form was being used [38].

3. Current policy on test strategy

With the collected dataset, we conducted a preliminary analysis of the weakness of the current strategy for the COVID-19 test, a static strategy prioritizing the travel and contact history that had been followed since the incipience of the pandemic.

3.1. Positive detection rate

Due to the huge gap between limited capacity and sharply increasing need of test, the government regular test policy had been following a static strategy prioritizing the travel and contact history, where all the cases with a contact history or travel history were compelled to be tested [39], while a great number of the others, even with clinical disease compatible with COVID-19, could not get tested in time.

One important way to understand if countries were doing test sufficiently was to look at the share of test returning positive results—known as the positive detection rate. Countries with a very high positive rate were unlikely to do test widely enough to find all cases [10]. The WHO has suggested a positive rate of around 3–12% as a general benchmark of an adequate test [40]. For all the test data collected in Lahore till 11 July, as shown in Fig. 2, a large gap between the positive detection number and test number was observed in the government tests, which was distinct from that of the private approach. Considering that the government test made up the majority of the overall test (210450 of 290437), it accounted for a lower positive detection rate of the overall test in Lahore. The overall positive detection rate was 14.39%. Among which for the private clinical test it was 30.90% whereas for the government test, the percentage was just as low as 8.11% (see Table 2). The discrepancy implied that although the statistical test results could indicate the sufficiency of the test, the selection of cases being tested, i.e. the test strategy could cause misleading results of positive detection rate.

3.2. Preliminary analysis of current strategy

To better understand the weakness of the current strategy, we assumed that the strategy was executed strictly on the whole dataset since 18-March, where only those with travel history or contact history got the chance of test, and the performance on each day was shown in

Fig. 1. A visualized description of data. (a) Confirmation, contact history, travel history and gender. (b) Wellness scale. (c) Age distribution.
Fig. 3. It was clear that at the starting days of the test, the positive detection rate could reach nearly 100%, which indicated that the current strategy performed perfectly at the incipient stage. Nevertheless, the positive rate kept declining with the pandemic developing in Lahore, and 40 days later it remained lower than 20%, which was performing even worse as time went by. Thus the current strategy had turned much more inapplicable than that of incipience. Obviously, it was not because the COVID-19 situation was becoming much better in Lahore, one of the most severely influenced countries in Asia by 11 July [10], but that the current test strategy, which had been almost static, were inappropriate for the pandemic at the current stage.

By conducting a significance analysis on the test result, it was shown how the effect of factors recorded changed over time seen as Fig. 4. It was surprisingly discovered that coefficients of travel history and contact history were all positive at the incipient stage, but they had already turned into negative as time went by, and remained negative in recent months. The analysis implied that travel history and contact history had transformed into protective factors from risk factors. The dataset collected could not give an accurate explanation of the behaviors or these two curves, but the possible reasons were that the pandemic had been out of control in Lahore and the lockdown was not strictly executed. As reported in work by Niud and Xu [27], that transmission by people with mild or no symptoms could undermine the effectiveness of the isolation strategy because of a reduced likelihood of isolating all cases and tracing all contacts. Although imported cases were the main source at the incipience, the cases generated domestically were far more serious now. On the other hand, the quarantine policy gave better protection to people with contact history, which reduced their chance of being infected than kept being exposed to the social environment. Anyway, it was proved that the current test strategy, which had been static since the beginning, no longer fitted the development of the pandemic at the present stage.

The transition of potential factors proved that a smart adjustment of the test strategy was critical in combatting the pandemic. At the start of the outbreak, many countries reserved testing for people who were symptomatic, and in Japan and Europe, testing generally focused on people with severe symptoms. However, their testing criteria had evolved and rules had been updated with the local and global situation and new scientific evidence [28]. An essential question was: when and how to adjust the test strategy? The thinking on this question by policymakers could never stop as the pandemic being evolved. The key was to learn from the past data dynamically and automatically. Therein, developing a smart strategy driven by past data was necessary.

4. The data-driven test strategy

Based on the analysis above, we believed that a satisfactory test strategy should be self-adaptive to the dynamic development of pandemic. Thus we proposed a data-driven strategy.

4.1. Data-driven strategy scheme

There were 4 general steps to conduct a test strategy on the whole
instances pending for COVID-19 test:

Step 1: Input the necessary information about each instance.
Step 2: Preliminary assessment of each instance.
Step 3: Priority ranking of all instances.
Step 4: Decision making according to the priority and test capacity.

Thereby, An executable strategy for COVID-19 test should at least include the 3 following components, which determined the quality of the strategy:

component I: Selection of feature vector \( \xi \).

component II: The prognostic model:

\[
p = f(\xi)
\]  

(1)

component III: Priority ranking criterion:

\[
P = [p_1, p_2, ..., p_n] \rightarrow \mathbb{R} = [r_1, r_2, ..., r_n, r_i \geq r_{i+1} \text{ for all } r_i, r_{i+1} \in \mathbb{R}]
\]  

(2)

To illustrate the generation of these three components in the proposed data-driven strategy, and the procedure of the strategy, the framework of the data-driven strategy was shown in Fig. 5.

For feature selection, Features such as age, gender, travel history, contact history, and wellness scale were considered and made up the initial feature vector \( \xi \) (Equation (3)), the input of the machine learning framework.

\[
\xi_i = [a_i, g_i, t_i, c_i, w_i]
\]  

(3)

For the initial feature vector \( \xi \), elements like travel history, contact history, and wellness scale \( (t, c, w) \) were widely used as the basis of test policy in many countries, while age and gender were basic factors for medical diagnosis. It should be noted that with more effective factors recorded in the COVID-19 test system, the feature vector \( \xi \) might have a larger dimension.

The prognostic model was constructed with logistic regression and trained by previous test data. Thus the final prognostic model was in the form described by Equations (4)–(6). As a classic machine learning model, The fitting coefficients could exhibit the influence of considered features, which was enlightening for the understanding of risk factors. The outcome of the logistic regression model could be consecutive between 0 and 1, i.e. the probability \( p \).

\[
f(\xi) = \omega^\top \xi
\]  

(4)

\[
s(x) = \frac{1}{1 + e^{-x}}
\]  

(5)

\[
y = s(f(\xi))
\]  

(6)

To evaluate the performance of a trained prognostic model, the ROC (receiver operating characteristic) curve and the index AUC (area under the receiver operating characteristic curve) were considered. The area under the ROC curve (often referred to as simply the AUC) was equal to the probability that a classifier would rank a randomly chosen positive instance higher than a randomly chosen negative one \([25]\), which was in line with our needs. Thereby, a closed-loop was formed for the optimization of feature selection and model construction.

Differently from the usual prognostic model in other medical studies using machine learning \([41–43]\), this work was not aimed at the prediction of a binary outcome by giving a threshold value, but to take the output of the prognostic model, i.e. the number \( p \) as the criterion for priority ranking. When the test capacity was given, those with a higher probability of being positive were more likely to be selected for the test.

4.2. Strategy evaluation and results comparison

We evaluated the performance of data-driven strategy on the collected data from 12 June to 11 July, and the data from 19 March to 12 June were hence used for model training. The size of the test dataset was 149k, while that of the training dataset was 107k. Readers kindly needed to distinguish the COVID-19 test from the term ‘test’ in machine learning in the following of the paper.

Fig. 6 showed three ROC curves of data-driven strategies trained by different factors. It was demonstrated that the curve on all the 5 factors \( (\xi_i = [a_i, g_i, t_i, c_i, w_i]) \) had the best performance (in red), and the

![Fig. 5. Framework of the data-driven strategy.](image)

![Fig. 6. ROC(receiver operating characteristic) curve for the data-drive strategy on data from 12 June to 11 July with different features considered. The trained prognostic model corresponding to the red ROC curve would be used in our strategy for the case study in the following.](image)
A common issue in data-driven problem was that what time period should be used for model training. The model trained by short term data might cause great uncertainty due to the stochasticity or the heterogeneity of data, while that trained by long term data might smooth out the behavior tendency when applied to recent data [48].

5. Adaptability on time and space

5.1. Temporal adaptability analysis

A widespread controversy was that the current strategy did not pay enough attention to those with wellness issues. In case 2 it was assumed that the test capacity was enlarged to 69479 to have all instances with wellness issues tested. It still turned out that the data-driven strategy had an overwhelming superiority against the current static strategy, with the positive detection amount 17284 against 6888, which might result from the transmission by people with mild or no symptoms.

In other words, following the data-driven strategy, only a much smaller test amount was needed to have the majority of positive cases detected (see Fig. 7). It should be noted that all the data were collected through real test, which meant that from 12 June to 11 July 119578 instances were tested in reality and 18170 of them were confirmed through the test. Assuming that the data-driven strategy was followed, then with 48.17% (57600/119578) of the original test amount, 89.35% (16234/18170) of total positive cases could be detected. Which implied a much more optimal utilization of test resources.

It should be noted that 62.09% (11282/18170) of detected positive cases had no symptoms, contact history, or travel history. This implied that the city failed to trace the transmission of Covid-19 at the time. The asymptomatic had become a major issue, and could only be detected through mass testing. From the aspect of policymaking, if positive cases were detected, but the transmission process could be traced, obviously, the most effective way to cut off the transmission was to test the whole population in certain relevant regions. Nevertheless, if the trace of transmission was lost, and the test capacity was far not enough to cover all the population (plenty of countries were actually facing the dilemma at the time [32]), a smart strategy was demanded to ensure that the test capacity was enlarged to 69479 to have all instances with wellness issues tested.

In the test dataset (‘test’ in terms of machine learning), 119578 instances were pending for test, and 18170 of them were positive. As shown in Table 3, merely 1463 positive instances were detected out of 57600 trials following the current strategy. In contrast, 16234 positive samples could be detected following the proposed data-driven strategy, which enhanced the performance by 1009.64% (Positive detection rate: 28.18% against 2.54%. Recall rate: 89.35% against 8.05%).

A widespread controversy was that the current strategy did not pay enough attention to those with wellness issues. In case 2 it was assumed that the test capacity was enlarged to 69479 to have all instances with wellness issues tested. It still turned out that the data-driven strategy had an overwhelming superiority against the current static strategy, with the positive detection amount 17284 against 6888, which might result from the transmission by people with mild or no symptoms.

In other words, following the data-driven strategy, only a much smaller test amount was needed to have the majority of positive cases detected (see Fig. 7). It should be noted that all the data were collected through real test, which meant that from 12 June to 11 July 119578 instances were tested in reality and 18170 of them were confirmed through the test. Assuming that the data-driven strategy was followed, then with 48.17% (57600/119578) of the original test amount, 89.35% (16234/18170) of total positive cases could be detected. Which implied a much more optimal utilization of test resources.

It should be noted that 62.09% (11282/18170) of detected positive cases had no symptoms, contact history, or travel history. This implied that the city failed to trace the transmission of Covid-19 at the time. The asymptomatic had become a major issue, and could only be detected through mass testing. From the aspect of policymaking, if positive cases were detected, but the transmission process could be traced, obviously, the most effective way to cut off the transmission was to test the whole population in certain relevant regions. Nevertheless, if the trace of transmission was lost, and the test capacity was far not enough to cover all the population (plenty of countries were actually facing the dilemma at the time [32]), a smart strategy was demanded to ensure that the test capacity was enlarged to 69479 to have all instances with wellness issues tested.

In the test dataset (‘test’ in terms of machine learning), 119578 instances were pending for test, and 18170 of them were positive. As shown in Table 3, merely 1463 positive instances were detected out of 57600 trials following the current strategy. In contrast, 16234 positive samples could be detected following the proposed data-driven strategy, which enhanced the performance by 1009.64% (Positive detection rate: 28.18% against 2.54%. Recall rate: 89.35% against 8.05%).

A widespread controversy was that the current strategy did not pay enough attention to those with wellness issues. In case 2 it was assumed that the test capacity was enlarged to 69479 to have all instances with wellness issues tested. It still turned out that the data-driven strategy had an overwhelming superiority against the current static strategy, with the positive detection amount 17284 against 6888, which might result from the transmission by people with mild or no symptoms.

In other words, following the data-driven strategy, only a much smaller test amount was needed to have the majority of positive cases detected (see Fig. 7). It should be noted that all the data were collected through real test, which meant that from 12 June to 11 July 119578 instances were tested in reality and 18170 of them were confirmed through the test. Assuming that the data-driven strategy was followed, then with 48.17% (57600/119578) of the original test amount, 89.35% (16234/18170) of total positive cases could be detected. Which implied a much more optimal utilization of test resources.

It should be noted that 62.09% (11282/18170) of detected positive cases had no symptoms, contact history, or travel history. This implied that the city failed to trace the transmission of Covid-19 at the time. The asymptomatic had become a major issue, and could only be detected through mass testing. From the aspect of policymaking, if positive cases were detected, but the transmission process could be traced, obviously, the most effective way to cut off the transmission was to test the whole population in certain relevant regions. Nevertheless, if the trace of transmission was lost, and the test capacity was far not enough to cover all the population (plenty of countries were actually facing the dilemma at the time [32]), a smart strategy was demanded to ensure that the test capacity was enlarged to 69479 to have all instances with wellness issues tested.
The average AUC versus different training time periods was shown as Fig. 8(a). Each point corresponded to the average AUC of 20 randomly sampled test periods with a length of 7 days. The shortest term of training datasets was just one day before the test period, while the longest term included the period within 1–60 days before the test period. Despite the fluctuation, no obvious tendency was observed. The average AUC trained by data of different periods was 0.701 ± 0.017. Thus the length of the time period for training would not make a great difference to decision making, and long-term data collection was not a prerequisite for the conduction of data-driven strategy. Therefore, the insufficiency of test datasets in the temporal dimension was not so troublesome for decision making.

Further analysis found that training time periods could influence the performance of the strategy slightly as the pandemic developed. Fig. 8(b) showed the AUC of different test time periods (‘test’ in terms of machine learning) from 12 June to 11 July over time respectively. On the one hand, a gap was observed between 23 June-27 June and 28 June-1 July, which implied a noticeable transition of potential factors around the time. On the other hand, for test periods (28 June-11 July) after the transition, a slight downward trend against the length of training time periods could be observed. The two observations indicated that after a noticeable transition of potential factors, a long training time period that included more data before the transition could bring a negative impact on the performance of the data-driven strategy. However, it should also be noted that despite the downward trend was observed under certain circumstances, compared with the relative gap of AUC among test windows, the fluctuation was inconspicuous. Thus from the aspect of the practice, smart refinement on the training time period might be helpful, but it was also acceptable even if the policymaker just skipped the selection and used a uniform length of training time periods due to the lack of long-term data.

Although AUC for each test data period remained quite stable over the change of training data period, the change of the test data period itself did influence the AUC value as seen in Fig. 9 (red solid line, each AUC value was the average of AUC values trained by data within ten days before the test period). The length of each test period was 5 days. The latest period was 7 July-11 July (corresponding to the leftmost point) and the furthest period was 1 May-5 May (corresponding to the rightmost point). Two rapid rises of AUC were interestingly discovered around 55–50 and 20-15 days before 11 July, i.e. 18 May-23 May and 21 June-26 June. While two synchronous rises of test capacity (blue dashed line) were found around the time surprisingly. It was speculated that the performance of the data-driven strategy was promoted by the booming of test capacity. The relationship needed further research in the future.

It was important for policymakers to keep following the latest data and update the database continuously in case of the transition of potential factors in the pandemic. While for regions with a slow reaction to the pandemic, which was usually in a lack of long term data, the data-driven strategy also had a positive effect. Besides, the reliability of the model would increase with the test amounts conducted in the near term, which indicated that a larger test capacity led to a better test efficiency following the strategy. Based on this fact, policymakers should not consider reserving any test resources for the reason of economics [30, 32].

5.2. Spatial adaptability analysis

Though the data-driven strategy was studied at the city scale, the relative features showed considerable distinction at different towns in Lahore seen as Fig. 10. To investigate if the driven data variation resulted from spatial variation had a significant influence on the strategy performance, ten subareas with distinct population density and economy were

---

Fig. 8. AUC of fixed test periods versus length of training time periods (‘test’ in terms of machine learning). (a) AUC averaged by all test periods. (b) AUC of different test periods from 13 June to 11 July.

Fig. 9. AUC for every five days as a test period (‘test’ in terms of machine learning) versus time. Each AUC value was the average of AUC values trained by all-time periods before the test day. The blue dashed line showed the change of test capacity. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
A cross-test ('test' in terms of machine learning) was conducted among the ten subareas, including the test using the overall dataset. In each run of the test, we used the data of one subarea for training, and test it on the data of another subarea. For the self-test, we split the data of a single subarea into training datasets and test datasets (size ratio 8:2) randomly and conduct the test for 20 times, the 20 outcomes of AUC were averaged as self-test result. The maximum AUC of each row in the cross test matrix laid distinctly on the diagonal line as shown in Fig. 11, which indicated that the spatial variation did play a role in model training. The strategy driven by the local data showed overwhelming superiority (AUC = 0.77 ± 0.04) than that driven by data of other subareas, including that driven by the overall data in Lahore (AUC = 0.62 ± 0.04).

From the perspective of management, an argument in some countries was that there was a need to develop a national testing policy instead of testing guidelines so that all states/provinces were able to prepare their own epidemic preparedness profile to geographically assess the population at risk [32]. However, according to our study, better management could be further achieved with the policymaking more refined on space, instead of just at the scale of states/provinces, but at the scale of subareas of cities. It was recommended that each subarea build its own COVID-19 data system. The policy-making on test strategy should distinguish subareas, and trained the model with the corresponding database respectively.

6. Conclusions and perspectives

The policy on test strategy under limited test capacity could make a great difference to the testing result, thus further affect the understanding and response to the spread of the pandemic, which did not receive adequate attention, however.

We gave a mathematical definition on test strategy, and took the current strategy of Lahore as an example to analyze how such static strategy failed to fit the development of pandemic-the risk factors at the incipient stage could have transited into protective factors as the pandemic developing.

A self-adaptive data-driven strategy based on machine learning was proposed, and its overwhelming superiority over the static strategy was demonstrated, where the positive detection rate was enhanced from 2.54% to 28.18%, and the recall rate was boosted from 8.05% to 89.35% under strictly limited test capacity. Besides, much more optimal utilization of test resources could be realized where 89.35% of total positive cases could be detected with 48.17% of the original test amount. The strategy showed self-adaptability to the development of pandemic, while the performance turned out to be rising with the enlargement of test capacity. It was found that long-term data collection was not a prerequisite for the conduction of data-driven strategy, thus the insufficiency of test datasets in the temporal dimension was not troublesome for decision making. The strategy driven by local data was also proved to be optimal than that driven by data of other or overall districts in Lahore. The main implications of this research on management and policymaking in the pandemic were concluded as follows:

a) It was important for policymakers to keep following the latest data and update the database continuously in case of the transition of potential factors in the pandemic.

b) The major significance of the data-driven strategy was that it helped the policymaker attaching importance to both the efficiency and objectivity of the test criteria in the mass testing when troubled with little trace of the transmission process and an insufficient test capacity.

c) Smart refinement on the training time period might be helpful, but it was also acceptable even if the policymaker just skipped the selection and used a uniform length of training time periods due to the lack of long term data at the incipient stage.

d) A larger test capacity led to a better test efficiency following the strategy. Policymakers should not consider reserving any test resources for the reason of economics.

e) Better management could be further achieved with the policymaking more refined on space. It was recommended that each subarea build its own COVID-19 data system. The policy-making on test strategy should distinguish subareas, and trained the model with the corresponding database respectively.

Due to the restriction of data, we could not give a complete assessment of all possible effective factors. We believed that with more effective factors considered, the performance of the data-driven strategy could be further promoted. There were also some findings discovered but could not be thoroughly explained in the regime of this study, such as the transition of travel history and contact history factors, the synchronization between the AUC value and test capacity. These findings could be further investigated in future research with more data collected.

We recommended a generalization of such a data-driven test strategy for a better response to the global developing pandemic. Besides, the construction of the COVID-19 data system should be more refined on space for local applications.
Permission declaration

The data used in this study were supplied by the District Health Authority, District Administration Lahore for research purposes.

Author contributions

Chuanli Huang and Min Wang contributed equally to this work.

CRediT authorship contribution statement

Chuanli Huang: Methodology, Formal analysis, Writing – original draft. Min Wang: Investigation, Validation, Formal analysis. Warda Rafafat: Conceptualization, Data curation. Salman Shabbir: Resources. Liping Lian: Funding acquisition. Jun Zhang: Writing – review & editing. Siuming Lo: Supervision. Weiguo Song: Conceptualization, Supervision.

Acknowledgment

The authors acknowledge the foundation support from the National Key Research and Development Program (2018YFC0807000), National Natural Science Foundation of China (71904006).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.seps.2021.101091.

References

[1] Polak SB, et al. A systematic review of pathological findings in COVID-19: a pathophysiological timeline and possible mechanisms of disease progression. Mod Pathol 2020;1–11.

[2] Verikios G. The dynamic effects of infectious disease outbreaks: the case of pandemic influenza and human coronavirus. Socio-Economic Planning Sciences; 2020. p. 100911.

[3] de Bruin YB, et al. Initial impacts of global risk mitigation measures taken during the combatting of the COVID-19 pandemic. Saf Sci 2020:128:8.

[4] Ocampo I, Yamagishi K. Modeling the lockdown relaxation protocols of the Philippine government in response to the COVID-19 pandemic: an intuitionistic fuzzy DEMATEL analysis. Socio-Economic Planning Sciences; 2020. p. 100911.

[5] Boysen-Hogrefe J, et al. German Economy Summer 2020–German economy faces sluggish recovery. 2020.

[6] Bonacini L, Gallo G, Patriarca F. Identifying policy challenges of COVID-19 in hardly reliable data and judging the success of lockdown measures. GLO Discussion Paper; 2020.

[7] De Bruin YB, et al. Risk management measures for chemicals in consumer products: documentation, assessment, and communication across the supply chain. J Expo Sci Environ Epidemiol 2007;17(1):55–66.

[8] Kerstens K, Shen ZJTF, change s. Using COVID-19 mortality to select among hospital plant capacity models: an exploratory empirical application to Hubei province 2021.16:120535.

[9] World-Health-Organization. Coronavirus disease 2019 (COVID-19) situation report – 71. March 2020. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200331-sitrep-71-covid-19.pdf?sfvrsn=436/d692b.4.

[10] Our-World-in-Data. Coronavirus(Covid-19) testing: statistics and research. 7 July 2020. https://ourworldindata.org/coronavirus-testing.

[11] Hellewell J, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. The Lancet Global Health; 2020.

[12] Tanne JH, et al. Covid-19: how doctors and healthcare systems are tackling the coronavirus (COVID-19) in China and will influence future Urban health policy internationally. A&I 2020;1(2):156–65.

[13] Santosh K. AI-driven tools for coronavirus outbreak: need of active learning and cross-population train/test models on multitudinal/multidimensional data. J Med Syst 2020;44(5):1–5.

[14] Pagani, G., et al., Prevalence of SARS-CoV-2 in an area of unrestricted viral circulation: Mass seroepidemiological screening in castiglione D’adda, Italy.

[15] Di G, et al. Simple and effective primary assessment of emergency patients in a COVID-19 outbreak area: a retrospective, observational study. Risk Manag Healthc Pol 2020;13:1253–60.

[16] World-Health-Organization. Advice on the use of point-of-care immunodiagnostic tests for COVID-19. 2020 April. https://www.who.int/news-room/commentaries/detail/advice-on-the-use-of-point-of-care-immunodiagnostic-tests-for-covid-19.

[17] Salathe M, et al. COVID-19 epidemic in Switzerland: on the importance of testing, contact tracing and isolation. Swiss Med Wkly 2020;150(11–12):w20225.

[18] Raoofi A, et al. COVID-19 pandemic and comparative health policy learning in Iran. Arch Iran Med 2020;23(4):220–34.

[19] Castro R, et al. COVID-19: a meta-analysis of diagnostic test accuracy of commercial assays registered in Brazil. The Brazilian Journal of Infectious Diseases; 2020.

[20] Senghore M, et al. Leveraging Africa’s preparedness towards the next phase of the COVID-19 pandemic. The Lancet Global Health; 2020.

[21] Niud Y, Xu FJ. Deciphering the power of isolation in controlling COVID-19 outbreaks. Lancet Global Health 2020;8(4):E452–3.

[22] European-Centre-for-Disease-Prevention-and-Control, Testing strategies for SARS-CoV-2.

[23] Centers-for-Disease-Control-and-Prevention, Overview of testing for SARS-CoV-2 (COVID-19). Oct. 20, 2021.

[24] Thomas S. Why the U.S. Needs to change its testing strategy to fight the coronavirus pandemic. 2020. Aug 13.

[25] Commission Recommendation (EU) 2020/1595 of 28. On COVID-19 testing strategies, including the use of rapid antigen tests. October 2020.

[26] Sarkar A, et al. Public health preparedness and responses to the COVID-19 pandemic in South Asia: a situation and policy analysis. 2020.

[27] European-Data-Portal, Widespread Testing. Different strategies across Europe. 06 Aug 2020.

[28] Shahid SS. Coronavirus: shortage of resources in Pakistan. 2020. Apr 30.

[29] Dunya-News. Govt. hospitals run out of coronavirus testing kits in Lahore. 09 June 2020.

[30] World-Health-Organization, reportRevised case report for confirmed novel coronavirus COVID-19.

[31] Government-of-the-Pakistan-Primary-&-Secondary-Healthcare-Department. Annexure 3: district health emergency response committee. 19 March 2020. https://pshealth.punjab.gov.pk/Upload/Downloads/03xjmny3.gdy20200515.pdf.

[32] World-Health-Organization. Annexure 14: patient reporting form. 2020. https://pshhealth.punjab.gov.pk/Upload/Downloads/03xjmny3.gdy20200515.pdf.

[33] Primary-&-Secondary-Healthcare-Department. COVID-19 NOTIFICATION/SOPS & guidelines. 2020. https://pshealth.punjab.gov.pk/Home/Covid19SopsGuidelines.

[34] World-Health-Organization. COVID-19 - virtual press conference. 30 March 2020. https://www.who.int/health-topics/coronavirus-(2019-ncov)/updates']/home/corona-virus/press-conference-full-30mar2020.pdf?xmlv=s6b8a4e2.

[35] Colubri A, et al. Machine-learning prognostic models from the 2014–16 Ebola outbreak: data-harmonization challenges, validation strategies, and mHealth applications. EClinicalMedicine 2019;11:54–64.

[36] Sakagianni A, et al. Setting up an easy-to-use machine learning pipeline for medical decision support: case study for COVID-19 diagnosis based on deep learning with CT scans. Stud Health Technol Inform 2020;272:13–4.

[37] Yu Z, et al. Rapid identification of COVID-19 severity in CT scans through classification of deep features. 2020.

[38] Urban health policy

[39] World-Health-Organization. Coronavirus disease 2019 (COVID-19): a pathophysiological timeline and possible mechanisms of disease progression. Mod Pathol 2020;1–11.

[40] Duffey RB, Zio E. Prediction of CoVid-19 infection, transmission and recovery rates: a new analysis and global societal contributions. Saf Sci 2020;128:8; 494–506.

[41] Alam Z, De G, Jones DS. Artificial intelligence (AI) provided early detection of the coronavirus (COVID-19) in China and will influence future Urban health policy internationally. A&I 2020;1:156–65.
Min Wang was born in Tongcheng city, Anhui Province, China in 1993. She is a physician received the Medical Bachelor’s degree from Jining Medical College in 2017 and clinical degree of medical Master from Soochow University in 2020. From 2012 to 2020, she has been working in the Third-grade A hospital. At present, she has mastered the diagnosis and treatment of common diseases of different clinical departments. Besides, She has published 2 scientific and technological core journals as the first author and has participated in university-level research project. Her research interests include standardized treatment of malignant tumors, nuclear emergency injury treatment and health protection against occupational diseases.

Warda Rafquat born in Pakistan did her sixteen years education in geographical information systems from Punjab University College of Information Technology, Lahore Pakistan. She worked with Urban Unit and Data Solutions Private Limited for three years. Currently doing her Masters from USTC Hefei. GIS, RS, ML are her interests and implementation of the fields for Safety and Management.

Salman Shabbir born in Pakistan did his sixteen years education in geographical information systems from Punjab University College of Information Technology, Lahore Pakistan. He worked with Urban Unit and Punjab Disaster Management Authority. Currently working with Punjab Information Technology Board. He worked in Punjab Dengue Program and worked under different projects for Punjab Government where he used GIS techniques.

Liping Lian is a postdoctoral fellow in Peking University Shenzhen Graduate School. She received her PhD at University of Science & Technology of China (2016) and City University of Hong Kong. She is interested in pedestrian evacuation and dynamic and publish six scientific papers in this field. She is responsible for one project from National Natural Science Foundation of China and one project from China Postdoctoral Science Foundation.

Prof. Jun Zhang is the professor of University of Science & Technology of China. He received his PhD at Peking University in 2001. His has been focusing on pedestrian and evacuation dynamics, capacity assessment on emergency treatments, and application of virtual reality technology. He is a communication expert of National Natural Science Foundation of China. His work have been supported by two projects from the National Natural Science Foundation of China, a project from National Key Research and Development Program of China (Grant No. 2018YFC0808600), a project from the Anhui Provincial Natural Science Foundation and a project from the Fundamental Research Funds for the Central Universities. Dr. Zhang has published more than 20 scientific papers.

Prof. SM Lo is a professor of the Department of Architecture and Civil Engineering, City University of Hong Kong. He received his PhD in Architecture from the University of Hong Kong. His main research interests include spatial planning for pedestrian movement, decision support system, fire protection engineering and evacuation, urban planning and construction management. He has been the principal investigator for many research grants including 14 Competitive Earmarked Research/ GRF Grants from the Research Grant Council of Hong Kong. His work have been supported by two projects from the Ministry of Science and Technology of China, one project supported by the Ministry of Science and Technology of China, three projects from the National Natural Science Foundation of China, and one project from the Ministry of Education of China. Moreover, Prof. Song has participated in many research projects supported by the National Basic Research Program of China (also known as the “973” Program), and the National Key Technologies R & D Program of China in his capacity as a key researcher. Prof. Song entered the list of recipients supported by the ‘New Century Talents Supporting Program’ of the Ministry of Education in 2004. Prof. Song has authored over 60 research papers, published one book, and owned copyrights of four computer softwares. Prof. Song is a member of International Association of Fire Safety Sciences(IFASS), member of China Association of Public Safety Science and Technology (2012-). He is also a reviewer of Risk Analysis, Physica A, Building and Environment, Ecological Modelling, IEEE Transactions on Intelligent Transportation Systems, International Journal of Remote Sensing etc.