INTRODUCTION

The coronavirus disease 2019 (COVID-19) outbreak caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was first sparked in Wuhan, the capital of Hubei province in China. From the epicentre, it subsequently spread to the entire country and now the globe. Since the declaration as a pandemic by WHO on 11 March 2020 (Green, 2020), it is infiltrating into every corner...
of the world, but local epidemics associated with this pandemic are highly dynamic. Some regions have developed as epicentres, whereas others may struggle with imported cases. There is currently intense debate and great confusion among political leaders, healthcare authorities and the general public on how to respond to the COVID-19 pandemic (The, 2020).

Since the outbreak in December, 2019, in Wuhan, the city rapidly grew into an epicentre. Because Wuhan is a major transportation hub located in the centre of China and the outbreak coincided with a massive population movement due to the Chinese lunar new year holiday, it quickly spread to the entire country (Chen et al., 2020). Since 23 January 2020, the central government ordered heavy control measures, including city lockdown, travel restriction and within-population quarantine. The pandemic was eventually under control by early March nationwide.

Extensive clinical, epidemiological and modelling studies have well-characterized the epidemic features of the epicentre Wuhan (Guan et al., 2020; Tian et al., 2020; Wu et al., 2020). In contrast, little attention has been paid to the initial case importation and subsequent epidemic control of the different parts of Mainland China outside Wuhan. In this study, we aim to gain insight into this respect by mathematical modelling. Classical epidemic compartmental models, such as the susceptible-infected-recovered (SIR) or susceptible-exposed-infected-recovered (SEIR) model, have been widely and proven to be useful for modelling COVID-19 epidemics (Tolles & Luong, 2020). These basic models are easy to compute, but also oversimplify the complexity of disease processes, the heterogeneity of target population/society and the diversity of control measures.

In this study, we aim to develop a mathematical model to recapitulate the SARS-CoV-2 transmission patterns from epicentre to other regions during the early stages of the outbreak. To better recapitulate the real-world complexity and heterogeneity, we constructed a modified multi-regional, hierarchical-tier susceptible-latent-infected/hospitalized-recovered (SLIHR) model by incorporating the population flow network data. With this model, we estimated the overall and individual epidemics in different provinces of Mainland China outside Wuhan. We further performed comprehensive data/information mining to understand the effectiveness of specific control measures in contributing to local epidemic control at provincial level.

2 | METHODS

2.1 | Model assumption

A two-stage, hierarchical-tier, multi-regional SEIR model for COVID-19 epidemic was developed based on the population flow network. The first tier of our model was an open epidemic transmission system that described the risk of case importation. Before January 24, COVID-19 spread freely in Wuhan, and the isolation for infected individuals was incomplete under insufficient medical resources. The actively infected and latently infected individuals who were neither hospitalized nor isolated had been imported to other provinces in Mainland China as the population flow and then spread locally. In addition, our model specified infected individuals who could not be hospitalized for isolation due to medical resource constraints in Wuhan at the early stage. This population largely contributed to case importation from Wuhan to other provinces of Mainland China. Simultaneously, the model considered the heterogeneity of infection risk between close contacts and general contacts. The second tier of our model was multiple independent closed epidemiological transmission systems describing the local spread of COVID-19. From January 23 to January 25, the Level 1 emergency response was activated in all provinces of Mainland China except Xizang Province. The COVID-19 epidemic was confined to multiple separate closed systems and localized transmission occurred within each province under strict interprovincial traffic restrictions.

We simulated the possible epidemic scenarios assuming that there were no control measures with localized transmission in provinces under the same basic reproduction number. The gaps between the number of probable infections without control measures and real-world confirmed cases in each province represent the effectiveness of provincial responses on controlling the COVID-19 epidemic.

2.2 | Data sources

The Chinese government implemented strict control measures. Large-scale screening was conducted to identify infected cases and contacts with confirmed cases were closely tracked, resulting in rapid containment of the epidemic in China (Xing et al., 2020). Thus, data of confirmed cases reported by China National Health Commission highly represent the real-world epidemic (Table S1). Population flow network was based on ‘Baidu Migration’ Big Data Platform, the largest database that reflects the size of the regional population inflow or outflow according to the geographic change of the users’ mobile devices in China (Baidu Online Network Technology (Beijing) Co. L. Baidu Map Smart). Data/Information on control measures of different provinces in Mainland China were comprehensively collected from local government documents, announcements and press conferences (Table S2).

2.3 | Mathematical model

A modified hierarchical-tier, multi-regional SLIHR (susceptible—latently infected—un-isolated actively infected—hospitalized and isolated—dead) model with interregional and interpersonal network was built (Figure 1). For interregional transmission network, we considered that China has developed advanced transportation networks, and the early outbreak was coincided with a massive population migration, because of the Chinese lunar new year holiday. In the free spreading stage, Wuhan and its neighbours form a star network structure with traffic connections, and the latently (without symptom) or actively (with symptoms) infected individuals were exported
from Wuhan in one direction. The initial COVID-19 cases of other regions in China were mainly imported from the epicentre. Thus, at the controlling stage, the initial values of multi-regional model were based on Wuhan population outflow network.

Considering the real-world situation of COVID-19 epidemics in China, we first developed an open SLIHR model (Figure 1a) for the epicentre Wuhan considering incomplete isolation and heterogeneous interactions. Transmission of SARS-CoV-2 occurs in the open system, and the total population of Wuhan varies with the cross-regional inflow and outflow of the population. The susceptible, the latently infected and infected but un-isolated individuals had been imported to other provinces of Mainland China. The probability of each subpopulation flowing out or staying in Wuhan was assumed equal. The infected individuals who were isolated, admitted to hospital or recovered could not flow out of Wuhan. In the controlling stage, SARS-CoV-2 was confined to several independent and closed systems under traffic restrictions after Level 1 emergency response was implemented. First-generation cases in other regions outside Wuhan depended on population outflow from Wuhan and the number of latent or active patients without isolation in Wuhan before travel ban.

For better understanding the epidemic spread, interpersonal transmission network was also established (Figure 1b). The real-world interpersonal contact network is not homogeneous but close to community structure where individuals are grouped by different interaction frequency, which directly result in heterogeneous transmission in our model. The nodes (individuals) in the interaction network are grouped. The nodes (individuals) in the group (members of the same family) have high frequency of contact, while the nodes (individuals) between the groups (daily contact acquaintances) have low frequency of contact.

In order to precisely capture the epidemic spread from epicentre to adjacent regions in Mainland China, we adopted parameters from widely cited studies on early epidemics in Wuhan and other regions in Mainland China (Chinazzi et al., 2020; Li et al., 2020). The details of our model were described in Appendix S1, and the parameters in our model were referred to widely cited literatures (Table S3). We first estimated the epidemics in Wuhan and the exportation
to other adjacent regions before travel ban. We then simulated the possible epidemic scenarios in other provinces except Wuhan under the assumption of no control measures. The heterogeneous effects of control measures in each province could be analysed by comparing the number of probably infected individuals and the real-world confirmed cases.

### 2.4 Classification of control measures

Based on collected data, we characterized their epidemic management systems in six dimensions and classified control measures into three levels based on the timeliness of implementation. The dimension of Level 1 emergency response reflects the speed in responding to the outbreak. Traffic restrictions represent the government’s ability to control interprovincial traffic and urban public transport in a timely manner. The dimension of mask wearing claim portrays the intensity and timeliness of personal protection in public places requested by local governments. Community prevention means that governments conduct population screening in the communities. Big data tools are used by governments for contact tracing, accurate forecasting and control. The dimension of work resumption is expressed as the rate of return-to-work. The timeline for each control measure was determined by the median time of all provinces, and each province is assigned a rating based on the chronological order in which that measure was implemented.

### 2.5 Scope and timeframe

The effectiveness and heterogeneity of epidemic containment across different parts of China mainly depend on the levels of case importation from epicentre and the subsequent control measures. Our network-based mathematical model is capable of recapitulating and integrating these multiple factors. Because control measures were ordered by the central government on January 23 and a temporary diagnostic method that is different from the regular RT-PCR case diagnosis was adopted on February 12 in Hubei province; therefore, we used the epidemiological data from January 24 to February 11 to assess the epidemic control.

### 2.6 Data visualization and statistic analysis

The mathematical models were solved by simulation methodology with the help of MATLAB 2016b (The MathWorks, Inc.). Epidemic maps were drawn using ArcGIS 10.1 (Environmental Systems Research Institute, Inc.). Kendall’s tau-b correlation coefficient was chosen to test the non-parametric correlation between ordered categorical variables and a continuous variable, where multiple mutual linear problems of independent variables were excluded by collinearity diagnosis. We used the Kendall’s tau-b correlation coefficient method to test the non-parametric correlation between the varied control measures and the number of reduced probable infections. Statistical analysis was performed using IBM SPSS Statistics version 24 (International Business Machines Corporation). \( p < .05 \) was considered as statistically significant, and all tests were two-tailed.

### 3 RESULTS

#### 3.1 Simulating the epidemics in Wuhan and other regions of Mainland China

We collected and categorized the number of daily confirmed new cases in Wuhan and other regions of Mainland China, respectively (Figure 2a). As expected, the first case in Wuhan appeared much earlier, but the epidemic curve was delayed as compared to that of Mainland China outside Wuhan. This indicates delayed case diagnosis during the early epidemic in Wuhan. The overall epidemic period was much shorter in other regions of Mainland China compared to that in Wuhan. Of note, the unusual escalation of case number on February 12 reported by China National Healthcare Commission was attributed to the temporary inclusion of the ‘clinically diagnosed’ cases without RT-PCR confirmation (Sun et al., 2020). This was intended to swiftly isolate and treat the large number of suspected cases in the epicentre, while the capacity of RT-PCR testing was limited at that time.

We first simulated the epidemic in Wuhan by an open SEIHR model considering incomplete isolation and heterogeneous interaction. The infected but unisolated individuals were then estimated based on population flow network (Figure 1). The probable infections in each province except Wuhan were simulated assuming no control measures implemented. Our simulation revealed that the number of latently infected individuals without symptom was 1,600, and the total number of infected individuals with symptom was 1,581 in Wuhan as of January 24. We estimated that 369 latent or active patients already left Wuhan before implementation of the travel ban, thus spreading to other regions outside Wuhan.

Our model simulation fits with the reported cases during the early stages of the epidemics, for example by February 13 in Wuhan (Figure 2b) and by February 1 outside of Wuhan (Figure 2c). More specifically, the simulated number of isolated (hospitalized) active cases in Wuhan was in agreement with the reported confirmed cases by January 24. However, by that early stage, screening capability and
hospital beds were insufficient, there were infected but unidentified cases difficult to be estimated. From January 25 to February 14, the total number of simulated actively infected individuals was consistent with the reported data (Figure 2b). Similarly, in other regions outside Wuhan, the simulated infections assuming no control measures were in line with the reported data in January. But since February, there was an increasing gap between estimated and reported numbers, which exactly indicates the effects of implemented control measures in limiting epidemic spread (Figure 2c).

3.2 | Simulating the overall effects on epidemic control in Mainland China outside of Wuhan

From January 23, the Chinese central government began to implement heavy control measures, including city lockdown, travel ban and within-population quarantine. Here, the effects of control measures were based on the reduced number of probable infections which was calculated by subtracting confirmed cases from the simulated probable infections assuming no control measures in each province.

We estimated that the probable number of actively infected individuals would exceed 17.4 (90% CI 8.1–39.1) million and the number of latent individuals would be 29.8 (90% CI 11.7–78.9) million in Mainland China outside Wuhan 1 month after January 24, if no control measures were implemented. In real world, control measures were universally implemented albeit at various levels across different provinces of China. Thus, the number of reported daily new cases peaked around 10 days after implementing control measures but declined steadily thereafter. The rebound on February 12 was attributed to the temporary adoption of the new clinical diagnose method (Figure 2d).

In general, the simulated results before the epidemic peak were highly in agreement with the reported confirmed cases, both for Wuhan and outside Wuhan in Mainland China. In contrast, the reported confirmed cases after the peak were far less than the simulated numbers, which reflected the effects of control measures in mitigating SARS-CoV-2 spread. Two weeks of implementing control measures were estimated to reduce the probable number of infections by 56,535 (90% CI 31,681–96,743) in Mainland China outside Wuhan. This constitutes a 73% (90% CI 60.2%–82.2%) reduction. One month after implementation, the number of infections was reduced by 99.8% (90 CI% 99.6–99.9), thus preventing 17.4 (90% CI 8.1–39.1) million people from infection.

3.3 | Heterogeneity in epidemic control across different provinces in Mainland China

After January 24, the movement of people across provinces was restricted due to travel ban. Thus, the local transmission of COVID-19 in each province is considered as an individual unit, as described by the second tier of our model framework (Figure 1). COVID-19 spread in each province was independent of each other, and the control measures presented heterogeneity. In this case, different effectiveness of provincial responses could be estimated by comparing the real-world confirmed cases data and the number of probable infections assuming no control measure.

We generated a country map with cross-sectional comparison that visualizes the distinct effectiveness in containing COVID-19 epidemics across different provinces in Mainland China (Figure 3). The different gaps between reported cases and simulated numbers assuming without control measures were also shown for each province. The larger the gap indicates the more effective in epidemic control in that province (Figure 4). To further compare the differences at different stages, we mapped the real-world COVID-19 spread among different provinces during 1 month after implementing control measures (Figure S1).

3.4 | The effectiveness of specific control measures on local epidemic containment at provincial level

To understand the effectiveness underlying the heterogeneity in epidemic control, we characterized their epidemic management systems in six dimensions and classified control measures into three levels based on the timeliness of implementation (Figure 5).

By correlation analysis, we found that Level 1 emergency response, community prevention and application of big data tools significantly correlated with the effectiveness of local epidemic containment (Table 1). Traffic restriction and mask wearing were universally and strictly implemented across the country, and therefore did not show statistically significant correction with the effectiveness of local epidemic containment (Table 1). We did not include the item work resumption in this analysis, as this occurred mainly at the end or post-epidemic, which could be relevant for possible second wave but not the primary epidemic.

4 | DISCUSSION

Effective mitigation of COVID-19 pandemic requires deep understanding of transmission dynamics and control measures for both epicentres and emerging regions primarily with imported cases. However, research tends to focus on epicentres but disregard other regions. In fact, emerging regions with imported cases are relatively easy to be contained, but neglectance bears high risk of growing into new epicentres.

Classical compartmental mathematical models are excellent for studying epidemics in relatively homogeneous settings (Chinazzi et al., 2020; Giordano et al., 2020; Salje et al., 2020). The COVID-19 epidemics in other regions of China outside Wuhan, however, involve initial case importation and subsequent local transmission and control measure implementation (Ji et al., 2020). An innovative aspect of this study is that we constructed a two-tier SLIHR model to accommodate this complexity and heterogeneity. The first tier
ZHENG ET AL.

considered an open SLIHR model of latent time lag, incomplete isolation, heterogeneous contacts and exported cases from Wuhan to other regions in Mainland China. The second tier captured the spread of COVID-19 in multiple closed and unassociated regions based on the pre-imported cases. In our model, we adopted epidemiological data of case numbers reported by China National Health Commission. There was debate on the accuracy of reporting case number at early stage of the outbreak in Wuhan (Wu et al., 2020), but we mainly focus on the epidemics outside the epicentre. Their reported data are highly accurate, because identification of imported cases and contact tracing were rigorously implemented in all provinces of Mainland China. Furthermore, we did not consider the death cases, because data reporting on death was scarce and inconsistent at early stage of the epidemic. More importantly, we focus on regions outside Wuhan where death rates were extremely low in general but with huge variations across different provinces (Ji et al., 2020).

Exporting cases from Wuhan to other regions of China is a heterogamous process mainly determined by population migration. We have incorporated the large population flow network data in our model. These data were retrieved from the ‘Baidu Migration’ big data platform which is based on the user’s mobile device geographic location changes and reflects the size of population regional inflow or outflow. There are many factors affecting the level and pattern of population flow, including geographic locations, transportation connections, socioeconomic status, population characteristics, social values and cultural norms. In Mainland China, there are 22 provinces, five autonomous regions and four municipalities, which are all at provincial level but have their own distinct features. For example, the Tibet Autonomous Region (Xizang) only had one imported case and it was immediately contained. This is clearly attributed to the geographic, population and cultural distinctions of the region that has minimal population movement between the epicentre. In contrast, Zhejiang province which is several hundred kilometres away from Wuhan had one of the highest numbers of imported cases. This was mainly related to advanced economic development that many people from this province are running businesses in Wuhan, and returned back for Chinese New Year festival. The high level of case importation has compromised their effectiveness of epidemic control, although Zhejiang province has implemented heavy measures with ample resources available (Qian et al., 2020).

Although the general policy of control measures was ordered by the central government, it is operated and coordinated at provincial level locally. This is the second level attributing to the heterogeneity of effectiveness in epidemic control among different provinces. For example, Heilongjiang, the northernmost province of the Northeast region, had limited number of imported cases, but their epidemic grew substantially. As we estimated, Heilongjiang has the lowest effectiveness in controlling COVID-19 epidemic.

In this study, we in-depth analysed the association of six dimensions of control measures with the effectiveness of epidemic control in different provinces. We did not observe significant correlation between traffic restrictions or mask wearing claim with effectiveness of local epidemic control. Travel restriction is essential for limiting COVID-19 spread. But the central government has already imposed strict lockdown across the country, which may explain that...
FIGURE 4 The effect of province-level responses in Mainland China excluding Wuhan. Simulated and confirmed case numbers were plotted for each province, and the effectiveness of local epidemic control was calculated as percentage. Because the epicentre Wuhan is the capital for Hubei province. Thus, the case number of Wuhan was excluded for calculating Hubei province. Data on confirmed cases were reported by China National Health Commission.
this policy dimension from local government may not have additional effect. Whether wearing face masks for the general population can protect against SARS-CoV-2 has long been debated, especially in the western world (Lazzarino et al., 2020). It now becomes clear that people wear masks not only protect themselves, but also protect others by limiting spread of respiratory droplets. Many countries, including from western world, have required or advised their citizens to wear masks in public places (Cheng et al., 2020). In contrast, the Chinese population is highly aware of the protective values of wearing face masks (Wang et al., 2020), and they spontaneously and universally adopted this measure even before the request from government. This may explain why the dimension of mask wearing claim from local government did not have additional impact in our study.

It is not surprising that Level 1 emergency response is significantly associated with epidemic control effectiveness. We also found community prevention and application of big data tools are significant factors. Both require resources, expertise and advanced economic status. The economic status and growth in China have geographic imbalances, and our results indicate that this appears to have an effect on local epidemic response. We call the authorizes to pay attention to this regional inequality and to ensure equal access to resources, advanced tools and technologies for enhancing outbreak preparedness across the country.

In summary, we developed a multi-regional, hierarchical-tier SLIHR model that is capable of recapitulating the complexity and heterogeneity of COVID-19 epidemics in China. We revealed insight into how COVID-19 was spread from the epicentre to other regions of Mainland China and characterized the key control measures that contributed to the effectiveness of local epidemic containment. These findings bear important implications for many countries or regions to understand and better respond to their local epidemics associated with this COVID-19 pandemic.

CONFLICT OF INTEREST
The authors do not have any disclosures to report.
AUTHOR CONTRIBUTIONS

ETHICAL APPROVAL
The authors confirm that the ethical policies of the journal, as noted on the journal’s author guidelines page, have been adhered to.

DATA AVAILABILITY STATEMENT
All data needed to evaluate the conclusions in the paper are present in the paper and/or the Appendix S1.

ORCID
Qiuew Pan https://orcid.org/0000-0001-9982-6184

REFERENCES

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.