science.sciencemag.org/cgi/content/full/science.abb6105/DC1

Supplementary Material for

An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China

Huaiyu Tian*, Yonghong Liu, Yidan Li, Chieh-Hsi Wu, Bin Chen, Moritz U. G. Kraemer, Bingying Li, Jun Cai, Bo Xu, Qiqi Yang, Ben Wang, Peng Yang, Yujun Cui, Yimeng Song, Pai Zheng, Quanyi Wang, Ottar N. Bjornstad, Ruifu Yang*, Bryan T. Grenfell*, Oliver G. Pybus*, Christopher Dye*

*Corresponding author. Email: tianhuaiyu@gmail.com (H.T.); christopher.dye@zoo.ox.ac.uk (C.D.); oliver.pybus@zoo.ox.ac.uk (O.G.P.); grenfell@princeton.edu (B.G.); ruifuyang@gmail.com (R.F.Y.);

Published 31 March 2020 as *Science* First Release
DOI: 10.1126/science.abb6105

This PDF file includes:

Materials and Methods

Epidemiological, demographic and geographical data

We collected data from the official reports of the health commission of 34 provincial-level administrative units and 342 city-level units. We recorded the date of the first reported case in all newly-infected cities, including daily reports from 31 December 2019 to 19 February 2020, the first 50 days of the epidemic. Only laboratory-confirmed cases of COVID-19 were used. Population sizes for each city in 2018 were obtained from the China City Statistical Yearbook (http://olap.epsnet.com.cn/). Using ArcGIS we calculated the great circle distance between Wuhan and each city reporting COVID-19 cases. The location of each city is geocoded by the latitude and longitude coordinates of the city centre. To make a comparison with the 2009 H1N1 Pandemic (2009-H1N1pdm), daily case data were collected from China Information System for Disease Control and Prevention (CISDCP) from 10 May 2009 to 30 April 2010, a total of more than 180,000 cases (*5*).

Human mobility data

Human movements were tracked with mobile phone data, through location-based services (LBS) employed by popular Tencent applications such as WeChat and QQ. Movement outflows from Wuhan City to other cities (i.e. records of the number of people leaving each day) by air, train and road, were obtained from the migration flows database (https://heat.qq.com/) (*19*) from 13 January 2017 to 21 February 2017 (Spring Festival travel 2017), from 1 February 2018 to 12 March 2018 (Spring Festival travel 2018), and from 1 January 2018 to 31 December 2018 (entire 2018). To reconstruct the movement outflow from Wuhan during the 2020 Spring Festival (from 11 January to 25 January, before the Chinese Lunar New Year), mobile phone data (provided by the telecommunications operators) were used together with the Baidu migration index (http://qianxi.baidu.com/); using both data sources gave the most accurate measure of movement volume. The expected movement outflows from Wuhan after the New Year festival from 26 January to 19 February, had there been no travel ban, were generated by using travel volumes for 2017 and 2018 and the recorded travel destinations prior to the shutdown in 2020. We assumed that the proportion of daily outflows from Wuhan to each of the other destinations in China was constant through time.

Association between the Wuhan city travel ban and the arrival time of COVID-19 in other cities

In order to quantify the association between the Wuhan travel shutdown (23 January 2020) and COVID-19 spread, we used data collected between 31 December 2019 and 28 January 2020. The association between distance, human movement, interventions and timing of COVID-19 spread was assessed by a linear regression. Among five possible regression models examined (Table S3), the model judged best by the Akaike Information Criterion) was:

$$E[Y_i] = \alpha + \beta_1 \log_{10}(TotalFlow_i) + \beta_2 \log_{10}(Pop_i) + \beta_3 Long_i + \beta_4 Lat_i + \beta_5 Shutdown_i$$

Dependent variable Y_i is the arrival time (day) of the first confirmed case in city i, a measure of the spatial spread of COVID-19. The β_j*'s* are the regression coefficients. α is the intercept. $TotalFlow_i$ represents the passenger volume from Wuhan to city i by airplane, train and road during the whole of 2018. Pop_i is the population of city i. Lat_i and $Long_i$ represent the latitude and longitude of city i. The binary dummy variable $Shutdown_i$ is used to identify whether the arrival time of COVID-19 in newly-infected city i is associated with the Wuhan travel ban. For each city, *Shutdown* was set to 0 for arrival before 23 January 2020 and 1 for arrival on or after 23 January 2020. The regression analysis was performed using the R version 3.4.0. All of the candidate models examined (Table S3) produced similar estimates for the estimated delay in the arrival time due to the shutdown.

Diagnostic analyses were performed to check whether model assumptions have been violated. "Residuals" in this section refers to the residuals of linear regression model for travel ban. Spatial coordinates (*Lat* and *Long*) and passenger volume from Wuhan (*TotalFlow*) were included to adjust for possible spatial and temporal autocorrelation, respectively. However, we performed additional checks to investigate whether there was spatial autocorrelation among the residuals, which could contribute to data dependency.

If there were spatial autocorrelation in the residuals, then the pairwise residual differences of cities closer together would tend to be smaller than those of cities far apart. In other words, we would expect a high correlation between the pairwise differences in residuals and pairwise distances of cities. Such a correlation was not evident (r=0.03).

Differential accessibility from Wuhan to other cities may contribute to data-dependence. Therefore, we examined the timing of peak inflow from Wuhan City and the arrival time of COVID-19 in each city. The result shows no

evident correlation between them (r=-0.07, P=0.25). In addition, we calculated the correlation between the residuals and peak time of Wuhan inflow during 11 to 23 January (15 days before Chinese New Year to the Wuhan shutdown). Again, we found no evident correlation (r=0.08).

We did not detect issues regarding heteroscedasticity (studentized Breuch-Pagan test P=0.20) nor normality (Shapiro-Wilk test P=0.06). The residuals generally lie close to the straight line in the Q-Q plot (Fig. S6). Therefore, it is not evident that the distribution of errors departs from a normal distribution.

The associations between transmission control measures and the number of cases reported during the first week of an outbreak in a new location

The Level 1 national emergency response required suspected and confirmed cases of COVID-19 to be isolated and reported immediately in all cities. Using data for 296 cities across China, we investigated the associations between the epidemic intensity and three transmission control measures: closure of entertainment venues and banning public gatherings (*B*); suspension of intra-city public transport (*S*); and prohibition of travel by any means to and from other cities (*P*). The timing of implementation was recorded for each control measure in each city, including the delay in implementation since 31 December 2019 (day 0 of the epidemic). Each city was regarded as implementing an intervention when the official policy was announced publicly (Table S1). Other transmission control measures included delineating control areas, closure of schools, isolation of suspected and confirmed cases, and the disclosure of information. We could not investigate the associations between these interventions and development of the epidemic because they were reportedly applied in all cities uniformly and without delay.

We perform an association analysis using Poisson regression to investigate how the interventions *B, S* and *P* were associated with the dependent (Poisson) variable Y_i—the total number of confirmed cases that were reported during the first seven days of the outbreak in city *i*. The analysis was performed by using the GLM function in the MASS package (*20*) in R (version 3.6.2) to construct the model:

$$\log(\mathrm{E}[Y_i]) = \alpha + \beta_1 M_{i,\mathrm{S}} + \beta_2 M_{i,\mathrm{P}} + \beta_3 M_{i,\mathrm{B}} + \beta_4 T_{i,\mathrm{S}} + \beta_5 T_{i,\mathrm{P}} + \beta_6 T_{i,\mathrm{B}} + \beta_7 A_i + \beta_8 D_i + \log(Q_i) + \log(F_i),$$

where population size (in millions) (Q_i) and inflow (in millions) from Wuhan (F_i) of city i are offset variables, while the distance to Wuhan (D_i) and the arrival time (A_i) of the infection are adjustments to control for confounding with other independent variables. The β_j's are regression coefficients. $M_{i,k}$ is a binary variable indicating whether or not control measure k is implemented in city i by the seventh day of the outbreak in that city. $M_{i,k} = 1$, if city i has implemented k before or during its first seven days of outbreak. $M_{i,k} = 0$, if city i has implemented k after its first seven days of outbreak or never implemented k. $T_{i,k}$ represents the timing of implementing control measure k in city i, where 31 December 2019 is day 0. For example, if city i implements control measure k on 27 January 2020, then $T_{i,k} = 27$. If $M_{i,k} = 0$, then $T_{i,k} = 0$. This set up enables the β coefficient of $T_{i,k}$ to only measure the association between Y and the timing variable of k when $M_{i,k} = 1$. D_i is the distance from city i to Wuhan City. A_i is the arrival time of the epidemic in city i (the date of the first confirmed case), where 31 December 2019 is day 0, and for example, if the date of the first confirmed case in city i is on 28 January 2020, then $A_i = 28$.

We detected heteroscedasticity in the Pearson residuals of the Poisson regression model. In attempt to remedy this we applied negative-binomial regression and quasi-poisson regression. Neither methods resolved the issue with heteroscedasticity. Therefore, we resorted to regression models that does not have dependent variables as counts.

To check and confirm the validity of results obtained with the Poisson regression model, we repeated the analysis with a log-linear model. The first step was to standardize case counts by dividing by the number of people in each city (incidence per capita) and the number of people arriving from Wuhan, giving dependent variable Y^*. The log-linear model is then:

$$\mathrm{E}[\log(Y_i^*)] = \alpha + \beta_1 M_{i,\mathrm{S}} + \beta_3 M_{i,\mathrm{B}} + \beta_4 T_{i,\mathrm{S}} + \beta_6 C_{i,\mathrm{B}} + \beta_7 A_i.$$

The subscripts of the coefficients (β_j) are consistent with a Poisson regression model. We found that the relationship between Y^* and the timing of B is nonlinear. To correct for this, we discretize the timing variable $T_{i,\mathrm{B}}$ to obtain an ordinal variable $C_{i,\mathrm{B}}$ with four categories, namely 1) 30 December 2019–22 January 2020, 2) 23 January 2020, 3) 24 January 2020 and 4) 25 January onwards. To preserve statistical power, proportional change of $\mathrm{E}[Y^*]$ was assumed when moving onto a later category. For example, the difference in $\mathrm{E}[\log(Y_i^*)]$ time categories 1 and 2 was assumed

the same as that between time categories 2 and 3.

To avoid heteroscedasticity, variables describing the distance from Wuhan, and the implementation and timing of *P* (prohibiting inter-city travel) were removed. Further exploration of the model showed that these variables did not help to explain the variation in Y^*. Table S3 presents the results of the log-linear regression analysis, which uphold the conclusions reached from the Poisson regression model.

As the purpose of the log-linear regression analysis is to validate the results from the Poisson regression analysis, we performed various diagnostic checks on the log-linear regression model. In summary, we did not find any issues that might invalidate the results. The details of the diagnostic analysis are presented below. "Residuals" in this section refers of the residuals of the log-linear regression model.

As with the regression analysis on the Wuhan city travel ban, we explored whether there is any spatial autocorrelation in the residuals. There was no evident correlation between the pairwise differences in residuals and those of cities (r=-0.09). We also examined whether the timing of Wuhan inflow induced dependence in the data. To this end, we evaluated the correlation between the residuals and peak time of inflow from Wuhan to each city during the period from 11 to 23 January. Such correlation is not evident (r=-0.04),

In terms of temporal correlation within the residuals, there was no evidence that the mean residuals with arrival time on day *j* associated with the residuals with arrival time on day *j+1* (*P*=0.80).

We found no evidence of heteroscedasticity in the residuals (studentised Breusch-Pagan test *P*=0.148). The Shapiro-Wilk test provides evidence against residual normality (*P*<0.01). However, in a sufficiently large dataset such as this one, the Shapiro-Wilk test is very sensitive to departure from normality but does not indicate the severity level of the departure. The Q-Q plot (Fig. S7) indicates a moderate departure from the normal distribution in the residuals, where the histogram indicates that there is a longer right tail (Fig. S7). As the distribution is not severely asymmetric, and our sample size is sufficiently large, this should not be an issue by the central limit theorem (*21, 22*).

Nevertheless, to ensure the robustness of the conclusions, we also evaluated bootstrap estimates of the regression coefficients, confidence interval and *P*-value (Table S3) using the R package boot (*23, 24*). The bootstrap estimates are very similar to the least squared estimates from log-linear regression model (data not shown).

Epidemic modelling

For each province, we estimated the effect of transmission control measures by fitting an SEIR model (*25*) to the number of new confirmed cases reported each day from each province using Bayesian Markov Chain Monte Carlo methods (*26*). The model is:

$$\frac{dS_i}{dt} = -I_i \frac{\beta C_{Segment} S_i}{N_i}$$

$$\frac{dE_i}{dt} = I_i \frac{\beta C_{segment} S_i}{N_i} - \delta E_i + \lambda_i$$

$$\frac{dI_i}{dt} = \delta E_i - \gamma I_i$$

$$\frac{dR_i}{dt} = \gamma I_i$$

where *S*, *E*, *I*, and *R* are the number of susceptible, exposed (latent), infectious, and removed individuals on day *t* in province *i*. This standard SEIR model makes some simplifying assumptions: for example, the human population is homogeneous (e.g. not stratified by age or sex), contacts between infectious and susceptible people are also homogeneous (e.g. not stratified by social group) and infection is fully immunizing (*1*). The model and its assumptions apply only to the first 50 days of the epidemic, the scope of the analysis in this study.

Even with these assumptions, the model describes the data quite accurately for the whole country (Fig. 4A), and reasonably well for each province, with the notable exceptions of Beijing and Shanghai, which had relatively few cases among their well-connected populations (Fig. 4A, Fig. S3). The accuracy of the model's description of

national data is remarkably good, given that the model fitting was done separately for each province, and then aggregated to obtain the national total.

The basic reproductive number of the model is $R_0 = \beta/\gamma$, where β is the per capita transmission rate per day and $1/\delta$ and $1/\gamma$ are, respectively, the mean latent and infectious periods.

Variable λ is the estimated number of cases imported from Wuhan City on day t:

$$\lambda_i(t) = \frac{I_w}{\rho_w P_w} T_i shutdown_w$$

I_w is the number of reported cases in Wuhan on day t, P_w is the Wuhan population size, and ρ_w is the proportion of all infected people (including infectious cases) reported in Wuhan. T_i is the number of people leaving Wuhan on day t travelling to province i, derived from data describing mobility 15 days before the Chinese Lunar New Year 2020. The binary variable *shutdown* is used to identify whether cases were or were not exported from Wuhan on or after 23 January 2020.

The role of control measures at different stages of the outbreak are captured by estimated parameter C (range 0-1), which reduces transmission and R_0 proportionally as a multiplicand of β. The timing and implementation of transmission control measures in the 342 cities and 31 provinces are shown in Fig. S4. Before 22 January 2020, there were no recorded interventions thus C_0=1. From 23 January onwards, provinces gradually scaled up Level 1 emergency responses (stage 1), with effects measured as C_1 (Fig. S4). Because the effects of control measures varied among provinces during the scale-up, C_1 was grouped into high C_{1_high}, medium C_{1_medium}, and low C_{1_low}. The allocation of provinces to groups was made by proposing several alternative hypotheses and testing each by model fitting (Table 3, Table S4). Stage 2 of control (C_2) began when more than 95% of cities in a province had implemented control measures, including the closure of entertainment venues, suspension of intra-city public transport or prohibition of travel by any means to and from other cities (see above). In Hubei Province (except Wuhan city), stage 2 included the use of shelter or "Fang Cang" hospitals from early February onwards.

Model fitting was performed using the Metropolis–Hastings Markov chain Monte Carlo (MCMC) algorithm with the MATLAB (version R2016b) toolbox DRAM (Delayed Rejection Adaptive Metropolis). Prior estimates of the mean and (Gaussian) variance of R_0, δ, and γ were derived from epidemiological surveys (*27*). There was no evidence to inform a prior for the reporting rate ρ, the proportion of cases that were reported among all latent and infectious individuals in Wuhan. Systematic surveys of infection (e.g. by serological testing) have not yet been reported. In the absence of any guiding data, ρ was given a prior uniform distribution between 0 and 1.

After a burn-in of 1 million iterations, we ran the MCMC simulation for a further 10 million iterations, sampled at every 1000th step to avoid auto-correlation. Trace plots and Gelman and Rubin diagnostics were used to judge convergence of the MCMC chains (Fig. S4). Each fitting exercise was repeated three times to test the robustness of results, which converged to the same estimates on each occasion (Fig. S5). We used the fitted SEIR model, with posterior estimates of parameter values, to simulate outbreaks outside Wuhan, with and without the Wuhan travel ban and with and without the national emergency response (Fig. 4B).

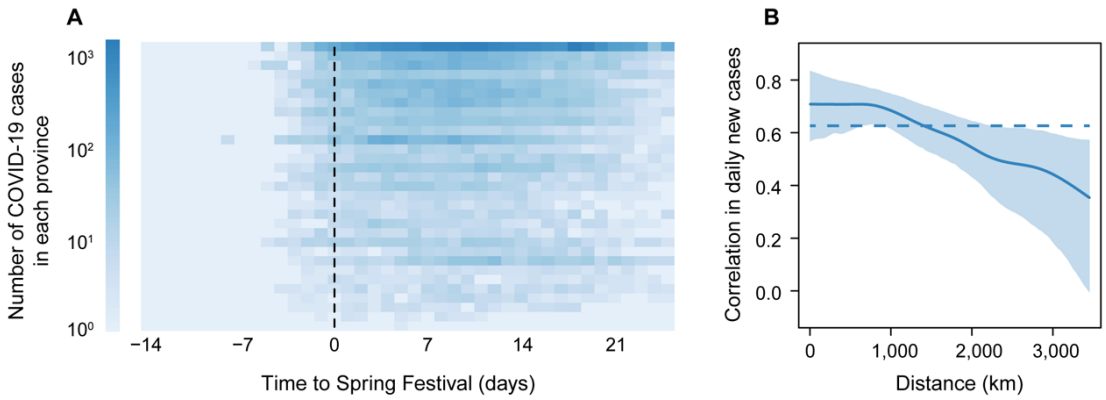

Fig. S1. Patterns of COVID-19 dispersal out of Wuhan (Hubei province) to other provinces by time and geographical distance. (A) Daily reports of confirmed cases from each of 31 provinces. Provinces are ranked by decreasing volume of people leaving Wuhan for other destinations, to elsewhere in Hubei Province (top) and to Tibet (bottom). (B) Synchrony of epidemics in different provinces in relation to distance between provinces. Synchrony is measured by the correlation between the number of cases reported in two provinces on each day, using a spatial non-parametric correlation function (*28*).

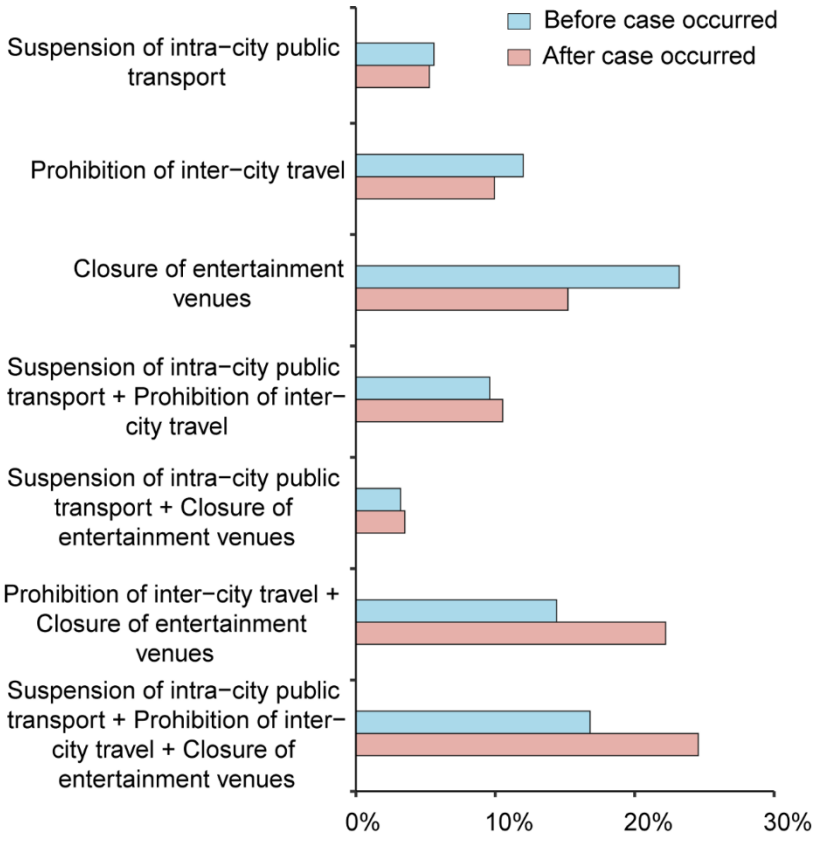

Fig. S2. Percentage of cities that implemented three kinds of transmission control measures before (blue), or on the same day or after (red), the first case was reported.

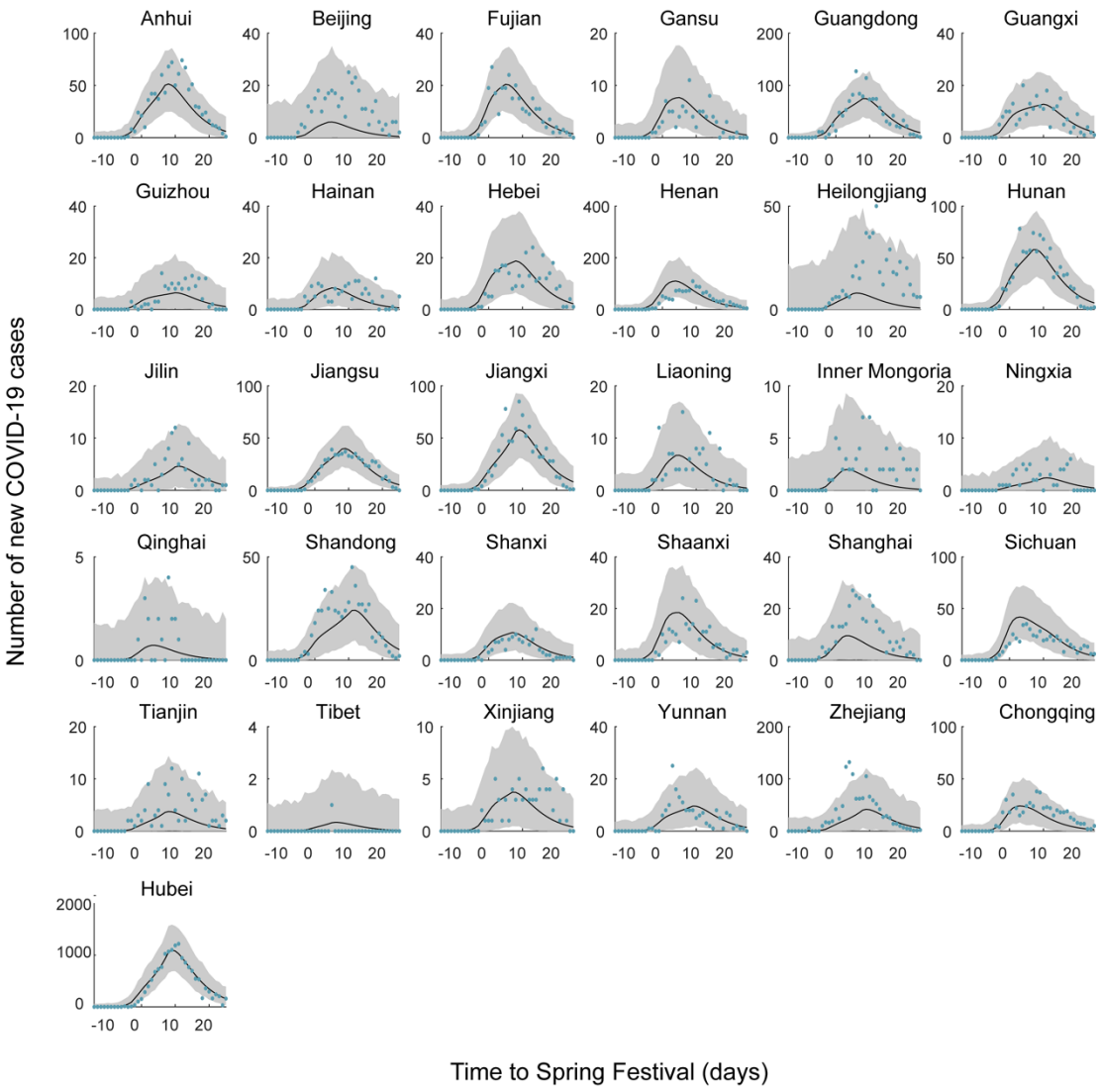

Fig. S3. Fits of the SEIR epidemic model to time series of reported cases from 31 provinces. The numbers of confirmed cases reported (points) and estimated (lines) each day in each provinces (Hubei excludes Wuhan city). Grey areas correspond to pointwise 95% prediction envelopes. The period covers the 40 days of the Spring Festival, from 15 days before to 25 days after the Chinese Lunar New Year. The Spring Festival travel ended on 19 February, day 50 of the epidemic.

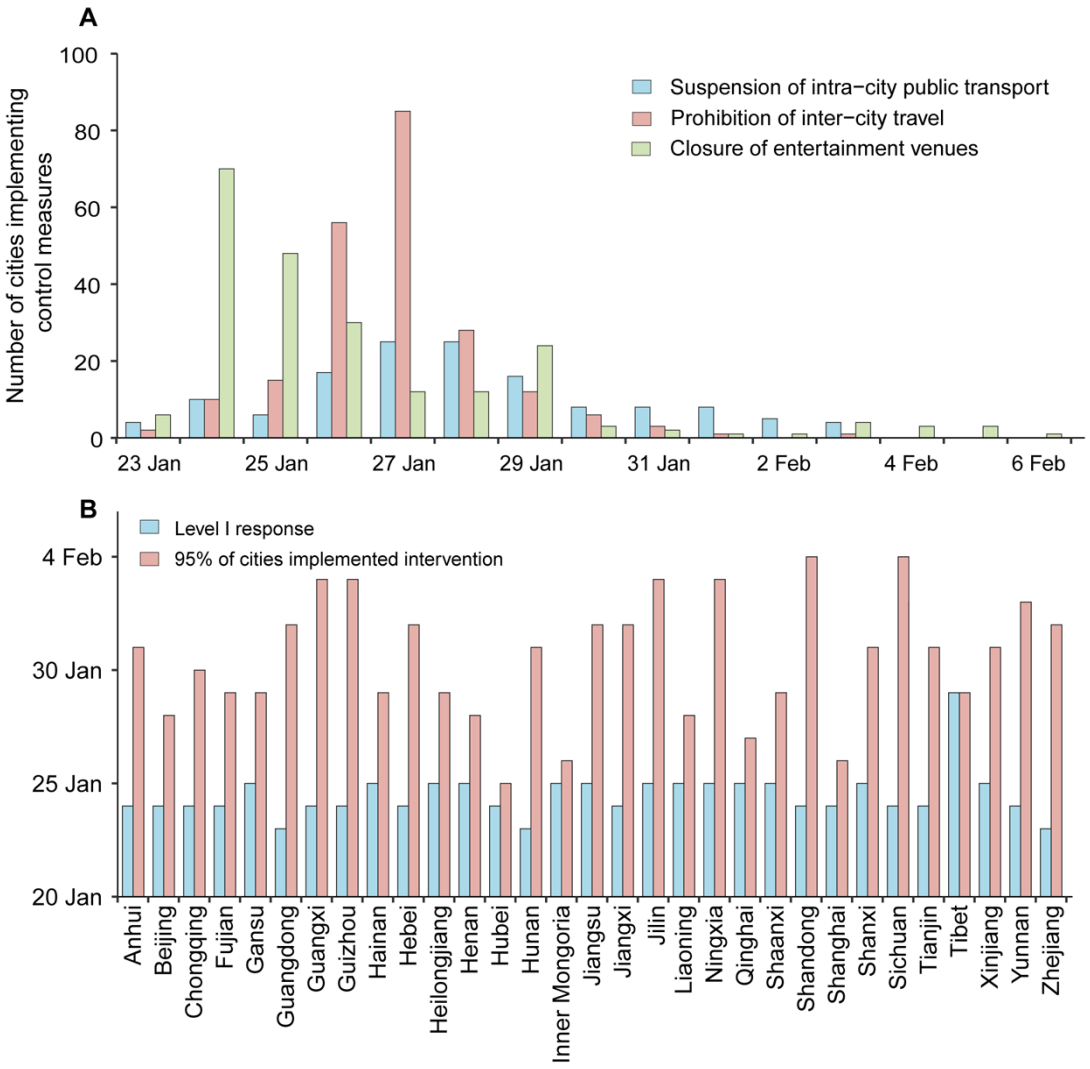

Fig. S4. (A) The number of cities implementing three interventions by date in 342 cities (see also Fig. S2). (B) Dates (vertical axis) on which the Level 1 emergency response began (blue, start of stage 1), and on which 95% of cities had implemented transmission control measures (red, end of stage 1, beginning of stage 2), in 31 provinces.

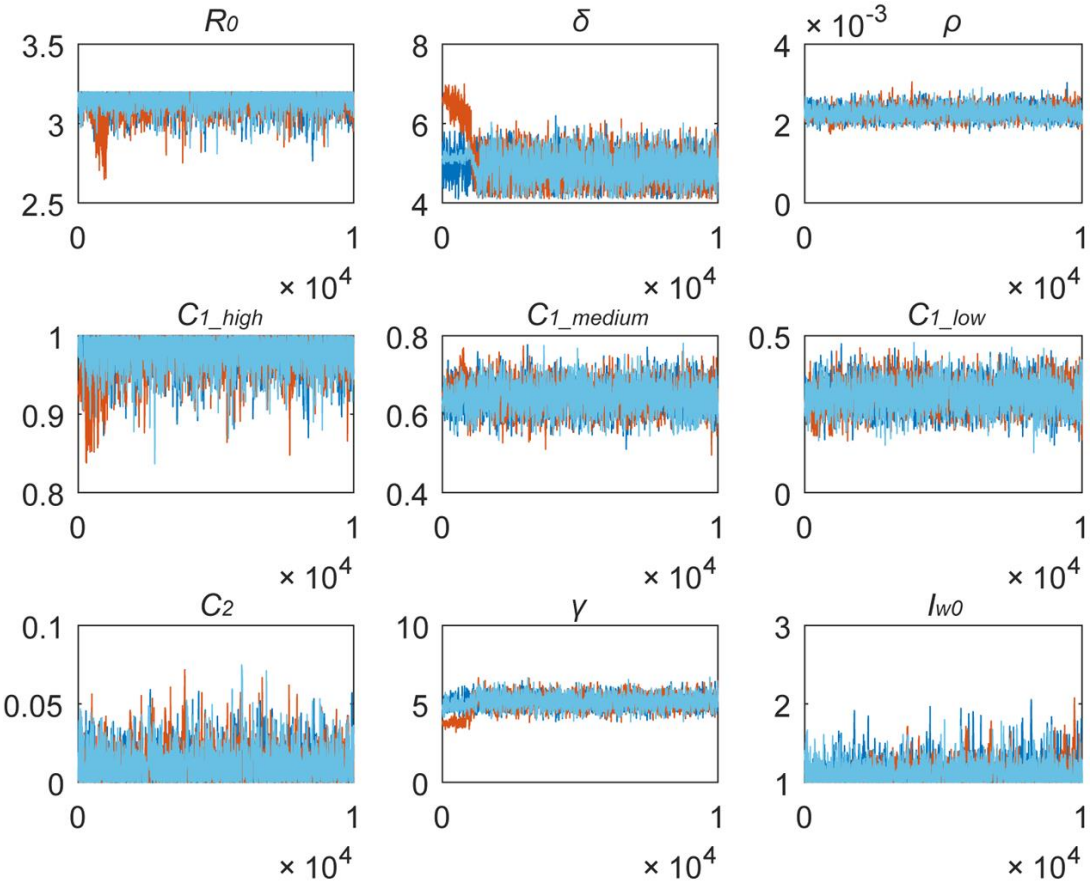

Fig. S5. Trace plots of parameter values for the epidemic model, estimated by Bayesian Markov Chain Monte Carlo (MCMC) methods. The three different colours represent three runs of the MCMC model, with one run (light blue) presented at the forefront.

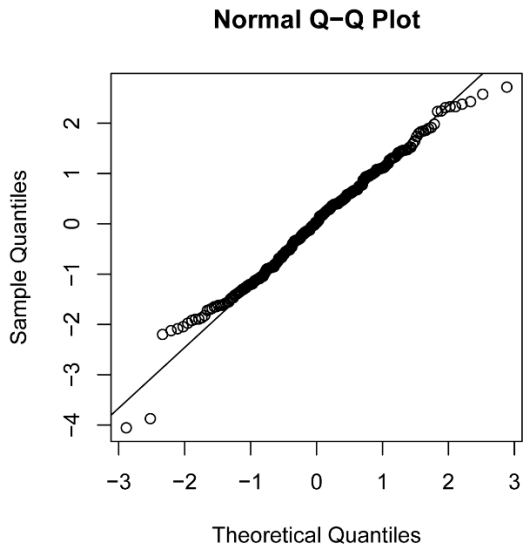

Fig. S6. Q-Q plot the residuals of the linear regression analysis on the association between the Wuhan travel ban and dispersion of COVID-19 across cities in China.

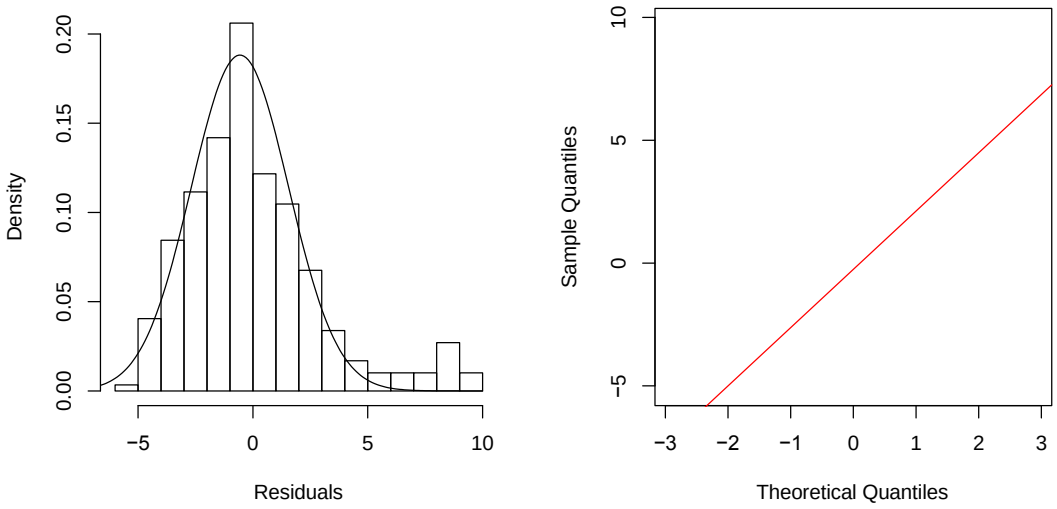

Fig. S7. Histogram and Q-Q plot the residuals of the log-linear regression analysis on the associations between the transmission control measures and incident per capita divided by inflow from Wuhan, across cities in China.

Table S1. Candidate statistical models used to study the association between the Wuhan city travel ban and the arrival time of COVID-19 in other cities (see Table 1 of the main text). The chosen model is shown in bold face.

Model	AIC*
Y=log10(TotalFlow) + log10(AirFlow) + log10(RoadFlow) + log10(TrainFlow) + log10(Pop) + log10(Dis) + Lat + Long + Shutdown	113.73
Y=log10(TotalFlow) + log10(AirFlow) + log10(RoadFlow) + log10(Pop) + log10(Dis) + Lat + Long + Shutdown	111.91
Y=log10(TotalFlow) + log10(AirFlow) + log10(Pop) + log10(Dis) + Lat + Long + Shutdown	110.63
Y=log10(TotalFlow) + log10(Pop) + log10(Dis) + Lat + Long + Shutdown	109.83
Y=log10(TotalFlow) + log10(Pop) + Lat + Long + Shutdown	108.57

* Akaike information criterion

Table S2. Summary of interventions and their timing across 342 cities (see Table 2 of the main text).

Level 1 response to major public health emergencies	Number of cities implementing control measures	Average lags (days) between implementation and 31 December 2019‡
Identify the affected area of a city*	342	0
Close schools*	342	0
Close entertainment venues and ban public gatherings	220	27.17 (2.82)
Isolate patients with infectious diseases*	342	0
Isolate suspected patients*	342	0
Suspend intra-city public transport (bus and subway)	136	29.00 (2.60)
Prohibit inter-city travel	219	27.86 (1.49)
Collect, evaluate, report and publish information on public health emergencies daily*	342	0
Assist subdistrict, township (town), neighbourhood and village committee staff*	342	25.32 (1.07)

*Interventions implemented immediately were not included in the regression analysis.

‡Summary statistics reported for timing are mean (standard deviation).

Table S3. Associations of the type and timing of transmission control measures with the number of cases reported during the first week of outbreak in a new location (city), estimated from a log-linear regression model. This analysis checks and confirms the robustness of results in Table 2 of the main text. As described in the main text, the prohibition of inter-city travel, the third intervention that was investigated in this study, did not significantly reduce the number of cases reported during the first week of city outbreaks.

Least square estimates			
Covariates	**Coefficient**	**95% CI**	**P**
(Intercept)	-1.18	(-4.96, 2.61)	0.54
Arrival time	0.27	(0.12, 0.42)	<0.01
Suspension of intra-city public transport			
Implementation	-12.71	(-21.77, -3.64)	<0.01
Timing	0.46	(0.13, 0.80)	<0.01
Closure of entertainment venues			
Implementation	-3.41	(-5.73, -1.10)	<0.01
Timing (discretised)	1.51	(0.66, 2.36)	<0.01
Bootstrap estimates			
Covariates	**Coefficient**	**95% CI**	**P**
(Intercept)	-1.18	(-4.45, 2.14)	0.49
Arrival time	0.27	(0.14, 0.40)	<0.01
Suspension of intra-city public transport			
Implementation	-12.66	(-21.37, -4.74)	<0.01
Timing	0.46	(0.17, 0.80)	<0.01
Closure of entertainment venues			
Implementation	-3.42	(-5.82, -0.86)	<0.01

Timing (discretised)	1.51	(0.57, 2.44)	<0.01

Table S4. Candidate models used to characterize the association between control measures and the new confirmed cases reported each day in different provinces (see Table 3 of the main text). Based on the deviance information criterion (DIC), the chosen model is shown in bold face.

Model	**DIC* (mean)** ‡	**DIC (median)**
Universal *C*, *γ*	117.55	38.97
Universal *C1*, *C2*, *γ*	78.81	28.41
C1_high*, *C1_medium*, *C1_low*, *C2*, *γ	55.66	27.43
C1_high, *C1_low*, *C2*, *γ*	62.31	28.13
C1_high, *C1_medium*, *C1_low*, *γ*	110.95	34.76

* Distribution for provinces

‡ The deviance information criterion

C_{1_high}, Effect of control in stage 1, high

C_{1_medium}, Effect of control in stage 1, medium

C_{1_low}, Effect of control in stage 1, low

C_2, Effect of control in stage 2

γ, Rate of removal of infectious cases before isolation

High, medium and low represent the efficacy of control measures in three groups of provinces.

References and Notes

1. N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, W. Tan; China Novel Coronavirus Investigating and Research Team, A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020). doi:10.1056/NEJMoa2001017 Medline

2. R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Y. Bi, X. Ma, F. Zhan, L. Wang, T. Hu, H. Zhou, Z. Hu, W. Zhou, L. Zhao, J. Chen, Y. Meng, J. Wang, Y. Lin, J. Yuan, Z. Xie, J. Ma, W. J. Liu, D. Wang, W. Xu, E. C. Holmes, G. F. Gao, G. Wu, W. Chen, W. Shi, W. Tan, Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020). doi:10.1016/S0140-6736(20)30251-8 Medline

3. F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E. C. Holmes, Y.-Z. Zhang, A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020). doi:10.1038/s41586-020-2008-3 Medline

4. P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020). doi:10.1038/s41586-020-2012-7 Medline

5. J. Cai, B. Xu, K. K. Y. Chan, X. Zhang, B. Zhang, Z. Chen, B. Xu, Roles of Different Transport Modes in the Spatial Spread of the 2009 Influenza A(H1N1) Pandemic in Mainland China. *Int. J. Environ. Res. Public Health* **16**, 222 (2019). doi:10.3390/ijerph16020222 Medline

6. C. Wang, P. W. Horby, F. G. Hayden, G. F. Gao, A novel coronavirus outbreak of global health concern. *Lancet* **395**, 470–473 (2020). doi:10.1016/S0140-6736(20)30185-9 Medline

7. S. Chen, J. Yang, W. Yang, C. Wang, T. Bärnighausen, COVID-19 control in China during mass population movements at New Year. *Lancet* **395**, 764–766 (2020). doi:10.1016/S0140-6736(20)30421-9 Medline

8. M. U. G. Kraemer, C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, L. du Plessis, N. R. Faria, R. Li, W. P. Hanage, J. S. Brownstein, M. Layan, A. Vespignani, H. Tian, C. Dye, O. G. Pybus, S. V. Scarpino, The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* eabb4218 (2020). doi:10.1126/science.abb4218

9. Chinadaily, “Tibet activates highest-level public health alert” (Jan 30, 2020). https://www.chinadaily.com.cn/a/202001/202029/WS202005e202318a202036a3101282172739c3101282172731.html.

10. B. T. Grenfell, O. N. Bjørnstad, J. Kappey, Travelling waves and spatial hierarchies in measles epidemics. *Nature* **414**, 716–723 (2001). doi:10.1038/414716a Medline

11. D. Brockmann, D. Helbing, The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–1342 (2013). doi:10.1126/science.1245200 Medline

12. A. Wesolowski, N. Eagle, A. J. Tatem, D. L. Smith, A. M. Noor, R. W. Snow, C. O. Buckee, Quantifying the impact of human mobility on malaria. *Science* **338**, 267–270 (2012). doi:10.1126/science.1223467 Medline

13. N. M. Ferguson, D. A. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, D. S. Burke, Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* **437**, 209–214 (2005). doi:10.1038/nature04017 Medline

14. K. E. Jones, N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman, P. Daszak, Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008). doi:10.1038/nature06536 Medline

15. D. M. Morens, G. K. Folkers, A. S. Fauci, The challenge of emerging and re-emerging infectious diseases. *Nature* **430**, 242–249 (2004). doi:10.1038/nature02759 Medline

16. C. Viboud, O. N. Bjørnstad, D. L. Smith, L. Simonsen, M. A. Miller, B. T. Grenfell, Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–451 (2006). doi:10.1126/science.1125237 Medline

17. A. Wesolowski, T. Qureshi, M. F. Boni, P. R. Sundsøy, M. A. Johansson, S. B. Rasheed, K. Engø-Monsen, C. O. Buckee, Impact of human mobility on the emergence of dengue epidemics in Pakistan. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11887–11892 (2015). doi:10.1073/pnas.1504964112 Medline

18. H. Tian, Y. Liu, Y. Li, C.-H. Wu, B. Chen, M. U. Kraemer, B. Li, J. Cai, B. Xu, Q. Yang, B. Wang, P. Yang, Y. Cui, Y. Song, P. Zheng, Q. Wang, O. N. Bjornstad, R. Yang, B. T. Grenfell, O. G. Pybus, C. Dye, Code for: An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Zenodo* (2020); doi:10.5281/zenodo.3727336

19. J. T. Wu, K. Leung, G. M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *Lancet* **395**, 689–697 (2020). doi:10.1016/S0140-6736(20)30260-9 Medline

20. B. Ripley, B. Venables, D. M. Bates, K. Hornik, A. Gebhardt, D. Firth, M. B. Ripley, Package 'MASS'. *CRAN Repository*, See http://cran.r-projectorg/web/packages/MASS/MASS.pdf (2013).

21. B. R. Kirkwood, J. A. Sterne, *Essential medical statistics*. (John Wiley & Sons, 2010).

22. D. E. Bailey, *Probability and Statistics*. (John Wiley & Sons, 1971).

23. A. Canty, B. Ripley, boot: Bootstrap R (S-Plus) functions. R package version 1.3-24. (2019).

24. A. C. Davison, D. V. Hinkley, *Bootstrap methods and their application*. (Cambridge university press, Cambridge, 1997).

25. R. M. Anderson, R. M. May, *Infectious Diseases of Humans: Dynamics and Control*. (Oxford Univ Press, Oxford, 1992).

26. A. Morton, B. F. Finkenstädt, Discrete time modelling of disease incidence time series by using Markov chain Monte Carlo methods. *J. R. Stat. Soc.* **54**, 575–594 (2005). doi:10.1111/j.1467-9876.2005.05366.x

27. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020). doi:10.1056/NEJMoa2001316 Medline

28. O. N. Bjørnstad, R. A. Ims, X. Lambin, Spatial population dynamics: Analyzing patterns and processes of population synchrony. *Trends Ecol. Evol.* **14**, 427–432 (1999). doi:10.1016/S0169-5347(99)01677-8 Medline