Supplementary Information

The next generation matrix for the SEIR model described in the main text was

$$\frac{\beta D_I}{N}\begin{bmatrix}\alpha_1 N_1 & \alpha_1 N_1 & \alpha_1 N_1\\ \alpha_2 N_2 & \alpha_2 N_2 & \alpha_2 N_2\\ \alpha_3 N_3 & \alpha_3 N_3 & \alpha_3 N_3\end{bmatrix}$$

and the basic reproductive number R_0 was the largest eigenvalue of this matrix, which is $\frac{\beta D_I}{N}[\alpha_1 N_1 + \alpha_2 N_2 + \alpha_3 N_3]$.

The number of new cases (onset) and the cumulative number of cases in age group *i* on day *d* was $\omega_{d,i} = \int_{d-1}^{d} A_{i,onset}(t)dt$ and $\Omega_{d,i} = \int_0^d A_{i,onset}(t)dt$, respectively. The cumulative number of death cases in age group *i* up to time *t* was $D_i(t) = \int_0^t A_{i,death}(u)du$. Let $\omega_d = \sum_{i=1}^{3}\omega_{d,i}$, $\Omega_d = \sum_{i=1}^{3}\Omega_{d,i}$ and $D(t) = \sum_{i=1}^{3} D_i(t)$ be the summation of the number of new cases, cumulative number of cases and cumulative number of deaths across the three age groups. Similarly, let $E(t) = \sum_{i=1}^{3} E_i(t)$ and $I(t) = \sum_{i=1}^{3} I_i(t)$ be the total latent and infectious individuals at time *t*.

We used θ to denote the set of parameters that were subject to inference (Table 2). The likelihood function was a product of several components associated with the data in Tables S1-S8:

$$L(\theta) = \prod_{k=1}^{8} L_k(\theta)$$

The formulation of each component was as follows.

1. The number of observed international case exportations on each day was assumed to be an imperfect Poisson observation of the number of infected travelers leaving Wuhan on that day who had or would develop symptoms. Let x_d be the observed number of such international case exportations on day *d* between 25 December 2019 ($D_{s,1}$) and 19 January 2020 ($D_{e,1}$) based on the data in Table S2. We assumed that travel behaviour was not affected by disease and hence such

case exportation occurred according to a non-homogeneous process with rate $\lambda(t) = P_{sym}\frac{L_{W,I}(t)}{N(t)}\big(E(t)+I(t)\big)$. Let the probability an infected traveler is detected in the destination country be P_{det}. The expected number of case exportation on the day d was $\lambda_d = P_{det}\int_{d-1}^{d}\lambda(u)du$ and hence $x_d \sim Poisson(\lambda_d)$. As such, the likelihood function associated with the data in Table S2 was

$$L_1(\theta) = \prod_{d=D_{s,1}}^{D_{e,1}} \frac{e^{-\lambda_d}{\lambda_d}^{x_d}}{x_d!}$$

2. Let y_d be the observed number of confirmed cases of COVID-19 in Wuhan with no epidemiologic links to Huanan Seafood Wholesale Market (which was presumed to be the index zoonotic source of the COVID-19 epidemic) on day *d* between 1 December 2019 ($D_{s,2}$) and 3 January 2020 ($D_{e,2}$) based on the data in Table S1 [2]. These cases were assumed to be a Poisson observation of the true number of newly symptomatic cases on that day, with a sensitivity ε that remained fixed over this time period. As such, assuming $y_d \sim Poisson(\varepsilon\omega_d)$ and the likelihood function for the data in Table S1 was

$$L_2(\theta) = \prod_{d=D_{s,2}}^{D_{e,2}} \frac{e^{-\varepsilon\omega_d}(\varepsilon\omega_d)^{y_d}}{y_d!}$$

3. We considered the test results of entry screening among expatriates upon returning to their countries from Wuhan on charter flights between 29 January 2020 ($D_{s,3}$) and 4 February 2020 $(D_{e,3})$. Let m_d^{all} be the number of such passengers on day *d* who were tested regardless of symptoms (e.g. Japan, Germany, South Korea, etc; Table S3) and m_d^{sym} be the number of such passengers on day *d* who were probably tested only if they showed symptoms (e.g. US, UK, Thailand, Australia, etc; Table S3). Let u_d^{all} and u_d^{sym} be the respective observed number of returning expatriates who were confirmed to be infected based on the data in Table S3. The

prevalence of infection and symptoms among travelers were assumed to reflect a representative binomial sample of the same quantities in the Wuhan population on their day of departure. The likelihood function associated with the data in Table S3 was

$$L_3(\theta) = \prod_{d=D_{s,3}}^{D_{e,3}} \binom{m_d^{all}}{u_d^{all}} q_d{}^{u_d^{all}} (1 - q_d)^{m_d^{all} - u_d^{all}} \binom{m_d^{sym}}{u_d^{sym}} P_S q_d{}^{u_d^{sym}} (1 - P_S q_d)^{m_d^{sym} - u_d^{sym}}$$

where $q_d = (E(d) + I(d))/N(d)$ was the proportion of individuals who were infected on day d and P_S was the proportion of infections that were symptomatic.

4. We assumed that all deaths from infection in Wuhan were confirmed. Let G be the cumulative number of death cases officially reported by Wuhan at the end of our study time horizon (time T). The likelihood function associated with this data was

$$L_4(\theta) = \frac{e^{-D(T)} D(T)^G}{G!}$$

5. We assumed that the age distribution of reported cases was a multinomial sampling process from the age distribution of true cases. Let c_i be the observed number of cases in age group i in Wuhan based on the data in Table S4. The likelihood function for the data in Table S4 was

$$L_5(\theta) = \frac{(c_1 + c_2 + c_3)!}{c_1!\, c_2!\, c_3!} \prod_{i=1}^{3} \left(\frac{\Omega_{T,i}}{\Omega_T}\right)^{c_i}$$

6. We assumed that the age distribution of reported deaths was a multinomial sampling process from the age distribution of true deaths. Let b_i be the observed number of death cases in age group i in Wuhan based on the data in Table S5. The likelihood function for the data in Table S5 was

$$L_6(\theta) = \frac{(b_1 + b_2 + b_3)!}{b_1!\, b_2!\, b_3!} \prod_{i=1}^{3} \left(\frac{D_i(T)}{D(T)}\right)^{b_j}$$

7. With regard to the data in Table S6, let A be the set of death cases whose onset dates were known, and B be the set comprising the remaining cases. Let v_j be the observed time delay between onset

and death for the *j*th case in *A*, and v_j^L be the observed time between hospital admission and death (which served as a lower-bound for the delay between onset and death) for the *j*th case in *B*. The likelihood function for the data in Table S6 was

$$L_7(\theta) = \prod_{j\in A} f_{onset-death}(v_j|\theta) \prod_{j\in B} \left(1 - F_{onset-death}(v_j^L|\theta)\right)$$

where $f_{onset-death}$ and $F_{onset-death}$ were the probability density function (pdf) and cumulative density function (cdf) of the time between onset and death (assumed to be gamma distributed with mean μ_D and standard deviation σ_D).

8. With regard to the data in Table S8, let *A* be the set of infector-infectee pairs for whom the serial interval (time elapsed between their onset dates) were known and *B* be the set comprising the remaining pairs for whom only the ranges of their serial intervals were known. Let s_j be the observed value of the serial interval for the *j*th pair in *A*, and (s_j^L, s_j^U) be the observed range of the serial interval for the *j*th pair in *B*. The likelihood function for the data in Table S8 was

$$L_8(\theta) = \prod_{j\in A} f_{SI}(s_j|\theta) \prod_{j\in B} \left(F_{SI}(s_j^U|\theta) - F_{SI}(s_j^L|\theta)\right)$$

where f_{SI} and F_{SI} were the pdf and cdf of the serial interval. Assuming that the serial interval and the generation time was quantitatively the same, the pdf of the serial interval was the convolution of the pdf of the latent period (f_E; exponentially distributed with mean D_E) and the pdf of the infectious period (f_I; exponentially distributed with mean D_I) in the SEIR model [23], i.e.

$$f_{SI}(t|\theta) = \int_0^t f_E(u|\theta)\, f_I(t-u|\theta)du$$

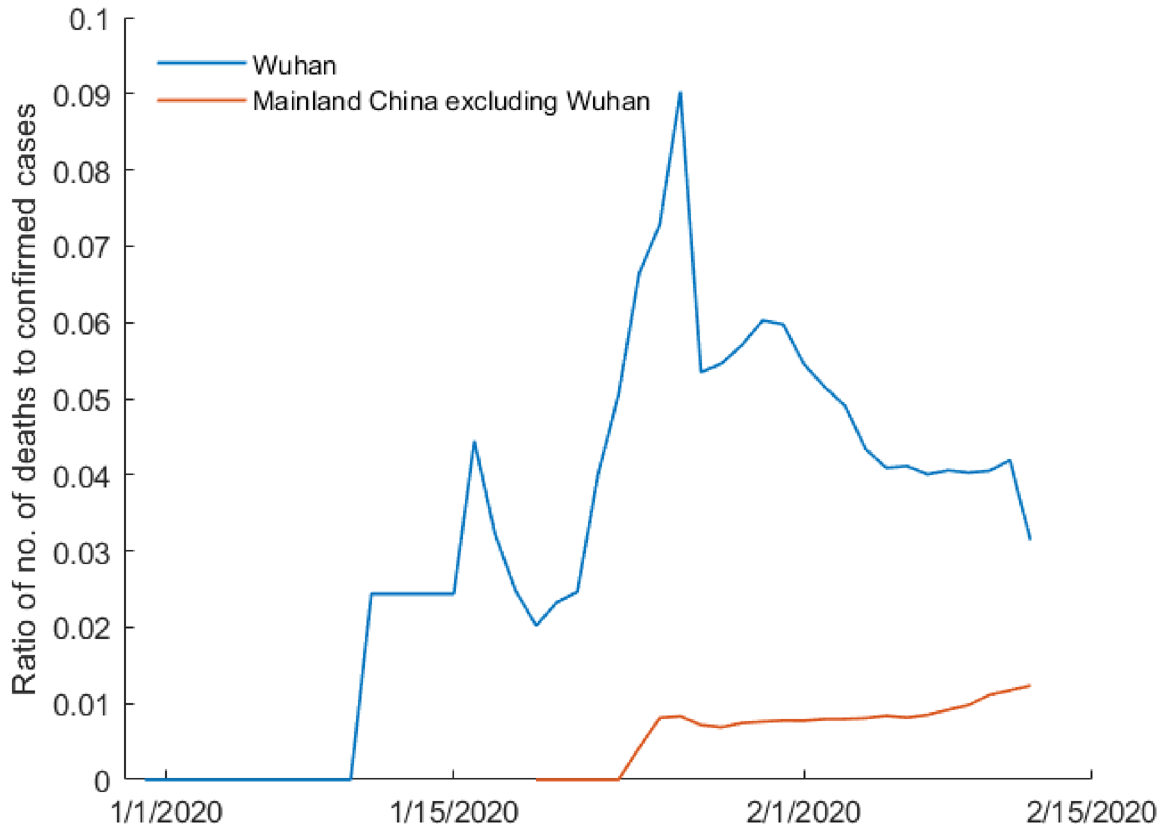

Figure S1. The ratio of no. of deaths to confirmed cases in Wuhan and in cities of mainland China other than Wuhan

Table S1. The number of confirmed cases of 2019-nCoV in Wuhan with no epidemiologic link to Huanan Seafood Wholesale Market between 1 Dec 2019 and 3 Jan 2020 [2]

Date	No. of confirmed cases with no epidemiologic link to Huanan Market
12/10/2019	1
12/11/2019	0
12/12/2019	0
12/13/2019	0
12/14/2019	1
12/15/2019	2
12/16/2019	2
12/17/2019	0
12/18/2019	0
12/19/2019	0
12/20/2019	1
12/21/2019	1
12/22/2019	2
12/23/2019	1
12/24/2019	0
12/25/2019	2
12/26/2019	0
12/27/2019	0
12/28/2019	1
12/29/2019	4
12/30/2019	3
12/31/2019	1
1/1/2020	6
1/2/2020	6
1/3/2020	10

Table S2. The details of cases exported from Wuhan to cities outside mainland China* via air travel on each day between 25 December 2019 and 19 January 2020 [1]

Country/region	City	Sex	Age	Date of arrival	Date of symptom onset	Date of confirmation
Thailand	Bangkok	F	61	1/8/2020	1/5/2020	1/12/2020
Thailand	Bangkok	F	74	1/13/2020	1/5/2020	1/17/2020
Thailand	Bangkok	M	68	1/19/2020	1/19/2020	1/19/2020
Thailand	Nakhon Pathom	F	73	1/1/2020	12/31/2019	1/22/2020
Vietnam	Hanoi	M	66	1/13/2020	1/17/2020	1/23/2020
Japan	Kanagawa	M	35	1/6/2020	1/3/2020	1/15/2020
Japan	Tokyo	M	40+	1/19/2020	1/14/2020	1/24/2020
Japan	Tokyo	F	30+	1/18/2020	1/21/2020	1/25/2020
Singapore	Singapore	M	56	1/19/2020	1/25/2020	1/28/2020
Singapore	Singapore	F	56	1/18/2020	-	1/27/2020
South Korea	Seoul	F	35	1/19/2020	1/17/2020	1/20/2020
United States	Snohomish	M	30+	1/15/2020	1/19/2020	1/20/2020
United States	Chicago	F	60	1/13/2020	1/22/2020	1/24/2020
France	Paris	M	31	1/18/2020	1/19/2020	1/25/2020
France	Paris	F	30	1/18/2020	1/23/2020	1/25/2020
Nepal	Kathmandu	M	32	1/5/2020	1/13/2020	1/24/2020
Australia	Sydney	M	43	1/19/2020	1/20/2020	1/25/2020
Australia	Sydney	M	35	1/6/2020	1/7/2020	1/25/2020
Australia	Melbourne	M	50+	1/19/2020	1/23/2020	1/25/2020
Sri Lanka	Colombo	F	43	1/19/2020	1/25/2020	1/27/2020
Taiwan	Central Taiwan	M	40+	1/12/2020	1/21/2020	1/31/2020

* Cases reported in Hong Kong and Macau were not included in the estimation of outbreak size because most of the patients transit in Shenzhen/Zhuhai before coming to Hong Kong or Macau.

Table S3. The number of ex-pats and confirmed cases who returned to their countries from Wuhan on chartered flights between 29 January and 4 February 2020*

Country or region	Date of arrival	Number of passengers	Number of confirmed cases	Number of asymptomatic cases
Japan	1/29/2020	206	4	1
United States	1/29/2020	201	0	-
Japan	1/30/2020	210	3	2
Singapore	1/30/2020	92	4	2
Portugal	1/30/2020	350	0	-
France	1/31/2020	200	0	-
Japan	1/31/2020	149	2	1
South Korea	1/31/2020	368	1	1
United Kingdom	1/31/2020	110	0	-
Germany	2/1/2020	124	2	0
South Korea	2/1/2020	333	1	0
Turkey	2/1/2020	42	0	-
Indonesia	2/1/2020	250	0	-
Bangladesh	2/1/2020	316	0	-
Sri Lanka	2/1/2020	33	0	-
India	2/1/2020	324	0	-
France	2/2/2020	254	0	-
United Kingdom	2/2/2020	11	0	-
Ukraine	2/2/2020	128	0	-
Belgium	2/2/2020	9	1	1
Saudi Arabia	2/2/2020	10	0	-
Kazakhstan	2/2/2020	83	0	-
Morocco	2/2/2020	167	0	-
Australia	2/3/2020	243	0	-
Taiwan	2/3/2020	247	1	-
Italy	2/3/2020	56	1	-
Malaysia	2/4/2020	107	2	2
Thailand	2/4/2020	138	1	-

[*] Japan, Singapore, South Korea, Germany, Belgium and Malaysia had tested all the passengers and a few pre-symptomatic or asymptomatic cases were confirmed. For other countries, it is unknown whether they had tested all the passengers or only those showing symptoms.

Table S4. The age distribution of 425 confirmed cases of 2019-nCoV during the early stages of the epidemic in Wuhan [2] and in mainland China [3-5]

Age	Wuhan no. of cases	Wuhan proportion	Mainland China no. of cases	Mainland China proportion
<15	0	0.0%	115	3.2%
15-44	84	19.8%	1625	45.4%
45-64	179	42.1%	1370	38.2%
>64	162	38.1%	472	13.2%
Total	425	100.0%	3582	100%

Table S5. The age distribution of 41 confirmed death cases of 2019-nCoV during the early stages of the epidemic in Wuhan [3-5]

Age	No. fatal cases	Proportion of fatal cases
<15	0	0.0%
15-44	2	4.9%
45-64	6	14.6%
>64	33	80.5%
Total	41	100.0%

Table S6. The daily number of laboratory-confirmed death cases of 2019-nCoV in Wuhan up to 11 February 2020 [6]

Date	Cumulative no. of cases	Cumulative no. of deaths
12/31/2019	27	0
1/3/2020	44	0
1/5/2020	59	0
1/8/2020	59	0
1/10/2020	41	0
1/11/2020	41	1
1/12/2020	41	1
1/13/2020	41	1
1/14/2020	41	1
1/15/2020	41	1
1/16/2020	45	2
1/17/2020	62	2
1/18/2020	121	3
1/19/2020	198	4
1/20/2020	258	6
1/21/2020	365	9
1/22/2020	425	17
1/23/2020	495	25
1/24/2020	572	38
1/25/2020	618	45
1/26/2020	698	63
1/27/2020	1590	85
1/28/2020	1905	104
1/29/2020	2261	129
1/30/2020	2639	159
1/31/2020	3215	192
2/1/2020	4109	224
2/2/2020	5142	265
2/3/2020	6384	313
2/4/2020	8351	362
2/5/2020	10117	414
2/6/2020	11618	478
2/7/2020	13603	545
2/8/2020	14982	608
2/9/2020	16902	681
2/10/2020	18454	748
2/11/2020	19558	820

Table S7. The time between onset and death or the time between admission and death for 41 death cases of 2019-nCoV in Wuhan [3]

Sex	Age	Date of symptom onset	Date of first admission	Death date	Time between onset to death	Time between admission to death
M	61	12/20/2019	12/27/2019	1/9/2020	20	13
M	69	12/25/2019	12/31/2019	1/15/2020	21	15
M	89	1/5/2020	1/8/2020	1/18/2020	13	10
M	89	1/13/2020	1/13/2020	1/19/2020	6	6
M	75	1/4/2020	1/11/2020	1/20/2020	16	9
F	69		1/14/2020	1/20/2020		6
M	66	1/10/2020	1/16/2020	1/20/2020	10	4
F	48	12/10/2019	12/10/2019	1/20/2020	41	41
M	86	1/2/2020	1/9/2020	1/21/2020	19	12
F	85		11/26/2019	1/21/2020		56
M	82	1/9/2020	1/14/2020	1/21/2020	12	7
M	81	1/9/2020	1/13/2020	1/21/2020	12	8
F	70		1/13/2020	1/21/2020		8
F	66	1/8/2020	1/19/2020	1/21/2020	13	2
M	66	12/22/2019	12/22/2019	1/21/2020	30	30
M	66		1/11/2020	1/21/2020		10
M	65	1/5/2020	1/11/2020	1/21/2020	16	10
M	53		1/13/2020	1/21/2020		8
M	84	1/6/2020	1/9/2020	1/22/2020	16	13
F	82	1/3/2020	1/3/2020	1/22/2020	19	19
M	81		1/18/2020	1/22/2020		4
F	80		1/11/2020	1/22/2020		11
M	73	12/30/2019	1/5/2020	1/22/2020	23	17
F	70		1/18/2020	1/22/2020		4
M	87	1/12/2020	1/19/2020	1/23/2020	11	4
F	82	1/10/2020	1/17/2020	1/23/2020	13	6
M	72	1/11/2020	1/18/2020	1/23/2020	12	5
M	70	1/15/2020	1/19/2020	1/23/2020	8	4
F	67	1/10/2020	1/12/2020	1/24/2020	14	12
M	65	1/12/2020	1/16/2020	1/23/2020	11	7
M	58	1/1/2020	1/1/2020	1/24/2020	23	23
M	36	1/6/2020	1/9/2020	1/23/2020	17	14
M	79	1/11/2020	1/17/2020	1/24/2020	13	7
M	78	1/13/2020	1/23/2020	1/24/2020	11	1
F	76		1/5/2020	1/24/2020		19
M	67	1/10/2020	1/15/2020	1/24/2020	14	9
M	58	1/3/2020	1/18/2020	1/24/2020	21	6
M	55	1/8/2020	1/19/2020	1/24/2020	16	5
F	73		1/23/2020	1/28/2020		5

F	73			1/26/2020		
F	31	1/21/2020	1/23/2020	1/26/2020	5	3

Table S8. Dates of symptom onset of infector-infectee pairs

Region	Age of index patient	Sex of index patient	Date of arrival of index patient	Date of onset of index patient	Age of secondary patient	Sex of secondary patient	Date of onset of secondary patient	Source of secondary patient
Wuhan	61	M		12/20/2019	57	F	12/25/2019	Li et al [2]
Wuhan	61	M		12/20/2019	31	F	12/29/2019	Li et al [2]
Wuhan	62	F		12/27/2019	64	M	1/3/2020	Li et al [2]
Wuhan	49	M		12/12/2019	78	M	12/19/2019	Li et al [2]
Wuhan	52	F		12/21/2019	25	F	12/24/2019	Li et al [2]
Huanggang	32	M		1/4/2020	28	F	1/11/2020	Li et al [2]
Shenzhen	65, 66, 37, 36	F, M, F, M	1/4/2020	1/1/2020 - 1/4/2020	63	F	1/8/2020	Chan et al [24]
Chicago	60+	F	1/13/2020	1/22/2020 - 1/23/2020	60+	M	1/28/2020	6th case in US
Taiwan	40+	M	1/12/2020	1/21/2020	40+	F	1/27/2020	9th case in Taiwan
Hong Kong	73, 72	F, M	1/22/2020	1/22/2020 - 1/25/2020	37	F	1/28/2020	11th case in Hong Kong
Hong Kong	39	M	1/23/2020	1/29/2020	72	F	2/1/2020	15th case in Hong Kong
Singapore	28	F	Local infection	1/29/2020	45	M	2/1/2020	27th case in Singapore
Singapore	28	F	Local infection	1/29/2020	44	F	2/2/2020	21st case in Singapore
Vietnam	23	F	1/17/2020	1/25/2020	42	F	1/31/2020	10th case in Vietnam
Vietnam	23	F	1/17/2020	1/25/2020	49	F	2/3/2020	11th case in Vietnam
Vietnam	23	F	1/17/2020	1/25/2020	16	F	2/3/2020	12th case in Vietnam
Hong Kong	60	M	Local infection	1/22/2020	28	F	1/30/2020	19th case in Hong Kong
Hong Kong	60, 28	M, F	Local infection	1/22/2020 - 1/30/2020	56	F	2/4/2020	20th case in Hong Kong
Malaysia	42	M	1/23/2020	1/29/2020	40	F	2/1/2020	14th case in Malaysia
Shenzhen	64	F	Local infection	1/4/2020	69	M	1/12/2020	12nd case in Shenzhen
Shenzhen	69	M	1/15/2020	1/12/2020	25	F	1/25/2020	188th case in Shenzhen
Shenzhen	64, 71	F, M	1/12/2020	1/12/2020 -1/15/2020	38	F	1/22/2020	39th case in Shenzhen
Shenzhen	73	M	1/19/2020	1/20/2020	49	M	1/24/2020	68th case in Shenzhen
Shenzhen	73	M	1/19/2020	1/20/2020	49	F	1/22/2020	69th case in Shenzhen
Shenzhen	62	F	1/22/2020	1/19/2020	43	F	2/4/2020	320th case in Shenzhen
Shenzhen	30	F	1/21/2020	1/22/2020	38	F	1/28/2020	108th case in Shenzhen
Shenzhen	61	M	1/22/2020	1/24/2020	35	F	1/26/2020	120th case in Shenzhen

Shenzhen	61	M	1/22/2020	1/24/2020	39	M	1/26/2020	121st case in Shenzhen
Shenzhen	61, 35, 39	M, F, M	1/24/2020	1/24/2020 - 1/26/2020	2	M	2/5/2020	327th case in Shenzhen
Shenzhen	61, 35, 39	M, F, M	1/22/2020	1/24/2020 - 1/26/2020	62	F	2/5/2020	345th case in Shenzhen
Shenzhen	64	F	1/22/2020	1/17/2020	37	F	1/28/2020	275th case in Shenzhen
Shenzhen	34	M	1/20/2020	1/20/2020	32	M	1/23/2020	100th case in Shenzhen
Shenzhen	34	M	1/20/2020	1/20/2020	32	M	1/25/2020	101st case in Shenzhen
Shenzhen	34	M	1/20/2020	1/20/2020	34	M	1/26/2020	102nd case in Shenzhen
Shenzhen	64	M	1/17/2020	1/18/2020	37	M	1/24/2020	169th case in Shenzhen
Shenzhen	78	M	1/16/2020	1/26/2020	78	F	1/30/2020	283rd case in Shenzhen
Shenzhen	59, 61	F, M	1/21/2020	1/24/2020	34	F	1/25/2020	253rd case in Shenzhen
Shenzhen	50	F	Local infection	1/26/2020	37	F	2/5/2020	354th case in Shenzhen
Shenzhen	42	M	Local infection	1/20/2020	6	M	2/6/2020	340th case in Shenzhen
Shenzhen	69	M	Local infection	1/21/2020	66	F	2/2/2020	335th case in Shenzhen

Table S9. Population age distribution in Wuhan [25]

Age	Nationwide Census 2010*	Proportion 2010	Household Survey 2017**	Proportion 2017
0-14	1249115	14.1%	1159753	13.6%
15-44	4285149	48.4%	3591038	42.1%
45-64	2527155	28.5%	2616976	30.7%
65+	795131	9.0%	1168750	13.7%
Total	8856550	100.0%	8536517	100.0%

*Nationwide Population Census included residents who had lived in Wuhan for at least 6 months prior to the Census.

**Household Survey only included residents whose household registrations (Hukou) were in Wuhan during the Survey. These residents might not live in Wuhan while the Survey was conducted.