CORONAVIRUS

The effect of human mobility and control measures on the COVID-19 epidemic in China

Moritz U. G. Kraemer[1,2,3]*, Chia-Hung Yang[4], Bernardo Gutierrez[1,5], Chieh-Hsi Wu[6], Brennan Klein[4], David M. Pigott[7], Open COVID-19 Data Working Group†, Louis du Plessis[1], Nuno R. Faria[1], Ruoran Li[8], William P. Hanage[8], John S. Brownstein[2,3], Maylis Layan[9,10], Alessandro Vespignani[4,11], Huaiyu Tian[12], Christopher Dye[1], Oliver G. Pybus[1,13]*, Samuel V. Scarpino[4]*

The ongoing coronavirus disease 2019 (COVID-19) outbreak expanded rapidly throughout China. Major behavioral, clinical, and state interventions were undertaken to mitigate the epidemic and prevent the persistence of the virus in human populations in China and worldwide. It remains unclear how these unprecedented interventions, including travel restrictions, affected COVID-19 spread in China. We used real-time mobility data from Wuhan and detailed case data including travel history to elucidate the role of case importation in transmission in cities across China and to ascertain the impact of control measures. Early on, the spatial distribution of COVID-19 cases in China was explained well by human mobility data. After the implementation of control measures, this correlation dropped and growth rates became negative in most locations, although shifts in the demographics of reported cases were still indicative of local chains of transmission outside of Wuhan. This study shows that the drastic control measures implemented in China substantially mitigated the spread of COVID-19.

The outbreak of coronavirus disease 2019 (COVID-19) spread rapidly from its origin in Wuhan, Hubei Province, China (*1*). A range of interventions were implemented after the detection in late December 2019 of a cluster of pneumonia cases of unknown etiology and identification of the causative virus, severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2), in early January 2020 (*2*). Interventions include improved rates of diagnostic testing; clinical management; rapid isolation of suspected cases, confirmed cases, and contacts; and, most notably, restrictions on mobility (hereafter called cordon sanitaire) imposed on Wuhan city on 23 January 2020. Travel restrictions were subsequently imposed on 14 other cities across Hubei Province, and partial movement restrictions were enacted in many cities across China. Initial analysis suggests that the Wuhan cordon sanitaire resulted in an average 3-day delay of COVID-19 spread to other cities (*3*), but the full extent of the effect of the mobility restrictions and other types of interventions on transmission has not been examined quantitatively (*4*–*6*). Questions remain over how these interventions affected the spread of SARS-CoV-2 to locations outside of Wuhan. Here, we used real-time mobility data, crowdsourced line list data of cases with reported travel history, and timelines of reporting changes to identify early shifts in the epidemiological dynamics of the COVID-19 epidemic in China, from an epidemic driven by frequent importations to local transmission.

[1]Department of Zoology, University of Oxford, Oxford, UK. [2]Harvard Medical School, Harvard University, Boston, MA, USA. [3]Boston Children's Hospital, Boston, MA, USA. [4]Network Science Institute, Northeastern University, Boston, MA, USA. [5]School of Biological and Environmental Sciences, Universidad San Francisco de Quito USFQ, Quito, Ecuador. [6]Mathematical Sciences, University of Southampton, Southampton, UK. [7]Institute for Health Metrics and Evaluation, Department of Health Metrics, University of Washington, Seattle, WA, USA. [8]Harvard T.H. Chan School of Public Health, Boston, MA, USA. [9]Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, UMR2000, CNRS, Paris, France. [10]Sorbonne Université, Paris, France. [11]ISI Foundation, Turin, Italy. [12]State Key Laboratory of Remote Sensing Science, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China. [13]Department of Pathobiology and Population Sciences, The Royal Veterinary College, London, UK.
***Corresponding author. Email: s.scarpino@northeastern.edu (S.V.S.); oliver.pybus@zoo.ox.ac.uk (O.G.P.); moritz.kraemer@zoo.ox.ac.uk (M.U.G.K.)** †Members of the Open COVID-19 Data Working Group are listed in the supplementary materials.

Human mobility predicts the spread and size of epidemics in China

As of 1 March 2020, 79,986 cases of COVID-19 were confirmed in China (Fig. 1A) (*7*). Reports of cases in China were mostly restricted to Hubei until 23 January 2020 (81% of all cases), after which most provinces reported rapid increases in cases (Fig. 1A). We built a line list dataset from reported cases in China with information on travel history and demographic characteristics (*8*). We note that the majority of early cases (before 23 January 2020; see the materials and methods) reported outside of Wuhan had known travel history to Wuhan (57%) and were distributed across China (Fig. 1B), highlighting the importance of Wuhan as a major source of early cases. However, initial testing was focused mainly on travelers from Wuhan, potentially biasing estimates of travel-related infections upward (see the materials and methods). Among cases known to have traveled from Wuhan before 23 January 2020, the time from symptom onset to confirmation was 6.5 days (SD = 4.2 days; fig. S2), providing opportunity for onward transmission at the destination. More active surveillance reduced this interval to 4.8 days (SD = 3.03 days; fig. S2) for those who traveled after 23 January 2020.

To identify accurately a time frame for evaluating early shifts in SARS-CoV-2 transmission in China, we first estimated from case data the average incubation period of COVID-19 infection [i.e., the duration between time of infection and symptom onset (*9*, *10*)]. Because infection events are typically not observed directly, we estimated the incubation period from the span of exposure during which infection likely occurred. Using detailed information on 38 cases for whom both the dates of entry to and exit from Wuhan were known, we estimated the mean incubation period to be 5.1 days (SD = 3.0 days; fig. S1), similar to previous estimates from other data (*11*, *12*). In subsequent analyses, we added an upper estimate of one incubation period (mean + 1 SD = 8 days) to the date of Wuhan shutdown to delineate the date before which cases recorded in other provinces might represent infections acquired in Hubei (i.e., 1 February 2020; Fig. 1A).

To understand whether the volume of travel within China could predict the epidemic outside of Wuhan, we analyzed real-time human mobility data from Baidu Inc., together with epidemiological data from each province (see the materials and methods). We investigated spatiotemporal disease spread to elucidate the relative contribution of Wuhan to transmission elsewhere and to evaluate how the cordon sanitaire may have affected it.

Among cases reported outside of Hubei province in our dataset, we observed 515 cases with known travel history to Wuhan and a symptom onset date before 31 January 2020, compared with only 39 cases after 31 January 2020, illustrating the effect of travel restrictions (Figs. 1B and 2A and fig. S3). We confirmed the expected decline of importation with real-time human mobility data from Baidu Inc. Movements of individuals out of Wuhan increased in the days before the Lunar New Year and the establishment of the cordon sanitaire, before rapidly decreasing to almost no movement (Fig. 2, A and B). The travel ban appears to have prevented travel into and out of Wuhan around the time of the Lunar New Year celebration (Fig. 2A) and likely reduced further dissemination of SARS-CoV-2 from Wuhan.

To test the contribution of the epidemic in Wuhan to seeding epidemics elsewhere in China, we built a naïve COVID-19 "generalized" linear model [GLM (*13*)] of daily case counts (see the materials and methods). We estimated the epidemic doubling time outside of Hubei to be 4.0 days (range across provinces, 3.6 to 5.0 days) and estimated the epidemic doubling time within Hubei to be 7.2 days, consistent with previous reports (*5*, *12*, *14*, *15*). Our model predicted daily case counts across all provinces with relatively high accuracy (as measured with a pseudo-R^2 from a negative binomial GLM)

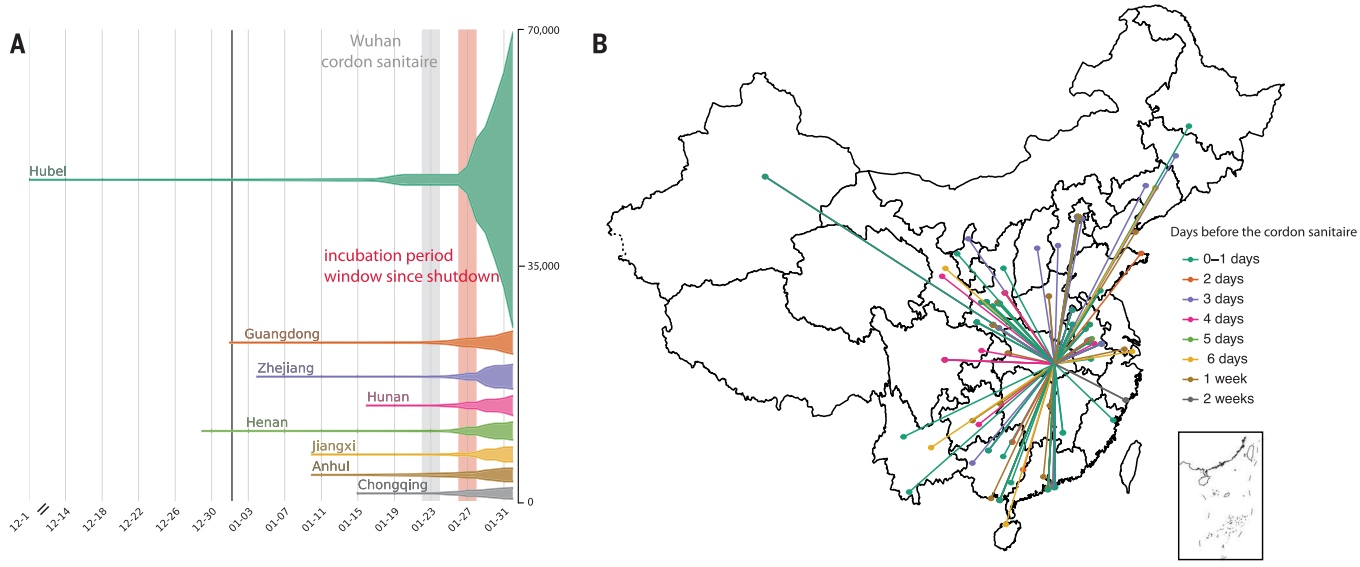

Fig. 1. Number of cases and key dates during the epidemic. (**A**) Epidemic curve of the COVID-19 outbreak in provinces in China. Bars indicate key dates: implementation of the cordon sanitaire of Wuhan (gray) and the end of the first incubation period after the travel restrictions (red). The black line represents the closure of the Wuhan seafood market on 1 January 2020. The width of each horizontal tube represents the number of reported cases in that province. (**B**) Map of COVID-19 confirmed cases (n = 554) that had reported travel history from Wuhan before travel restrictions were implemented on 23 January 2020. Colors of the lines indicate date of travel relative to the date of travel restrictions.

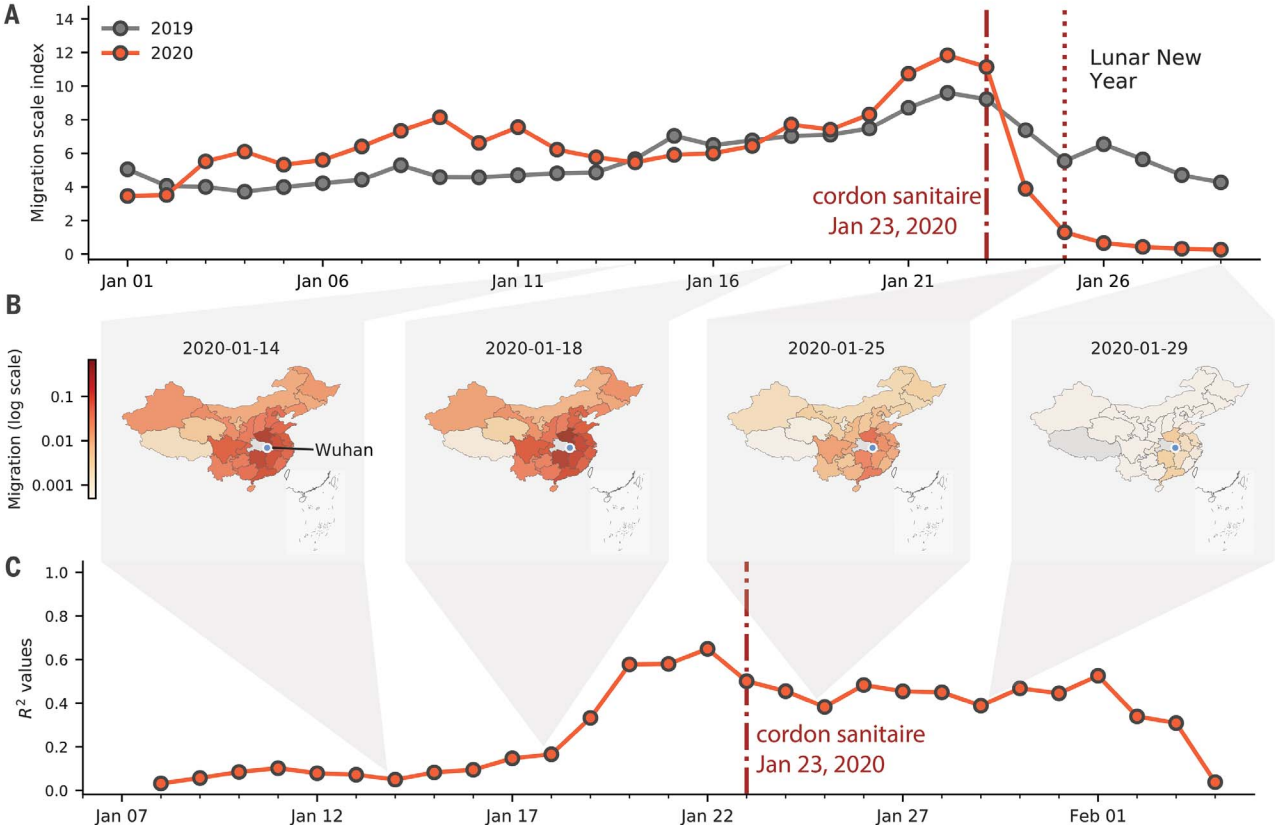

Fig. 2. Human mobility, spread, and synchrony of the COVID-19 outbreak in China. (**A**) Human mobility data extracted in real time from Baidu Inc. Travel restrictions from Wuhan and large-scale control measures started on 23 January 2020. Gray and red lines represent fluxes of human movements for 2019 and 2020, respectively. (**B**) Relative movements from Wuhan to other provinces in China. (**C**) Timeline of the correlation between daily incidence in Wuhan and incidence in all other provinces, weighted by human mobility.

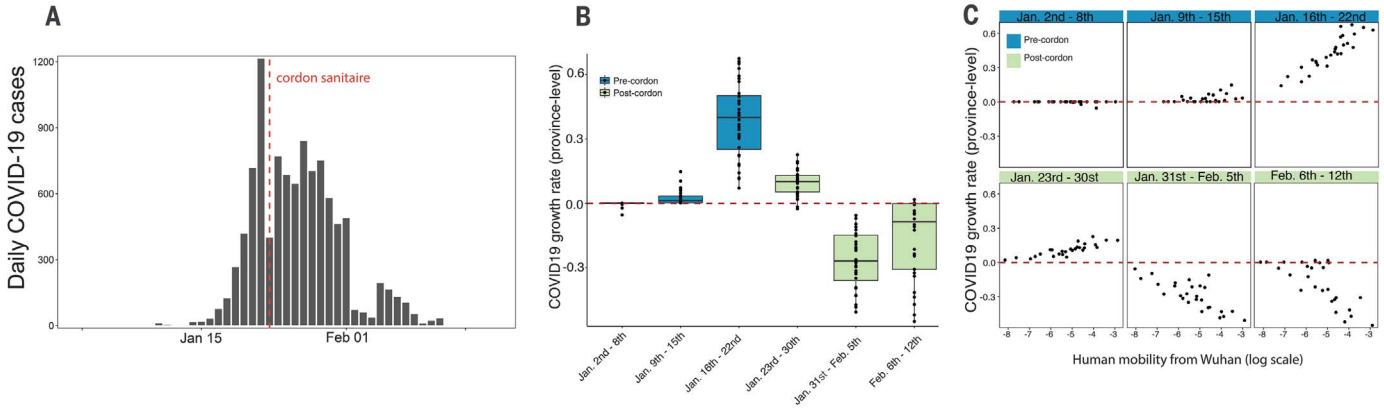

Fig. 3. Human mobility explains the early epidemic growth rate in China. (**A**) Daily counts of cases in China. (**B**) Time series of province-level growth rates of the COVID-19 epidemic in China. Estimates of the growth rate were obtained by performing a time-series analysis using a mixed-effects model of lagged, log linear daily case counts in each province (see the materials and methods). Above the red line are positive growth rates and below are negative rates. Blue indicates dates before the implementation of the cordon sanitaire and green after. (**C**) Relationship between growth rate and human mobility at different times of the epidemic. Blue indicates before the implementation of the cordon sanitaire and green after.

throughout early February 2020 and when accounting for human mobility (Fig. 2C and tables S1 and S2), consistent with an exploratory analysis (*6*).

We found that the magnitude of the early epidemic (total number of cases until 10 February 2020) outside of Wuhan was very well predicted by the volume of human movement out of Wuhan alone (R^2 = 0.89 from a log-linear regression using cumulative cases; fig. S8). Therefore, cases exported from Wuhan before the cordon sanitaire appear to have contributed to initiating local chains of transmission, both in neighboring provinces (e.g., Henan) and in more distant provinces (e.g., Guangdong and Zhejiang) (Figs. 1A and 2B). Further, the frequency of introductions from Wuhan were also predictive of the size of the early epidemic in other provinces (controlling for population size) and thus the probability of large outbreaks (fig. S8).

After 1 February 2020 (corresponding to one mean + one SD incubation period after the cordon sanitaire and other interventions were implemented), the correlation of daily case counts and human mobility from Wuhan decreased (Fig. 2C), indicating that variability among locations in daily case counts was better explained by factors unrelated to human mobility, such as local public health response. This suggests that whereas travel restrictions may have reduced the flow of case importations from Wuhan, other local mitigation strategies aimed at halting local transmission increased in importance later.

We also estimated the growth rates of the epidemic in all other provinces (see the materials and methods). We found that all provinces outside of Hubei experienced faster growth rates between 9 January and 22 January 2020 (Fig. 3, A and B, and fig. S4b), which was the time before travel restrictions and substantial control measures were implemented (Fig. 3C and fig. S6); this was also apparent from the case counts by province (fig. S6). In the same period, variation in the growth rates is almost entirely explained by human movements from Wuhan (Fig. 3C and fig. S9), consistent with the theory of infectious disease spread in highly coupled metapopulations (*16*, *17*). After the implementation of drastic control measures across the country, growth rates became negative (Fig. 3B), indicating that transmission was successfully mitigated. The correlation of growth rates and human mobility from Wuhan became negative; that is, provinces with larger mobility from Wuhan before the cordon sanitaire (but also larger number of cases overall) had more rapidly declining growth rates of daily case counts. This could be due partly to travel restrictions but also to the fact that control measures may have been more drastic in locations with larger outbreaks driven by local transmission (for more details, see "Current role of imported cases in Chinese provinces" section).

The travel ban coincided with increased testing capacity across provinces in China. Therefore, an alternative hypothesis is that the observed epidemiological patterns outside of Wuhan were the result of increased testing capacity. We tested this hypothesis by including differences in testing capacity before and after the rollout of large-scale testing in China on 20 January 2020 [the date that COVID-19 became a class B notifiable disease (*18*, *19*)] and determined the impact of this binary variable on the predictability of daily cases (see the materials and methods). We plotted the relative improvement in the prediction of our model (on the basis of normalized residual error) of (i) a model that includes daily mobility from Wuhan and (ii) a model that includes testing availability (for more details, see the materials and methods). Overall, the inclusion of mobility data from Wuhan produced an improvement in the model's prediction [delta-Bayesian information criterion > 250 (*20*)] over a naïve model that considers only autochthonous transmission with a doubling time of 2 to 8 days (Fig. 3B). Of the 27 provinces in China reporting cases through 6 February 2020, we found that the largest improvements in prediction for 12 provinces could be achieved using mobility only (fig. S5). In 10 provinces, both testing and mobility improved the model's prediction, and in only one province (Hunan) was testing the most important factor improving model prediction (fig. S5). We conclude that laboratory testing during the early phase of the epidemic was critical; however, mobility out of Wuhan remained the main driver of spread before the cordon sanitaire. Large-scale molecular and serological data will be important to investigate further the exact magnitude of the impact of human mobility compared with other factors.

Current role of imported cases in Chinese provinces

Because case counts outside of Wuhan have decreased (Fig. 3B), we can further investigate the current contribution of imported cases to local epidemics outside of Wuhan by investigating case characteristics. Age and sex distributions can reflect heterogeneities in the risk of infection within affected populations. To investigate meaningful shifts in the epidemiology of the COVID-19 outbreak through time, we examined age and sex data for cases from different periods of the outbreak and from individuals with and without travel from Wuhan. However, details of travel history exist for only a fraction of confirmed cases, and this information was particularly scant for some provinces (e.g., Zhejiang and Guangdong). Therefore, we grouped confirmed cases into four categories: (I) early cases (i.e., reported before 1 February 2020) with travel

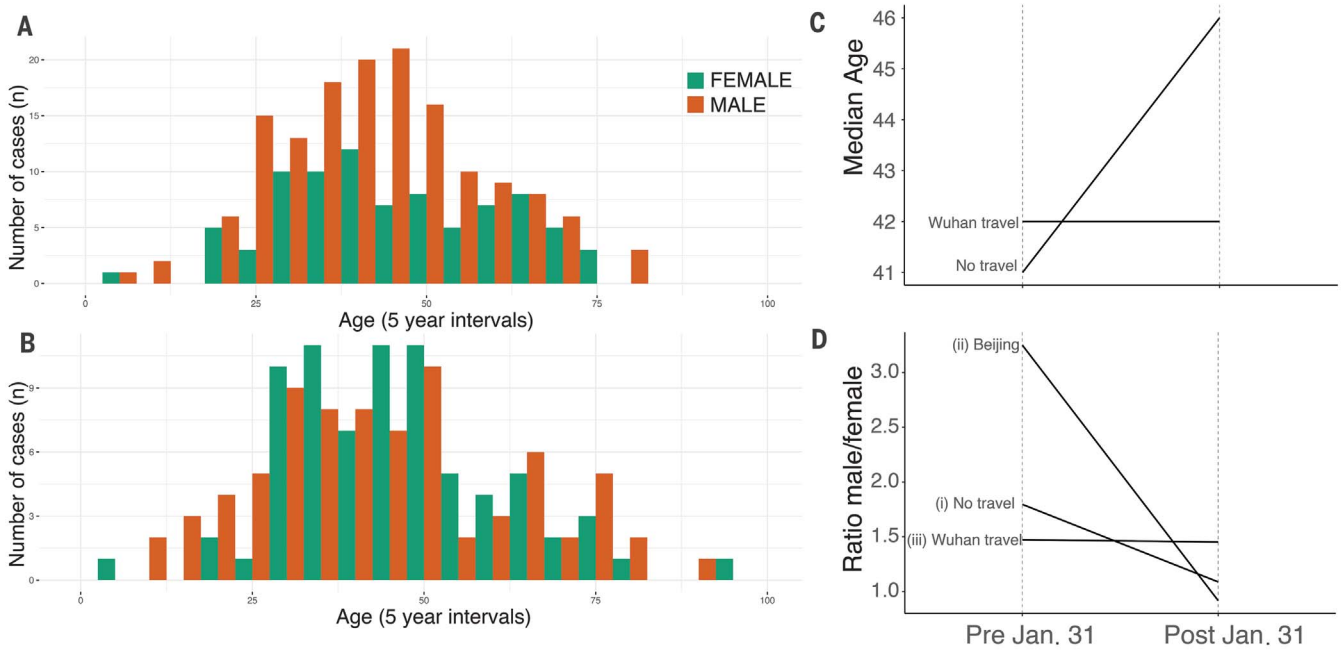

Fig. 4. Shifting age and sex distributions through time. (**A**) Age and sex distributions of confirmed cases with known travel history to Wuhan. (**B**) Age and sex distributions of confirmed cases that had no travel history to Wuhan. (**C**) Median age for cases reported early (before 1 February) and those reported later (between 1 and 10 February). Full distributions are shown in fig. S7. (**D**) Change through time in the sex ratio of (i) all reported cases in China with no reported travel history, (ii) cases reported in Beijing without travel history, and (iii) cases known to have traveled from Wuhan.

history, (II) early cases without travel history, (III) later cases (i.e., reported between 1 February and 10 February 2020) with travel history, and (IV) later cases without travel history.

Using crowdsourced case data, we found that cases with travel history (categories I and III) had similar median ages and sex ratios in both the early and later phases of the outbreak (age 41 versus 42 years; 50% interquartile interval: 32.75 versus 30.75 and 54.25 versus 53.5 years, respectively; P value > 0.1, 1.47 versus 1.45 males per female, respectively; Fig. 4D and fig. S7). Early cases with no information on travel history (category II) had a median age and sex ratio similar to those with known travel history (age 42 years; 50% interquartile interval: 30.5 to 49.5, P value > 0.1; 1.80 males per female; Fig. 4D). However, the sex ratio of later cases without reported travel history (category IV) shifted to ~1:1 (57 male versus 62 female, X^2 test, P value < 0.01), as expected under a null hypothesis of equal transmission risk [Fig. 4, A, B, and D; see also (*21*, *22*) and the materials and methods], and the median age in this group increased to 46 (50% interquartile interval: 34.25 to 58, t test: P value < 0.01; Fig. 4, A to C, and fig. S7). We hypothesize that many of the cases with no known travel history in the early phase were indeed travelers who contributed to disseminating SARS-CoV-2 outside of Wuhan. The shift toward more equal sex ratios and older ages in nontravelers after 31 January 2020 confirms the finding that epidemics outside of Wuhan were then driven by local transmission dynamics. The case definition changed to include cases without travel history to Wuhan after 23 January 2020 (see the materials and methods).

Discussion

Containment of respiratory infections is particularly difficult if they are characterized by relatively mild symptoms or transmission before the onset of symptoms (*23*, *24*). Intensive control measures, including travel restrictions, have been implemented to limit the spread of COVID-19 in China. Here, we show that travel restrictions are particularly useful in the early stage of an outbreak when it is confined to a certain area that acts as a major source. However, travel restrictions may be less effective once the outbreak is more widespread. The combination of interventions implemented in China was clearly successful in mitigating spread and reducing local transmission of COVID-19, although in this work it was not possible to definitively determine the impact of each intervention. Much further work is required to determine how to balance optimally the expected positive effect on public health with the negative impact on freedom of movement, the economy, and society at large.

REFERENCES AND NOTES

1. S. Chen, J. Yang, W. Yang, C. Wang, T. Bärnighausen, *Lancet* **395**, 764–766 (2020).
2. N. Zhu *et al.*, *N. Engl. J. Med.* **382**, 727–733 (2020).
3. H. Tian *et al.*, *Science* 10.1126/science.abb6105 (2020).
4. Z. Du *et al.*, *Emerg. Infect. Dis.* **26** (2020).
5. J. T. Wu, K. Leung, G. M. Leung, *Lancet* **395**, 689–697 (2020).
6. S. Zhao *et al.*, *J. Travel Med.* 10.1093/jtm/taaa022 (2020).
7. World Health Organization (WHO), *Coronavirus Disease 2019 (COVID-19) Situation Report – 27* (2020); https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200216-sitrep-27-covid-19.pdf?sfvrsn=78c0eb78_2.
8. B. Xu *et al.*, *Sci. Data* **7**, 106 (2020).
9. S. Cauchemez *et al.*, *Lancet Infect. Dis.* **14**, 50–56 (2014).
10. J. Lessler *et al.*, *Lancet Infect. Dis.* **9**, 291–300 (2009).
11. J. A. Backer, D. Klinkenberg, J. Wallinga, *Euro Surveill.* **25**, 20–28 (2020).
12. Q. Li *et al.*, *N. Engl. J. Med.* NEJMoa2001316 (2020).
13. T. J. Hastie, D. Pregibon, "Generalized linear models" in *Statistical Models in S*, J. M. Chambers, T. J. Hastie, Eds. (Wadsworth & Brooks/Cole, 1992), pp. 195–246.
14. J. Riou, C. L. Althaus, *Euro Surveill.* **25**, 1–5 (2020).
15. A. R. Tuite, D. N. Fisman, *Ann. Intern. Med.* (2020).
16. M. J. Keeling, O. N. Bjørnstad, B. T. Grenfell, "Metapopulation dynamics of infectious diseases" in *Ecology, Genetics and Evolution of Metapopulations*, I. Hanski, O. E. Gaggiotti, Eds. (Elsevier, 2004), pp. 415–445.
17. D. J. Watts, R. Muhamad, D. C. Medina, P. S. Dodds, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11157–11162 (2005).
18. World Health Organization, *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)* (WHO, 2020); https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf.
19. Z. Wu, J. M. McGoogan, *JAMA* **2019**, 17–20 (2020).
20. K. P. Burnham, D. R. Anderson, *Sociol. Methods Res.* **33**, 261–304 (2004).
21. Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 145–151 (2020).
22. E. Goldstein, V. E. Pitzer, J. J. O'Hagan, M. Lipsitch, *Epidemiology* **28**, 136–144 (2017).
23. C. Fraser, S. Riley, R. M. Anderson, N. M. Ferguson, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6146–6151 (2004).

24. R. Li *et al*., *Science* 10.1126/science.aba9757 (2020).
25. M. U. G. Kraemer *et al*., Open COVID-19 Data Working Group, L. du Plessis *et al*., Code for: The effect of human mobility and control measures on the COVID-19 epidemic in China. Zenodo (2020); https://doi.org/10.5281/zenodo.3714914.

ACKNOWLEDGMENTS

We thank all individuals who are collecting epidemiological data of the COVID-19 outbreak around the world. **Funding:** H.T., O.G.P., and M.U.G.K. acknowledge support from the Oxford Martin School. M.U.G.K. is supported by a Branco Weiss Fellowship. B.G. is supported by a Universities of Academic Excellence Scholarship Program of the Secretariat for Higher Education, Science, Technology, and Innovation of the Republic of Ecuador (grant no. ARSEQ-BEC-003163-2017). N.R.F. is supported by a Sir Henry Dale Fellowship. W.P.H. is supported by the National Institute of General Medical Sciences (grant no. U54GM088558). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. **Author contributions:** M.U.G.K., O.G.P., and S.V.S. developed the idea and research. M.U.G.K. and S.V.S. wrote the first draft of the manuscript, and all other authors discussed results and edited the manuscript. M.U.G.K., B.G., S.V.S., D.M.P., and the Open COVID-19 Data Working Group collected and validated epidemiological data. R.L. and M.U.G.K. collected intervention data. C.-H.Y., B.K., and S.V.S. collected and processed human mobility data. **Competing interests:** S.V.S. is on the advisory board for BioFire Diagnostics Trend Surveillance, which includes paid consulting. A.V. reports past grants and personal fees from Metabiota Inc. outside of the submitted work. The remaining authors declare no competing interests. **Data and materials availability:** Code and data are available on the following GitHub repository: https://github.com/Emergent-Epidemics/covid19_cordon and permanently on Zenodo (*25*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6490/493/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 and S2
List of Members of the Open COVID-19 Data Working Group
References (*26*–*39*)

3 March 2020; accepted 23 March 2020
Published online 25 March 2020
10.1126/science.abb4218

Science

The effect of human mobility and control measures on the COVID-19 epidemic in China

Moritz U. G. Kraemer, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M. Pigott, Open COVID-19 Data Working Group, Louis du Plessis, Nuno R. Faria, Ruoran Li, William P. Hanage, John S. Brownstein, Maylis Layan, Alessandro Vespignani, Huaiyu Tian, Christopher Dye, Oliver G. Pybus and Samuel V. Scarpino

Science **368** (6490), 493-497.
DOI: 10.1126/science.abb4218originally published online March 25, 2020

Tracing infection from mobility data

What sort of measures are required to contain the spread of severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19)? The rich data from the Open COVID-19 Data Working Group include the dates when people first reported symptoms, not just a positive test date. Using these data and real-time travel data from the internet services company Baidu, Kraemer *et al.* found that mobility statistics offered a precise record of the spread of SARS-CoV-2 among the cities of China at the start of 2020. The frequency of introductions from Wuhan were predictive of the size of the epidemic sparked in other provinces. However, once the virus had escaped Wuhan, strict local control measures such as social isolation and hygiene, rather than long-distance travel restrictions, played the largest part in controlling SARS-CoV-2 spread.

Science, this issue p. 493

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.