

RESEARCH ARTICLE SUMMARY

CORONAVIRUS

Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2

Kaiyuan Sun*†, Wei Wang†, Lidong Gao†, Yan Wang, Kaiwei Luo, Lingshuang Ren, Zhifei Zhan, Xinghui Chen, Shanlu Zhao, Yiwei Huang, Qianlai Sun, Ziyang Liu, Maria Litvinova, Alessandro Vespignani, Marco Ajelli, Cécile Viboud‡, Hongjie Yu*‡

INTRODUCTION: The role of transmission heterogeneities in severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) dynamics remains unclear, particularly those heterogeneities driven by demography, behavior, and interventions. To understand individual heterogeneities and their effect on disease control, we analyze detailed contact-tracing data from Hunan, a province in China adjacent to Hubei and one of the first regions to experience a SARS-CoV-2 outbreak in January to March 2020. The Hunan outbreak was swiftly brought under control by March 2020 through a combination of nonpharmaceutical interventions including population-level mobility restriction (i.e., lockdown), traveler screening, case isolation, contact tracing, and quarantine. In parallel, highly detailed epidemiological information on SARS-CoV-2-infected individuals and their close

contacts was collected by the Hunan Provincial Center for Disease Control and Prevention.

RATIONALE: Contact-tracing data provide information to reconstruct transmission chains and understand outbreak dynamics. These data can in turn generate valuable intelligence on key epidemiological parameters and risk factors for transmission, which paves the way for more-targeted and cost-effective interventions.

RESULTS: On the basis of epidemiological information and exposure diaries on 1178 SARS-CoV-2-infected individuals and their 15,648 close contacts, we developed a series of statistical and computational models to stochastically reconstruct transmission chains, identify risk factors for transmission, and infer the infectiousness profile over the course of a typical

infection. We observe overdispersion in the distribution of secondary infections, with 80% of secondary cases traced back to 15% of infections, which indicates substantial transmission heterogeneities. We find that SARS-CoV-2 transmission risk scales positively with the duration of exposure and the closeness of social interactions, with the highest per-contact risk estimated in the household. Lockdown interventions increase transmission risk in families and households, whereas the timely isolation of infected individuals reduces risk across all types of contacts. There is a gradient of increasing susceptibility with age but no significant difference in infectivity by age or clinical severity. Early isolation of SARS-CoV-2-infected individuals drastically alters transmission kinetics, leading to shorter generation and serial intervals and a higher fraction of presymptomatic transmission. After adjusting for the censoring effects of isolation, we find that the infectiousness profile of a typical SARS-CoV-2 patient peaks just before symptom onset, with 53% of transmission occurring in the presymptomatic phase in an uncontrolled setting. We then use these results to evaluate the effectiveness of individual-based strategies (case isolation and contact quarantine) both alone and in combination with population-level contact reductions. We find that a plausible parameter space for SARS-CoV-2 control is restricted to scenarios where interventions are synergistically combined, owing to the particular transmission kinetics of this virus.

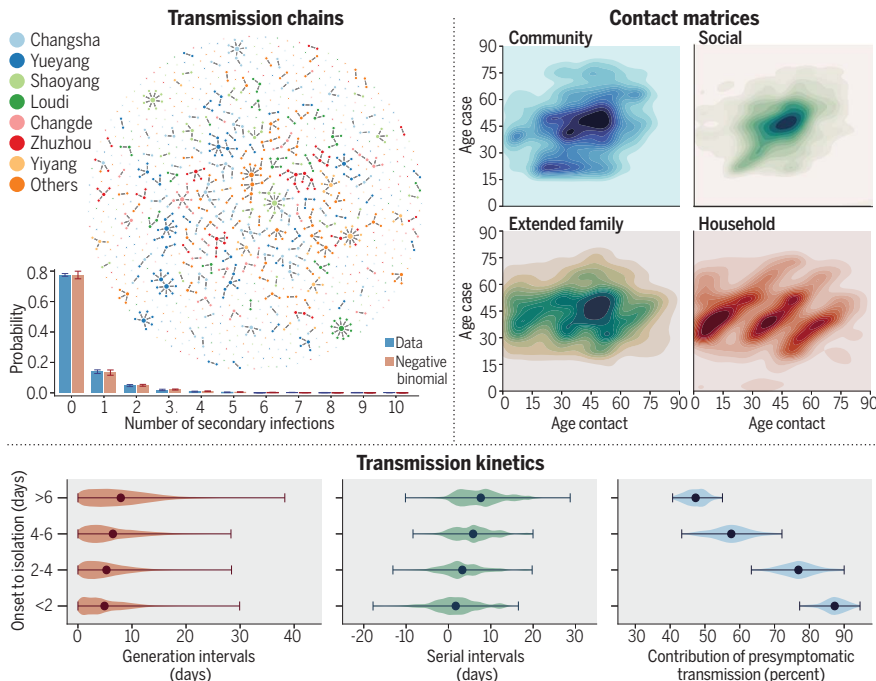
CONCLUSION: There is considerable heterogeneity in SARS-CoV-2 transmission owing to individual differences in biology and contacts that is modulated by the effects of interventions. We estimate that about half of secondary transmission events occur in the presymptomatic phase of a primary case in uncontrolled outbreaks. Achieving epidemic control requires that isolation and contact-tracing interventions are layered with population-level approaches, such as mask wearing, increased teleworking, and restrictions on large gatherings. Our study also demonstrates the value of conducting high-quality contact-tracing investigations to advance our understanding of the transmission dynamics of an emerging pathogen. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: kaiyuan.sun@nih.gov (K.S.); yhj@fudan.edu.cn (H.Y.)

†These authors contributed equally to this work.
‡These authors contributed equally to this work.

This is an open-access article distributed under the terms of the Creative Commons Attribution license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Cite this article as K. Sun et al., *Science* 371, eabe2424 (2021). DOI: 10.1126/science.abe2424

READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abe2424>



Transmission chains, contact patterns, and transmission kinetics of SARS-CoV-2 in Hunan, China, based on case and contact-tracing data from Hunan, China.

(Top left) One realization of the reconstructed transmission chains, with a histogram representing overdispersion in the distribution of secondary infections. (Top right) Contact matrices of community, social, extended family, and household contacts reveal distinct age profiles. (Bottom) Earlier isolation of primary infections shortens the generation and serial intervals while increasing the relative contribution of transmission in the presymptomatic phase.

RESEARCH ARTICLE

CORONAVIRUS

Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2

Kaiyuan Sun^{1*}†, Wei Wang²†, Lidong Gao³†, Yan Wang², Kaiwei Luo³, Lingshuang Ren², Zhifei Zhan³, Xinghui Chen², Shanlu Zhao³, Yiwei Huang³, Qianlai Sun³, Ziyang Liu³, Maria Litvinova^{4,5}, Alessandro Vespignani^{6,5}, Marco Ajelli^{4,6}, Cécile Viboud¹†, Hongjie Yu^{2*}†

A long-standing question in infectious disease dynamics concerns the role of transmission heterogeneities, which are driven by demography, behavior, and interventions. On the basis of detailed patient and contact-tracing data in Hunan, China, we find that 80% of secondary infections traced back to 15% of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) primary infections, which indicates substantial transmission heterogeneities. Transmission risk scales positively with the duration of exposure and the closeness of social interactions and is modulated by demographic and clinical factors. The lockdown period increases transmission risk in the family and households, whereas isolation and quarantine reduce risks across all types of contacts. The reconstructed infectiousness profile of a typical SARS-CoV-2 patient peaks just before symptom presentation. Modeling indicates that SARS-CoV-2 control requires the synergistic efforts of case isolation, contact quarantine, and population-level interventions because of the specific transmission kinetics of this virus.

Although it has been well documented that the clinical severity of COVID-19 increases with age (1–5), information is limited on how transmission risk varies with demographic factors, clinical presentation, and contact type (6–12). Individual-based interventions such as case isolation, contact tracing, and quarantine have been shown to accelerate case detection and interrupt transmission chains (13). However, these interventions are typically implemented in conjunction with population-level physical distancing measures, and their effects on contact patterns and transmission risk remains difficult to separate (14–24). A better understanding of the factors driving severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmission is key to achieving epidemic control while minimizing societal cost, particularly as countries relax physical distancing measures.

Hunan, a province in China adjacent to Hubei, where the COVID-19 pandemic began, experienced sustained SARS-CoV-2 transmission in late January and early February 2020,

followed by a quick suppression of the outbreak by March 2020. As in many other provinces in China, epidemic control was achieved by layering interventions targeting SARS-CoV-2 cases and their contacts with population-level physical distancing measures. In this study, we reconstruct transmission chains among all identified SARS-CoV-2 infections in Hunan, as of 3 April 2020, on the basis of granular epidemiological information collected through extensive surveillance and contact-tracing efforts. We identify the demographic, clinical, and behavioral factors that drive transmission heterogeneities and evaluate how interventions modulate the topology of the transmission network. Further, we reconstruct the infectiousness profile of SARS-CoV-2 over the course of a typical infection and estimate the feasibility of epidemic control by individual- and population-based interventions.

We analyze detailed epidemiological records for 1178 SARS-CoV-2-infected individuals and their 15,648 close contacts—representing 19,227 separate exposure events—compiled by the Hunan Provincial Center for Disease Control and Prevention. Cases were identified between 16 January and 3 April 2020; primary cases were captured by passive surveillance, contact tracing, or traveler screening and then were laboratory confirmed by reverse transcription polymerase chain reaction (RT-PCR). Individuals who were close contacts of the primary cases were followed for at least 2 weeks after the last exposure to the infected individual. Before 7 February 2020, contacts were tested only if they developed symptoms during the quarantine period. After 7 February 2020, RT-PCR testing was required for all con-

tacts, and specimens were collected at least once from each contact during quarantine, regardless of symptoms. Upon positive RT-PCR test results, infected individuals were isolated in dedicated hospitals, regardless of their clinical severity, and their contacts were quarantined in medical observation facilities. The case ascertainment process is visualized in fig. S1.

The dataset includes 210 epidemiological clusters representing 831 cases, with an additional 347 sporadic cases (29%) unlinked to any cluster (see supplementary materials and methods for more details). For each cluster, we stochastically reconstruct transmission chains and estimate the timing of infection most compatible with each patient's exposure history. We analyze an ensemble of 100 reconstructed transmission chains to account for uncertainties in exposure histories (Fig. 1 visualizes one realization of the transmission chains, and fig. S2A illustrates variability in the topology of the aggregation of 100 realizations of transmission chains).

We observe between zero and four generations of transmission, with the largest cluster involving 20 SARS-CoV-2-infected individuals. The number of secondary infections ranges from 0 to 10, with a distribution of secondary infections best characterized by a negative binomial distribution with mean $\mu = 0.40$ [95% confidence interval (CI): 0.35 to 0.47] and variance $\mu(1 + \mu/k) = 0.96$ (95% CI: 0.74 to 1.26), where $k = 0.30$ (95% CI: 0.23 to 0.39) is the dispersion parameter (Fig. 1). We find that 80% of secondary infections can be traced back to 15% of SARS-CoV-2-infected individuals, which indicates substantial transmission heterogeneities at the individual level. We can also assess geographic diffusion within Hunan province and find that the majority of transmission events occur within the same prefecture (94.3%; 95% CI: 93.7 to 95.0%), with occasional spread between prefectures (5.7%; 95% CI: 5.0 to 6.3%).

Characterizing SARS-CoV-2 transmission heterogeneities at the individual level

To dissect the individual transmission heterogeneities and identify predictors of transmission, we analyze the infection risk among a subset of 14,622 individuals who were close contacts of 870 SARS-CoV-2 patients. This dataset excludes primary cases whose infected contacts reported a travel history to Wuhan. The dataset represents 74% of all SARS-CoV-2 cases recorded in the Hunan patient database. Contacts of these 870 patients have been carefully monitored so that 17,750 independent exposure events have been captured.

We start by characterizing variation in transmission risk across the diverse set of 17,750 exposures. We study how the per-contact transmission risk varies with the type

¹Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA. ²School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China. ³Hunan Provincial Center for Disease Control and Prevention, Changsha, China.

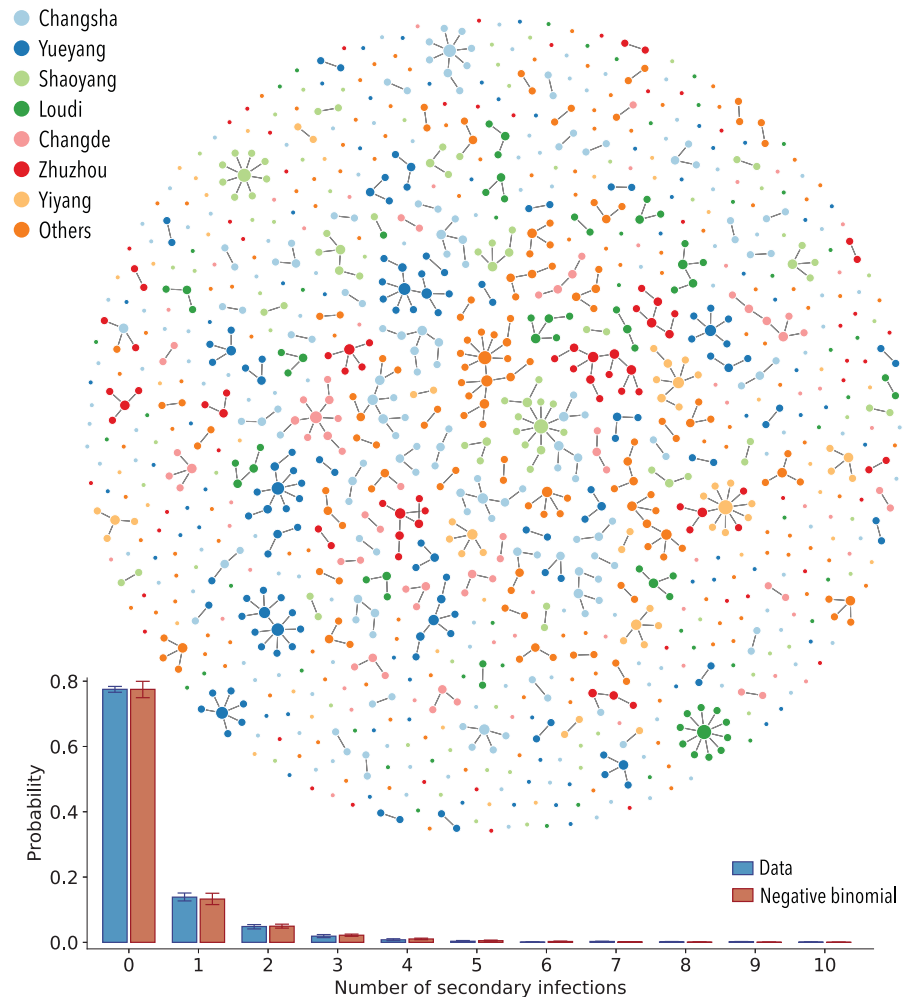
⁴Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA. ⁵ISI Foundation, Turin, Italy. ⁶Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA, USA.

*Corresponding author. Email: kaiyuan.sun@nih.gov (K.S.); yhj@fudan.edu.cn (H.Y.) †These authors contributed equally to this work. ‡These authors contributed equally to this work.

Fig. 1. SARS-CoV-2 transmission chains. (Top)

One realization of the reconstructed transmission chains among 1178 SARS-CoV-2–infected individuals in Hunan province. Each node in the network represents a patient infected with SARS-CoV-2, and each link represents an infector–infectee relationship. The color of the node denotes the reporting prefecture of the infected individuals.

(Bottom) Distribution of the number of secondary infections. Blue bars represent the ensemble averaged across 100 stochastic samples of the reconstructed transmission chains. Orange bars represent the best fit of a negative binomial distribution to the ensemble average. Vertical lines indicate 95% CIs across 100 samples (of both data and the models' fitting results). Some confidence intervals are narrow and not visible on the plot. For sensitivity analysis, we also fit the distribution with geometric and Poisson distributions. On the basis of the Akaike information criterion (AIC), the negative binomial distribution fit the data the best (average AIC score for negative binomial distribution: 1902; for geometric distribution: 1981; and for Poisson distribution: 2259).



of exposures, exposure duration, exposure timing, and physical distancing intervention, after adjusting for demographic, clinical, and travel-related factors. Exposures are grouped into five categories on the basis of contact type—i.e., household, extended family, social, community, and health care (table S2)—with the duration of exposure approximated by the time interval between the initial and final dates of exposure. To gauge the impact of physical distancing on transmission risks, we further stratify exposures by the date of occurrence, with 25 January 2020 marking the beginning of lockdown in Hunan [based on Baidu Qianxi mobility index (25); fig. S3A, insert]. To address putative variation in infectiousness over the course of infection, we distinguish whether exposures overlap with the date of symptom onset of a primary case, a period associated with high viral shedding. We use a mixed-effects multiple logistic regression model (GLMM-logit) to quantify the effects of these factors on the per-contact risk of transmission (see table S3 for a detailed definition of all risk factors and summary statistics).

On the basis of the point estimates of the regression (see fig. S3A for regression results), we find that household contacts pose the highest risk of transmission followed by extended family, social, and community contacts, in agreement with a prior study (12). Health care contacts have the lowest risk, which suggests that adequate protective measures were adopted by patients and health care staff in Hunan. Notably, the impact of physical distancing differs by contact type (Table 1): The risk of transmission in the household increases during the lockdown period, likely because of increased contact frequency at home as a result of physical confinement. By contrast, the transmission risk decreases for community and social contacts during lockdown, possibly because of the adoption of prudent behaviors such as mask wearing, hand washing, and coughing-sneezing etiquette. We find that longer exposures are riskier, with 1 additional day of exposure increasing the transmission risk by 10% (95% CI: 5 to 15%). Further, transmission risk is higher around the time of symptom presentation of the primary case (Table 1). Additionally, susceptibility to infec-

tion (defined as the risk of infection given contact with a primary case) varies by age: Children aged 0 to 12 years are significantly less susceptible than individuals aged 26 to 64 years (odds ratio 0.41; 95% CI: 0.26 to 0.63), and patients older than 65 years are significantly more susceptible (odds ratio 1.39; 95% CI: 1.02 to 1.91). By contrast, we find no statistical support for age difference in infectivity (fig. S3A). These results are in agreement with previous findings (12, 26, 27).

For each of the 17,750 contact exposure events, we estimate the probability of transmission using the point estimate of the baseline odds and the odds ratios from the GLMM-logit regression (fig. S3A). In Fig. 2A, we plot the distribution of transmission probabilities for household, extended family, social, and community contacts separately. The average per-contact transmission probability is highest for household contacts (7.2%; 95% CI: 1.2 to 19.6%) followed by family (1.7%; 95% CI: 0.4 to 5.6%) and social contacts (0.9%; 95% CI: 0.2 to 2.7%), whereas the risk is lowest for community contacts (0.4%; 95% CI: 0.1 to 1.1%). These transmission probabilities reflect the joint

Table 1. SARS-CoV-2 transmission risk in Hunan by contact type, duration of exposure, and whether the exposure window contains the date of symptom onset of the primary case—a period of intense viral shedding. Risk is further stratified by the date of implementation of social distancing interventions in Hunan, which is 25 January 2020. The regression model is adjusted for demographic characteristics of the cases and their contacts, clinical symptoms, and travel history. Details are provided in the materials and methods, and the full results of the regression, including additional risk factors, are shown in fig. S3.

Risk factors		Odds ratio	95% CI
Household contacts	Before 25 January 2020	2.20***	(1.39, 3.49)
	After 25 January 2020	3.79***	(2.47, 5.79)
Extended family contacts	Before 25 January 2020	1.00	Reference
	After 25 January 2020	0.94	(0.60, 1.46)
Social contacts	Before 25 January 2020	0.63	(0.37, 1.06)
	After 25 January 2020	0.41**	(0.21, 0.78)
Community contacts	Before 25 January 2020	0.37**	(0.19, 0.74)
	After 25 January 2020	0.20*	(0.05, 0.71)
Health care contacts	Before 25 January 2020	0.15*	(0.03, 0.68)
Duration of exposure (days)	After 25 January 2020	0.10*	(0.01, 0.90)
Symptom onset within exposure window (yes)		1.49*	(1.09, 2.04)

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

effect of duration of exposure (Fig. 2B) superimposed on differences in transmission risk by type of contact (fig. S3A). Although confidence intervals on risk estimates are broad, there is statistical support for separating out contacts in five categories and including a time covariate to capture the effect of the lockdown rather than collapsing the contact data into fewer categories (table S4). By contrast, there is no statistical support for a more complex model that considers a different effect of contact duration by type of contact (table S4). It is worth noting that the per-contact transmission probabilities were estimated in a situation of intense interventions and high population awareness of the disease, and thus, they may be not generalizable elsewhere.

The number of contacts is also a key driver of individual transmission potential and varies by contact type. Figure 2C presents the contact degree distribution, defined as the number of distinct contacts per individual. We find that the distributions of individual contact degree are overdispersed with dispersion parameter $0 < k < 1$ across all contact types. Furthermore, household ($k = 0.72$) and extended family ($k = 0.64$) contacts are less dispersed than social ($k = 0.19$) and community ($k = 0.14$) contacts, which suggests that contact heterogeneities are inversely correlated with the closeness of social interactions. Figure 2D visualizes the age-specific contact patterns between the primary cases and their contacts, demonstrating diverse mixing patterns across different types of contact. Specifically, household contacts present the canonical three-bands pattern,

where the diagonal illustrates age-assortative interactions, and the two off-diagonals represent intergenerational mixing (28, 29). Other contact types display more diffusive mixing patterns by age. We also observe that among all primary cases, young and middle-aged adults have the most social contacts (Fig. 2E).

Next, we summarize the overall transmission potential of an individual by calculating the cumulative contact rate (CCR) of all primary cases. The CCR captures how contact opportunities vary with demography, temporal variation in the infectiousness profile, an individual's contact degree, and interventions (see section 4.3 in the materials and methods for detailed definition). After adjusting for age, sex, clinical presentation, and travel history to Wuhan, we find that physical distancing measures increase CCRs for household and extended family contacts and decrease (although not statistically significantly) CCRs for social and community contacts (Fig. 2E). By contrast, faster case isolation universally reduces CCRs, decreasing transmission opportunities across all contact types (Fig. 2E).

Characterizing the natural history of SARS-CoV-2 infection by strength of interventions

We have characterized SARS-CoV-2 transmission risk factors and have shown that individual- and population-based interventions have a differential impact on contact patterns and transmission potential. Next, we use our probabilistic reconstruction of infector-infectee pairs to further dissect transmission kinetics and project the impact of interventions on SARS-CoV-2 dynamics. On the basis of the recon-

structed transmission chains, we estimate a median serial interval of 5.3 days, with an interquartile range (IQR) of 2.7 to 8.3 days, which represents the time interval between symptom onset of an infector and that of his or her infectee (fig. S7, B and D). The median generation interval—defined as the interval between the infection times of an infector and his or her infectee—is 5.3 days, with an IQR of 3.1 to 8.7 days (fig. S7, A and C). We estimate that 63.4% (95% CI: 60.2 to 67.2%) of all transmission events occur before symptom onset, which is comparable to findings from other studies (6–8, 10–13, 18, 30, 31). However, these estimates are affected by the intensity of interventions; in Hunan, isolation and quarantine were in place throughout the epidemic.

Case isolation and contact quarantine are meant to prevent potentially infectious individuals from contacting susceptible individuals, effectively shortening the infectious period. As a result, we would expect right censoring of the generation and serial interval distributions (32). Symptomatic cases represent 86.5% of all SARS-CoV-2 infections in our data; among these patients, we observe longer generation intervals for cases isolated later in the course of their infection (Fig. 3A). The median generation interval increases from 4.0 days (IQR, 1.9 to 7.3 days) for cases isolated 2 days since symptom onset to 7.0 days (IQR, 3.6 to 11.3 days) for those isolated >6 days after symptom onset ($P < 0.001$; Mann-Whitney U test). We observe similar trends for the serial interval distribution (Fig. 3B). The median serial interval increases from 1.7 days (IQR, -1.6 to 4.8 days) for cases isolated <2 days after symptom onset to 7.3 days (IQR, 3.4 to 10.8 days) for those isolated >6 days after symptom onset ($P < 0.001$; Mann-Whitney U test).

Faster case isolation restricts transmission to the earlier stages of infection, thus inflating the contribution of presymptomatic transmission (Fig. 3C). The proportion of presymptomatic transmission is estimated at 87.3% (95% CI: 79.8 to 93.4%) if cases are isolated within 2 days of symptom onset, whereas this proportion decreases to 47.5% (95% CI: 41.4 to 53.3%) if cases are isolated >6 days after symptom onset ($P < 0.001$; Mann-Whitney U test).

Next, we adjust for censoring caused by case isolation and reconstruct the infectiousness profile of a typical SARS-CoV-2 patient in the absence of interventions. To do so, we characterize changes in the timeliness of case isolation over time in Hunan. Figure S8 shows the distributions of time from symptom onset to isolation during three different phases of epidemic control, coinciding with major changes in COVID-19 case definition (phase I: before 27 January; phase II: 27 January to 4 February; and phase III: after 4 February; fig. S3) (33). In phase I, 78% of cases were detected through passive surveillance; as a result, most cases

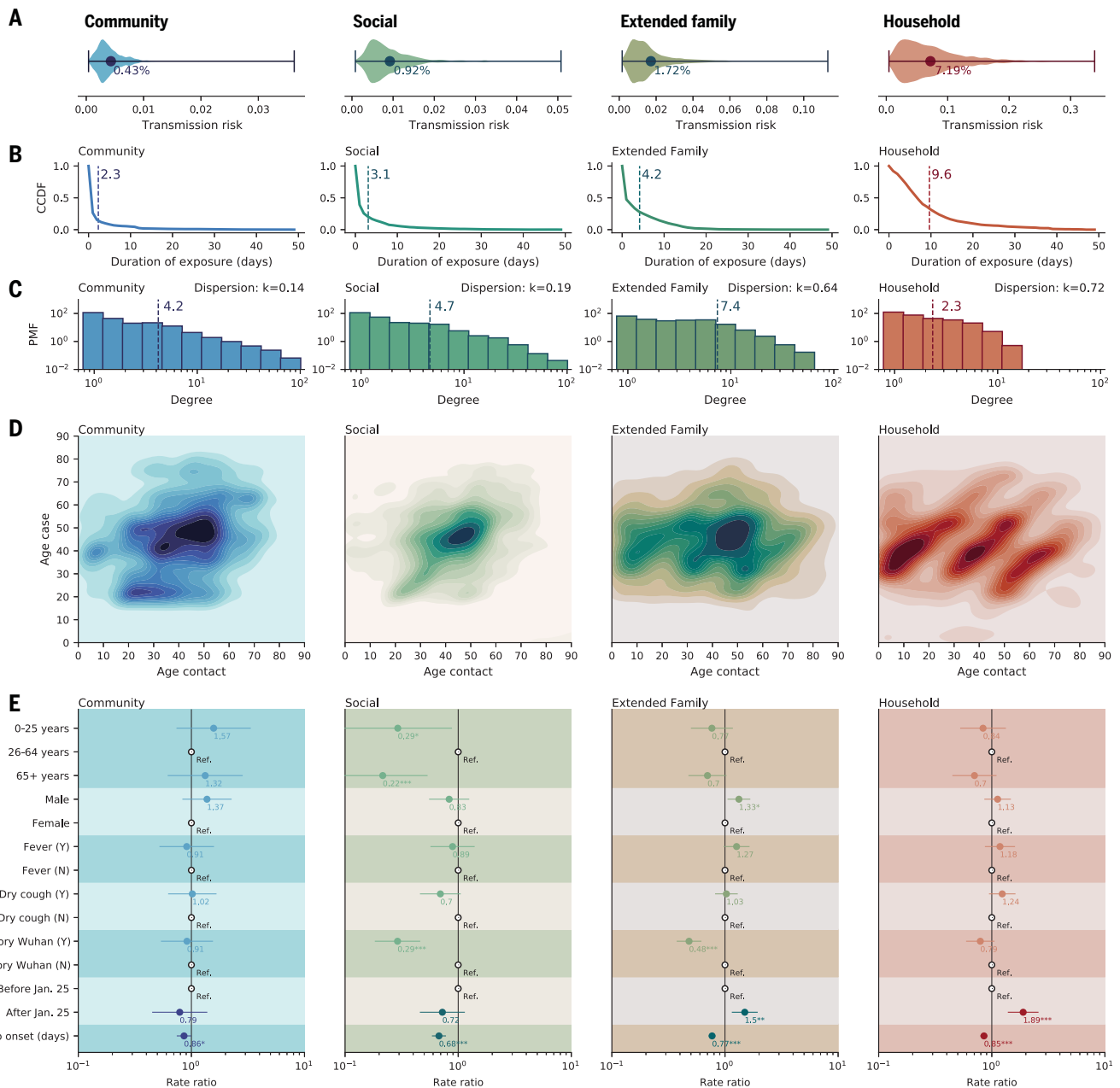


Fig. 2. Heterogeneity in contact rates of SARS-CoV-2 cases and impact of interventions, separated by contact type. Columns from left to right represent community contacts (e.g., public transportation, food, and entertainment), social contacts, extended family contacts, and household contacts. **(A)** Violin plots representing the distribution of per-contact transmission probability by contact type, adjusted for all other covariates in fig. S3 (probability expressed in percentage; x axis). **(B)** Complementary cumulative distribution function (CCDF) (y axis) for duration of exposure (i.e., the probability that exposure is longer or equal to a certain value). Dashed vertical lines indicate average values. Household contacts last the longest, and as expected, contact duration decreases as social ties loosen. **(C)** The distribution of the number of distinct contacts (degree distribution) of the primary cases for each contact type. The y axis indicates probability mass function (PMF). The dashed vertical lines indicate average values. The

dispersion parameter k is calculated on the basis of the relationship $\sigma^2 = \frac{\mu}{1+\mu/k}$, where μ and σ^2 are the mean and variance of the number of distinct contacts. Values of $k < 1$ indicate overdispersion. **(D)** Age distribution of SARS-CoV-2 case-contact pairs (contact matrices). **(E)** Rate ratios of negative binomial regression of the CCRs against predictors including the infector's age, sex, presence of fever or cough, Wuhan travel history, whether symptom onset occurred before social distancing was in place (before or after 25 January 2020), and time from isolation to symptom onset. CCRs represent the sum of relevant contacts over a 1-week window centered at the date of the primary case's symptom onset. Dots and lines indicate point estimates and 95% CIs of the rate ratios, respectively, and numbers below the dots indicate the numerical value of the point estimates. Ref., reference category. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Downloaded from https://www.science.org on October 26, 2021

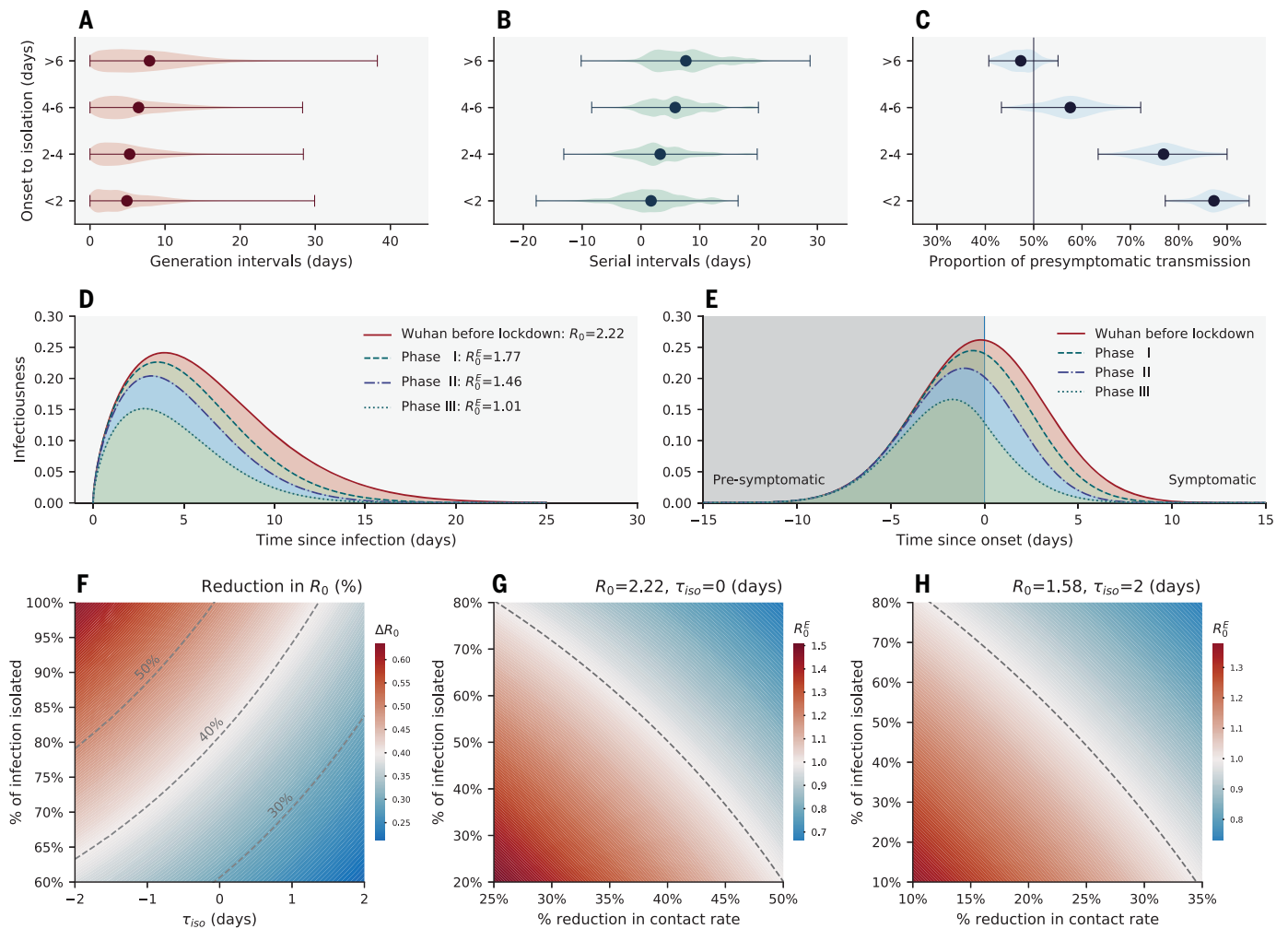


Fig. 3. The impact of interventions on SARS-CoV-2 transmission dynamics.

(A) Violin plot of the generation interval distributions stratified by time from symptom onset to isolation or presymptomatic quarantine, based on an ensemble of 100 realizations of the sampled transmission chains. (B) Same as (A) but for the serial interval distributions. (C) Same as (A) but for the fraction of presymptomatic transmission, among all transmission events, with vertical line indicating 50% of presymptomatic transmissions. In (A) to (C), dots represent the mean, and whiskers represent minimum and maximum. (D) Estimated average (across 100 realizations of sampled transmission chains) transmission risk of a SARS-CoV-2-infected individual since time of infection under four intervention scenarios: the red solid line represents an uncontrolled epidemic scenario modeled after the early epidemic dynamics in Wuhan before lockdown, and the dashed lines represent scenarios where quarantine and case isolation are in place and mimic phases I, II, and III of epidemic control in Hunan. The shapes of these curves match those of the generation interval distributions in each scenario, and the areas under the curve are equal to the ratios of the baseline/effective basic reproduction numbers (R_0/R_0^E). (E) Same as in (D) but with time since symptom onset on the x axis [colors are the same as in (D)].

The vertical line represents symptom onset. (F) Reduction (percentage) in the basic reproduction number as a function of mean time from symptom onset (or from peak infectiousness for asymptomatic cases) to isolation τ_{iso} (x axis) and fraction of SARS-CoV-2 infections being isolated (y axis). The distribution of onset to isolation follows a normal distribution with mean τ_{iso} and standard deviation of 2 days. The dashed lines indicate 30, 40, and 50% reductions in R_0 under interventions. (G) Effective basic reproduction number as a function of population-level reduction in contact rates (i.e., through physical distancing; expressed as a percentage, x axis) and isolation rate (fraction of total infections detected and further isolated). We assume baseline basic reproduction number $R_0 = 2.19$ and a normal distribution for the distribution from onset to isolation with a mean of 0 days and a standard deviation of 2 days. The dashed line represents the epidemic threshold, $R^E = 1$. The blue area indicates the region below the epidemic threshold (namely, controlled epidemic), and the red area indicates the region above the epidemic threshold. (H) Same as in (G) but assuming $R_0 = 1.57$ (a more optimistic estimate of R_0 in Wuhan, adjusted for reporting changes) and a normal distribution for the distribution from onset to isolation with a mean of 2 days and a standard deviation of 2 days.

were isolated after symptom onset [median time from onset to isolation, 5.4 days (IQR, 2.7 to 8.2 days); fig. S8A]. By contrast, in phase III, 66% of cases were detected through active contact tracing, which shortened the median time from onset to isolation to -0.1 days (IQR, -2.9 to 1.8 days; fig. S8C). Timeliness of iso-

lation is intermediate in phase II. We use mathematical models (detailed in the materials and methods) to dynamically adjust the serial interval distribution for censoring, and we apply the same approach to the time interval between symptom onset of a primary case and onward transmission (fig. S10). These

censoring-adjusted distributions can be rescaled by the basic reproduction number R_0 to reflect the risk of transmission of a typical SARS-CoV-2 case since the time of infection or since symptom onset (Fig. 3, D and E). We find that in the absence of interventions, infectiousness near the time of symptom onset (fig. S10D).

This is consistent with our regression analysis, where the higher risk of transmission is near symptom onset (Table 1).

Evaluating the impact of individual- and population-based interventions on SARS-CoV-2 transmission

Next, we use the estimated infectiousness profile of a typical SARS-CoV-2 infection (Fig. 3, D and E) to evaluate the impact of case isolation on transmission. We first set a baseline reproduction number R_0 for SARS-CoV-2 in the absence of control. Results from a recent study (33) suggest that the initial growth rate in Wuhan was 0.15 day^{-1} in raw case data (95% CI: 0.14 to 0.17), although the growth rate could be substantially lower (0.08 day^{-1}) if changes in case definition are considered. Conservatively, we consider the upper value of the growth rate at 0.15 day^{-1} together with our generation interval distribution adjusted for censoring (fig. S10C) to estimate R_0 . We obtain a baseline reproduction number $R_0 = 2.19$ (95% CI: 2.08 to 2.36) using the renewal equation framework (34). This represents a typical scenario of unmitigated SARS-CoV-2 transmissibility in an urban setting. The reconstructed infectiousness profile in the absence of control is shown in solid red lines in Fig. 3, D and E, with respect to time of infection and symptom onset, respectively. Notably, we find that SARS-CoV-2 infectiousness peaks slightly before symptom onset (-0.1 days on average), with 87% of the overall infectiousness concentrated within ± 5 days of symptom onset and 53% of the overall infectiousness in the presymptomatic phase (Fig. 3E).

Next, we evaluate the impact of case isolation on transmission by considering three different intervention scenarios mimicking the timeliness of isolation in the three phases of the Hunan epidemic control. We further assume that 100% of infections are detected and isolated and that isolation is fully protective (i.e., there is no onward transmission after the patient has been isolated or quarantined). The infectiousness profiles of the three intervention scenarios are shown in dashed lines in Fig. 3, D and E. We find that the basic reproduction number decreases in all intervention scenarios, but the projected decrease is not sufficient to interrupt transmission (Fig. 3D; $R_0^E = 1.75$ for phase I, $R_0^E = 1.46$ for phase II, and $R_0^E = 1.01$ for phase III, where R_0^E is the effective basic reproduction number).

We further relax the assumption of 100% case detection and isolation and relate changes in the basic reproduction number to two independent parameters measuring the strength of interventions: the effectiveness of case isolation and contact quarantine (measured as the fraction of total infections isolated) and the timeliness of isolation (measured as the delay from

symptom onset to isolation; phase diagram in Fig. 3F). Dashed lines in Fig. 3F illustrate 30, 40, and 50% of reduction in R_0 . To reduce R_0 by half (the minimum amount of transmission reduction required to achieve control for a baseline $R_0 \sim 2$), 100% of infections would need to be isolated even if individuals were isolated as early as the day of symptom onset. In practice, epidemic control is unrealistic if case isolation and quarantine of close contacts are the only measures in place.

Our data support the idea that case isolation and quarantine of close contacts are effective in reducing SARS-CoV-2 transmission, especially if these interventions occur early in the infection. To achieve epidemic control, however, these interventions need to be layered with additional population-level measures, including increased teleworking, reduced operation in the service industry, or broader adoption of face masks. The synergistic effects of these interventions are illustrated in Fig. 3G. We find that a 30% reduction in transmission from population-level measures would require a 70% case detection rate to achieve epidemic control, assuming that cases can be promptly isolated on average upon symptom presentation. Notably, a 30% reduction in transmission could also encompass the benefits of residual population-level immunity from the first wave of COVID-19, especially in hard-hit regions (35, 36). As a sensitivity analysis, we further consider a more optimistic scenario with a lower baseline $R_0 = 1.56$, corresponding to an epidemic growth rate of 0.08 day^{-1} (95% CI: 0.06 to 0.10) in Wuhan (33), which is adjusted for reporting changes. As expected, control is much easier to achieve in this scenario: If detected SARS-CoV-2 infections are effectively isolated on average 2 days after symptom onset, a 25% population-level reduction in transmission coupled with a 42% infection isolation rate is sufficient to achieve control (Fig. 3H).

Discussion

Detailed information on 1178 SARS-CoV-2-infected individuals along with their 15,648 contacts allowed us to dissect the behavioral and clinical drivers of SARS-CoV-2 transmission, to evaluate how transmission opportunities are modulated by individual- and population-level interventions, and to characterize the typical infectiousness profile of a case. Informed by this understanding, particularly the importance of presymptomatic transmission, we have evaluated the plausibility of SARS-CoV-2 control through individual- and population-based interventions.

Health care contacts posed the lowest risk of transmission in Hunan, which suggests that adequate protective measures against SARS-CoV-2 were taken in hospitals and medical observation centers (Table 1). The average risk

of transmission scales positively with the closeness of social interactions: The average per-contact risk is lowest for community exposures (including contacts in the public transportation system and at food and entertainment venues), intermediate for social and extended family contacts, and highest in the household. The average transmission risk in the household is further elevated when intense physical distancing is enforced, and the risk is also elevated for contacts that last longer. These lines of evidence support the idea that SARS-CoV-2 transmission is facilitated by close proximity, confined environment, and high frequency of contacts.

Regression analysis indicates a higher risk of transmission when an individual is exposed to a SARS-CoV-2 patient around the time of symptom onset, in line with our reconstructed infectiousness profile. These epidemiological findings are in agreement with viral shedding studies (6, 37–40). We estimate that overall in Hunan, 63% of all transmission events were from presymptomatic individuals, in concordance with other modeling studies (6, 7, 10, 12, 41). However, the estimated presymptomatic proportion is affected by case-based measures, including case isolation and contact quarantine. We estimate that the relative contribution of presymptomatic transmission drops to 52% in an uncontrolled scenario where case-based interventions are absent.

Case isolation reduces the effective infectious period of SARS-CoV-2-infected individuals by blocking contacts with susceptible individuals. We observe that faster isolation significantly reduces CCRs across contact types (Fig. 2E). We also observe shorter serial and generation intervals and a larger fraction of presymptomatic transmission when individuals are isolated faster (Fig. 3, A to C). By contrast, population-level physical distancing measures have differential impacts on CCRs—decreasing CCRs for social and community contacts, while increasing CCRs in the household and family contacts. As a result, strict physical distancing confines the epidemic mostly to families and households (see also fig. S7). The precise impact of physical distancing on transmission is difficult to separate from that of individual-based interventions. However, our analysis suggests that physical distancing changes the topology of the transmission network by affecting the number and duration of interactions. Notably, the topology of the household contact network is highly clustered (42), and theoretical studies have shown that high clustering hinders epidemic spread (43, 44). These higher-order topological changes could contribute to reducing transmission beyond the effects expected from an overall reduction in CCRs. In parallel, the effectiveness of physical distancing measures on reducing COVID-19 transmission has been

demonstrated in empirical data from China (24, 45) and elsewhere (46).

We have explored the feasibility of SARS-CoV-2 epidemic control against two important metrics related to case isolation and contact quarantine: the timeliness of isolation and the infection detection rate (Fig. 3F). For a baseline transmission scenario compatible with the initial growth phase of the epidemic in Wuhan, we find that epidemic control solely relying on isolation and quarantine is difficult to achieve. Layering these interventions with moderate physical distancing makes control more likely over a range of plausible parameters—a situation that could be further improved by residual immunity from the first wave of SARS-CoV-2 activity (35, 36). Successful implementation of contact tracing requires a low level of active infections in the community, as the number of contacts to be monitored is several folds the number of infections (~13 contacts were traced for each SARS-CoV-2-infected individual in Hunan). The timing of easing of lockdown measures should align with the capacities of testing and contact-tracing efforts relative to the number of active infections in the community. In parallel, technology-based approaches can also facilitate these efforts (7, 47).

Overall, we find that case isolation and quarantine successfully blocked transmission to close contacts in Hunan, with an estimated 4.3% of transmission occurring after SARS-CoV-2 patients were isolated. In this setting, all SARS-CoV-2 infections were managed under medical isolation in dedicated hospitals regardless of clinical severity, and contacts were quarantined in designated medical observation centers. Self-regulated isolation and quarantine at home, however, may not be as effective, and a higher proportion of onward transmission should be expected.

Several caveats are worth noting. We could not evaluate the risk of transmission in schools, workplaces, conferences, prisons, or factories, as no contacts in these settings were reported in the Hunan dataset. Our study is likely underpowered to assess the transmission potential of asymptomatic individuals given the relatively small fraction of these infections in our data (13.5% overall and 22.1% of infections captured through contact tracing). There is no statistical support for decreased transmission from asymptomatic individuals (fig. S3A), although we observe a positive, but nonsignificant, gradient in average transmission risk with disease severity. Evidence from viral shedding studies is conflicted; viral load appears to be independent of clinical severity in some studies (6, 22, 38, 48), whereas others suggest that there is faster viral clearance in asymptomatic individuals (49).

Another limitation relates to changes in testing practices for contacts of primary cases. Testing was initially limited to contacts exhibit-

ing symptoms, and this condition was relaxed after 7 February. The early testing scheme may lead to underestimation of susceptibility in children, as younger individuals are less likely to develop SARS-CoV-2 symptoms (50). However, sensitivity analyses indicate that the age gradient of susceptibility is preserved even after stratification for changes in testing protocol. Further, our finding of lower susceptibility to infection among children under 12 years of age relative to adults remains stable in the period with comprehensive testing (fig. S4). Overall, the contribution of asymptomatic infections to transmission remains debated but has profound implications on the feasibility of control through individual-based interventions. Careful serological studies combined with virologic testing in households and other controlled environments are needed to fully resolve the role of asymptomatic infections and viral shedding on transmission.

Detailed contact-tracing data illuminate heterogeneities in SARS-CoV-2 transmission driven by biology and behavior and modulated by the impact of interventions. Notably, and in contrast to SARS-CoV-1, the ability of SARS-CoV-2 to transmit during the host's presymptomatic phase makes it particularly difficult to achieve epidemic control (51). Our risk factor estimates can provide evidence to guide the design of more-targeted and sustainable mitigation strategies, and our reconstructed transmission kinetics will help calibrate further modeling efforts.

Materials and methods summary

We combined individual-level data on 1178 SARS-CoV-2 infections with detailed diaries of exposures collected through contact-tracing efforts in Hunan, China, to stochastically reconstruct transmission chains and infer infection times. Reconstructed transmission chains had to be compatible with highly resolved individual-level data on symptom onset dates, daily records of exposure to infected contacts, and travel history to high-risk regions. On the basis of the reconstructed transmission chains, we characterize the distribution of key SARS-CoV-2 transmission parameters—including the number of secondary cases, the generation and serial intervals, and the interval from infection or symptom onset to isolation—at different stages of the epidemic. To further understand the drivers of transmission heterogeneity and the dispersion in the number of secondary cases, we study the degree distribution of SARS-CoV-2-infected individuals, the duration of exposures, and the age-specific contact patterns between infectors and infectees, separately by contact type (household, family, community, transportation, and health care). We also use logistic regression analysis to model the per-contact risk of transmission, with contact type and duration,

symptoms, demographic factors, and different periods of the outbreak as covariates. Missing data are addressed through multivariate imputation algorithms. We conduct sensitivity analyses to test the robustness of regression results.

We next use our data to model the synergistic effects of case-based and population-level interventions on transmission. We reconstruct the average infectiousness profile of a SARS-CoV-2 infection, after adjusting for the truncation effects of case isolation. On the basis of the estimated infectiousness profile, we use mathematical models to estimate the effect of layered interventions on transmission (measured as changes in the effective reproduction number). We consider different intensities of population-level physical distancing, case detection, and timeliness of isolation or quarantine. A full description of the materials and methods is provided in the supplementary materials.

REFERENCES AND NOTES

1. W. J. Guan *et al.*, 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020). doi: [10.1056/NEJMoa2002032](https://doi.org/10.1056/NEJMoa2002032); pmid: [32109013](https://pubmed.ncbi.nlm.nih.gov/32109013/)
2. J. T. Wu *et al.*, Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat. Med.* **26**, 506–510 (2020). doi: [10.1038/s41591-020-0822-7](https://doi.org/10.1038/s41591-020-0822-7); pmid: [32284616](https://pubmed.ncbi.nlm.nih.gov/32284616/)
3. G. Grasselli *et al.*, Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. *JAMA* **323**, 1574–1581 (2020). doi: [10.1001/jama.2020.5394](https://doi.org/10.1001/jama.2020.5394); pmid: [32250385](https://pubmed.ncbi.nlm.nih.gov/32250385/)
4. R. Verity *et al.*, Estimates of the severity of coronavirus disease 2019: A model-based analysis. *Lancet Infect. Dis.* **20**, 669–677 (2020). doi: [10.1016/S1473-3099\(20\)30243-7](https://doi.org/10.1016/S1473-3099(20)30243-7); pmid: [32240634](https://pubmed.ncbi.nlm.nih.gov/32240634/)
5. K. Sun, J. Chen, C. Viboud, Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: A population-level observational study. *Lancet Digit. Heal.* **2**, e201–e208 (2020). doi: [10.1016/S2589-7500\(20\)30026-1](https://doi.org/10.1016/S2589-7500(20)30026-1); pmid: [32309796](https://pubmed.ncbi.nlm.nih.gov/32309796/)
6. X. He *et al.*, Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020). doi: [10.1038/s41591-020-0869-5](https://doi.org/10.1038/s41591-020-0869-5); pmid: [32296168](https://pubmed.ncbi.nlm.nih.gov/32296168/)
7. L. Ferretti *et al.*, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936 (2020). doi: [10.1126/science.abb6936](https://doi.org/10.1126/science.abb6936); pmid: [32234805](https://pubmed.ncbi.nlm.nih.gov/32234805/)
8. A. Kimball *et al.*, Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility - King County, Washington, March 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 377–381 (2020). doi: [10.15585/mmwr.mm6913e1](https://doi.org/10.15585/mmwr.mm6913e1); pmid: [32240128](https://pubmed.ncbi.nlm.nih.gov/32240128/)
9. W. E. Wei *et al.*, Presymptomatic Transmission of SARS-CoV-2 - Singapore, January 23-March 16, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 411–415 (2020). doi: [10.15585/mmwr.mm6914e1](https://doi.org/10.15585/mmwr.mm6914e1); pmid: [32271722](https://pubmed.ncbi.nlm.nih.gov/32271722/)
10. Z. Du *et al.*, Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. *Emerg. Infect. Dis.* **26**, 1341–1343 (2020). doi: [10.3201/eid2606.200357](https://doi.org/10.3201/eid2606.200357); pmid: [32191173](https://pubmed.ncbi.nlm.nih.gov/32191173/)
11. E. Lavezzo *et al.*, Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature* **584**, 425–429 (2020). doi: [10.1038/s41586-020-2488-1](https://doi.org/10.1038/s41586-020-2488-1); pmid: [32604404](https://pubmed.ncbi.nlm.nih.gov/32604404/)
12. S. Hu, W. Wang, Y. Wang, M. Litvinova, K. Luo, L. Ren, Q. Sun, X. Chen, G. Zeng, J. Li, L. Liang, Z. Deng, W. Zheng, M. Li, H. Yang, J. Guo, K. Wang, X. Chen, Z. Liu, H. Yan, H. Shi, Z. Chen, Y. Zhou, K. Sun, A. Vespignani, C. Viboud, L. Gao, M. Ajelli, H. Yu, Infectivity, susceptibility, and risk factors associated with SARS-CoV-2 transmission under intensive contact tracing in Hunan, China. *medRxiv* 2020.07.23.20160317 [Preprint]. 3 November 2020. . doi: [10.1101/2020.07.23.20160317](https://doi.org/10.1101/2020.07.23.20160317)
13. Q. Bi *et al.*, Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: A retrospective cohort study. *Lancet Infect. Dis.* **20**, 911–919

Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2

Kaiyuan SunWei WangLidong GaoYan WangKaiwei LuoLingshuang RenZhifei ZhanXinghui ChenShanlu ZhaoYiwei HuangQianlai SunZiyan LiuMaria LitvinovaAlessandro VespignaniMarco AjelliCécile ViboudHongjie Yu

Science, 371 (6526), eabe2424. • DOI: 10.1126/science.abe2424

Time and intimacy drive transmission

A minority of people infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmit most infections. How does this happen? Sun *et al.* reconstructed transmission in Hunan, China, up to April 2020. Such detailed data can be used to separate out the relative contribution of transmission control measures aimed at isolating individuals relative to population-level distancing measures. The authors found that most of the secondary transmissions could be traced back to a minority of infected individuals, and well over half of transmission occurred in the presymptomatic phase. Furthermore, the duration of exposure to an infected person combined with closeness and number of household contacts constituted the greatest risks for transmission, particularly when lockdown conditions prevailed. These findings could help in the design of infection control policies that have the potential to minimize both virus transmission and economic strain.

Science, this issue p. eabe2424

View the article online

<https://www.science.org/doi/10.1126/science.abe2424>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of think article is subject to the [Terms of service](#)