Household Transmission of SARS-CoV-2 along the Evolution of Pandemic

Yuan Zhang  
Renmin University of China

Chong You  
Peking University

Xin Gai  
Peking University

Hao Zhang  
Yichang Municipal Center for Disease Control and Prevention

Michael Lydeamore  
Monash University

Peng Yin  
Chinese Center For Disease Control and Prevention

Jinlei Qi  
Chinese Center For Disease Control and Prevention

Maigeng Zhou  
Chinese Center For Disease Control and Prevention

Jianhua Liu  
Yichang Municipal Center for Disease Control and Prevention

Xiao-Hua Zhou (✉ azhou@math.pku.edu.cn)  
Peking University

Article

Keywords: COVID-19, Household Transmission, Erdös–Rényi Random Graph, Dynamic Model, Second Attack Rate, Effective Reproductive Number

Posted Date: May 3rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2768212/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Over the past three years, we have gained some understanding of the transmission mechanisms of COVID-19. One of the key findings that experts have agreed on is that household transmission is an important pathway for the pandemic. However, most studies on the transmission patterns of COVID-19 focus on the community transmission only, while the equally important study on household transmission has lagged behind. We developed a stochastic dynamic model motivated by the cluster growth algorithm in Erdös–Rényi Random Graph to differentiate the COVID-19 transmission within households from that in the community by noting only a small fraction of the total susceptible population, replenished dynamically by the community transmission events, are indeed vulnerable to household transmission. Thus the model allowed us to the role and characteristics of household transmission within the full framework of virus transmission, beyond the intrinsic characteristics of household transmission. It was then applied to a comprehensive individual-level pandemic dataset collected in Yichang, China. Our findings showed that household transmission accounted for 25.1% and 38.5% of total infections before and during the lockdown, respectively, and that 80.9% of infections were unavoidable. Our model suggests that household-level contact tracing could have reduced the number of infections by over 50% and advanced the clearance date of active infection by 72 days. This model can be used to fit COVID-19 data outside Yichang or other infectious diseases, though modifications might be needed.

Introduction

The COVID-19 pandemic has severely impacted the lives of many since the first reported case back in December 2019. Although mass vaccination programs have been introduced in many parts of the world, there is no clear sign that the COVID-19 could be fully controlled by vaccination alone. This pandemic poses grievous challenges for both the economy and public health, and it is thus important to understand the dynamics of the COVID-19 transmission mechanism and its potential impact to develop more effective prevention strategies and policies. In particular, the intra-household transmission of COVID-19 has received increasing attention due to its significant role in virus spread, particularly when strict public health and social measures were imposed. As reported by the WHO-China Joint Mission, the transmission of the SARS-CoV-2 virus mainly occurred in families in China\(^1\), highlighting the importance of studying the household transmission not just for COVID-19 but future other infectious diseases. While existing studies on the household transmission of COVID-19 have focused on second attack rate (SAR), risk factors for susceptibility and infectivity, and household reproductive number based on either observed estimates\(^1\)–\(^4\), regression models\(^5,6\), chain-binomial transmission models\(^7,8\), or meta-analysis\(^3,9\), the investigation of the dynamic role of household transmission along the evolution of pandemic is still limited in its scope and validity. Such limitations are mainly due to difficulties in collecting complete and individual-level data and identifying transmission chains within a family, which are required to be known in most of the aforementioned studies. Note that the majority of current household transmission studies assumed that all secondary cases were infected by the index case in a household, while the index case
was set to be the case who first developed symptoms or first tested positive in a household, which might not hold true in reality.

To address the limitations of the previous studies, we propose a novel stochastic dynamic model, called the ER-Household model. This model draws inspiration from the cluster growth algorithm in Erdős–Rényi Random Graph\(^{10}\) to adequately depict household transmission within the full transmission mechanism by noticing that the transmission in a household would end in a short period due to the limited size of household members. This model allows us to study the role and characteristics of household transmission within the full framework of virus transmission, rather than focusing solely on the household transmission like the existing studies were doing. The model was applied to a comprehensive individual-level pandemic dataset collected in Yichang, Hubei, China, where a complete episode of an outbreak was well recorded from 23 January to 27 February 2020 (diagnosis date). We estimated the community transmission rate, the household transmission rates for different household sizes and the impacts of interventions on both community and household transmission which were inferred directly from the proposed dynamic model. Based on the inferred dynamic model, we further assessed the time-varying contribution of household transmission to the pandemic, SAR, as well as the overall and household effective reproductive numbers in Yichang. The effectiveness of the community-level containment measure was also accessed. To the best of our knowledge, this study is the first to investigate household transmission based on a stochastic dynamic model that incorporate the unique features of COVID-19 and containment measures in China.

**Method**

**Data and definitions**

The epidemiological investigation was performed by Yichang Centers for Disease Control and Prevention (CDC), China. The data covered all confirmed cases in Yichang city (including all 11 designated hospitals for COVID-19) and diagnosed between 23 January 2020 and 27 February 2020. During this time, various containment measures were implemented under the Level I public health emergency response, includin (1) the strict entry and exit screening and traffic control implemented on 24 January 2020; (2) lockdown from 25 January; (3) mask-wearing policy; (4) epidemiological investigation and contact tracing; and (4) awareness and campaign.

A three-level audit system (county-, city- and provincial-level audit) was used to ensure the efficiency and accuracy of data collection. For each confirmed case, their demographic information, dates of symptoms onset and diagnosis, clinical classifications, laboratory test results, exposure history in the 14 days before symptom onset, hospital admission, progression and prognosis were recorded; see Table 1 for more details. Different editions of the National Health Commission of China Guidelines for Diagnosis and Management of COVID-19 were used at different stages; especially, clinical diagnosis was used during the implementation of the fourth edition. However, this would not affect our model as the related model parameter \(D_q\) was calculated directly from our dataset (see Supplement A for detailed explanation for
This data recorded a complete episode of an outbreak in Yichang without overwhelmed medical resources and hence was a good reflection of the true situation of the epidemic in Yichang at the time. This study was approved by the ethics committee of the Yichang CDC, and de-identified data were used for the analysis. Note that, Yichang CDC, the provider of the data involved in this research, has confirmed that the data used in this study was collected during the contact tracing campaign in 2020. According to the Law of the People's Republic of China on Prevention and Control of Infectious Diseases and the Regulations on Public Health Emergencies in China, the reporting of such information is a statutory obligation, an informed consent is not applicable during the contact tracing process.

Table 1
Demographic and clinical characteristics of confirmed cases in Yichang, China, who were diagnosed between 23 January 2020 and 27 February 2020. The clinical classification here represents the most severe clinical classification ever achieved during the disease course.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>4/269(2%)</td>
<td>19/662(3%)</td>
</tr>
<tr>
<td>20–59</td>
<td>194/269(72%)</td>
<td>398/662(60%)</td>
</tr>
<tr>
<td>≥60</td>
<td>71/269(26%)</td>
<td>245/662(37%)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>122/269(45%)</td>
<td>329/662(50%)</td>
</tr>
<tr>
<td>Male</td>
<td>147/269(55%)</td>
<td>333/662(50%)</td>
</tr>
<tr>
<td>Household size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One household member (single household)</td>
<td>8/269(3%)</td>
<td>36/662(5%)</td>
</tr>
<tr>
<td>2–3 household members (small household)</td>
<td>142/269(53%)</td>
<td>366/662(55%)</td>
</tr>
<tr>
<td>≥4 household members (large household)</td>
<td>119/269(44%)</td>
<td>260/662(40%)</td>
</tr>
<tr>
<td>Clinical classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>25/269(9%)</td>
<td>130/662(20%)</td>
</tr>
<tr>
<td>Ordinary</td>
<td>191/269(71%)</td>
<td>386/662(58%)</td>
</tr>
<tr>
<td>Severe</td>
<td>33/269(12%)</td>
<td>102/662(15%)</td>
</tr>
<tr>
<td>Critical</td>
<td>20/269(8%)</td>
<td>44/662(7%)</td>
</tr>
<tr>
<td>Death rate</td>
<td>9/269(3%)</td>
<td>27/662(4%)</td>
</tr>
</tbody>
</table>
**Model**

We modelled the outbreak in Yichang across two periods in accordance with the lockdown policy implemented in Yichang from 25 January 2020, that is, 11 January – 24 January (before lockdown) and 25 January – 24 February (after lockdown). We assumed a constant population size of 4 million in Yichang\(^\text{11}\). Note that we did not include the data from 4 January – 10 January due to high possibility of recall bias in the early period. The full dynamics of COVID-19 in Yichang was modelled with consideration of the transmission capacity in different households, and more importantly, the fact that for a given household, household transmission is only possible when there is at least one susceptible individual and one infectious individual within the household. By modelling households of different size, we could incorporate the fact that a household may quickly deplete its susceptible members, and the resulting spread of disease within the household would rapidly “burn out”, giving rise to a household-level “herd immunity”. Inspired by the celebrated Erdös–Rényi (ER) Random Graph and its cluster growth algorithm\(^\text{10}\), we proposed the so-called ER-household model, where only a small fraction of the total susceptible population, replenished dynamically by the community transmission events, are indeed vulnerable to household transmission. We then fitted the data of symptoms onset stratified by the household size: single, small household (2–3 household members) and large household (\(\geq 4\) household members). Note that such grouping in household size is consistent with other studies, such as Madewell (2020)\(^\text{9}\).

The proposed model includes 21 compartments:

- \(S_{(\cdot)}\) : Susceptible population for community transmission only and from a particular household category specified in parentheses (note that such household category specification applies to all compartments below)
- \(S^\text{act}_{(\cdot)}\) : Susceptible population for household transmission only
- \(E_{(\cdot)}\) : Exposed population
- \(P_{(\cdot)}\) : Pre-symptomatic population
- \(A_{(\cdot)}\) : Asymptomatic/mild population
- \(AA_{(\cdot)}\) : Asymptomatic/mild population but ascertainable due to extra awareness caused by, for example, contact and tracing, diagnosis or symptoms development of their co-inhabitants
- \(I_{(\cdot)}\) : Symptomatic population

where

\[
(\cdot) = \begin{cases} 
(sing) & \text{if household size } = 1 \\
(smalla) & \text{if household size is between 2 and 3} \\
(larga) & \text{if household size is equal to or larger than 4.}
\end{cases}
\]
Note the compartments $S_{*}^{act}$ and $AA_{*}$ do NOT apply to single households as no household transmission is allowed in a single household. See Fig. 1 for illustration of the ER-household model. The dynamics of these compartments are described in Supplement A. It is worth to point out that $S_{*}^{act}$ was introduced to model the potential situation where susceptible members within household might be quickly depleted. It was motivated by ER Random Graph aforementioned, that is, only when a new community transmission occurred, the population in $S_{*}^{act}$ would increase by the corresponding household size subtracting one, see Fig. 1B. All parameters of interests in the dynamic were estimated by Markov Chain Monte Carlo (see Supplement E for details) and further converted into the contribution of household transmission to the pandemic, SAR, the community and household effective reproductive numbers.

**Parameter settings and initial states**

Table S1 in Supplement B provides a list of parameter settings in the proposed ER-household model. The initial populations in the compartments were assigned as follows.

1. The initial populations in $AA_{(small)}$ and $AA_{(large)}$ were set to 0 as no contact tracing was implemented in the very early stage of the outbreak.

2. Considering the small number of cases reported from single households, the initial population of $E_{(single)}$, $P_{(single)}$, $I_{(single)}$ and $A_{(single)}$ should be small, and were set to be 1, 1, 1 and 0 respectively.

3. To our knowledge, there was no shortage of medical resources in Yichang during the study period, hence the initial population of $A_{(small)}$ and $A_{(large)}$ should be small and were set to be one here.

4. The number of confirmed cases was 5 as of 10 January 2020, and more than 3 new cases were reported every day for most of the following 10 days. As a result, $E_{(small)}$, $P_{(small)}$, $I_{(small)}$, $E_{(large)}$, $P_{(large)}$ and $I_{(large)}$ should not be too small and were set to be 9.

5. Considering that the first day in the model is 11 January 2020 which was the early stage of the pandemic, the initial populations of $S_{(single)}$, $S_{(small)}$ and $S_{(large)}$ were set to be the population sizes in Yichang from single, small and large household respectively. Based on the percentage of household sizes in Hubei Province (Yichang is a city in Hubei) and population size in Yichang from the sample survey of population in Hubei, the initial populations of the aforementioned compartments are 240,000, 1,840,000 and 1,920,000 respectively.

See sensitivity analyses in Supplement C for different choices of initial values and prefixed parameters.

**Estimation of SAR, effective reproductive numbers and contribution of household transmission**
The SAR was estimated as the proportion of all household infections excluding their primary cases to the total number of susceptible in households with at least one infection. Suppose there were accumulatively $n$ households with at least one active virus carrier, and then excluding primary cases, the total number of susceptible to household transmission was $N = \sum_{i=1}^{n} S_i - n$, where $S_i$ is the size of the $i$th household. If $K$ individuals were finally infected through household transmission, the SAR can be estimated through $K/N$. Note that the compartmental nature of the proposed model had limited the estimation of SAR to its averaged version. This number can be more directly and accurately calculated in an individual based dynamic model.

We denoted $R_C(t)$ as the community effective reproductive number at time $t$, that is, for an active virus carrier who got exposed at time $t$, the average number of secondary cases outside the household he/she passed the pathogen to throughout his/her entire infectious period. The household effective reproductive number at time $t$, $R_H(t)$, is defined similarly to $R_C(t)$ but the transmission was restricted within the household. The overall effective reproductive number is the sum of $R_C(t)$ and $R_H(t)$. See Supplement F for more detailed formulae.

The contribution of household transmission over the two study periods was calculated as the ratio of the household infections to the total infections in the first and second period respectively.

**RESULTS**

From 23 January 2020 to 27 February 2020, a total of 931 COVID-19 cases were reported in Yichang, among which 44 cases were living alone, 508 were from small households (2–3 household members) and 379 were from large household (≥ 4 household members), see Method section for more detailed demographic information. Combined with the 7th national census data and the sample survey of Hubei province\textsuperscript{11,12}, different incident rates were detected among different household sizes ($\chi^2 = 25.99, P < 5 \times 10^{-6}$), see Table 2. Based on the proposed ER-household model, we hence modelled the complete episode of the pandemic using the data stratified by three household categories aforementioned. Our model fit the observed data well (Fig. 2) and estimated the total number of all infections, including asymptomatic and unascertained cases, to be 1296 (95% Bayesian credible interval:1037-1847, unless specified otherwise, all parenthetical ranges refer to the 95% Bayesian credible interval hereafter) up to 24 February 2020. The number of active infections peaked at 840 (658–1202) on 24 January and dropped to 32 (10–72) on 24 February. The clearance of all active infections would occur on 27 March 2020 (6 March to 22 May). The estimates of all parameters of interest is presented in Table S4 in Supplement C, and the simulations on assessing the effects of non-pharmaceutical interventions (NPIs) and vaccines, as well as on new virus strains are in Supplement D. Characteristics of the epidemic were estimated and detailed as follows.
Table 2

Incident rates among different household sizes. Population sizes for the different household
categories were calculated based on the 7th national census data in Yichang (performed in
2020) and the sample survey of Hubei province (performed in 2019). Note that the total
population size in Yichang was obtained from the 7th national census data, while the
population proportions of single, small and large household were obtained from sample survey
of Hubei province in 2019. It is assumed that the population proportions of single, small and
large household in Yichang were consistent to that of Hubei as Yichang is a city in Hubei Province.

<table>
<thead>
<tr>
<th></th>
<th>Single Household</th>
<th>Small Household</th>
<th>Large Household</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population Size</td>
<td>238,653</td>
<td>1,854,814</td>
<td>1,924,140</td>
<td>4,017,607</td>
</tr>
<tr>
<td>Incident Size</td>
<td>44</td>
<td>508</td>
<td>379</td>
<td>931</td>
</tr>
<tr>
<td>Incident Rate</td>
<td>1.84‰</td>
<td>2.74‰</td>
<td>1.97‰</td>
<td>2.32‰</td>
</tr>
</tbody>
</table>

**Contribution of household transmission**

We confirmed that household transmission played a substantial role in the pandemic. Household
transmission accounted for 25.1% (12.9%-41.4%) of the total infected population before the lockdown,
and 38.5% (22.7%-54.4%) after.

**Household SAR**

The household secondary attack rate (SAR) is an important metric of household transmission, defined as
the probability that an infected case transmits the pathogen to any susceptible household member during
infectious period. The estimated household SAR based on the ER-household model was 14.4%
(8.1%-24.7%) in Yichang, which is well consistent with other studies, such as, 16.6% in a meta-analysis\(^9\),
15.6% (15.2–16.0%) in the first wave of the pandemic at Wuhan\(^8\) and 17.1% (13.3–21.8%) in
Guangzhou\(^7\).

**Transmission rate**

After the implementation of the lockdown policy on 25 January 2020, the community transmission rate
decreased significantly from 0.49 (0.39–0.55) to 0.04 (0.03–0.06), corresponding to a 91.84% reduction.
With the assumption of constant household transmission rates before and after lockdown, the
transmission rates within small and large households were 0.4 (0.18–0.69) and 0.17 (0.03–0.33)
respectively. This is to say the transmission rate within large households was not even half that of small
households, which seems counterintuitive, but the same conclusion was made in House (2021)\(^6\) and
Madewell (2020)\(^9\).
Effective reproductive numbers

We estimated the time-varying effective reproductive numbers ($R$) over the period from 4 January to 24 February 2020 (see Fig. 3). The effective reproductive number can be split into two components: effective household reproductive number ($R_H$) and effective community reproductive number ($R_C$). The estimated $R_C$ in Yichang decreased from 3.43 (2.75–3.89) on 11 January 2020 to 2.85 (2.28–3.26) on 24 January 2020 which was the day before the lockdown, then with a dramatic reduction of 91.2% to 0.250 (0.17–0.32) on the first day of lockdown on 25 January. It then went relatively stable in the following 30 days. The estimated $R_H$ had a similar pattern to $R_C$ but with a less apparent descent on the day of lockdown (large-scale contact tracing was also implemented together with lockdown). It decreased from 1.14(0.63–1.79) on 11 January 2020 to 0.90 (0.49–1.42) on 24 January 2020, then dropped substantially to 0.16 (0.06–0.27) on 25 January, an 82.2% reduction. Note that the sum of $R_H$ and $R_C$ reproduces the standard overall effective reproductive number.

Sensitivity and robustness analyses

We performed a series of sensitivity analyses with different choices of initial values and parameter settings. For sensitivity analyses (S1) – (S15), we assessed the model sensitivity on the pre-fixed parameters of the latent period and the ratio of the transmission rate of presymptomatic/asymptomatic cases to that of symptomatic cases. Analyses (S16) – (S23) were to assess the robustness of the model estimates due to the different choices of the initial values of compartments. We finally assessed the effect of recall error on the robustness of the model by fitting a new data generated through randomly adding or subtracting the onset date by one day for all cases (except for those cases whose onset date was on the same day as the diagnosis date). From Table S4, we can see that our estimates were relatively robust to the different choice of prefixed parameters and initial values. See Supplement C for more details.

DISCUSSION

Based on our proposed ER-household model, we found that the total number of infections in Yichang was 1296 (1037–1847) by 24 February 2020, the number of active infections peaked at 840 (658–1202) on 24 January and the clearance of all active infections occurred on 27 March 2020 (6 March to 22 May). The community transmission rate decreased significantly from 0.49 (0.39–0.55) to 0.04 (0.03–0.06) corresponding to a 91.84% reduction after the implementation of strict lockdown measures from 25 January 2020, while the transmission rates within small and large households were 0.4 (0.18–0.69) and 0.17 (0.03–0.33) respectively. It is worth to note that the transmission rate within large households was not even half that of small households. This might be because, in China, small households are more likely to be composed of nuclear families where members are more tightly knitted comparing to large households, and larger households might have more space for infected household members to quarantine. The household and community effective reproductive numbers were estimated over time and
declined significantly before and after the implementation of containment measures. We assessed that the household transmission accounted for 25.1% (12.9%-41.4%) and 38.5% (22.7%-54.4%) of the total infections before and after the lockdown. Note that most of the current studies on household transmission were based on household data only and were not able to estimate the contribution of the household transmission in the evolution of the pandemic. Here the increase in this percentage was due to the overall decrease in non-household transmission due to the containment measures, rather than being an actual increase in the amount of household transmission. Our household SAR estimate of 14.4% (8.1%-24.7%) in Yichang was consistent with other studies implying that, comparing to community transmission, an active infection within household might impose a significantly higher risk of subsequent transmission.

In particular, the model suggested 80.9% of infections were inevitable even if the containment measures imposed in Yichang from 24 January 2020 would cut all routes of community transmission and could only advance the clearance of all active infections by 12 days, which implies that the community-level containment measures in Yichang at the time was sufficiently effective, and further escalation of community-level control measures might be unnecessary. Furthermore, if additional household-level containment measures, such as mask wearing and social distancing inside households, were adopted on top of community-level containment measures, namely, the household transmission rate declines as much as community transmission rate in the second period (an average reduction of ~ 90%), the total infections could only be reduced by 6.02% up to 24 February 2020. Our model suggested that the household-level contact tracing did reduce the size of pandemic significantly by more than 50% up to 24 February 2020 and advanced the clearance date of active infection by 72 days. This provided evidence supporting that prompt contact tracing followed by quarantine at designated (non-household) sites is one of the most effective NPIs against household transmission. This model also suggested that widespread vaccination may prolong the pandemic, although in doing so reduce the final attack rate. This is because the vaccine for COVID-19 is more effective at preventing symptoms than preventing infections. Therefore, under the protection of vaccines, infected cases are much more likely to be asymptomatic and thus would go undetected and not quarantined, leading to further transmission. However, the introduction of vaccines overall is a clear positive, with reduced overall infection numbers, and although not modelled, a far lower number of individuals experiencing severe outcomes.

We acknowledge that our proposed model has certain limitations. First, we assumed that only one primary case in each household was induced by community transmission, while in reality it is possible that a couple in one household were infected at same time by their common friend for example. However, by recognizing the particularly low attack rate of 0.025% in Yichang, such possibility is unlikely in Yichang. Second, we did not consider population movement at Yichang before the lockdown (25 February 2020) in the model, which might lead to an overestimated reproductive number in the first period. Nonetheless, it is worth to note, unlike Wuhan, Yichang is not a transportation hub, population mobility is relatively small and hence impact of this mobility on the model is likely to be small. Third, the extended family was not considered in this model due to unavailability of such information. Fourth, reinfection due
to waning immunity and newly emerged variants was not considered in our model, but given the short period of this particular episode of the pandemic in Yichang, reinfection should be negligible. Lastly, our model did not account for individual variation in infectiousness beyond household size, such as age, socioeconomic status, geographical region and superspreading events. An individualized model is needed to incorporate such individual variation, and this is our planned future work in the next step.

Conclusion

In summary, our proposed ER-household model investigated the dynamic role of household transmission along the evolution of the pandemic based a well-recorded complete episode of an outbreak in Yichang, China. This model could effectively differentiate transmission within the household from that in the community due to the limited size of household members, and hence reveal the significant role of (1) community-level control measure in reducing the pandemic size, (2) household transmission to the total infections, and (3) contact tracing policy and the resulting quarantine at assembly sites against household transmission. Our findings would suggest that additional attention for co-inhabitants of high exposure risk workers, such as, healthcare delivery and support staffs, public service staffs, is needed due to the high SAR of COVID-19 and prompt isolation at assembly sites is essential following effective contact tracing. This suggestion should be also extended to any household-like scene, such as a work place, a school. The ER-household model can be used to fit COVID-19 data outside Yichang or other infectious diseases, though modifications might be needed.

Declarations

Ethics approval and consent to participate

This study was approved by the ethics committee of the Yichang Municipal Center for Disease Control and Prevention (CDC), and de-identified data was used for the analysis. This study does not involve any experiment. The manuscript focused on developing a new mathematical model to infer and evaluate the role played by household transmission during an epidemic. The model was then validated with meta-population data collected by local authority in their campaign of controlling the COVID epidemic. Yichang CDC, the provider of the data involved in this research, confirmed that the data used in this study was collected during the contact tracing campaign in 2020. According to relative laws in China, the reporting of such information is a statutory obligation. Hence, as a result, an informed consent is not applicable during the contact tracing process.

Authors' contributions

YZ: study design, data analysis, results interpretation and writing.

CY: study design, results interpretation and writing.

XG: study design, data analysis and results interpretation.
ML: results interpretation and writing.

HZ, PY, JQ, MZ and JL: data collection.

XHZ: overall study design.

Availability of data and materials

The datasets analysed during the current study are not publicly since the original dataset includes family information that might have the potential to track specific patients, but are available from the corresponding author on reasonable request and agreement.

Competing interests

The authors declare that they have no competing interests.

Funding

This research is supported by National Natural Science Foundation of China grant 82041023 and the Bill & Melinda Gates Foundation (INV-016826). Both of the funds are to support the analysis, interpretation of data and the manuscript writing.

Supporting information captions

Further information on method and results are available in the Supplementary Information for “Household Transmission of SARS-CoV-2 along the Evolution of Pandemic”.

References


Figures
Figure 1

Illustration of the ER-household model. (A) Relationship between different compartments. (B) Demonstration for a household transmission. Step 1: a household member got infected through community transmission (we assume that only one member in a household can be infected through community transmission). Step 2: the rest of the household member would be then susceptible for household transmission, that is, the population in $S_{act}^{(small)}$ would be increased by two for a household of
three. Step 3: (i) if contact tracing was NOT implemented, all household members might be eventually infected; (ii) if contact tracing was implemented, they could be isolated at hospitals or other facilities before all were infected.

Figure 2
Fitting of the number of the reported cases by onset date in Yichang and fitting in single, small and large households, the shaded area is the 95% credible intervals, and the colored points are the mean values based on 200,000 MCMC samples.

**Figure 3**

Estimation of $R_H$, $R_C$ and $R$ based on the ER-household model, starting from 11 January 2020.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- draftSRsup.docx