The role of children in the spread of COVID-19: Using household data from Bnei Brak, Israel, to estimate the relative susceptibility and infectivity of children

Itai Dattner[†1], Yair Goldberg[†2], Guy Katriel[†3], Rami Yaari[†4], Nurit Gal[4], Yoav Miron[4], Arnona Ziv[4], Yoram Hamo[4], Amit Huppert[4,5]

[1]University of Haifa, Israel
[2]Technion - Israel Institute of Technology, Israel
[3]ORT Braude College of Engineering, Israel
[4]The Gertner Institute for Epidemiology and Health Policy Research, Israel
[5]Tel Aviv University, Israel

[†] Contributed equally

Summary

Background

One of the significant unanswered questions about COVID-19 epidemiology relates to the role of children in transmission. In this study we estimate susceptibility and infectivity of children compared to those of adults. Understanding the age-structured transmission dynamics of the outbreak provides precious and timely information to guide epidemic modelling and public health policy.

Methods

Data were collected from households in the city of Bnei Brak, Israel, in which all household members were tested for COVID-19 using PCR. To estimate relative transmission parameters in the absence of data on who infected whom, we developed an estimation method based on a discrete stochastic dynamic model of the spread of the epidemic within a household. The model describes the propagation of the disease between household members allowing for susceptibility and infectivity parameters to vary among two age groups. The parameter estimates are obtained by a maximum likelihood method, where the likelihood function is computed based on the stochastic model via simulations.

Findings

Inspection of the data reveals that children are less likely to become infected compared to adults (25% of children infected over all households, 44% of adults infected over all households, excluding index cases), and the chances of becoming infected increases with age. An interesting exception is that infants up to age one year are more likely to be infected than children between one and four. Using our modelling approach, we estimate that the susceptibility of children (under 20 years old) is 45% [40%, 55%] of the susceptibility of adults. The infectivity of children was estimated to be 85% [65%,110%] relative to that of adults.

Interpretation

It is widely observed that the percentage of children within confirmed cases is low. A common explanation is that children who are infected are less likely to develop symptoms than adults, and thus are less likely to be tested. We estimate that children are less susceptible

to infection, which is an additional factor explaining their relatively low rate of occurrence within confirmed cases. Moreover, our results indicate that children, when infected, are somewhat less prone to infect others compared to adults; however, this result is not statistically significant.

The resulting estimates of susceptibility and infectivity of children compared to adults are crucial for deciding on appropriate interventions, and for controlling the epidemic outbreak and its progress. These estimates can guide age-dependent public health policy such as school closure and opening. However, while our estimates of children's susceptibility and infectivity are lower than those of adults within a household, it is important to bear in mind that their role in the spread of COVID-19 outside the household, is also affected by different contact patterns and hygiene habits.

Introduction

The COVID-19 pandemic, which emerged in Wuhan, China during December 2019, has now spread over the world. Extreme measures have been taken worldwide in response to the outbreaks, among them, extended school and workplace closures. Guiding such extreme public health policies crucially depends on understanding the effect of age-structure on the epidemic dynamics. In particular, susceptibility and infectivity are two critical aspects to consider when studying population heterogeneity in the context of infectious diseases. At this stage of the epidemic, it has become clear that the clinical characteristics of the disease among children are different from those of adults[1] yet the role of children in transmitting and spreading is not clear [2]. In addition, several studies report descriptive statistics, such as the lower percentage of children diagnosed relative to their share in the population.

A key question then, is whether the above-noted difference between children and adults is the result of lower susceptibility of children to infection, or, as hypothesized by Ludvigsson [1], due to milder (or no) symptoms displayed by children, which, based on common testing policy, leads to under-detection. These explanations are non-exclusive. Moreover, little information is available regarding the ability of those children who are already infected to infect others. Deeper understanding of this issue has the potential to affect future policies to optimally mitigate the outbreaks.

Following is a short overview of related studies. In a report by the Centers for Disease Control and Prevention (CDC) [3], only 291 of 2572 children who were infected with SARS-CoV-2 were symptomatic, though this may be due to poor reporting. It has also been reported that most COVID-19 cases in children are mild, even though serious COVID-19 illness resulting in hospitalization still occurs in this age group. Zhang et al. [4] conclude that children are less likely to be infected compared to adults by about 60%, while Bi et al. [5] conclude that children were as likely to be infected as adults. Dong et al. [6] report that children of all ages appeared susceptible to COVID-19, with no significant gender difference. Kelvin et al [7] suggest that there is clear evidence that children are susceptible to SARS-CoV-2 infection, but frequently do not display notable disease symptoms, raising the possibility that children could be facilitators of viral transmission. Cai et al's [8] analysis of 10 children diagnosed with COVID-19, states that one cannot neglect the potential risk of transmission from the infected child-to-adult contacts, based on one patient. A study from New South Wales schools in Australia [9] based on both virus and antibody testing, suggests that children are not the primary drivers of COVID-19 spread in schools or in the community. According to Zimmerman et al., [10] the importance of children in transmitting the virus remains uncertain.

Viner et al. [11] state that evidence of COVID-19 transmission through child-to-child contact or through schools is not yet available, but that household transmission has an important role in the outbreak. Preliminary results from an ongoing research of the National Institute for Public Health and the Environment in the Netherlands (RVIM) [12] shows no indications that children younger than 12 years were the first to be infected within the household, and suggest that patients under 20 years play a much smaller role in the spread than adults and the elderly. Young children were less likely to test positive for SARS-CoV-2 than adolescents or adults, based on an Icelandic study [13]. In this population, using a PCR survey with random sampling (not symptom based), no children up to age 10 were found to be infected with SARS-CoV-2 as compared with 0·8% of children over age 10. Concurrently, in the targeted test survey described in the same study, the risk of children being ill was about half that of adults (6·7% versus 13·7%).

Methods

Sources of data

This study was based on data collected from the city of Bnei Brak (population 213,046) which is one of the most densely populated cities in Israel. Most of its residents are ultra-orthodox Jews, with large households and young population (approximately 51% under the age of 20) [15]. We used the COVID-19 PCR test results and epidemiological investigations from the Israeli COVID-19 database, performed in Bnei Brak until May 2, 2020. In addition, in order to map households, we used the municipality database of Bnei Brak residents born before May 25, 2020.

The original inclusion criteria, met by 637 households, were households with at least 2 members, in which all household members were tested and at least one member had tested positive to COVID-19. The 637 households comprise a total of 3,353 people of which 1,510 COVID-19 tested positive. A histogram of household sizes in the Bnei Brak data is displayed in Figure 1. In Figure 2 we present a histogram of the number of positives per household size in the Bnei Brak data set.

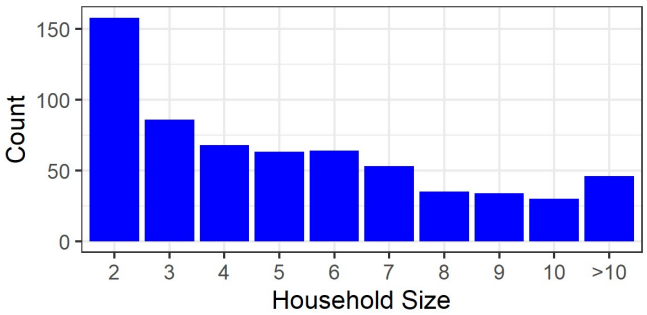

Figure 1: Histogram of household sizes in the Bnei Brak data set.

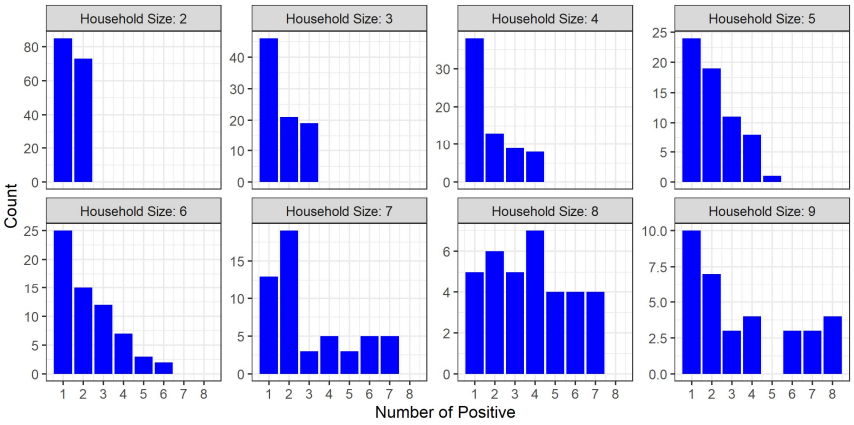

Figure 2: Histogram of the number of positives per household size in the Bnei Brak data set (Households of size greater than or equal to 10 are not shown).

A list of symptoms and the starting date of symptoms were self-reported. As part of the epidemiological investigations, those who tested positive were asked whether they had symptoms, and the date on which these symptoms appeared. The list of symptoms included fever, cough, shortness of breath, abdominal pains, headache, diarrhea, chills, sore-throat, muscle-pain, vomiting, other respiratory difficulties and additional symptoms, such as smell or taste problems, weakness, etc. For some cases, no onset date of symptoms was reported. In cases for which the reported symptoms onset date of symptoms of an individual was over four weeks prior to the first test (12 cases), the onset date of symptoms was discarded. Overall, 1,243 of the 1,510 positive cases (82%) had a valid onset date of symptoms.

The reported onset dates of symptoms and test dates were used to discern the observed epidemic time-period for each household. For each household, the first indicative date was set as the first onset date of symptoms or the first positive test date of a household member (whichever came first). The last indicative date was set as the last test date of a household member.

The observed household epidemic duration was determined as follows: 1) If the first indicative date was an onset date of symptoms, then the observed epidemic duration was set as the difference between the first and last indicative dates plus five days, which is the mean incubation period [16]. 2) If the first indicative date was a positive test date, then the observed epidemic duration was set as the difference between the first and last indicative dates plus ten days, which is the mean time from infection to detection as estimated from Israeli epidemiological investigation data. See Figure 3 for a histogram of the obtained household epidemic duration in the data set.

The reported onset dates of symptoms and test dates were also used to discern the index case in each household. Positive household members whose first indicative date was five or less days from the minimum indicative date in that household were considered suspected index cases with equal probability. In total there were 75 (out of 637) such cases of which there was more than one suspected index case. Thus, for example, if there were three positive household members whose first indicative date was within five days of the first indicative date in the household, each of the three were given a probability of being an index case of 1/3 (and the rest of the household were given probability 0). For each household, the probability

that the index case is an adult (child) was obtained by summing the probability of being an index case for all the adults (children) in that household.

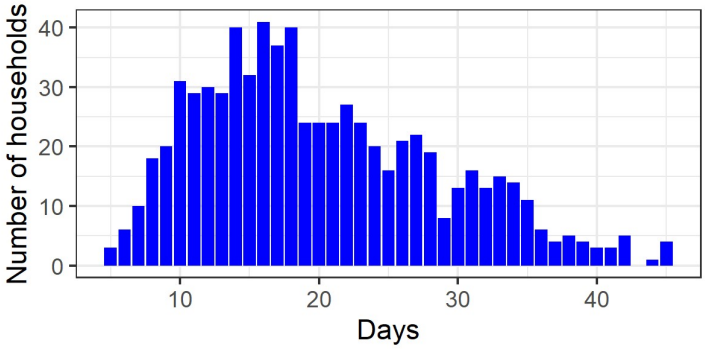

Figure 3: Histogram of the duration of observed epidemics in households of the Bnei Brak data set.

Until the end of May, the policy in Israel was to approve testing for people who have been in close contact with someone who has tested positive for COVID-19 or who has returned from abroad and, in both cases, has at least one of the symptoms in the list. However, in a significant number of cases, the tests were not executed according to the policy (29·1% of positives had symptoms but no close contact or being abroad), or without mention of the reason (3·1% of positives and 13·3% of negatives).

Estimating susceptibility and infectivity of children and adults through modelling

In order to assess differences in susceptibility and infectivity among different age groups, we use a mathematical model allowing for these differences, and fit it to the observational data on infection in the households. Our data does not include information about who infected whom, nor dates of infection. Note that the onset of symptoms and testing dates in our data are used only to identify the index case and the household epidemic duration. We use only aggregate numbers of infected individuals in the two age groups in the different households. The key point is that these data on outcomes of the many "household outbreaks" contain valuable information concerning the infectivity and susceptibility parameters, which can be extracted by a model-fitting approach: Only certain ranges of values of these parameters will generate outcomes which are consistent with those observed in reality.

We use a stochastic dynamic model for a household outbreak. Time is indexed by the discrete variable t (in days). We denote by $S_a(t)$ and $S_c(t)$ the number of adults and children who are still susceptible on day t, respectively. The notation $i_a(t)$ and $i_c(t)$ stands for the number of adults, and children who become infected on day t, respectively. The dynamic equations are

$$i_a(t) \sim Bin\left(S_a(t), 1 - e^{-\sum_{\tau=1}^{T} P_\tau[\beta_{aa} i_a(t-\tau) + \delta\beta_{aa} i_c(t-\tau)]}\right),$$

$$i_c(t) \sim Bin\left(S_c(t), 1 - e^{-\sum_{\tau=1}^{T} P_\tau[\gamma\beta_{aa} i_a(t-\tau) + \delta\gamma\beta_{aa} i_c(t-\tau)]}\right),$$

where $P_\tau, \tau \in \{1, \dots, T\}$ is the generation-time distribution, set to be a discretized version of a gamma distribution with a mean of 4·5 days, and a standard deviation of 2·5 days. This mean generation time is based on the mean intervals between symptom onset from ~2600 pairs of

known infector-infectee in the data set of confirmed cases in Israel and is also compatible with findings from other studies [17,18].

The parameter β_{aa} stands for the transmission rate among adults, γ is the susceptibility of children relative to that of adults, and δ is the infectivity of children relative to that of adults. The number of susceptible adults and children on day t is given by the equations,

$$S_a(t+1) = S_a(t) - i_a(t), \qquad S_c(t+1) = S_c(t) - i_c(t)\,.$$

In words, the relative susceptibility of two individuals is defined as the ratio of their probabilities being infected per unit time, when exposed to the same infectious factor. The relative infectivity of two individuals is defined as the ratio of the probabilities per unit time that these individuals generate infection, when making contact with individuals who have identical susceptibilities. See Supplementary Material for a detailed description of the model.

By fixing the number of individuals of each age group in a household, the age group to which the index case belongs, and the transmission parameters, one can generate simulations of such a household outbreak. For each such simulation, we record the number of individuals of each age group who were infected in the time period corresponding to the household considered (determined by symptom onset and testing dates as described above). Since the model is stochastic, different realizations of such a simulation will lead to different outcomes, where an outcome is defined as the number of adults infected and the number of children infected in the household. The probability distribution over the finite set of possible outcomes is approximated for each of the households, by running 1000 simulations, for each of a range of transmission parameter values on a grid with a resolution of 0·05. These probability distributions, which are dependent on the transmission parameters, enable us to compute the likelihood function corresponding to the outcomes in each of the households in our empirical data. The total likelihood is then the product of the likelihoods for all households. This likelihood is a function of the transmission parameters (β_{aa}, γ, δ) and may thus be used to estimate the transmission parameters using maximum likelihood. Note that since our likelihood function is computed using simulation, our estimation method is what is known as "simulated maximum likelihood" [19].

To test the ability of our estimation procedure to identify parameters, we carried out a simulation study in which household outbreaks with known parameters were generated in a collection of households of the same type as those in the data set, and our method was used to estimate the parameters. A similar procedure was used to obtain parametric bootstrap confidence intervals, by generating 1000 simulated data sets using the parameter estimated from the real data, and re-estimating the parameters.

An R software package applying our methodology in a computationally efficient way is available online [20], allowing other researchers to estimate relative susceptibility and infectivity of children and adults given an appropriate dataset.

Results

Inspection of the data reveals that chances of becoming infected increase with age, up to around age 20, and seems to remain more or less constant thereafter (see Figure 4).

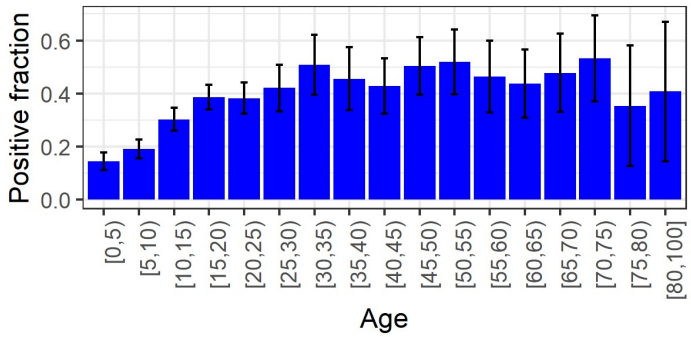

Figure 4: Fraction of positives by age-group in the data set, excluding index cases. Binomial confidence intervals were calculated using the normal approximation.

We divide the population into a children's group (0-19 including) and an adult group (20+), which is also consistent with RIVM research cutoff [12]. Using this cutoff, we obtain 1,809 adults of whom 998 were found to be positive to COVID-19 (55%) and 1,544 children of whom 512 were positive to COVID-19 (33%). Excluding index cases, which in most cases were adults (see Table 1 in the Supplementary Material), 44% of adults were infected over all compared to 25% of the children. Interestingly, children under the age of one seem to be more likely to be infected than children between one and four (Figure 5). Of the 998 positive adults in the data set, 875 reported having symptoms (88%). In comparison, 368 of the 512 positive children (72%) reported having symptoms.

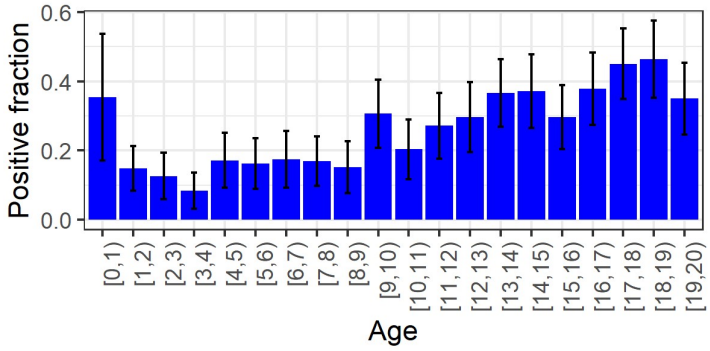

Figure 5: Fraction of positives by age in children, excluding index cases. Binomial confidence intervals were calculated using the normal approximation.

Using our modelling approach, we estimate that the relative susceptibility of children (γ) is 45% [40%, 55%]. The relative infectivity (δ) of children was estimated to be 85% [65%, 110%]. The adult-adult transmission parameter β_{aa} was estimated as 0·3 [0·25, 0·35]. The ranges reported are based on parametric bootstrap confidence intervals. Figure 6 displays level curves of the likelihood as a function of the susceptibility and infectivity parameters, for three values of the adult-adult transmission parameter β_{aa}. Figure 7 shows the model fit to

the observations within households of different sizes. The fit obtained using the model is much better than the fit obtained using a naïve model that ignores secondary infections within households (see Figure S2 in the Supplementary Material).

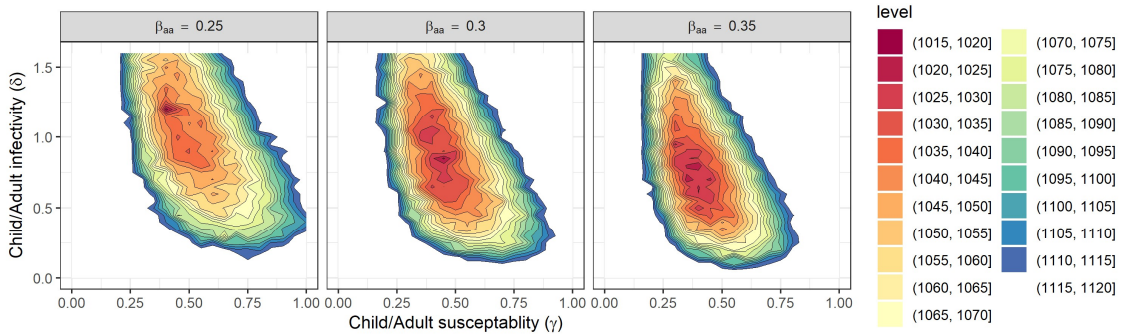

Figure 6: Likelihood level curves for the parameters γ, δ, and three values of β_{aa}. For smaller values of β_{aa} the maximal-likelihood estimates of parameters γ and δ are larger.

We performed sensitivity analysis to examine the effect of various assumptions on the results. These included sensitivity to the assumed generation-time distribution, the assumed age-group of the index case in households in which there was some doubt regarding the index case's age-group, and to the assumed observed duration of the epidemic in the households. In general, our results seem to be robust to reasonable variations in all of these attributes. Full description of the sensitivity analyses appears in the Supplementary Material.

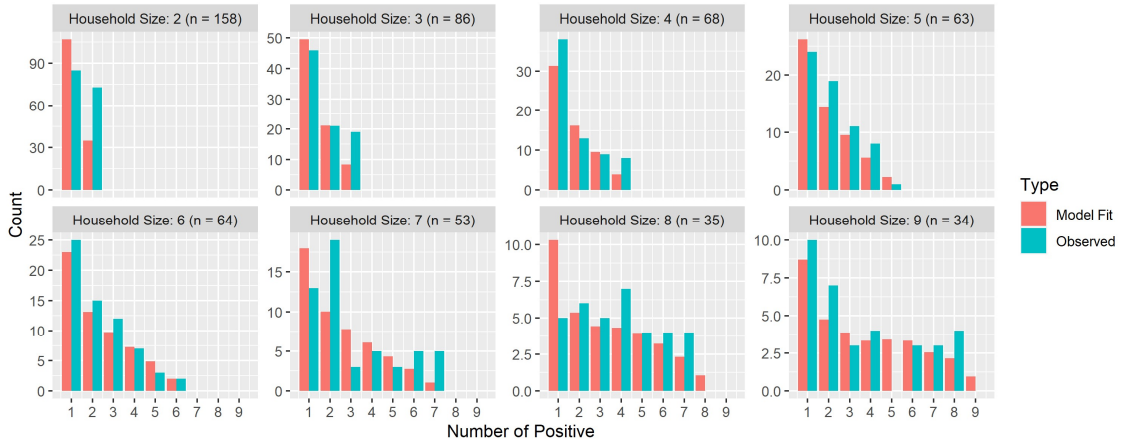

Figure 7: Best model fit to the observed data in the Bnei Brak data set, aggregated according to the household size (fit to households of size greater than or equal to 10 are not shown).

Discussion

Currently, one of the most significant unanswered questions about COVID-19 transmission relates to the role of children in the spread of infection. In this study we address this

knowledge gap with the aim of quantifying susceptibility and infectivity of children compared to adults.

We fitted a stochastic age-of-infection type model, allowing us to take into account the generation-time distribution. Although the propagation of the disease within a household was not reported, a simulated likelihood approach enabled us to fit such a model using only aggregated data of infected individuals.

The estimation results indicate that the role of children in the transmission of infection is less prominent than that of adults: children are less susceptible than adults (relative susceptibility 45% [40%, 55%]), and their infectivity may be somewhat lower as well (relative infectivity 85% [65%, 110%]). The data were more informative regarding the relative susceptibility of children than regarding their relative infectivity, as indicated by much wider confidence intervals for the relative infectivity in comparison to those for the relative susceptibility. Data containing more index cases in the children's group would provide more information about children's infectivity. In order to provide a full explanation for the fact that this dataset cannot provide a more accurate estimate of the relative infectivity of children, understanding the variance of the estimators is required, a task which is beyond the scope of this work and will be undertaken elsewhere.

As we have noted in the Introduction, the fact that the fraction of children among the confirmed cases has been found to be low in many countries can be accounted for by two (nonexclusive) hypotheses: (1) Children display milder symptoms than adults when infected, so are less likely to be tested, (2) Children are less susceptible to infection than adults. Our results lend support to the second hypothesis and suggest that lower susceptibility of children to infection could indeed play a large role in explaining the epidemiological pattern noted.

The result concerning the lower susceptibility of children raises the question of possible biological mechanisms that could account for such an effect. A recent study has found evidence suggesting the presence some residual immunity in people not previously exposed to SARS-CoV-2, in the form of SARS-CoV-2-reactive CD4+ T cells, attributed to circulating ‘‘common cold’’ coronaviruses [21]. It is possible that this form of partial protection is more common in children since infection rates with seasonal coronaviruses are higher in children [22]. The fact that in our data set, children under the age of one have higher rates of infection with SARS-CoV-2 compared to children between one and four, is consistent with the hypothesis that partial immunity to SARS-CoV-2 could be related to past exposure to seasonal coronaviruses.

We note that while our estimates of children's susceptibility and infectivity are lower than those of adults within a household, it is important to bear in mind that their role in the spread of COVID-19 outside the household is also affected by different contact patterns and hygienic habits.

Summarizing, our findings shed light on empirical observations gathered worldwide regarding the role of children in the spread of disease, and can contribute to better modelling of the epidemic dynamics, devising control measures and guiding public health policy. Our methodology can be applied to other household studies. In particular, it could be employed to the results of serological tests in households, where all members of the households are tested.

References

1. Ludvigsson JF. Systematic review of COVID-19 in children show milder cases and a better prognosis than adults. *Acta Paediatrica* 2020; 109: 1088-1095.
2. Hay JA, Haw DJ, Hanage WP, Metcalf CJE and Mina MJ. Implications of the Age Profile of the Novel Coronavirus. *Harvard University's DASH repository* 2020.
3. Centers for Disease Control and Prevention (CDC). Coronavirus Disease 2019 in Children—United States, February 12–April 2, 2020. *Morbidity and Mortality Weekly Report* 2020; 69: 422-426.
4. Zhang J, Litvinova M, Liang Y, et al. Age profile of susceptibility, mixing, and social distancing shape the dynamics of the novel coronavirus disease 2019 outbreak in China. *medRxiv* 2020.
5. Bi Q, Wu Y, Mei S, et al. Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts in Shenzen, China: a retrospective cohort study. *Lancet* 2020. DOI: https://doi.org/10.1016/S1473-3099(20)30287-5
6. Dong Y, Mo X, Hu Y, et al. Epidemiology of COVID-19 among children in China. *Pediatrics* 2020; 145. DOI: https://doi.org/10.1542/peds.2020-0702
7. Kelvin AA, Halperin S. COVID-19 in children: the link in the transmission chain. *Lancet Infectious Dis.* 2020; 20:633-634.
8. Cai J, Xu J, Lin D, et al. A Case Series of children with 2019 novel coronavirus infection: clinical and epidemiological features. *Clinical Infectious Diseases* 2020; DOI: https://doi.org/10.1093/cid/ciaa198.
9. National Centre for Immunization Research and Surveillance (NCIRS). COVID-19 in schools – the experience in NSW. New South Wales Government, Australia, 2020.
10. Zimmermann P, Curtis N. Coronavirus Infections in Children Including COVID-19: An Overview of the Epidemiology, Clinical Features, Diagnosis, Treatment and Prevention Options in Children. *Pediatr Infect Dis*. 2020: 39: 355-368.
11. Viner RM, Russell SJ, Croker H, et al. School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. *Lancet Child & Adolescent Health* 2020; 4: 397-404. DOI: https://doi.org/10.1016/S2352-4642(20)30095-X
12. National Institute for Public Health and the Environment (RVIM). Children and COVID-19. Ministry of Health, Welfare and Sport, Netherlands, 2020. https://www.rivm.nl/en/novel-coronavirus-covid-19/children-and-covid-19 (accessed June 2, 2020).
13. Gudbjartsson DF, Helgason A, Jonsson H, et al. Spread of SARS-CoV-2 in the Icelandic Population. *New England Journal of Medicine*, 2020; DOI: 10.1056/NEJMoa2006100
14. World Health Organization. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). 2020.
15. Central Bureau of Statistics. Local authorities in Israel: Bnei Brak. Israel, 2016. https://old.cbs.gov.il/publications18/local_authorities16_1722/pdf/70_6100.pdf (accessed June 2, 2020).
16. Lauer SA, Grantz KH, Bi Q, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine* 2020; 172:577-582.

17. Du Z, Xu X, Wu Y, et al. Serial interval of COVID-19 among publicly reported confirmed cases. *Emerg Infect Dis* 2020; 26. DOI: https://doi.org/10.3201/eid2606.200357
18. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. *Int J Infect Dis.* 2020; 93:284–6.
19. Diggle PJ, and Gratton RJ. Monte Carlo methods of inference for implicit statistical models. *Journal of the Royal Statistical Society: Series B (Methodological)* 1984; 46:193-212.
20. Goldberg Y, Yaari R. Simulated likelihood for household epidemics (R package). Available from: https://github.com/yairgoldy/sl4hm (accessed June 2, 2020)
21. Grifoni A, Weiskopf D, Ramirez SI, et al. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* 2020; 181:1-13.
22. Biggs HM, Killerby ME, Vet MB, et al. Human Coronavirus Circulation in the USA, 2014–2017. *Open Forum Infect Dis.* 2017; 4:S311–S312.

Ethics statement
An exemption from institutional review board approval was determined by the Israeli Ministry of Health as part of an active epidemiological investigation, based on use of anonymous data only and no medical intervention.

Declaration of interests
All authors declare no competing interests.

Data sharing
The computer code (in R languages) for the data analysis can be downloaded from [20].

Acknowledgments
The authors would like to thank Boaz Lev and Baruch Velan for their support, encouragement, and fruitful discussions.