# Using social contact data to predict and compare the impact of social distancing policies with implications for school re-opening

Ellen Brooks-Pollock[1,2*], Jonathan M. Read[3], Angela McLean[4], Matt J. Keeling[5,6], Leon Danon,[2,7,8].

[1]Bristol Veterinary School, University of Bristol, BS40 5DU.
[2]NIHR Health Protection Research Unit in Behavioural Science and Evaluation, Population Health Sciences, Bristol Medical School, University of Bristol, BS8 2BY.
[3]Lancaster Medical School, University of Lancaster.
[4]Department of Zoology, University of Oxford, OX1 3SZ.
[5,6]Mathematics Institute and School of Life Sciences, University of Warwick, CV4 7AL.
[7] CEMPS, University of Exeter.
[8]The Alan Turing Institute.

## Abstract

### Background

Social distancing measures, including school closures, are being used to control SARS-CoV-2 transmission in many countries. Once "lockdown" has driven incidence to low levels, selected activities are being permitted. Re-opening schools is a priority because of the welfare and educational impact of closures on children. However, the impact of school re-opening needs to be considered within the context of other measures.

### Methods

We use social contact data from the UK to predict the impact of social distancing policies on the reproduction number. We calibrate our tool to the COVID-19 epidemic in the UK using publicly available death data and Google Community Mobility Reports. We focus on the impact of re-opening schools against a back-drop of wider social distancing easing.

### Results

We demonstrate that pre-collected social contact data, combined with incidence data and Google Community Mobility Reports, is able to provide a time-varying estimate of the reproduction number ($R$). From an pre-control setting when R=2.7 (95%CI 2.5, 2.9), we estimate that the minimum reproduction number that can be achieved in the UK without limiting household contacts is $0.45\ (95\%CI: 0.41 - 0.50)$; in the absence of other changes, preventing leisure contacts has a smaller impact ($R = 2.0, 95\%CI: 1.8 - 2.4$) than preventing work contacts ($R = 1.5, 95\%CI: 1.4 - 1.7$). We find that following lockdown (when $R$=0.7 (95% CI 0.6, 0.8)), opening primary schools in isolation has a modest impact on transmission $R = 0.83\ (95\%CI: 0.77 - 0.90)$ but that high adherence to other measures is needed. Opening secondary schools as well as primary school is predicted to have a larger overall impact ($R = 0.95, 95\%CI: 0.85 - 1.07$), however transmission could still be controlled with effective contact tracing.

### Conclusions

Our findings suggest that primary school children can return to school without compromising transmission, however other measures, such as social distancing and contract tracing, are required to control transmission if all age groups are to return to school. Our tool provides a mapping from policies to the reproduction number and can be used by policymakers to compare the impact of social-easing measures, dissect mitigation strategies and support careful localized control strategies.

## Introduction

The reproduction number, or the 'R number' has become a central statistic being used to characterise the transmission of novel severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2). Early estimates of the reproduction number, which is the average number of secondary cases due to a single case, range between 2.5 and 3.5[1,2], indicating that at least 2 out of every 3 transmission events need to be prevented in order to avoid an outbreak or control an ongoing epidemic. In the United Kingdom (UK), social distancing restrictions, including school closures, introduced on 23 March 2020, led to an overall reproduction number less than 1 and a subsequent decline in the daily number of cases and deaths. It is therefore important to quantify the effect of interventions and their easing on the reproduction number. It is uncertain how the relaxation of these restrictions, especially the physical return to school of the school-age population, will affect the transmission of the virus, though contact tracing and isolation of discovered cases is anticipated to mitigate some of the impact.

The reproduction number of close contact infections such as SARS-CoV-2 depends critically on social contact behaviour. Questionnaire surveys, enumerate and describe face-to-face contacts an individual had on a given day, are the most direct way of assessing the potential for spread in a population[3]. Several such surveys have quantified the behaviour of the UK population prior to the pandemic in 2020[4–6]. A more recent survey in the UK estimated that post 23 March 2020 the number of social interactions was 2.9 contacts per person per day[7].

Social distancing measures, such as the closure of schools and workplaces and mandatory reduction in social interactions, while effective at preventing transmission, have severe economic and psychological effects, and of particular concern is their impact on children[8]. Age-specific behavioural patterns mean that social distancing measures affect age groups differently. In normal circumstances, the majority of social contact hours for persons over 60 years old occur at home while only a quarter of their social contact hours are associated with leisure activities outside the home. In contrast, nearly 60% of twenty to thirty-year olds' social contact hours are at work[5]. Crucially, nearly half of children's social contact hours are made within a school setting, meaning that school closures have a major impact on the social experience of young people. In this study, we use social contact data[5], including an additional targeted survey of children, to quantify the impact of re-opening schools on the reproduction number in the UK[9].

## Results

We used publicly available death data from the UK[10] to estimate an exponential growth rate of 0.23 (95% CI 0.22, 0.24) deaths per day between 13 March 2020 and 23 March 2020. This corresponds to a reproduction number of 2.7 using a mean serial interval of 7.5 days [11–13].

We combine this estimate of the reproduction number prior to lockdown with social contact data to estimate a transmission probability per contact hour of 0.002 $hour^{-1}$, see Materials and Methods.

Following lockdown, we use Google Community Mobility Reports as a proxy for the percentage reduction in active work, leisure and travel contacts. With a 65% reduction in work contacts, a 75% reduction in leisure contacts and a 95% reduction in school contacts, the reproduction number is reduced to 0.7 (95% CI 0.6, 0.8) (figure 1), which is consistent with direct estimates[7].

Eliminating all but households results in a reproduction number of 0.45 (95% CI, 0.41, 0.50), but this estimate does not include contacts outside the home arising from essential services.

Adding work and school contacts to household contacts, with no leisure or other contacts, increases the reproduction number to 2.0 ($95\% CI: 1.8 - 2.4$). Adding leisure contacts to household contacts, while preventing work and school contacts, increases the reproduction number to 1.5 ($95\% CI: 1.4 - 1.7$).

Tracing and isolating social contacts of symptomatic cases so they do not transmit onwards is beneficial but does not control transmission in isolation[14]. The impact of contact tracing increases as social distancing measures are eased (figure 2). If all children are in school, then when 20% of normal contacts are present, the reproduction number is close to 1. In this scenario even modest contact tracing is enough to control transmission. The added benefit of tracing 10 contacts per index case over 5 contacts per case is minimal, because very few people have more than 5 contacts. With 60% of contacts outside the home present, schools could fully re-open with effective contact tracing in place. If contact patterns return to pre-covid levels, then contact tracing on its own is unlikely to be able to control transmission without other measures in place.

From the lockdown baseline reproduction number estimated above of 0.7, we investigate the impact of school opening scenarios on the reproduction number. Figure 3 illustrates the predicted reproduction number under scenarios in which schools are open. The shaded regions represent different policies under the assumptions that children are half as infectious as adults and there is no immunity in the population[15].

We find that if no other social contacts outside the home increase apart from those occurring within primary schools, then opening primary schools is consistent with a reproduction number less than 1, $R = 0.83$ ($95\% CI: 0.77 - 0.90$) (Fig. 3A). However, even a modest increase in contacts outside home and school, relative to post-lockdown levels, would push the reproduction number back above 1. In the absence of substantial population-level immunity, the additional opening of secondary schools is likely to bring transmission close to epidemic growth in the population ($R = 0.95, 95\% CI: 0.85 - 1.07$). In general, higher adherence to other social distancing measures is required as more children return to school.

We predict that contact tracing could increase the options for opening schools (Figs 3B and 3C). We assume that a given proportion of all contacts are successfully traced, self-isolate, and that their contribution to the reproduction number is effectively zero. Under a scenario similar to the situation in early June 2020, where 20% of contacts are effectively traced and isolated, a larger proportion of pupils could return to school while still limiting transmission.

If 60% of contacts of symptomatic cases were traced and isolated, we estimate that schools could fully re-open while maintaining control of transmission, as long as at least 50% of other contacts are prevented ($R = 0.84, 95\% CI: 0.76 - 0.92$). In this scenario, other forms of social distancing, including working from home and eliminating leisure contacts, would still be required if schools were to be fully open before a pharmaceutical solution is found.

Finally, we consider the impact of the relative infectiousness of children when re-opening schools. The most pessimistic scenario, where children are as infectious as adults, corresponds to the scenarios considered in figure 3. If children are less infectious than adults then re-opening primary and secondary schools has a smaller impact on the reproduction number, but the impact of increasing other contacts outside home and school settings remains the same.

**Discussion**

In this paper, we demonstrate that a combination of early death counts and social contact data provide sufficient information to estimate the potential impact of combinations of social distancing measures on the reproduction number for COVID-19 in the United Kingdom. We

focus on the effects of school closures/re-openings on COVID-19 transmission in the context of other control measures. Our findings suggest that that high adherence to social distancing outside school settings is needed to maintain epidemic control. Opening primary schools has a modest impact on R, while opening secondary schools is predicted to have a larger overall impact; a combination of reopening both would result in a loss of epidemic control.

Our findings support the use of contact tracing as a key part of epidemic control; however, tracing needs to be highly effective. After the introduction of Test and Trace system in the UK only 20% of social contacts of cases were successfully traced and isolated within 48 hours, though this has substantially increased over time[16]. While tracing 20% of contacts has a positive impact on the reproduction number, it is insufficient to prevent epidemic growth if all schools are fully open.

The greater risk arises from contact with people outside the home and school contexts. It is likely that reopening of schools will also lead to an increase in contacts made outside school, due to caregivers returning to work and interactions between parents. A strength of this analysis is its predictive value of the effect of combined interventions. Using metrics of adherence to social distancing measures, such as Google mobility or contemporary social contact surveys, it is possible to map the country's progression across figure 1, and therefore estimate the effect of policy changes on the reproduction number and hence the population attributable fraction of cases due to multiple combined interventions[17].

This analysis was made possible by pre-existing detailed social contact data. Social contact patterns have been used to characterise the potential for disease transmission in a population[18], design vaccine and control programmes for infectious diseases including influenza[19], meningitis[20] and now COVID-19[7]. However, in most settings, such data are out-of-date or not available. Given their proven value, we argue that regular, representative social contact surveys should be become a routine part of epidemic control and preparedness.

## Materials and Methods

Data description

The Social Contact Survey surveyed 5,861 individuals in the UK in 2010 about their social contacts during a single day[5]. Participants were recruited using three approaches: a paper survey sent to people in the post, an online survey and an online survey aimed specifically at school-aged children. Participants were asked to complete demographic information about themselves including age, occupation and about their social contacts on the previous day. Participants were asked to report the number of people they met, the duration of the contact (<10 minutes, 10 to 59 minutes, 1 to 4 hours, 4+ hours), the context (home, work/school, travel, other/leisure), and whether the contact involved touch, e.g. a handshake, hug or kiss. To ease the ability to report large number of contacts per day, participants could report contacts as individual contact or groups of contacts; this methodology better captures the right-hand tail of the degree distribution. Participants were also asked about transitive interactions between contacts, reported in[21].

Estimating the Reproduction Number from social contact data

We use an individual-based approach to calculate a reproduction number of each of the participants of the Social Contact Survey study[9]. The reproduction number for an individual is given by

$$R_{ind} = \tau \sum_{i=1}^{k} n_i d_i \tag{1}$$

where $k$ is the number of contact events reported by each participant, $n_i$ is the number individuals in that contact (participants could report groups of similar contacts), $d_i$ is the duration of the contact and $\tau$ is the probability of transmission. Because we do not have ages of contacts, this is an ego-centric estimate of R, and does not include local depletion of susceptibles.

The population-wide reproduction number, $R_t$, is calculated using the age-adjusted mean of the squared individual reproduction numbers, i.e.

$$R_t = \frac{\sum_{j=1}^{N} \alpha_j \varepsilon_j \left(R_{ind}^{j}\right)^2}{\sum_{j=1}^{N} \alpha_j}, \tag{2}$$

where $N$ is the number of participants in the Social Contact Survey, $0 \leq \varepsilon_j \leq 1$ is the relative infectiousness of children relative to adults. $\alpha_j$ is the age-specific weighting for participant $j$, estimated to match the age distribution in the UK population, calculated as the ratio of the proportion of individuals aged $a$ in the UK, $P_{UK}(a)$, to the Social Contact Survey sample, $P_{SCS}(a)$,

$$\alpha_j = \frac{P_{UK}(a_j)}{P_{SCS}(a_j)}. \tag{3}$$

The uncertainty associated with the reproduction number is estimated by bootstrapping the contact data, weighted by age, using the boot function in R. We report the bootstrapped mean and 95% percentile confidence intervals.

*Model calibration*

The model can be calibrated using incidence data when the social contact patterns are known. Here, we calibrated the model to the exponential growth phase of the epidemic in the UK prior to the introduction of widespread mass social distancing on 23 March 2020. We estimated the growth rate, $g$, from death data between 13 March 2020 and 30 March 2020, then calculated the reproduction number as $R = \exp\,(gS)$ where $S$ is the serial interval.

*Estimating the reproduction number following lockdown on 23 March 2020*

Google has made community mobility reports[22] available for the period during COVID-19 transmission from 15 February 2020. The Google mobility reports provide a point estimate for the percentage change in number of visits to and length of stay at places categorized as grocery and pharmacy, parks, transit stations, retail and recreation, residential, and workplace. The median percentage change is relative to the median value for the same day of the week for the period between 3 January 2020 and 6 February 2020[22].

We mapped the context reported in the Social Contact Survey onto the Google mobility data categories as home is equivalent to residential, work/school to workplace, other/leisure to retail and recreation and travel to transit. We assumed that 100% of contacts were active during the week of 18 March 2020. We then used Google mobility estimate of the percentage of contacts that were active in subsequent weeks.

To simulate $x\%$ of contacts in a given context being active, we take a random sample without replacement of a proportion $(1 - x/100)$ of all contacts for that context according to age group. The selected contacts are flagged with a comply flag $c_i$ equal to 1. The reduced individual reproduction number is given by:

$$R_{t_{ind}} = \tau \sum_{i=1}^{k} \delta_{c_i 1} n_i d_i \tag{4}$$

Where $\delta_{c_i 1}$ equals zero if $c_i = 1$ and one otherwise. We estimate the mean and 95% confidence intervals for the reproduction number by sampling contacts then bootstrapping contacts, weighted by age, 2000 times and taking the percentile confidence interval.

We assumed that no age groups have pre-existing immunity against COVID-19 and all age groups are equally infectious.

*Forward simulating social distancing measures and school closures*

To simulate school closures, we remove all contacts for the relevant school aged children that have "school" listed as the context for the contact. To capture the 2% of children who are still attending school, we re-instate a random sample of the removed contacts. We simulate other contacts being active as above.

Contact tracing

We modelled contact tracing from symptomatic index cases. We assumed that an age-specific proportion of index cases were symptomatic, where index cases under 18 years old had a 25% chance of being symptomatic, then assuming a linear increase with age in the chance of symptoms up to 75% for people over 80 years old[15]. For each contact, we drew a random number to determine if the index case was symptomatic, and therefore eligible for contact tracing. We assumed that either 20% or 80% of contacts were traced and isolated before becoming infectious.

**References:**


1. Read JM, Bridgen JR, Cummings DA, Ho A, Jewell CP. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. medRxiv. 2020; 2020.01.23.20018549. doi:10.1101/2020.01.23.20018549
2. Challen RJ, Tsaneva K, Pitt M, Edwards T, Gompels L, Lucasa L, et al. Estimates of regional infectivity of COVID-19 in the United Kingdom following imposition of social distancing measures. medRxiv. 2020; 2020.04.13.20062760. doi:10.1101/2020.04.13.20062760
3. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 2008;5: e74. doi:10.1371/journal.pmed.0050074
4. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 2008;5: e74. doi:10.1371/journal.pmed.0050074
5. Danon L, Read JM, House TA, Vernon MC, Keeling MJ. Social encounter networks: characterizing Great Britain. Proc Biol Sci. 2013;280: 20131037. doi:10.1098/rspb.2013.1037

6. Klepac P, Kucharski AJ, Conlan AJ, Kissler S, Tang M, Fry H, et al. Contacts in context: large-scale setting-specific social mixing matrices from the BBC Pandemic project. 2020 [cited 17 Apr 2020]. doi:10.1101/2020.02.16.20023754

7. Jarvis CI, Van Zandvoort K, Gimma A, Prem K, Klepac P, Rubin GJ, et al. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. BMC Med. 2020;18: 124. doi:10.1186/s12916-020-01597-8

8. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. The Lancet. Lancet Publishing Group; 2020. pp. 912–920. doi:10.1016/S0140-6736(20)30460-8

9. Keeling MJ, Grenfell BT. Individual-based perspectives on R0. J Theor Biol. 2000;203: 51–61. doi:10.1006/jtbi.1999.1064

10. Public Health England. Coronavirus (COVID-19) in the UK. [cited 4 Jul 2020]. Available: https://coronavirus.data.gov.uk/?_ga=2.143412149.1288329299.1593890063-170644305.1585303313

11. Zhao S, Cao P, Gao D, Zhuang Z, Cai Y, Ran J, et al. Serial interval in determining the estimation of reproduction number of the novel coronavirus disease (COVID-19) during the early outbreak. J Travel Med. 2020;2020: 1–3. doi:10.1093/jtm/taaa033

12. Chan Y-WD, Flasche S, Lam T-LT, Leung M-HJ, Wong M-L, Lam H-Y, et al. Transmission dynamics, serial interval and epidemiology of COVID-19 diseases in Hong Kong under different control measures. Wellcome Open Res. 2020;5: 91. doi:10.12688/wellcomeopenres.15896.1

13. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. New England Journal of Medicine. 2020. pp. 1199–1207. doi:10.1056/NEJMoa2001316

14. Keeling MJ, Hollingsworth TD, Read JM. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). J Epidemiol Community Health. 2020; jech-2020-214051. doi:10.1136/jech-2020-214051

15. Davies NG, Klepac P, Liu Y, Prem K, Jit M, Eggo RM. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nat Med. 2020 [cited 15 Jul 2020]. doi:10.1038/s41591-020-0962-9

16. Weekly NHS Test and Trace bulletin, England: 18 to 24 June 2020 - GOV.UK. [cited 4 Jul 2020]. Available: https://www.gov.uk/government/publications/nhs-test-and-trace-statistics-england-18-june-to-24-june-2020/weekly-nhs-test-and-trace-bulletin-england-18-24-june-2020

17. Brooks-Pollock E, Danon L. Defining the population attributable fraction for infectious diseases. Int J Epidemiol. 2017;46: 976–982. doi:10.1093/ije/dyx055

18. Read JM, Lessler J, Riley S, Wang S, Tan LJ, Kwok KO, et al. Social mixing patterns in rural and urban areas of Southern China. Proc R Soc B Biol Sci. 2014;281. doi:10.1098/rspb.2014.0268

19. Baguelin M, Flasche S, Camacho A, Demiris N, Miller E, Edmunds WJ. Assessing Optimal Target Populations for Influenza Vaccination Programmes: An Evidence Synthesis and Modelling Study. Leung GM, editor. PLoS Med. 2013;10: e1001527.

doi:10.1371/journal.pmed.1001527

20. Christensen H, Hickman M, Edmunds WJ, Trotter CL. Introducing vaccination against serogroup B meningococcal disease: An economic and mathematical modelling study of potential impact. Vaccine. 2013;31: 2638–2646. doi:10.1016/J.VACCINE.2013.03.034

21. Danon L, Read JM, House TA, Vernon MC, Keeling MJ. Social encounter networks: Characterizing great Britain. Proc R Soc B Biol Sci. 2013;280. doi:10.1098/rspb.2013.1037

22. Google LLC. Google COVID-19 Community Mobility Reports. [cited 25 Jun 2020]. Available: https://www.google.com/covid19/mobility/

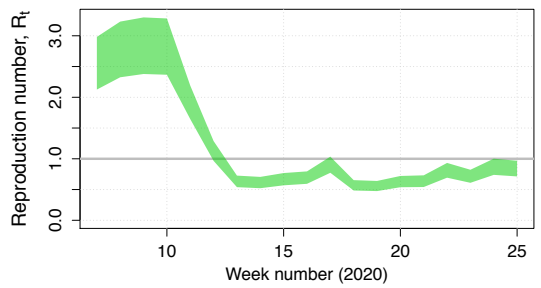


Figure 1: The time-varying reproduction number in the UK, estimated using incidence death data prior to lockdown, Social Contact Survey data and Google Community Mobility Reports.

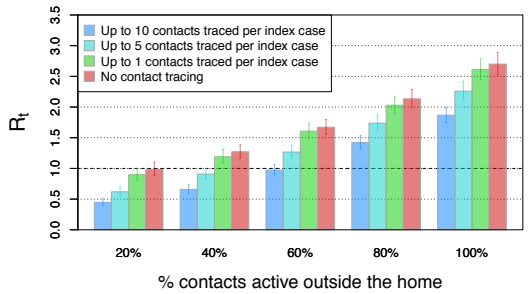


Figure 2: The impact of tracing and isolating contacts of symptomatic cases when all children are back at school, and assuming that children are half as infectous as adults.

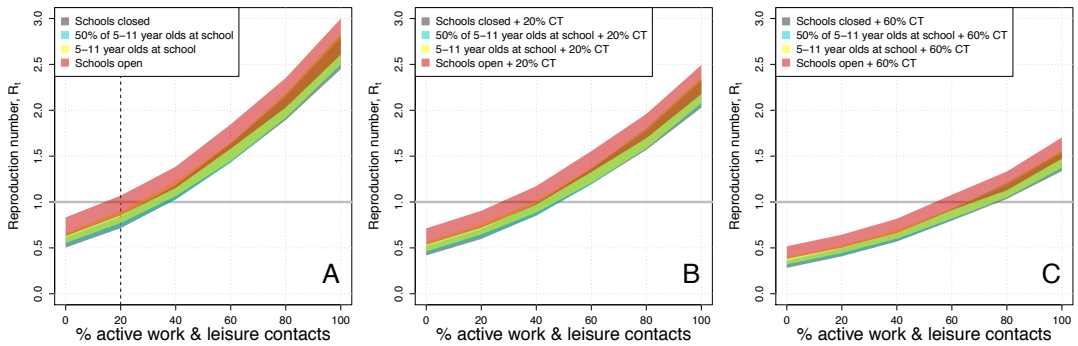


Figure 3: The reproduction number as a function of the percentage of active work and leisure contacts under different contact tracing and school closure scenarios. The shaded regions indicate the impact of school reopening scenarios, and the panels illustrate the impact with (A) no contact tracing, the vertical dotted line denotes lockdown characteristics (B) 20% of contacts traced, and; (C) 60% of contacts traced. The width of the ribbons indicates 95% confidence intervals. We have assumed here that children are half as infectious as adults.