Check for updates

RESEARCH ARTICLE

REVISED The effectiveness of social bubbles as part of a Covid-19 lockdown exit strategy, a modelling study [version 2; peer review: 2 approved]

Trystan Leng [ID][1], Connor White [ID][1], Joe Hilton [ID][1], Adam Kucharski[2], Lorenzo Pellis[3], Helena Stage[3], Nicholas G. Davies[2], Centre for Mathematical Modelling of Infectious Disease 2019 nCoV Working Group, Matt J. Keeling [ID][1], Stefan Flasche [ID][2]

[1]The Zeeman Institute for Systems Biology & Infectious Disease Epidemiology Research, University of Warwick, Coventry, UK
[2]Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, London, UK
[3]Department of Mathematics, University of Manchester, Manchester, UK

Abstract

Background: During the coronavirus disease 2019 (COVID-19) lockdown, contact clustering in social bubbles may allow extending contacts beyond the household at minimal additional risk and hence has been considered as part of modified lockdown policy or a gradual lockdown exit strategy. We estimated the impact of such strategies on epidemic and mortality risk using the UK as a case study.

Methods: We used an individual based model for a synthetic population similar to the UK, stratified into transmission risks from the community, within the household and from other households in the same social bubble. The base case considers a situation where non-essential shops and schools are closed, the secondary household attack rate is 20% and the initial reproduction number is 0.8. We simulate social bubble strategies (where two households form an exclusive pair) for households including children, for single occupancy households, and for all households. We test the sensitivity of results to a range of alternative model assumptions and parameters.

Results: Clustering contacts outside the household into exclusive bubbles is an effective strategy of increasing contacts while limiting the associated increase in epidemic risk. In the base case, social bubbles reduced fatalities by 42% compared to an unclustered increase of contacts. We find that if all households were to form social bubbles the reproduction number would likely increase to above the epidemic threshold of R=1. Strategies allowing households with young children or single occupancy households to form social bubbles

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers	
	1	**2**
version 2 (revision) 29 Mar 2021		
version 1 10 Sep 2020	✓ report	✓ report

1. **Jessica Enright** [ID], University of Glasgow, Glasgow, UK

2. **Martin C. J. Bootsma** [ID], University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
Utrecht University, Utrecht, The Netherlands

Any reports and responses or comments on the article can be found at the end of the article.

increased the reproduction number by less than 11%. The corresponding increase in mortality is proportional to the increase in the epidemic risk but is focussed in older adults irrespective of inclusion in social bubbles.

Conclusions: If managed appropriately, social bubbles can be an effective way of extending contacts beyond the household while limiting the increase in epidemic risk.

Keywords

Contact clustering, Social bubble, Exit strategy, Covid-19
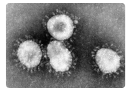
This article is included in the Coronavirus (COVID-19) collection.

Corresponding author: Trystan Leng (t.leng@warwick.ac.uk)

Author roles: Leng T: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **White C**: Data Curation, Software, Writing – Review & Editing; **Hilton J**: Data Curation, Writing – Review & Editing; **Kucharski A**: Writing – Review & Editing; **Pellis L**: Writing – Review & Editing; **Stage H**: Writing – Review & Editing; **Davies NG**: Writing – Review & Editing; **Keeling MJ**: Conceptualization, Investigation, Methodology, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Flasche S**: Conceptualization, Investigation, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Background

In the UK, similar to many other countries, the introduction of stringent physical distancing measures in March 2020 in response to the coronavirus disease 2018 (COVID-19) pandemic has reduced the transmission of severe acute respirator syndome coronavirus 2 (SARS-CoV-2) and alleviated the burden on the healthcare system[1,2]. However, this reduction has come at great economic, societal, and wider health costs[3–6]. As infection incidence has declined, countries have begun to ease restrictions in an attempt to reduce the societal and economic burden of lockdown. However, with infection incidence beginning to rise again in some countries, and with the spectre of a potential second epidemic wave in the Winter of 2020, countries must now strike a balance between easing restrictions and ensuring that the epidemic remains under control[7–11].

Multiple options, that could in combination form an exit strategy, have been proposed to allow easing of restrictions. These include: the widespread use of rapid, potentially app-based, contact tracing in combination with rapid testing and self-isolation[12,13]; expanded random testing to increase detection of asymptomatic infection[14–16]; strict quarantining of travellers on arrival[17,18]; and the use of face masks in high-risk environments[19–22]. Another potential component of a lockdown exit strategy that could allow for greater social interaction is the clustering of contacts beyond the household, commonly referred to as the social bubble or the 'double bubble' strategy[23–27]. Under this strategy, households are allowed to form a cohesive unit with one other household, generating a 'social bubble'; this allows individuals to increase their close, physical social interactions beyond their household while potentially limiting the risk of infection through the exclusivity of the bubble. A similar strategy has been implemented in some countries, including New Zealand and Germany, and is currently part of the lockdown exit strategy in the UK[28].

While physical distancing has placed additional pressures on society as a whole, some households are likely to be disproportionately more at risk of social isolation. Many adults in the UK will have been able to partially shift social contacts online and since 11th May (and 1st June) have been allowed to socialise outdoors with a maximum of one (and subsequently up to five) others while adhering to distancing guidelines[29]. However, such social contact replacements can be more difficult for children, for whom verbal interaction is only a small part of their communication with peers. Further, their carers have often had to balance working from home, childcare and homeschooling, generally without being able to access a support network from family, friends or professional childminders[30]. Single occupancy and single parent households have also likely been disproportionately affected as the complete absence of social face-to-face interactions for many months may impact mental wellbeing[31,32]. To address this, households with one adult have been allowed to form a 'support bubble' with another household in England and Northern Ireland since 13th June[33].

Here, using mathematical models, we assess the likely increase in transmission generated by various plausible social bubble strategies and use the UK as a case study. In particular, we compare the impact of limiting bubbles to those households who would benefit most (single occupancy households and those with young children) with allowing all households to form bubbles. We assess these changes in terms of both the increase in transmission (as characterised by the reproductive ratio, R) and short-term increase in fatalities.

Methods

Population

The model's synthetic population was created by generating individuals who are residents of one of 10,000 households. The size of the individual households, as well as the age distribution within households, was sampled to match that observed in the most recent census in England and Wales in 2011 (Figure 1)[34].

We used data from the 2011 census of England and Wales to construct a distribution of age-stratified household compositions in terms of 10-year age bands. Each household composition consisted of the number of individuals in each age band belonging to the household. We assigned probabilities to each composition observed in the census data based on the frequency of its appearance, and then used these probabilities to construct our simulated household population. This gave us a synthetic population whose age structure was comparable with that of England and Wales and whose household compositions reflected the observed correlations between the ages of household occupants. In particular, this formulation should realistically capture the generational structure of households in England and Wales, which we expect to be an important factor in transmission across age classes.

Transmission model

The transmission dynamics are set to simulate the status of COVID-19 interventions during 'lockdown' in the UK in May 2020 or future lockdowns to mitigate a second wave, in particular simulating contacts that are substantially reduced and largely household-based, with schools, non-essential retail, and leisure facilities closed. This is achieved through stochastic simulation of infections spreading through an interconnected population by generation; connections are captured by a matrix, A, which defines the probability that infection can pass between any two individuals in the population over an individual's entire infectious period. A is composed of the sum of two matrices H and B, which capture within-household and within-bubble transmission respectively (Figure 1). We can

Figure 1. top panel: schematic of model structure and its stratification into different household sizes with three components of transmission dynamics, community transmission, bubble transmission and household transmission; left panel: household size distribution for all households in England and Wales, for those households with at least one child younger than 20-years-old and for those with at least one child younger than 10-years-old (about primary school age and younger). Right panel: illustrative transmission probability matrix A, composed of household and bubble contacts and including community transmission.

then use the matrix A to drive forwards the stochastic dynamics using a next generation approach. To this, we add random (mean-field) transmission between individuals in the population to simulate the risk that infection in the wider community poses to the household and the social bubble, and vice versa. Transmission rates within the household and the wider community are matched to observed data on secondary attack rates and population-scale R estimates. We assume that households are adhering to current restrictions and social distancing, and therefore largely act as a coherent and largely isolated unit. We therefore assume that the risk of a household acquiring infection from the community is independent of its number of

occupants as observed in a cross-sectional serological study for SARS-CoV-2 in Germany in March and April 2020[35].

We assume that susceptibility to infection as well as transmissibility of infection can be age dependent, hence transmission rates across contacts depend on the age of both individuals.. There are two conflicting bodies of evidence about the potential role of children. Firstly, it has been observed that children are more likely to experience mild or no symptoms, and as such may have a lower transmission rate[36,37]. Secondly, cases with more severe symptoms are likely to self-isolate reducing their effective infectious period, therefore children that are asymptomatic (or

mildly symptomatic) may continue to transmit for longer[35]. In our base parameterisation, we assume that children are 50% as susceptible to infection as adults or elderly adults, but assume that transmissibility is independent of age; this echoes the assumptions of a previous model[9], but an alternative parameterisation based on other work[8] is considered as part of sensitivity analysis. We assume that transmission across close contacts (household or bubble) depends on the interaction between two individuals, and that the amount of 'interaction' an individual has with a close contact is inversely proportional to the number of that type of close contacts that person has[35,38,39]. Under this assumption, transmission within households and across households who share a bubble is frequency dependent.

Throughout, we compare the baseline model without the additional social interactions via bubbles (C1) with different ways in which bubbles could be allowed to form (scenarios 1–6, see below). To assess the effectiveness of social bubbles, compared to increasing contacts in an unclustered fashion, we consider two further comparison scenarios. In C2, individuals make the same number of infectious contacts with the population as in Scenario 6, but these are chosen randomly across the population. In C3, individuals' infectious contacts are resampled every generation. Scenario 6, C2, and C3 therefore represent fixed and clustered, fixed and unclustered, and variable additional contacts respectively.

Technical model summary

In this subsection we describe the model and its underlying assumptions in detail. Table 1 describes our notation, while Table 2 contains the formulae for transmission rates within our model. Key model parameters and assumptions are described below in Table 3.

Table 1. Model notation.

Symbol	Meaning
$C(i)$	Age-dependent susceptibility scaling factor of individual i
$T(i)$	Age-dependent transmissibility scaling factor of individual i
$N_H(i)$	Number of individuals in an individual i's household
τ_H	Baseline transmission rate across household contact
$\rho_H(i,j)$	Household transmission rate from individual j to individual i
τ_B	Baseline transmission rate across bubble contact, $\tau_B = k\tau_H$ where $k \in [0\ 1]$
$\rho_B(i,j)$	Bubble transmission rate from individual j to individual i
ε	Baseline mean-field transmission rate
$\varepsilon(i)$	Mean-field transmission rate to an individual i
$I_g(i)$	Infection status of an individual i at generation g. $I_g(i) = 0$ if i is susceptible at generation g. $I_g(i) = 1$ if i has been infected by generation g (so includes recovered individuals)
$S_g(i)$	$1 - I_g(i)$, i.e. the susceptibility status of an individual i at generation g.
N	Size of population.

Table 2. Transmission rates.

Transmission rate	Model formulation
Household, $\rho_H(i,j)$	$\dfrac{2T(j)C(i)\tau_H}{N_H(i)-1}$ if i and j are within the same household, 0 otherwise
Bubble, $\rho_B(i,j)$	$T(j)C(i)\left(\dfrac{\tau_B}{N_H(i)} + \dfrac{\tau_B}{N_H(j)}\right)$ if i and j are within the same bubble (but not the same household), 0 otherwise
Mean-field, $\varepsilon(i)$	$\dfrac{C(i)\epsilon}{N_H(i)} \dfrac{\sum_j T(j)(I_g(j) - I_{g-1}(j))N_H^{-1}(j)}{\sum_j N_H^{-1}(j)}$

Table 3. Key model parameters and assumptions.

Parameter	Description	Value (base case)	Value (Sensitivity)	Source
	Household structure and age distribution			34
τ_H	50% of the transmission rate for an adult, in a two person household	0.345 (20% SAR)	0.155 (10% SAR) 0.86 (40% SAR)	Calibrated to 40,41
τ_B	Transmission rate for an adult within the bubble	0.5 τ_H	1 τ_H 0.1 τ_H	*assumption*
	Relative transmissibility of a child and older adult vs adults	1 and 1	0.64 and 2.9	8,9
	Relative susceptibility of a child and older adult vs adults	0.5 and 1	0.79 and 1.25	8,9
	Infection fatality rate	In 10y age bands		42
R_e	Net reproduction number	0.8	0.7, 0,9	43
ε	Rate of infection from the community	1.13 (20% SAR)	1.29 (10% SAR) 0.925 (40% SAR)	Calibrated to R_e *given* τ_H

The expressions for household and bubble transmission rates are derived by assuming that transmission between close contacts depends on frequency of interaction between those individuals. We decompose interaction between individuals i and j into interaction led by i and interaction led by j. The amount of interaction led by i (or j) depends on the number of other close contacts i (or j) has. The total amount of interaction, hence the rate of transmission, is given by summing these interactions. For household contacts i and j, both individuals have the same number of household contacts, hence leading to the above formula, equivalent to the standard frequency dependence of transmission assumption. Across bubble contacts, interaction depends on the size of both households. To obtain equations for density-dependent transmission, the dividing N_H terms are omitted from the equations for $\rho_H(i,j)$ and $\rho_B(i,j)$.

Specific transmission rates also depend upon the transmissibility of the individual j transmitting infection ($T(j)$), and upon the susceptibility of the individual i receiving infection ($C(i)$), which are dependent on the age classes of individuals j and i.

The mean-field transmission to an individual, as well as an individual's contribution to mean-field infection, is inversely proportional to the number of individuals in their household, as we assume that a household acts as a coherent and largely self-contained unit when interacting with the population at large. $\varepsilon(i)$ also depends on the susceptibility of i, determined by their age class. The force of infection from the general population is given by $\sum_j T(j)(I_g(j) - I_{g-1}(j))N_H^{-1}(j)$, i.e.

the new infections in generation g, scaled by both the relative transmissibility of newly infected individuals and by their interaction with the general population determined by their household size.

By considering transmission as a Poisson process, we obtain the elements of the probability matrices H and B, the matrices of within household and within bubble transmissions respectively, by taking $H(i,j) = 1\text{-}e^{-\rho H(i,j)}$ and $B(i,j) = 1\text{-}e^{-\rho B(i,j)}$. A non-zero element within the matrix H (or B) indicates that the corresponding individuals are within the same household (or bubble). We obtain the overall probability matrix for the population by taking $A = H+B$.

In order to simulate an epidemic, we begin by randomly sampling the probability matrix A. Doing so, we retain only the infectious connections between individuals that will lead to an infection. These sampled probability matrices therefore represent *potential* transmission networks (though the exact transmission network for a given simulation will depend upon who is initially infected). Because these sampled matrices refer to transmission events rather than contacts, this matrix may be unsymmetric. We refer to the sampled matrix as A'. $A'(i,j) = 1$ denotes that individual j will infect individual i with probability 1, given individual j is infected. We initiate each simulation with the required number of infectious individuals for 1% of the population to be infected by generation 4. Initially infected individuals are chosen with probability proportional to their mean-field interaction, i.e. inversely proportional to their household size. Letting

$\underline{\mathbf{I}}_{\mathbf{g}}$ be the vector of infection statuses of individuals in generation g, we obtain the next generation by $\underline{\mathbf{I}}_{g+1} = \text{sign}((A'+Id) \times \underline{\mathbf{I}}_g)$, where Id is the identity matrix, and where sign() is an element-wise function equal to 1 for each positive element and 0 otherwise. Via this matrix multiplication, every newly infected individual in generation g infects all of their infectious contacts that generation. Here the identity matrix is added to impose that individuals do not become susceptible again after one generation, while the sign function is used to impose that individuals cannot be infected more than once. This process can be iterated until equilibrium is reached, and the epidemic has ended. To this, we also add mean-field transmission. Each generation, the number of new infections is calculated in order to calculate $\varepsilon(i)$ for each susceptible individual i, who is infected from mean field transmission with probability $1 - e^{\varepsilon(i)}$ each generation.

Recovery from infection is not explicitly modelled in the simulation, but rather is implicitly built into the structure of the model. If an individual i is infected in generation g, they will infect all of their transmission contacts in generation $g+1$ via the matrix multiplication. They also only contribute to community infection in generation $g+1$. While individual i remains 'infected' (with value 1), they no longer play any role in the infection dynamics, nor can they be reinfected. Hence, the simulation model assumes that individuals are infectious for one generation, before recovering with immunity.

Results are averages obtained from simulations of 1000 epidemics for 10 different sampled epidemic networks, hence results are averages of 10000 simulations.

Outcome metrics
We calculate two key metrics for the epidemiological impact of interventions in our household model with extended social contacts, which relate to epidemic risk and adverse health measures.

The net reproduction number (R) is a measure of risk of (increased) transmission that may eventually result in an exponential increase in infections and hence the need for stricter control measures if exceeding the epidemic threshold ($R>1$). R is defined as the number of secondary infections generated by a typical case. In models incorporating household structure, the typical case is effectively an average over the probability that such a case is the first, second, third or later generation case within the household[44]. Following the principle of Pellis et al.[45,46], we determine R numerically as the ratio of the number of new infections in the fifth to the fourth model generation, adjusted to account for the partial depletion of susceptibles. Specifically, defining $R(g)$ as:

$$R(g) = \frac{\sum_i I_{g+1}(i)}{\sum_i I_g(i)} \times \frac{N}{\sum_i S_g(i)}$$

we take $R(4)$ as R. In all simulations this provided sufficient time for the average state of infectious individuals to have stabilised at a value that persists over several generations (Figure 2).

Our second metric is the relative mortality (i.e. number of deaths), compared to the baseline model (C1) of isolated households; this provides a measure of adverse health impacts as a result of increased contact rates in the respective scenarios. We use age stratified infection fatality rates (IFR) estimated from repatriation flights early in the COVID-19 pandemic[42,47] to predict the mortality risk in the five generations following model burn in (i.e. from the fourth to ninth model generation - approximately the second month after social bubbles were initiated). Each simulation is initiated with the required number of infectious individuals for 1% of the population to be infected by generation 4, in order for the fatalities following this generation to be meaningfully compared.

Parameterisation
To parameterise the COVID-19 transmission dynamics in the model we need to define the infection dynamics within a household, within a bubble and from the community. To parameterise the within household transmission we assume that, in line with observations from contact tracing while accounting for some underreporting[40,41,48], the secondary household attack rate (SAR_{HH}) is 20%. This is achieved by tuning the transmission rate
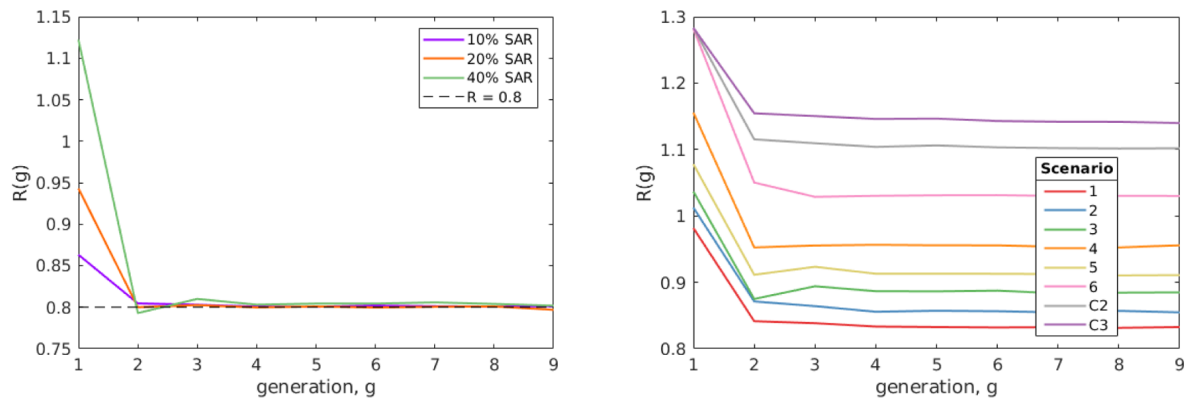
Figure 2. Numerical exploration of R by generation. Left shows examples of the method for which ε was fitted to satisfy $R(4) = 0.8$ under our baseline parameters for different values of SAR_{HH}. Right shows $R(g)$ by generation from each of our scenarios from our baseline assumption. In both plots, $R(g)$ decreases over the first few generations, before reaching an equilibrium value that persists over multiple generations.

(τ_H) between household members to achieve this average attack rate. We subsequently assume that community transmission is such that, in combination with household transmission, the model generates an overall reproduction number of 0.8, similar to estimates from mid-May 2020 in the UK[43,49]. Further, as a base-case, we assume that transmission between households within the same bubble is 50% lower than that within a household, i.e. $\tau_B = \frac{1}{2} * \tau_H$. In our base case parameterisation a 3.75-fold increase in community contacts yielded a reproduction number of about 2.5; this is in line with an approximate 70% reduction in contacts during lockdown and a reproduction number of about 2.5 in the early phase of the pandemic with hardly any distancing measures in place[49].

We additionally assume that all eligible households would take up the opportunity to expand their contacts and enter into a social bubble with one other household, and that they would adhere to the exclusivity of this bubble. The impact of only partial uptake is explored in our results, and the impact of non-adherence, incorporated by allowing 50% of eligible households to form an additional social bubble, is explored in our sensitivity analyses 5

Model parameterisation
All analyses were done in MATLAB 2019B[50] and R[51] v3.6.3 and are available via GitHub[52]. The R packages reshape2 v1.4.3, tidyverse v1.3.0 , and gridExtra v0.8.1 are required to generate plots. The analysis performed using MATLAB can be replicated using the open MATLAB compatible software GNU Octave, with minimal adjustments, outlined in the GitHub's ReadMe[52].

Scenarios modelled
We considered a number of contact clustering strategies of how bubbles could be allowed under any relaxation to lockdown measures:

1) Allow households with children younger than 10-years-old (about primary school age or younger) to pair up

2) Allow households with children younger than 20-years-old to pair up

3) Allow single occupancy households to pair up with another single occupancy household

4) Allow adults who live alone or with dependent children only to pair up with another household of any size in a 'support bubble'

5) A combination of Scenarios 1 and 3

6) Allow all households to pair up with one other household

All these scenarios assume that the pairing will occur at random between permitted households.

We compare the above scenarios against three counterfactuals that do not include social bubbles. These allows us to elucidate the impact of

C1) Perfect adherence to the current household-only contact strategy (other than the background transmission risk from the community)

C2) All individuals increase their number of contacts so that population level force of infection matches that of Scenario 6. Contacts are unclustered and chosen at random across the population but stay the same over time.

C3) All individuals increase their number of contacts so that population level force of infection matches that of Scenario 6. Contacts are unclustered and chosen at random across the population and are re-sampled at each generation.

Counterfactuals 2 and 3, maintain the same level of additional contacts outside the home as social bubbles but change how they are distributed.

C2 is obtained by taking the sampled infectious contacts from Scenario 6, then rewiring these sampled contacts. Doing so keeps the number of secondary infections from any individual constant across both scenarios. As the sampled contact matrices represent transmission networks, edges may be either directed or undirected. Directed and undirected links are rewired separately, so that C2 has the same number of undirected links (*i* would infect *j* and *j* would infect *i*) as in Scenario 6. C3 is obtained, like in C2, by sampling infectious contacts from Scenario 6 then rewiring these sampled contacts. However, in this situation, all links are treated as directed, and hence the number of undirected links diminishes, reflecting that an individual chooses new bubble contacts each generation. This rewiring reflects that edges are resampled each generation. We also use this method to produce scenario specific counterfactual scenarios, e.g. C2 and C3 for Scenario 1, to assess the effectiveness of bubbling.

Sensitivity analyses
Other than the previously described base case we performed a number of univariate sensitivity analyses to test the robustness of our findings to the underlying assumptions. Specifically, we assume that the current R is 0.7 or 0.9 instead of 0.8[43]; that the secondary attack rate in the household is 10% or 40% instead of 20%[41]; that transmission between individuals in the same bubble (but different households) is 10% or 100% of that within a household instead of 50%; that the risk of a household to get infected with SARS-CoV-2 from the community increases with increasing household size instead of being independent; that transmission across close contacts is density-dependent instead of frequency-dependent; that 50% of bubbles do not adhere to the recommendations but also form bubbles with an additional household rather than perfect adherence; that households including an individual over 70-years-old do not form bubbles; and that the relative susceptibility to infection of children and older adults compared to adults is 79% and 125% while the relative transmissibility is 64% and 290%, respectively[8,53].

We model non-adherence to the strategy by allowing 50% of eligible households to enter into close contact with an additional household. Doing so means that bubbles are no longer mutually exclusive, and that chains of transmission could potentially span many households. Letting B_2 denote the probability

matrix of additional bubbles through non-adherence, A is now obtained by the sum of H, B, and B_2.

Results

Households

From the 2011 census of England and Wales, the average size of a household was 2.36 persons. Considering households with at least one child under 10-years-old, the average household size increases to 3.89 persons, and 30.4% of the population live in such households. For households with at least one child under 20-years-old, the average household size is 3.73 persons, and 49.5% live in such households. In total, 37% of households are occupied by someone over the age of 60 years, and 50% of single occupancy households were occupied by such older adults. Single occupancy households comprise 30.2% of households. There is limited multi-generational mixing, with only 3.6% of households having both a child aged under 10 years and an adult aged over 60 years.

Impact of social bubble strategies on epidemic risk

Assuming an initial reproduction number of 0.8, perfect adherence to the recommended social bubble strategy and that all eligible households indeed pair up, we find that strategies that exclusively target single-person households (scenario 3) or households with young children (scenario 1) do not increase transmission substantially (R of 0.83 and 0.89 respectively in the base case scenario); their combination (scenario 5) is also predicted to only marginally increase transmission in the community (R of 0.91) (Figure 2). For these two targeted strategies, even under conservative assumptions ($SAR_{HH} = 40\%$, $\tau_H = \tau_B$), the increase in transmission is unlikely to lead to substantial spread of COVID-19 (R of 0.95 and 0.91 for scenario 1 and 3, respectively).

However, allowing all households to form bubbles (scenario 6) is estimated to increase the reproduction number to 1.02, and hence beyond the critical threshold value of 1 for the base case scenario (Figure 3).

Generally, the fewer households that were deemed eligible for expanding their social bubble under a specific strategy, the smaller the average household size of those involved. and the smaller the risk of transmission within the bubble, the smaller the increase in transmission as a result. The impact of social bubbles also depends on uptake - we find that R increases sublinearly with uptake (Figure 4).
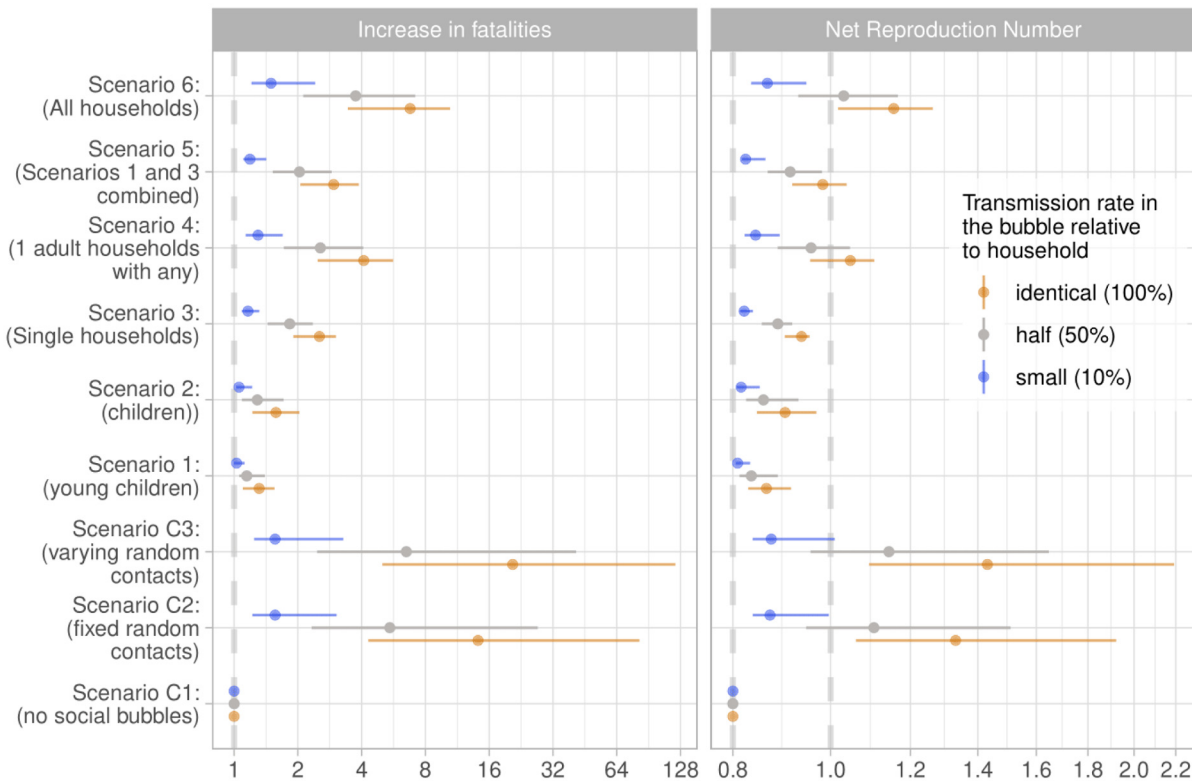
**Figure 3. Estimated reproduction number and increase in fatalities for the considered scenarios under the assumption that all eligible households pair up and thereby form exclusive social bubbles and that transmission rates within a social bubble are the same as within the household.** Central estimates are assuming SAR_{HH}=20% and the upper and lower limits represent the respective 10% and 40% assumption.

The impact bubbles have on epidemic risk depends upon the levels of transmission within the population prior to introducing a bubble strategy. We find that the impact of bubble strategies on transmission scales linearly with the prior R value, *but* for some strategies with a gradient > 1, meaning that the higher the level of community transmission within the population the larger the increase in R from allowing social bubbles (Figure 5).

The impact of social bubble strategies on mortality risk
The average age in the households eligible to form social bubbles in scenarios 1 to 6 was 21.8, 25.6, 58.1, 40.2, 32.2, and 39.4 years, hence the average infection fatality risk in an average household member implementing such a strategy was 0.09%, 0.14%, 2.36%, 1.05%, 0.74%, and 0.93%. In all scenarios, the increased number of contacts lead to both excess infections and fatalities. Excess risk for infection compared with no
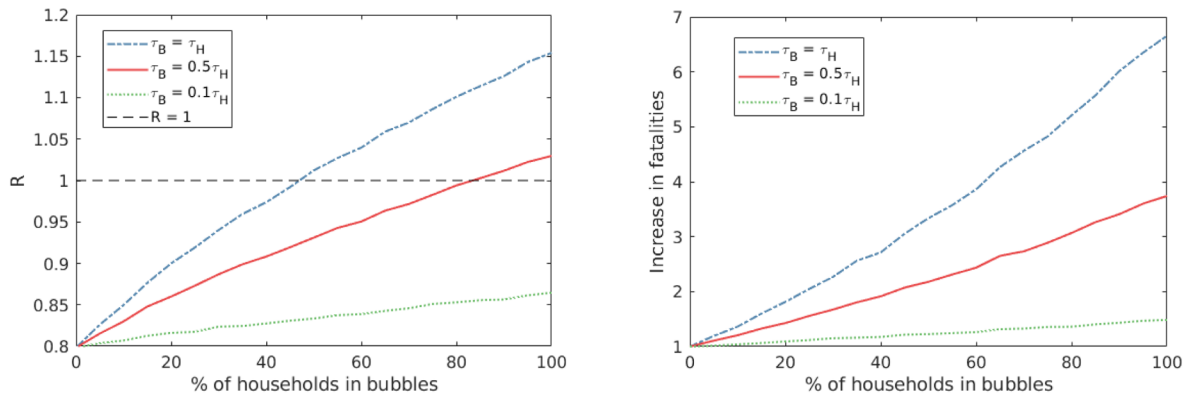
Figure 4. The impact of uptake on R and fatality. Here we consider the impact varying levels of uptake has on the reproduction number, R, and relative mortality. We consider this for our baseline parameters, at varying levels of transmission across bubble contacts ($\tau_B = \tau_H$ in blue, $\tau_B = 0.5\ \tau_H$ in red, $\tau_B = 0.1\ \tau_H$ in green). We observe that R scales sublinearly with uptake, with the gradient of increase dependent on transmission rate across bubble contacts.
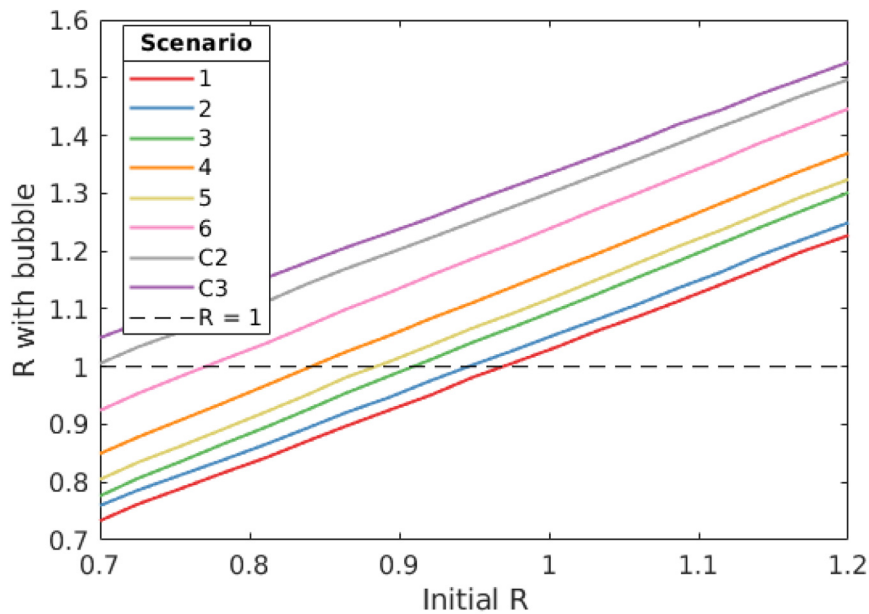
Figure 5. The relationship between initial R and R under different bubble scenarios. Here we consider the impact different bubble strategies have on the reproduction number, R. We consider this for our baseline parameters. For scenarios 3 to 6, we find that R with bubbles increases linearly with initial R with a gradient above 1, meaning that as initial R increases, the greater the increase bubble strategies have on R.

social bubbles (Scenario C1) was seen in households implementing the social bubbles as well as those households who were not eligible, although, as expected, the relative risk for infection was higher in eligible households (Figure 6). The resulting excess mortality risk depended highly on the estimated epidemic risk but also on the average age of the affected households.

For example, while social bubbles among households with young children (Scenario 1) saw the similar increases in infections to increases in deaths (with a risk ratio of 1.13 and 1.14 for infections and deaths respectively), social bubbles targeting single occupancy households saw a larger increase in deaths than infections (with a risk ratio of 1.26 and 1.83 for infections and deaths respectively) due to the older targeted demographic. In scenarios targeting families, the mortality risk was largely attributed to households not eligible to form social bubbles (Figure 6).

Effectiveness of social bubbles
The forming of social bubbles was effective at reducing the infection and thereby the mortality risk compared to strategies that increased contacts in a less clustered way: under base case assumptions all households forming social bubbles (Scenario 6) reduced the mortality risk by 30.9% and 42.4% compared to adding the same amount of contacts randomly (Scenario C2) and time varying (Scenario C3).

In general, the added benefit of social bubbles increases with the higher proportion of eligible households, alongside targeting riskier populations. For example, social bubbles for households with young children (Scenario 1) reduced mortality risk by 4.2% and 8.1% compared to those households increasing

contacts randomly and time-varying. In contrast, allowing households with one adult to form a support bubble with another household (Scenario 4) results in 51% of the households entering into a bubble, and leads to a 27.7% and 39.3% reduced mortality risk compared to those households increasing their contacts randomly and time-varying respectively (Figure 7).

Sensitivity analyses
We tested the robustness of our findings to a number of alternative assumptions governing the spread of SARS-CoV-2 and the implementation of the social bubble strategy. Within the tested parameter space, the alternative assumptions did not qualitatively change our findings. The two main factors that increased or decreased the epidemic risk were an initial value of R closer to 1 when implementing the strategy and a much higher than typically observed secondary household attack rate. However, for Scenarios 1 and 3, in neither of the univariately tested parameterisations did R exceed 1 (Figure 8). The assumptions on age stratified susceptibility and transmissibility were conservative for strategies focussed on households with children and were optimistic for single-person households; and vice versa for the assumption that the risk for community transmission was independent of household size. The epidemic risk from social bubbles is further reduced if within bubble transmission is reduced to 10% of that within household transmission. The alternative assumptions surrounding close-contact transmission and community transmission changed the ordering of risk of social bubble strategies. Under our baseline assumption, Scenario 1, targeting families with young children, resulted in the lowest increase in R; under these alternative assumptions, Scenario 3, targeting single-occupancy households, resulted in the lowest increase in R. The corresponding tornado diagrams for all other scenarios are included as extended data[52].
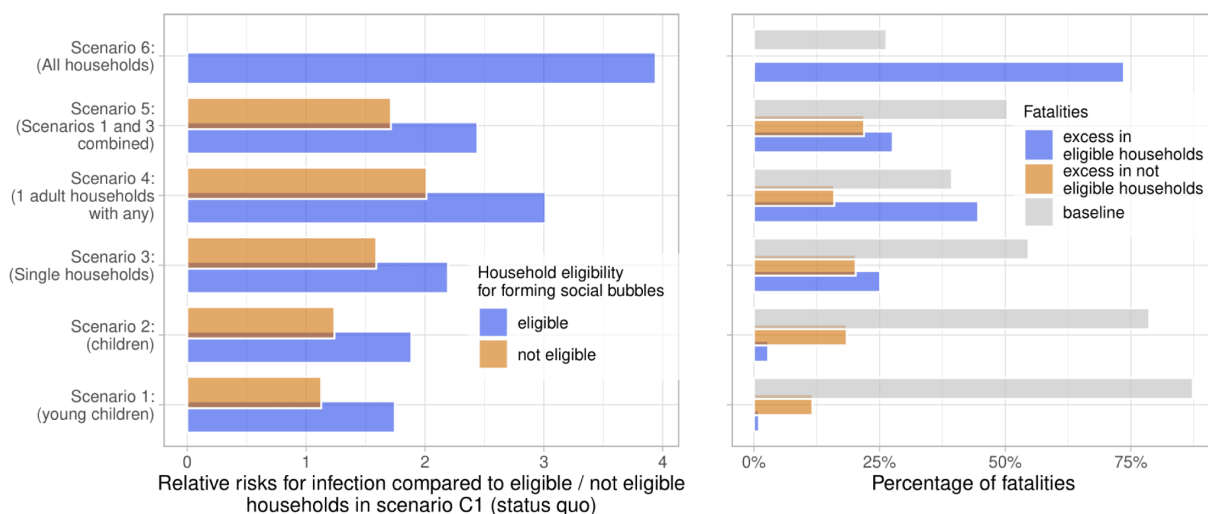
Figure 6. Relative risk of infection and fatality. Left panel: the relative risks for infection in the considered scenarios if compared to the status quo with no social bubbles (Scenario C1), stratified into the risks in households eligible and not eligible for forming social bubbles. Right panel: the population attributable fraction of fatalities in the considered scenarios. The overall mortality risk is stratified into the baseline risk, and the excess risk from forming social bubbles in both eligible and non eligible households.
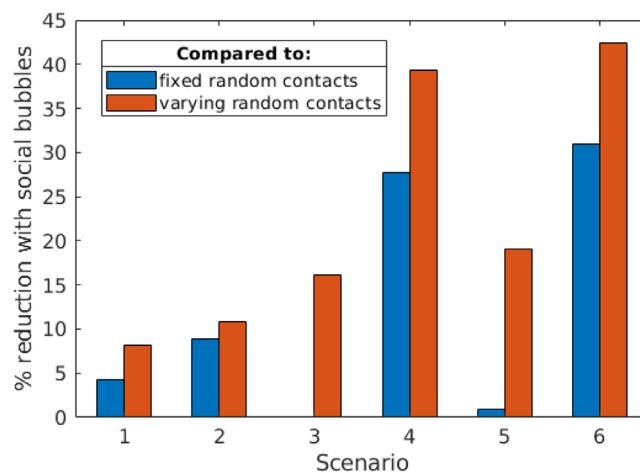
Figure 7. Scenario specific effectiveness of social bubbles. Here we compared the effectiveness of social bubbles in reducing mortality risk, when compared to other ways of increasing social contacts - where individuals from eligible households either make fixed random contacts (blue) or varying random contacts every generation (orange). In each comparison, individuals make the same number of infectious contacts, so the reduction in fatalities can be attributed to the clustering implied by social bubbles.

The effectiveness of social bubbles also varied according to the underlying parametric assumptions. Assuming our alternative assumptions around susceptibility and infectivity, the effectiveness of social bubbles was as large as a 46.1% and a 58.5% reduction in mortality risk compared to adding the same amount of contacts randomly (Scenario C2) and time varying (Scenario C3). Under our most conservative assumptions, the reductions in mortality risk compared to C2 and C3 were 87.2 and 91.3%.

Alongside the parameter sensitivity scenarios considered, we also considered each scenario where older adults were shielded and excluded from being allowed to form a bubble as a sensitivity analysis. This only has a small impact on the effect of social bubbles for families with children (Scenarios 1 and 2), because of the small amount of multi-generational mixing between households in the UK, but does reduce R for social bubbles for single occupancy households or all households (Scenarios 3–6). While shielding older individuals does decrease overall mortality risk, such a strategy still impacts older individuals; bubbling strategies increase overall cases, which in turn increases risk to older individuals through community transmission.

Another potential strategy would be to allow all households with two or fewer adults to form a bubble with households of any size - an extension to the current situation in England. However, as 87.7% of households have two or fewer adults, such a policy would result in 98% of households forming bubbles, and hence such a policy would have largely the same impact as allowing all households to form bubbles.

Discussion

We found that contact clustering, or the forming of social bubbles that join two households, can allow increased social contacts beyond the households while limiting additional risk of transmission. In the base case social bubbles reduced the mortality risk by 42% compared to a scenario that increased contacts by the same amount but without clustering thereof, and risk reduction under some alternative parameterisations was even higher. Allowing all households to form social bubbles may increase R above its epidemic threshold and hence lead to an increase in cases. A strategy that sees only those at potentially the highest need for an extension of their contacts beyond the household (families with young children and single-person households) should lead to a limited increase in epidemic risk (less than 11% individually and less than 15% in combination), which remained below the epidemic threshold in most scenarios considered. The epidemic risk can be further reduced if the transmission risk within the bubble is minimised. As the number of contacts and R increase with a social bubble strategy, so does the risk of adverse health outcomes. We find that adverse health outcomes are largely proportional to the epidemic risk, but will disproportionately affect households with older adults independently of their clustering behaviour.

Stringent physical distancing policies in many countries have reduced R from about 2.5 to just under 1[11,49,54]. This provides the opportunity to risk a small amount of additional contacts without necessarily experiencing an increase in COVID-19 cases, if crossing the epidemic threshold can be avoided. Here, we investigate the effectiveness of social bubbles as a potential option to ease the social impact of the lockdown without increasing transmission risks. However, while we consider the impact of social bubbles in isolation, such a policy would only be one part of a multi-variable exit strategy[28]. Hence, our comparisons of alternative bubble strategies against the epidemic threshold should be interpreted cautiously and in consideration of the other changes to behaviour, such as re-opening of non-essential retail and travel. It is likely that these other activities will combine
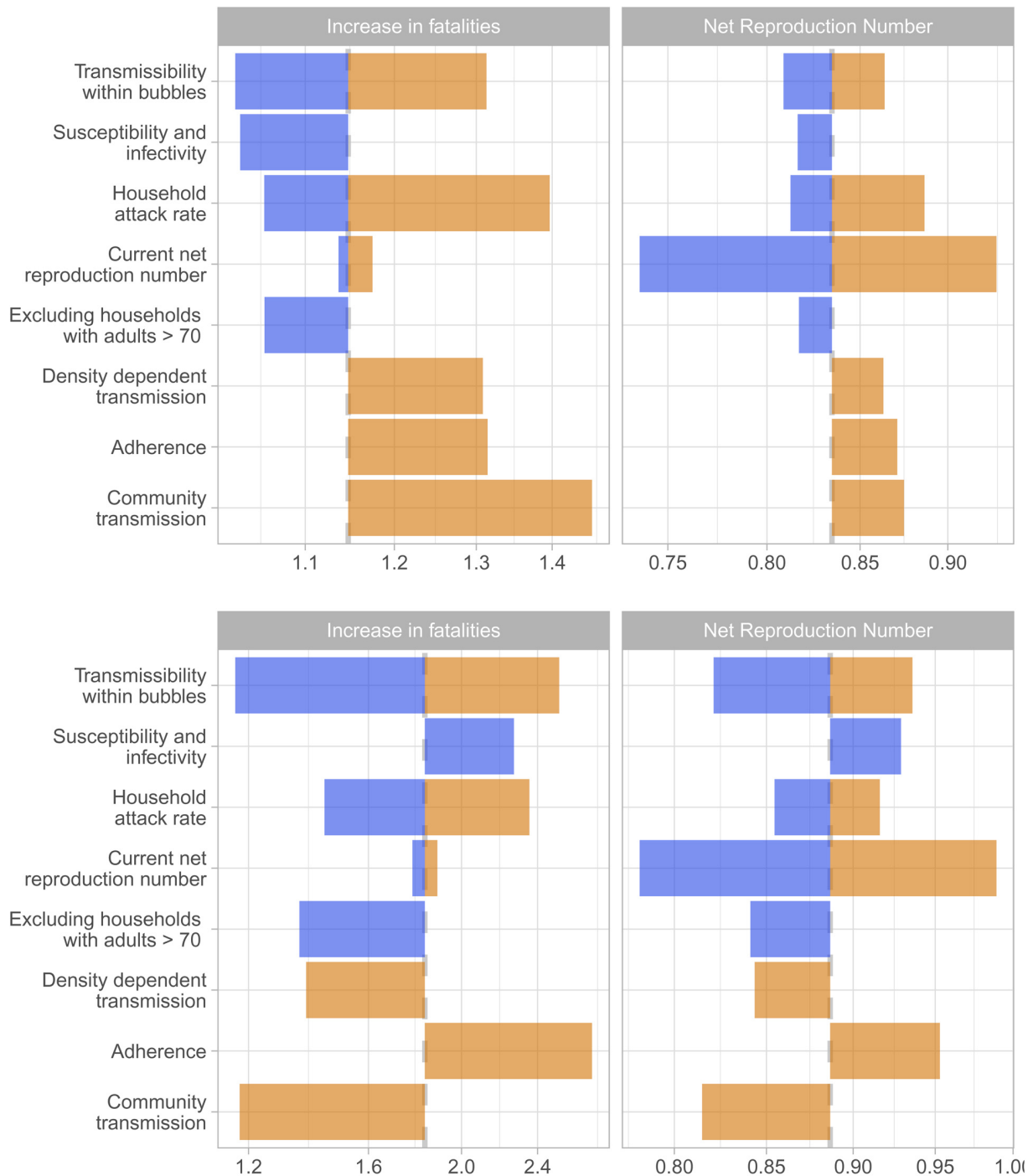
Figure 8. Sensitivity analyses. The tornado diagram shows a univariate sensitivity analysis on the expected increase in fatalities and the net reproduction number for scenario 1 (above) i.e. allowing households with young children to pair up, and for scenario 3 (below), i.e. allowing single occupancy households to pair up. The color coding is based on factors determining higher risk (orange) and lower risk (blue) for Scenario 1. The base case estimate is indicated through the dashed grey vertical line. The sensitivity scenarios are (from top to bottom): transmission across individuals of households sharing a bubble is 90% or 0% lower than that within a household instead of 50%; the relative susceptibility to infection of children and older adults compared to adults is 79% and 125% while the relative transmissibility is 64% and 290%; the secondary attack rate in the household is 10% or 40% instead of 20%; R_e is 0.7 or 0.9 instead of 0.8; that households including an adult over 70-years-old are excluded from forming bubbles; density-dependent transmission across close contacts instead of frequency-dependent transmission; that 50% of bubbles do not adhere to the recommendations but pair up with an additional household; and that the risk of a household to get infected from the community is proportional to the household size instead of being the same across households.

in a non-linear manner so that bubbling in a context of children returning to school might affect results.

Countries including Germany and New Zealand have implemented strategies similar to those considered here. In Bavaria, Germany, in early May and before the reopening of schools and nurseries, up to three households could form exclusive groups to share childcare amongst them[55]. Even during their highest national alert level, level 4 "Lockdown", New Zealand permitted people living alone to pair up with a "lockdown buddy" and key workers to identify "childcare buddies". New Zealand moved to level 3 in their COVID-19 alert system, "Restrict", on 27 April 2020, which included the advice to residents to stay within their household bubbles but permitted expansion of such to reconnect with close family, bring in caregivers or support isolated people[56]. A subsequent survey found that among respondents the highest increase in the quality of life by far would not be brought by re-opening of schools, shops, churches or fitness centres, but by allowing households to re-connect[57]. It also found that in going to alert level 3 only 50% of households took up the opportunity to expand their social bubbles and that there was high awareness of the importance of the exclusivity of the bubbles, with only 7.5% of bubbles reporting to have had contacts outside their bubble.

We identify three key risks to the success of social bubbles that may increase their epidemiological risk: potential lack of adherence, a higher than observed secondary household attack rate and being too close to the epidemic threshold. If the risk perception of the population changes as a result of allowing parts of the population to form social clusters, a lack of adherence to the exclusivity of the bubbles could lead to rebuilding of contact networks that in turn lead to the epidemic threshold being crossed. We find that some degree of non-adherence would not necessarily hinder the success of the strategy, but communication of the strategy is likely to be key. For example, in New Zealand, the social bubbles were not framed as a relaxation of social distancing rules but rather as a source of support for those who are at a higher risk of social isolation or with needs for care, including childcare[57]. We find that if the secondary household attack rate is substantially higher than assumed in our base case the epidemic risk is elevated close to the epidemic threshold. There remains some uncertainty surrounding the household attack rate of COVID-19, with high household attack rates observed in some instances. However, our base case assumptions are in line with an increasingly consistent picture emerging in the contemporary academic literature[41]. Also, superspreading events have been raised as a potentially important source for sustained transmission of SARS-CoV-2, which would further imply a rather low secondary household attack rate in most instances[58,59]. However, household attack rates may vary between different types of household, and may be larger for some households with unusual network structures[60], such as large student households. Similarly, if the R is very close to its epidemic threshold, an increase in contacts, even if clustered, could result in an increase in cases. Hence, careful monitoring of such is needed to assess the feasibility of expanding social bubbles.

An expansion of contacts into social bubbles will naturally lead to some increase in transmission in comparison to perfect adherence to the recommendation to restrict to all but essential contacts outside the household. However, such adherence may decline as a result of extended periods of time in lockdown and lead to an expansion of contacts that are unclustered, potentially leading to long chains of transmission. To illustrate such a scenario, we include alternative comparisons for the strategy that allows all households to form social bubbles (Scenario 6). We considered strategies that would have the same overall increase in transmission as in that scenario but where either the contacts or not cluster but stay fixed over time (Scenario C2) or where contacts are not clustered and vary over time (Scenario C3). We show that the clustering reduces the epidemic and reduces the number of infections and subsequent fatalities by 30.9% and 42.4% in the base case and even more in some of the parametric sensitivity analysis. Hence social bubbles, if given as a guidance to households who are struggling to cope with the lockdown, may give these households a safer alternative and thereby help to reduce the epidemic and mortality risk. This may particularly be the case for households with single parents or parents who cannot easily work from home; in such circumstances allowing social bubbles may help increase equity in the impact of the lockdown.

As in any epidemiological modelling study, we must make some assumptions surrounding transmission. Firstly, we assumed that transmission across close contacts was frequency dependent, informed by previous studies that indicate the probability of infection across two specific members of the same household decreases with household size for COVID-19 (and other communicable diseases like influenza). However, there remains uncertainty surrounding the nature of close-contact transmission for COVID-19, and the assumption of frequency dependence may not accurately capture transmission across all settings. Secondly, we assumed that the risk of a household acquiring infection from the community is independent of its number of occupants. While this assumption may be appropriate for some households (e.g. families where one adult leaves the household to do shopping), it may not hold true in other contexts (e.g. student households comprising of largely independent individuals). Because of this we tested the sensitivity of our results to these assumptions by alternatively considering close contact transmission as density dependent, and allowing the risk of household infection to increase with household size. These alternative assumptions do not qualitatively change our findings, but do change the ordering of the risk of social bubble strategies. In particular, under either of these alternative assumptions, social bubble strategies targeting single occupancy households result in a lower increase in R than strategies targeting families with children. Reports on the antibody prevalence in England found household size associated with antibody prevalence, suggesting that household size may play a role in the probability of acquiring infection[61]; a more detailed understanding of the nature of close contact and community transmission may help inform more precise evaluations of the effectiveness of social bubbles. We do not consider the risk of community transmission depending on bubble size. However,

if bubbles were to act as cohesive units, and as a consequence reduce their interaction with the community, this may further increase the effectiveness of social bubble strategies.

Our analyses have a number of limitations. Firstly, we only assessed the risk of extending social bubbles but not the benefits. As of June 2020 in England, social contact beyond the immediate household is restricted to virtual contact or contact in open spaces with up to five individuals while keeping 2 metres apart. In other words, one can have a conversation. While conversations are a large part of the social contacts of adults they have little role in the social interactions of young children. Hence the benefit of extending bubbles for children is likely disproportionately higher. Furthermore, clustering contacts into social bubbles is likely to ease contact tracing, which is an integral part of both containment and lockdown exit strategies. We considered social bubbles against the background of a lockdown, particularly where schools are closed. As lockdown measures are eased and schools are gradually re-opened forming social bubbles that largely overlap with societal contacts (e.g. forming social clusters with families that have children going to the same class) is likely further reducing the additional epidemic risk from social bubbles. By the same token, as lockdown eases, social bubbles aligned with other types of contacts, such workplace contacts or sport team contacts, would likely improve the efficacy of any bubbling strategy[27]. We also did not include the possibility to form bigger social bubbles that would cluster together three or more households. While this has been implemented in other countries, the complexity of creating an exclusive cluster of three or more households could lead to a loss of adherence. We did not consider further heterogeneity within society that may affect both risk of transmission and adverse health outcomes. For example, about 20% of the working population is classified as key workers and will have an increased risk for infection from the community, while adverse health outcomes have disproportionately affected individuals of low socioeconomic status. Further, we did not consider age-homogeneous mixing when pairing up households into bubbles, which may lead to a further layer of clustering and thereby reduce the mortality risks associated with the bubble strategy.

Conclusions

Our analyses highlight the continued need for social distancing despite a social bubble strategy being an effective way to expand contacts while limiting the risk of a resurgence of cases. Recommending social bubbles only for those who particularly struggle with a lockdown, while minimising opportunities for spread through prioritising outdoor settings for gatherings and adhering to distancing recommendations as much as possible, may strike an effective balance between minimising the impact on mental health and minimising the risk of a resurgence of cases. With the increased number of local lockdowns and the risk of a second wave in the Winter of 2020, social bubbles may again become a vital tool to provide social interactions to those that need it most whilst keeping R below one.

Data availability
Source data
The synthetic population for the study was constructed by from the 2011 (27 March) census, from the table '*SN 8637 - 2011*

Census Ad Hoc Household Composition (Age Groups) Safeguarded Tables (Lower Layer Super Output Area): England and Wales. This table is safeguarded, and can only be accessed by approved researchers. However, we have included our synthetic population in our repository as a Matlab workspace, 'FullCensusHouseholdWorkspace.mat', which can be used to regenerate results and to derive population level statistics. Alternatively, Office for National Statistics data for the age distribution of households in ten year age bands containing six individuals or fewer are available as dataset CT1088, (https://www.ons.gov.uk/search?:uri=search&:uri=search&q=CT1088*). Households of size six or less account for 98.2% of the households in England and Wales, and contain 97.8% of their combined population, so the loss accuracy induced by this cutoff is likely to be minimal.

Underlying data
The simulation model and analyses from this study are available via GitHub https://github.com/tsleng93/SocialBubble/tree/v1.1

Archived code at time of publication: http://doi.org/10.5281/zenodo.4605475[52]

To generate the underlying data for the plots in Figures 3, 8, and the extended data, users should run the code 'MainCode.m'. After running this, users should then run 'DataMaker.m' to produce the underlying .csv files.

To generate the underlying data[52] for the plots in Figure 6, users should run the code 'figure6Code.m', which runs the code and produces the underlying .csv files.

The R code 'plots.r' can be run as it is in order to generate the exact plots of Figures 1, 3, 6, 8 and the extended data[52]. To generate analogous plots from regenerated data, the .csv files should be replaced with those generated from 'MainCode.m' and 'figure7Code.m'

To generate the data and plots for Figure 2, Figure 4, Figure 5, and Figure 6 users should run the code 'figure2Code.m', 'figure4Code.m', etc.

Extended data
Zenodo: tsleng93/SocialBubble: SocialBubble. http://doi.org/10.5281/zenodo.4013702[52]

This project contains the following extended data:
- main_tornado_Scenario_2.pdf (univariate sensitivity analysis for Scenario 2)
- main_tornado_Scenario_4.pdf (univariate sensitivity analysis for Scenario 4)
- main_tornado_Scenario_5.pdf (univariate sensitivity analysis for Scenario 5)
- main_tornado_Scenario_6.pdf (univariate sensitivity analysis for Scenario 6)
- main_tornado_Scenario_C2.pdf (univariate sensitivity analysis for Scenario C2)
- main_tornado_Scenario_C3.pdf (univariate sensitivity analysis for Scenario C3)

Acknowledgements

References

1. Flaxman S, Mishra S, Gandy A, *et al.*: **Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries.** Imperial College London. 2020.
    Reference Source

2. Jit M, Jombart T, Nightingale ES, *et al.*: **Estimating number of cases and spread of coronavirus disease (COVID-19) using critical care admissions, United Kingdom, February to March 2020.** *Euro Surveill.* 2020; **25**(18): 2000632.
    PubMed Abstract | **Publisher Full Text**

3. Cluver L, Lachman JM, Sherr L, *et al.*: **Parenting in a time of COVID-19.** *Lancet.* 2020; **395**(10231): e64.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

4. Forman R, Atun R, McKee M, *et al.*: **12 Lessons Learned from the Management of the Coronavirus Pandemic.** *Health Policy.* 2020; **124**(6): 577–580.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Appleby J: **Tackling covid-19: are the costs worth the benefits?** *BMJ.* 2020; **369**: m1496.
    PubMed Abstract | **Publisher Full Text**

6. Staniscuaski F, Reichert F, Werneck FP, *et al.*: **Impact of COVID-19 on academic mothers.** *Science.* 2020; **368**(6492): 724.
    PubMed Abstract | **Publisher Full Text**

7. Gilbert M, Dewatripont M, Muraille E, *et al.*: **Preparing for a responsible lockdown exit strategy.** *Nat Med.* 2020; **26**(5): 643–4.
    PubMed Abstract | **Publisher Full Text**

8. Keeling MJ, Hill EM, Gorsich EE, *et al.*: **Predictions of COVID-19 dynamics in the UK: short-term forecasting and analysis of potential exit strategies.** *medRxiv.* 2020; 2020.05.10.20083683.
    Publisher Full Text

9. Davies NG, Kucharski AJ, Eggo RM, *et al.*: **Effects of non-pharmaceutical interventions on COVID-19 cases, deaths and demand for hospital services in the UK: a modelling study.** *Lancet Public Health.* 2020; **5**(7): e375–e385.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

10. **Coronavirus (COVID-19): framework for decision making - further information.** gov.scot. (accessed 30 May 2020).
    Reference Source

11. Ferguson N, Laydon D, Nedjati Gilani G, *et al.*: **Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand.** Imperial College London. 2020.
    Publisher Full Text

12. Ferretti L, Wymant C, Kendall M, *et al.*: **Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing.** *Science.* 2020; **368**(6491): eabb6936.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. Kucharski AJ, Klepac P, Conlan AJK, *et al.*: **Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study.** *Lancet Infect Dis.* 2020; **20**(10): 1151–1160.
    PubMed Abstract | **Publisher Full Text**

14. He X, Lau EHY, Wu P, *et al.*: **Temporal dynamics in viral shedding and transmissibility of COVID-19.** *Nat Med.* 2020; **26**(5): 672–5.
    PubMed Abstract | **Publisher Full Text**

15. Liu Y, Centre for Mathematical Modelling of Infectious Diseases nCoV Working Group, Funk S, *et al.*: **The contribution of pre-symptomatic infection to the transmission dynamics of COVID-2019 [version 1; peer review: 3 approved].** *Wellcome Open Res.* 2020; **5**: 58.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

16. Peto J, Alwan NA, Godfrey KM, *et al.*: **Universal weekly testing as the UK**

COVID-19 lockdown exit strategy. *Lancet.* 2020; **395**(10234): 1420–1421.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

17. Clifford S, Pearson CAB, Klepac P, *et al.*: **Effectiveness of interventions targeting air travellers for delaying local outbreaks of SARS-CoV-2.** *J Travel Med.* 2020; **27**(5): taaa068.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Kraemer MUG, Yang CH, Gutierrez B, *et al.*: **The effect of human mobility and control measures on the COVID-19 epidemic in China.** *Science.* 2020; **368**(6490): 493–7.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. Clase CM, Fu EL, Joseph M, *et al.*: **Cloth Masks May Prevent Transmission of COVID-19: An Evidence-Based, Risk-Based Approach.** *Ann Intern Med.* 2020; **173**(6): 489–491.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

20. Lustig SR, Biswakarma JJH, Rana D, *et al.*: **Effectiveness of Common Fabrics to Block Aqueous Aerosols of Virus-like Nanoparticles.** *ACS Nano.* 2020; **14**(6): 7651-7658.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

21. Worby CJ, Chang HH: **Face mask use in the general population and optimal resource allocation during the COVID-19 pandemic.** *Nat Commun.* 2020; **11**(1): 4049.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. Liang M, Gao L, Cheng C, *et al.*: **Efficacy of face mask in preventing respiratory virus transmission: a systematic review and meta-analysis.** *Travel Med Infect Dis.* 2020; **36**: 101751.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

23. Cheetham J: **The 'social bubble' approach to lockdown easing.** BBC News. 2020. (accessed 12 May 2020).
Reference Source

24. Hinsliff G: **Social distancing isn't going to end soon. So how do we live with it? Gaby Hinsliff.** *The Guardian*. (accessed 12 May 2020).
Reference Source

25. Jones A: **Two households will be able to meet up in the garden under new Government plans.** *The Telegraph*. 2020. (accessed 28 May 2020).
Reference Source

26. Flasche S: **Of hairclips and coronavirus - how contact clustering may allow a partial lockdown exit for young kids.** LSHTM. (accessed 28 May 2020).
Reference Source

27. Block P, Hoffman M, Raabe IJ, *et al.*: **Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world.** *Nat Hum Behav.* 2020; **4**(6): 588–596.
PubMed Abstract | **Publisher Full Text**

28. **Our Plan to Rebuild: The UK Government's COVID-19 recovery strategy.** 2020; 31.
Reference Source

29. **Staying alert and safe (social distancing).** GOV.UK. (accessed 30 May 2020).
Reference Source

30. Minello A: **The pandemic and the female academic.** *Nature.* 2020.
PubMed Abstract | **Publisher Full Text**

31. **Living alone in the coronavirus crisis.** Healthwatch Hillingdon. (accessed 28 May 2020).
Reference Source

32. Feys F, Brokken S, De Peuter S: **Risk-benefit and cost-utility analysis for COVID-19 lockdown in Belgium: the impact on mental health and wellbeing**.
Publisher Full Text

33. gov.org: **Meeting people from outside your household**. (accessed 25 Jun 2020).
Reference Source

34. Office for National Statistics: **2011 Census Ad Hoc Household Composition (Age Groups) Safeguarded Tables (Lower Layer Super Output Area): England and Wales**. [data collection]. UK Data Service. SN: 8637, 2020.
Publisher Full Text

35. Streeck H, Schulte B, Kuemmerer BM, *et al.*: **Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event**. Infectious Diseases (except HIV/AIDS), 2020.
Publisher Full Text

36. Held L: **A discussion and reanalysis of the results reported in Jones *et al.*** 2020.
Reference Source

37. Heavey L, Casey G, Kelly C, *et al.*: **No evidence of secondary transmission of COVID-19 from children attending school in Ireland, 2020.** *Euro Surveill.* 2020; **25**(21): 2000903.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

38. Cauchemez S, Carrat F, Viboud C, *et al.*: **A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data.** *Stat Med.* 2004; **23**(22): 3469–87.
PubMed Abstract | **Publisher Full Text**

39. Simpson REH: **Infectiousness of communicable diseases in the household (measles, chickenpox, and mumps).** *Lancet.* 1952; **2**(6734): 549–54.
PubMed Abstract | **Publisher Full Text**

40. Bi Q, Wu Y, Mei S, *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.* 2020; **20**(8): 911–919.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

41. Koh WC, Naing L, Rosledzana MA, *et al.*: **What do we know about SARS-CoV-2 transmission? A systematic review and meta-analysis of the secondary attack rate, serial interval, and asymptomatic infection**. Infectious Diseases (except HIV/AIDS). 2020.
Publisher Full Text

42. Verity R, Okell LC, Dorigatti I, *et al.*: **Estimates of the severity of coronavirus disease 2019: a model-based analysis.** *Lancet Infect Dis.* 2020; **20**(6): 669–677.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

43. **The R number in the UK**. GOV.UK. (accessed 26 May 2020).
Reference Source

44. Ball F, Sirl D, Trapman P: **Threshold behaviour and final outcome of an epidemic on a random network with household structure.** *Adv Appl Probab.* 2009; **41**(3): 765–96.
Publisher Full Text

45. Pellis L, Ball F, Trapman P: **Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of R_0.** *Math Biosci.* 2012; **235**(1): 85–97.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

46. Pellis L, Cauchemez S, Ferguson NM, *et al.*: **Systematic selection between age and household structure for models aimed at emerging epidemic predictions.** *Nat Commun.* 2020; **11**(1): 906.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

47. Meyerowitz-Katz G, Merone L: **A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates.** *medRxiv.* 2020; 2020.05.03.20089854.
Publisher Full Text

48. Van Vinh Chau N, Lam VT, Dung NT, *et al.*: **The natural history and transmission potential of asymptomatic SARS-CoV-2 infection**. Infectious Diseases (except HIV/AIDS). 2020.
Publisher Full Text

49. Jarvis CI, Van Zandvoort K, Gimma A, *et al.*: **Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK**. *BMC Med.* 2020; **18**(1): 124.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

50. **MATLAB**. Natick, Massachusetts: The MathWorks Inc. 2020.

51. R Core Team: **R: A Language and Environment for Statistical Computing.** Vienna, Austria: R Foundation for Statistical Computing, 2019.
Reference Source

52. Leng T, Flasche S: **tsleng93/SocialBubble: SocialBubble v1.1 (Version v1.1)**. *Zenodo.* 2020.
http://www.doi.org/10.5281/zenodo.4605475

53. Keeling MJ, Tildesley MJ, Atkins BD, *et al.*: **The impact of school reopening on the spread of COVID-19 in England**. (accessed 5 Jun 2020).
Publisher Full Text

54. Abbott S, Hellewell J, Thompson RN, *et al.*: **Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts [version 1; peer review: awaiting peer review].** *Wellcome Open Res.* 2020; **5**: 112.
Publisher Full Text

55. **Kursschwenk bei Kinderbetreuung: Söder greift Grünen-Idee auf**. BR24. 2020. (accessed 1 Jun 2020).
Reference Source

56. New Zealand Governtment: **COVID-19 Alert System**. Unite COVID-19. (accessed 1 Jun 2020).
Reference Source

57. Long NJ, Aikman PJ, Appleton NS, *et al.*: **Living in bubbles during the coronavirus pandemic: insights from New Zealand**. 2020. (accessed 31 May 2020).
Reference Source

58. Endo A, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Abbott S, *et al.*: **Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China [version 1; peer review: 1 approved, 1 approved with reservations].** *Wellcome Open Res.* 2020; **5**: 67.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

59. Riou J, Althaus CL: **Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020.** *Euro Surveill.* 2020; **25**(4): 2000058.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

60. Potter GE, Hens N: **A penalized likelihood approach to estimate within-household contact networks from egocentric data.** *J R Stat Soc Ser C Appl Stat.* 2013; **62**(4): 629–648.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

61. Ward H, Atchison C, Whitaker M, *et al.*: **Antibody prevalence for SARS-CoV-2 following the peak of the pandemic in England: REACT2 study in 100,000 adults.** *medRxiv.* 2020.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✓ ✓

Version 1

Reviewer Report 08 February 2021

https://doi.org/10.21956/wellcomeopenres.17747.r42171

✓ **Martin C. J. Bootsma** (iD)

[1] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
[2] Department of Mathematics, Faculty of Sciences, Utrecht University, Utrecht, The Netherlands

The authors have developed a model with which they can investigate what the effect is if two household are allowed to form a bubble. The authors analyse this for different groups of households which are allowed to form bubbles and compare this to scenarios without bubbles and if the additional contacts do not occur in bubbles but randomly.

I think the models are suitable for the research question. As the code is provided by the researchers. the results can be replicated and other research questions could be addressed by the model as well.

I have only minor points, mainly things that could be clarified to make it easier for readers to understand the methodology.

On page 3 "connections are captured by a matrix, A, which defines the probability that infection can pass between any two individuals in the population. A". I think that the authors should specify what the nature of this probability is and that the model uses generations and that the probabilities refer to probabilities of transmission during the infectious period (and not a probability per day, for example).

I was surprised by the assumption "transmission within households and across households who share a bubble is frequency dependent." And wondered whether it would be reasonably to perform a sensitivity analysis to this (assuming a density dependent transmission rate), as the risk of transmission within a household and in bubbles in some cases, say sitting around one table, may not be frequency dependent.

In table 1: regarding $\rho H(i,j)$ specify what the unit of time is for the rate (duration infectious period).

The authors may consider to use a different symbol than C which already in use as the age-dependent sensitivity.

In Table 2, check the use of \tau and T

The authors claim to have performed a sensitivity analysis for the case "that the risk of a household to get infected with SARS-CoV-2 from the community increases with increasing household size instead of being independent". I did not see the results in the manuscript or the online-material. Did I miss this?

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Mathematical modeler of infectious disease dynamics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 15 Mar 2021
Trystan Leng, University of Warwick, Coventry, UK

We thank both reviewers for their considered and detailed reviews of our manuscript. We respond to each of the points raised by reviewers below, and have also revised the main body
of the manuscript where appropriate.

Reviewer 2:
- Comment: On page 3 "connections are captured by a matrix, A, which defines the probability that infection can pass between any two individuals in the population. A". I

think that the authors should specify what the nature of this probability is and that the model uses generations and that the probabilities refer to probabilities of transmission during the infectious period (and not a probability per day, for example).

Response: We agree that this should be clarified. We have amended the above sentence to (added text in bold:)

"This is achieved through stochastic simulation of infections spreading through an interconnected population **by generation** ; connections are captured by a matrix, A, which defines the probability that infection can pass between any two individuals in the population over **an individual's entire infectious period .**"

○ Comment: I was surprised by the assumption "transmission within households and across households who share a bubble is frequency dependent." And wondered whether it would be reasonably to perform a sensitivity analysis to this (assuming a density dependent transmission rate), as the risk of transmission within a household and in bubbles in some cases, say sitting around one table, may not be frequency dependent.

○ Response: Our decision to assume frequency-dependent transmission for close contacts was based upon studies indicating that the probability of infection across two specific members of the same household decreases with household size for COVID-19 (and other communicable diseases like influenza). However, we acknowledge that for COVID-19 the nature of within-household transmission remains relatively uncertain, and that whether transmission within households is frequency or density dependent may depend on the specific context of the household. We have performed this sensitivity analysis - finding that the assumption of density-dependence does not qualitatively change our results (although does change the ordering of scenarios - with density dependent transmission, scenario 3 (single-occupancy household bubbles) now results
in the smallest increase in R. We have adjusted the text in the methods, the results (Figure 8), and the discussion to indicate we have considered density-dependent transmission. The added text in the discussion is included in bold below:

"As in any epidemiological modelling study, we must make some assumptions surrounding transmission. Firstly, we assumed that transmission across close contacts was frequency dependent, informed by previous studies that indicate the probability of infection across two specific members of the same household decreases with household size for COVID-19 (and other communicable diseases like influenza). However, there remains uncertainty surrounding the nature of close-contact transmission for COVID-19, and the assumption of frequency dependence may not accurately capture transmission across all settings. Secondly, we assumed that the risk of a household acquiring infection from the community is independent of its number of occupants. While this assumption may be appropriate for some households (e.g. families where one adult leaves the household to do shopping), it may not hold true in other contexts (e.g. student households comprising of largely independent individuals).. Because of this we tested the sensitivity of our results to these assumptions by alternatively considering close contact transmission as density dependent, and

> **allowing the**
> **risk of household infection to increase with household size. These alternative**
> **assumptions do not qualitatively change our findings, but do change the**
> **ordering of the risk of social bubble strategies. In particular, under either of**
> **these alternative assumptions, social bubble strategies targeting single**
> **occupancy households result in a lower increase in R than strategies targeting**
> **families with children. Reports on the antibody prevalence in England found**
> **household size associated with antibody prevalence, suggesting that household**
> **size may play a role in the probability of acquiring infection; a more detailed**
> **understanding of the**
> **nature of close contact and community transmission may help inform more**
> **precise evaluations of the effectiveness of social bubbles."**

- ○ Comment: The authors may consider to use a different symbol than C which already in use as the age dependent sensitivity.
  Response: Thanks for this suggestion. We have changed C in row 6 of table 1 to k.

- ○ Comment: In Table 2, check the use of \tau and T
  Response: We apologise for this typesetting error. Capital T should refer to the transmissibility of an individual i, whereas transmission rates should be τ's - this has been amended and notation has now been used consistently throughout the manuscript.

- ○ Comment: The authors claim to have performed a sensitivity analysis for the case "that the risk of a household to get infected with SARS-CoV-2 from the community increases with increasing household size instead of being independent
  Response: This appears as the 'Community transmission' row of the tornado diagrams of the sensitivity analyses. We now discuss the sensitivity analyses that change the ordering of scenarios (the alternative assumptions surrounding community transmission
  and density dependent transmission) in Results subsection 'sensitivity analyses' (added text in bold:)

  "The alternative assumptions surrounding close-contact transmission and
  community transmission changed the ordering of risk of social bubble
  strategies.
  Under our baseline assumption, Scenario 1, targeting families with young
  children, resulted in the lowest increase in R; under these alternative
  assumptions, Scenario 3, targeting single-occupancy households, resulted in the
  lowest increase in R."

Competing Interests: I declare no competing interests.

Reviewer Report 14 October 2020

https://doi.org/10.21956/wellcomeopenres.17747.r40386

Jessica Enright 🆔

School of Computing Science, University of Glasgow, Glasgow, UK

This paper describes a careful consideration of the potential for between-household social bubbling to contribute to Covid-19 control during the current (2020) pandemic. The main narrative comparison is between a lockdown scenario where most contact is within a household as compared to several different bubbling arrangements, but a variety of scenarios are considered beyond bubbling to assess bubbling efficacy. The authors conclude that within their modelling framework a variety of bubbling arrangements would feasibly keep the reproductive number below 1 (thus preventing exponential growth of infections). This is a valuable piece of work that I hope has contributed to the evidence for policymaking during emergence from lockdown after the first wave of the pandemic in the UK. To my understanding, a scenario closest to Scenario 4 is what was actually implemented within most of the UK.

The authors have done a good job of incorporating the available knowledge about the epidemiology of Covid-19 at the time they wrote this paper. Given the pace of developments it would have been impossible to incorporate all emerging information, and certainly more knowledge has emerged since the publication of this work that interacts with the work and conclusions here - I don't think it would be fair to request the authors try to incorporate these subsequent developments.

To my best understanding, the model is based on a generation matrix combining household and bubble contacts that is subsampled using transmission probabilities to give contacts that would result in infection. In addition there is mean-field infection across the population. Parameters are from a combination of published sources, adjustments to produce observed R values, and values chosen for sensitivity analysis.

The counterfactual approach to comparison in which the number of infectious contacts are maintained but are rewired outside of bubbles or households in C2 and C3 is useful, and it is unsurprising that these counterfactuals results in more overall infection. I was initially slightly puzzled by the use of directed edges here (given that a single contact could transmit in either direction), but on reflection it makes sense in the context of the subsampling of the matrix to give essentially a transmission network rather than a contact network. The authors may wish to include a sentence or two clarifying this choice.

I would also have been interested in seeing an analysis of bubbling value in a setting with a steadily increasing mean-field risk of transmission, as an attempt to model a general opening-up and increase in activity. However, I realise this is beyond the scope of this work, and I note that the authors discuss this in the Discussion, noting the likelihood of a 'multi-variable exit strategy' in which bubbles might combine with other restriction easements.

I wanted to highlight an assumption that I think is important for this model, and that may not

apply in all situations: the idea that "We therefore assume that the risk of a household acquiring infection from the community is independent of its number of occupants". While I think this is well-supported by citations and is appropriate within the setting considered here, I just wanted to highlight that this may not hold true in larger households composed of multiple unrelated adults - e.g. in student households. In the REACT-2 reports of antibody prevalence (https://www.medrxiv.org/content/10.1101/2020.08.12.20173690v2[1]) after several months of the pandemic in England, individuals in larger households were more likely to have antibodies - of course this could also be potentially explained by a fixed within-household SAR regardless of household size, and the many risk-factors captured within REACT-2 correlated with living in a large household. The authors include a violation of this assumption in the sensitivity analyses, and if I have understood correctly it doesn't qualitatively change the findings. I also found it interesting that the authors chose to have the mean-field transmission depend on an individual's household size, but not on bubble size - would we expect bubble interactions to decrease interactions in the community that would contribute to the general spread modelled as mean-field here? Am I correct in thinking that moderating the mean-field transmission including bubble size would have increased the efficacy of bubbles?

The authors note that the epidemic risk was sensitive to a much higher than observed secondary attack rate within a household - I wanted to note that this parameter is still somewhat uncertain, though their estimate is very reasonably based in observation. Again (and this does not detract from the application of this work to a general population), this may differ in very unusual e.g. large young-adult student households due to differing network effects within the household (for example, as described in https://rss.onlinelibrary.wiley.com/doi/10.1111/rssc.12011[2]).

There are many possible extensions to this work that could take it beyond the emerging-from-lockdown scenario - the authors mention that in the case of schools being open bubbles might be most effective if they aligned with those school contacts, and presumably that would also hold for other types of contacts (work, sport, etc.). It may also be worth considering bubbles as an option during second-wave restrictions, but with the important warning that the authors found that their model outcomes were sensitive to higher R values at the point of implementation of the strategy.

I was pleased to see that the code associated with this work is all available online and is nicely organised, reasonably documented, and CC0 licensed, making the results reproducible and the work re-usable.

I have a few smaller queries around possible typesetting errors:
- In the section 'Sensitivity Analysis' the following appears: "1when" - likely a missing space?

- In the section 'Sensitivity Analysis' the following appears: "underlying. parametric assumptions" - likely an extra full-stop?

- I have some confusion around the typesetting or naming of variables in expressions in Table 2:

    - In the 'Bubble' transmission rate, we see something that looks like τ_B, but I think that this does not appear in Table 1 - does it correspond to T_B in Table 1?

    - Similarly in the 'Household' transmission rate, I see $c(i)$ and τ_H - do these correspond

to C(i) and T_H in Table 1?

- ○ I think that the susceptibility term defined in Table 1 might not appear elsewhere, though it is mentioned in words - I could of course have missed it! If it does not appear, should it? If not, the authors might consider removing it from Table 1?

References
1. Ward H, Atchison C, Whitaker M, Ainslie K, et al.: Antibody prevalence for SARS-CoV-2 in England following first peak of the pandemic: REACT2 study in 100,000 adults. *medRxiv*. 2020. Publisher Full Text
2. Potter GE, Hens N: A penalized likelihood approach to estimate within-household contact networks from egocentric data.*J R Stat Soc Ser C Appl Stat*. 2013; **62** (4): 629-648 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: I have not recently published with any of the authors, but did recently work together with an author of this work at a VKEMS working group on opening higher education spaces in the UK during the pandemic.

Reviewer Expertise: Network modelling of infectious disease, graphs and networks

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 15 Mar 2021
Trystan Leng, University of Warwick, Coventry, UK

We thank both reviewers for their considered and detailed reviews of our manuscript. We

respond to each of the points raised by reviewers below, and have also revised the main body
of the manuscript where appropriate.

Reviewer 1:

○ Comment: I was initially slightly puzzled by the use of directed edges here (given that a single contact could transmit in either direction), but on reflection it makes sense in the context of the subsampling of the matrix to give essentially a transmission network rather than a contact network. The authors may wish to include a sentence or two clarifying this choice.
Response: Thank you for highlighting this. We have now clarified this in the main body of the test, adding in the sentences in bold:

"Doing so, we retain only the infectious connections between individuals that will lead to an infection. **These sampled probability matrices therefore represent *potential* transmission networks (though the exact transmission network for a given simulation will depend upon who is initially infected). Because these sampled matrices refer to transmission events rather than contacts, this matrix may be unsymmetric.**"

And later

"C2 is obtained by taking the sampled infectious contacts from Scenario 6, then rewiring these sampled contacts.Doing so keeps the number of secondary infections from any individual constant across both scenarios. **As the sampled contact matrices represent transmission networks, edges may be either directed or undirected.** "

○ Comment: I would also have been interested in seeing an analysis of bubbling value in a setting with a steadily increasing mean-field risk of transmission, as an attempt to model a general opening-up and increase in activity. However, I realise this is beyond the scope of this work, and I note that the authors discuss this in the Discussion, noting the likelihood of a 'multi-variable exit strategy' in which bubbles might combine with other
restriction easements.
Response: We agree that such an analysis is beyond the scope of this current work. However, we believe that Figure 5 provides some intuition as to what would happen in such a setting. Figure 5 shows that bubbling scenarios cause an approximately constant increase to R, demonstrating that over the range of mean-field transmissions considered in Figure 5 bubble strategies have a similar impact on R.

○ Comment: I wanted to highlight an assumption that I think is important for this model, and that may not apply in all situations: the idea that "We therefore assume that the risk of a household acquiring infection from the community is independent of its number of occupants". While I think this is well-supported by citations and is appropriate within the setting considered here, I just wanted to highlight that this may not hold true in larger

households composed of multiple unrelated adults - e.g. in student households. In the REACT-2 reports of antibody prevalence (https://www.medrxiv.org/content/10.1101/2020.08.12.20173690v21) after several months of the pandemic in England, individuals in larger households were more likely to have antibodies - of course this could also be potentially explained by a fixed within-household SAR regardless of household size, and the many risk-factors captured within REACT-2 correlated with living in a large household. The authors include a

violation of this assumption in the sensitivity analyses, and if I have understood correctly it doesn't qualitatively change the findings.

Response: Thank you for highlighting this. We have updated our discussion to highlight that the assumption of external infection being independent of household size may not apply to all contexts, and have included the suggested reference. In our sensitivity analyses we consider the situation where a household's probability of acquiring infection from the community scales with household size. This does not qualitatively change our findings, but does impact on the order of which bubbling scenario results in the smallest increase in R (for example, Scenario 3 now results in the smallest increase in R , as

Scenario 3 focuses on single-occupancy households). In reality, the risk of external infection to a household is probably between these two extremes, and may vary from household to household. We have also included a fixed within-household SAR regardless of household size (i.e. density dependent transmission) in our sensitivity analysis. Again, this does not qualitatively change our results, but does impact the ordering of scenarios. (Added text in bold below:)

"As in any epidemiological modelling study, we must make some assumptions surrounding transmission. Firstly, we assumed that transmission across close contacts was frequency dependent, informed by previous studies that indicate the probability of infection across two specific members of the same household decreases with household size for COVID-19 (and other communicable diseases like influenza). However, there remains uncertainty surrounding the nature of close-contact transmission for COVID-19, and the assumption of frequency dependence may not accurately capture transmission across all settings. Secondly, we assumed that the risk of a household acquiring infection from the community is independent of its number of occupants. While this assumption may be appropriate for some households (e.g. families where one adult leaves the household to do shopping), it may not hold true in other contexts (e.g. student households comprising of largely independent individuals).. Because of this we tested the sensitivity of our results to these assumptions by alternatively considering close contact transmission as density dependent, and allowing the risk of household infection to increase with household size. These alternative assumptions do not qualitatively change our findings, but do change the ordering of the risk of social bubble strategies. In particular, under either of these alternative assumptions, social bubble strategies targeting single occupancy households result in a lower increase in R than strategies targeting families with children. Reports on the antibody prevalence in England found household size associated with antibody prevalence, suggesting that household

size may play a role in the probability of acquiring infection; a more detailed understanding of the nature of close contact and community transmission may help inform more precise evaluations of the effectiveness of social bubbles."

○ Comment: I also found it interesting that the authors chose to have the mean-field transmission depend on an individual's household size, but not on bubble size would we expect bubble interactions to decrease interactions in the community that would contribute to the general spread modelled as mean-field here? Am I correct in thinking that moderating the mean-field transmission including bubble size would have increased the efficacy of bubbles?
Response: We chose to focus on scenarios where implementing bubble strategies would strictly increase transmission. If a household's mean-field transmission depends on the size of their bubble, the introduction of a social bubble strategy may in some instances decrease transmission (the increase in within-bubble transmission may be less than the decrease in community infection). You are correct in saying that this would have increased the efficacy of bubbles, and we have now noted this in the discussion (added text in bold below):

"We do not consider the risk of community transmission depending on bubble size. However, if bubbles were to act as cohesive units, and as a consequence reduce their interaction with the community, this may further increase the effectiveness of social bubble strategies."

○ Comment: The authors note that the epidemic risk was sensitive to a much higher than observed secondary attack rate within a household - I wanted to note that this parameter is still somewhat uncertain, though their estimate is very reasonably based in observation. Again (and this does not detract from the application of this work to a general population), this may differ in very unusual e.g. large young-adult student households due to differing network effects within the household (for example, as described in https://rss.onlinelibrary.wiley.com/doi/10.1111/rssc.120112)
Response: Thank you for highlighting this. We have updated our discussion to highlight
the uncertainty that remains in determining this parameter, and the heterogeneity in attack rates that may occur in different settings. (Added text in bold below):

"**There remains some uncertainty surrounding the household attack rate of COVID-19, with** high household attack rates observed in some instances. However, our base case assumptions are in line with an increasingly consistent picture emerging in the contemporary academic literature. Also, superspreading events have been raised as apotentially important source for sustained transmission of SARS-CoV-2, which would
further imply a rather low secondary household attack rate in most instances. **However,
household attack rates may vary between different types of household, and may be larger for some households with unusual network structures, such as large student households."**

○ Comment: There are many possible extensions to this work that could take it beyond the emerging-from lockdown scenario - the authors mention that in the case of schools being open bubbles might be most effective if they aligned with those school contacts, and presumably that would also hold for other types of contacts (work, sport, etc.). It may also be worth considering bubbles as an option during second-wave restrictions, but with
the important warning that the authors found that their model outcomes were sensitive to higher R values at the point of implementation of the strategy.
Response: We have included in the discussion that social bubbles aligned with other types of contacts, such as sport or workplace, would likely improve the efficacy of any bubbling strategy (Added text in bold below):

"As lockdown measures are eased and schools are gradually re-opened forming social bubbles that largely overlap with societal contacts one (e.g. forming social clusters with families that have children going to the same class) is likely further reducing the additional epidemic risk from social bubbles. **By the same token, as lockdown eases,
social bubbles aligned with other types of contacts, such workplace contacts or sport team contacts, would likely improve the efficacy of any bubbling strategy.**
"

Minor comments:

○ Comment: In the section 'Sensitivity Analysis' the following appears: "1when" - likely a missing space?
Response: This has now been amended.

○ Comment: In the section 'Sensitivity Analysis' the following appears: "underlying. parametric assumptions" - likely an extra full-stop?
Response: This has now been amended.

○ Comment: In the 'Bubble' transmission rate, we see something that looks like τ_B, but I think that this does not appear in Table 1 - does it correspond to T_B in Table 1? Similarly in the 'Household' transmission rate, I see $c(i)$ and τ_H - do these correspond to $C(i)$ and T_H in Table 1?
Response: We apologise for the typesetting errors in notation. The correct notation for household and transmission rates is τ_H and τ_B, and the correct notation the susceptibility scaling factor is $C(i)$. Tables have been amended and now notation is applied consistently throughout the paper.

○ Comment: I think that the susceptibility term defined in Table 1 might not appear elsewhere, though it is mentioned in words - I could of course have missed it! If it does not appear, should it? If not, the authors might consider removing it from Table 1?
Response: The susceptibility term appears in the equation for $R(g)$

Competing Interests: I declare no competing interests.