Modelling SARS-COV2 Spread in London: Approaches to Lift the Lockdown – technical appendix

1. The Model

The model is described by the following equations:

\[ S(t) = -\beta(t)S(t)(aI_a(t) + I_s(t) + \xi I_j(t)) \]

\[ E(t) = \beta(t)S(t)(aI_a(t) + I_s(t) + \xi I_j(t)) - kE(t) \]

\[ I_a(t) = (1-p)kE(t) - \gamma_a I_a(t) \]

\[ I_s(t) = pkE(t) - qI_s(t) \]

\[ I_j(t) = qI_s(t) - (\psi + \gamma_s + \mu_s)I_j(t) \]

\[ I_h(t) = \psi I_s(t) - (\gamma_h + \mu_h)I_h(t) \]

\[ R(t) = \gamma_a I_a(t) + \gamma_s I_s(t) + \gamma_h I_h(t) \]

\[ D(t) = \mu_s I_s(t) + \mu_h I_h(t) \]

**Age:** The model was modified to add age-dependent transmission heterogeneity. We sourced a contact matrix \(C\) for UK physical and non-physical contacts [30] and then adjusted it [31] to the current numbers of London residents sourced from the Office on National Statistics (ONS) [32] in the following age-groups: 0-14, 15-59 and 60+. We also adjusted the rate of hospitalization and mortality taking age in respect to hospitalization rates according to [9] and calibration to NHS data on COVID-19 deaths in London [33]. We obtained death rates by age strata calculated assuming that 50% of individuals in intensive care die [9] and that a percentage of 60+ not in intensive care can also die.

Moreover, to study the possible interactions with commuters not living in London, we modified the transmission component of the model as

\[ \beta(t)S(t) \sum_{j=1}^{3} C_{ij}(a(I_{Aj}(t)) + f_{A}(t)) + f_{S}(t) + I_{Sj}(t) + \xi I_{ij}(t)) \]

Where \(i, j = 1, \ldots, 3\) represent the three age groups and \(f_{A}(t)\) and \(f_{S}(t)\) are exponential functions

\[ f_j(x) = K_j(a e^{bx} + c e^{bx}) \]

calculated by inferring the prevalence of asymptomatic and symptomatic cases in England by running the model in the general English population (excluding London) and then applying these rates to the maximum number of people commuting daily to London from outside the city [32] according to their age group \((K_i)\). Data from 2017 show that the total number of in-commuters (i.e. people working in London but living anywhere except London) was 934,000.

**Borough Level Analysis:** To further address geographical heterogeneity, we disaggregated the London population by borough of residence, as PHE provides data of daily cases of COVID-19 disaggregated by local authority. In previous work [8] we showed a correlation between the use of the London underground network and the prevalence of influenza-like illnesses. Here we used data from Transport for London (TfL) [34] to create a contact matrix weighting contacts between different boroughs.
We used a two-weeks sample of trips that took place in October 2015 in the UK capital by underground and national rail service. Data include entry and exit stations and the day and time the trip took place. For the purpose of this work, we initially considered only trips starting and ending in one of the 33 London boroughs, obtaining a total of 937,134 trips. We then considered also all trips originating from outside the 33 London boroughs, specifically: Buckinghamshire, Essex, Hertfordshire, Kent and Surrey. These trips were aggregated in one single origin district in order to consider all commutes from outside the city. Because PHE reports notification data by aggregating cases from City of London and Hackney together, we subsequently obtained a total of 32 London boroughs plus one extra borough, with the latter representing the boroughs outside London.

We aggregated all trips by borough of arrival and borough of departure and finally calculated the average number of daily trips between each pair of boroughs which, divided by the average total number of trips departing in each station, allowed us to generate a matrix of weights and re-define the transmission component of the model as

$$\beta S_i \sum_{j=1}^{33} c_{ij}(\alpha I_{Aj} + I_{ij} + \xi I_{Sj})$$

where $i=1,...,33$ represents the borough currently observed and $j=1,...,33$ the borough it is paired with.

Because boroughs have different numbers of hospitals and available resources, we assumed that the notification parameter $d$ differs by boroughs, thus we calibrated the model to each borough-specific dataset of notified COVID-19 cases according to this parameter. We only considered data from 9th March until 21st March because, for cases notified in that time window, contagion happened before implementation of any restriction on movements. TfL reported that, since then, the number of people using the underground and rail services plummeted by around 95% [35].

### Parameters and data sources

Information on the total size of the population and its age structure come from London’s general demographic data. Transmission events occur through contacts between susceptible and infectious individuals, we assume these contacts are age-specific according to a contact matrix $C$ [36-41] and adjusted to the current population [31]. If transmission is studied locally, matrix $C$ is adapted using local transportation data to include also contacts between different neighbourhoods.

Infection-related parameters are inferred from current COVID-19 literature and technical reports [24]: the incubation period is assumed to be 5.1 days long, infectiousness is assumed to occur from 12 hours prior to the onset of symptoms for those that are symptomatic and from 4.6 days after infection in those that are asymptomatic. Symptomatic individuals are assumed 50% more infectious than asymptomatic individuals and constitute two third of all cases.

Parameters related to self-isolation, hospitalization and mortality are UK specific.

The model is calibrated to data on notified cases and deaths from Public Health England (PHE) using least-square fit and is coded in MATLAB R2019a.

All parameters for the age-stratified model can be found in Table S1.

### 2. Calibration

We initialised our model on 9th March and calibrated it to notification data from PHE. We assumed three different transmission rates $\beta_1$ and $\beta_2$ and $\beta_3$ representing the three trends in London’s COVID-19 outbreak following government decision on contacts restriction among the general population, on 9th, 17th and 23rd March respectively.

The model was first run considering London’s population as homogeneous, i.e. all individuals have the same probability of infection with no added heterogeneity.
The model was calibrated against number of cumulative notified cases (which include also recovered individuals) and deaths. These are represented in the model as the totality of hospitalised cases plus a small percentage of non-hospitalised as some GPs still test their patients for monitoring purposes.

Calibration in the age-stratified population model respect to notified cases (top) and deaths (bottom). $R^2$ values of 0.99 and 0.90 respectively (Figure S1).

![Graph](image)

**Figure S1.** Model calibration against notifications (top) and deaths (bottom).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>Effective transmission rate</td>
<td>$1.90 \times 10^{-07}$</td>
<td>calibration</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.81 \times 10^{-07}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$5.98 \times 10^{-08}$</td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Reduction of infectiousness in asymptomatic people</td>
<td>0.5</td>
<td>[24]</td>
</tr>
<tr>
<td>$\xi$</td>
<td>Relative infectiousness of isolated people</td>
<td>0.04</td>
<td>calibration</td>
</tr>
<tr>
<td>$\alpha\xi$</td>
<td>Relative infectiousness of asymptomatic isolated people</td>
<td>0.02</td>
<td>scenario</td>
</tr>
<tr>
<td>$k$</td>
<td>Progression from exposed to infectious initial stage</td>
<td>1/4.6</td>
<td>[24, 42, 43]</td>
</tr>
<tr>
<td>$p$</td>
<td>Proportion of symptomatic cases</td>
<td>0.66</td>
<td>[24, 25]</td>
</tr>
<tr>
<td>$q$</td>
<td>Progression from symptomatic unaware to self-isolated</td>
<td>1/1.5</td>
<td>[24]</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
<td>Source</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>$\psi$</td>
<td>Proportion of cases that require hospitalization by age group</td>
<td>0.0017, 0.0440, 0.2270</td>
<td>[24]</td>
</tr>
<tr>
<td>$\gamma_a$</td>
<td>Recovery rate for asymptomatic individuals</td>
<td>1/6.5</td>
<td>[24]</td>
</tr>
<tr>
<td>$\gamma_s$</td>
<td>Recovery rate for symptomatic individuals</td>
<td>1/6.5</td>
<td>[24]</td>
</tr>
<tr>
<td>$\gamma_h$</td>
<td>Recovery rate for hospitalised individuals</td>
<td>1/10.4</td>
<td>[24]</td>
</tr>
<tr>
<td>$\mu_s$</td>
<td>Mortality rate of symptomatic non-hospitalised cases</td>
<td>0</td>
<td>calibration</td>
</tr>
<tr>
<td>$\mu_h$</td>
<td>Mortality rate of hospitalised by age group</td>
<td>0.025, 0.027, 0.510</td>
<td>[24, 25] and calibration</td>
</tr>
</tbody>
</table>

**Table S1. Model parameters.**

Figure S2 shows modelling results on the total number of infections according to uncertainty around the transmission parameter $\beta$. The peak of infection falls between early April and mid-August.

![Infections](image)

**Figure S2** Numbers of total infections (all ages, including asymptomatic, symptomatic and hospitalised) when uncertainty on transmission rate $\beta$ is taken into account ($5.70 \times 10^{-08}$ to $6.65 \times 10^{-08}$, with a 95% confidence bound).

Results obtained when running the age-disaggregated model for 60 days can be seen in Figure S3a-c.
Figure S3 COVID-19 in London’s population disaggregated by age groups. Figure S3a shows numbers of infections and deaths in the 0-14 age group. Figure S3b shows numbers of infections and deaths in the 15-59 age group. Figure S3c shows numbers of infections and deaths in the 60+ age group. (Y axes scales are different for each graph.)

**Borough calibration:**

Table S2 reports R² numbers of calibration against notified cases, figures S4a-w can be found below.
<table>
<thead>
<tr>
<th>Boroughs</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barking and Dagenham</td>
<td>0.90</td>
</tr>
<tr>
<td>Barnet</td>
<td>0.88</td>
</tr>
<tr>
<td>Bexley</td>
<td>0.96</td>
</tr>
<tr>
<td>Brent</td>
<td>0.99</td>
</tr>
<tr>
<td>Bromley</td>
<td>0.98</td>
</tr>
<tr>
<td>Camden</td>
<td>0.88</td>
</tr>
<tr>
<td>Croydon</td>
<td>0.96</td>
</tr>
<tr>
<td>Ealing</td>
<td>0.97</td>
</tr>
<tr>
<td>Enfield</td>
<td>0.97</td>
</tr>
<tr>
<td>Greenwich</td>
<td>0.93</td>
</tr>
<tr>
<td>Hackney + City of London</td>
<td>0.94</td>
</tr>
<tr>
<td>Hammersmith and Fulham</td>
<td>0.89</td>
</tr>
<tr>
<td>Haringey</td>
<td>0.95</td>
</tr>
<tr>
<td>Harrow</td>
<td>0.98</td>
</tr>
<tr>
<td>Havering</td>
<td>0.97</td>
</tr>
<tr>
<td>Hillingdon</td>
<td>0.96</td>
</tr>
<tr>
<td>Hounslow</td>
<td>0.99</td>
</tr>
<tr>
<td>Islington</td>
<td>0.93</td>
</tr>
<tr>
<td>Kensington and Chelsea</td>
<td>0.90</td>
</tr>
<tr>
<td>Kingston upon Thames</td>
<td>0.94</td>
</tr>
<tr>
<td>Lambeth</td>
<td>0.97</td>
</tr>
<tr>
<td>Lewisham</td>
<td>0.98</td>
</tr>
<tr>
<td>Merton</td>
<td>0.95</td>
</tr>
<tr>
<td>Newham</td>
<td>0.96</td>
</tr>
<tr>
<td>Redbridge</td>
<td>0.97</td>
</tr>
<tr>
<td>Richmond upon Thames</td>
<td>0.98</td>
</tr>
<tr>
<td>Southwark</td>
<td>0.98</td>
</tr>
<tr>
<td>Sutton</td>
<td>0.79</td>
</tr>
<tr>
<td>Tower Hamlets</td>
<td>0.94</td>
</tr>
<tr>
<td>Waltham Forest</td>
<td>0.98</td>
</tr>
<tr>
<td>Wandsworth</td>
<td>0.96</td>
</tr>
<tr>
<td>Westminster</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Table S2. Values of $R^2$ for each borough-specific calibration.
Results

General model

$\hat{y} = 0.015 + 0.005x$

Coefficient with 95% confidence intervals

$\hat{b} = (0.008, 0.005)$

Goodness of fit:

$R^2 = 0.983$

$F$-value $= 12.13$

Adjusted $R^2 = 0.983$

RMSE $= 0.257$

Table of fits

<table>
<thead>
<tr>
<th>Fit name</th>
<th>Data</th>
<th>Fit type</th>
<th>SSE</th>
<th>R-square</th>
<th>Adjust R-sq</th>
<th>RMSE</th>
<th>F-value</th>
<th>P-value</th>
<th>Validation</th>
<th>Validation SSE</th>
<th>Validation RMSE</th>
</tr>
</thead>
</table>
3. Scenarios technical details

Retrospective Borough Analysis: To model the isolation of RBKC from the rest of the city, we set to 0 the entries of the weighted contact matrix $c$ representing people moving in and out of the district. To introduce lockdown inside the district (but not on the rest of the city) we set $\beta = \beta_3$ for the component describing transmission happening inside the borough, and $\beta = \beta_1$ outside.
Results of the borough analysis are illustrated in Figure S5 a-b.

(a) Results of the borough analysis are illustrated in Figure S5 a-b.

(b) Results of the borough analysis are illustrated in Figure S5 a-b.

**Figure S5.** Impact of isolating only the RBKC on the borough’s epidemic (a) and on the whole city epidemic (b). The solid line shows the number of total infections (asymptomatic, symptomatic, and hospitalised) in the borough with no lockdown in place. The first dotted line (:) shows the number of total infections if the borough’s residents had been isolated from the rest of the city but were still able to circulate inside the district. The second dotted line (--) shows the number of total infections if full lockdown at home had immediately been imposed on the whole district.

**Scenario 1**

Extended lockdown for 547 days: $\beta = \beta_3$

Lockdown lifted on the 61st day: $\beta = \beta_1$

**Scenario 2**

Lockdown lifted on 61st day with social distancing measures recommended: $\beta = \beta_2$

Universal testing: we assume that weekly universal testing entails detection of asymptomatic cases which can then move to an isolated asymptomatic compartment $I_{AI}$ at rate of $1/3.5$.

Weekly universal testing: asymptomatic isolation rate $1/3.5$

Universal testing twice a week: asymptomatic isolation rate $1/2.3$

Universal testing three times a week: asymptomatic isolation rate $1/1.75$

We here assume a perfect test of accuracy (sensitivity and specificity) 100%.

Because asymptomatic cases are assumed to recover in 6.5 days, detected asymptomatic remain infectious and isolated for an additional 3 days, after which they recover.
The transmission component of the model is thus modified as

$$\beta(t)S_i \sum_{j=1}^{3} C_{ij}(\alpha I_{Aj}(t) + f_{Aj}(t)) + \xi \alpha I_{Aj}(t) + f_{Sj}(t) + I_{Sj}(t) + \xi I_{Ij}(t))$$

Parameters are defined in Table S1. Results of scenario 2 are illustrated in Figure S7 a-b.

**Figure S7** Numbers of infections (a) and deaths (b) in London following the end of lockdown on 8th May, according to different scenarios of Universal testing. Dotted line (:) represents the scenario of weekly universal testing, dotted line (-) the scenario of testing the whole population twice a week, and the dotted line (--) the scenario of testing the whole population three times a week. These are compared against the baseline scenario of continued lockdown (solid line).

**Scenario 3**

A lockdown targeting the 60+ population only from day 54+ means

$$S_i \sum_{j=1}^{3} C_{ij}\beta_{ij}(t)[\alpha (I_{Aj}(t) + f_{Aj}(t)) + \xi \alpha I_{Aj}(t) + f_{Sj}(t) + I_{Sj}(t) + \xi I_{Ij}(t)]$$

Where $\beta_0 = \beta_1$, and $\beta_0 = \beta_3$ only when transmission happens from, or towards, the age group 60+.

Results of scenario 3 are illustrated in Figure S8.
**Figure S8.** Numbers of daily infections (red), cumulative deaths in the whole population (magenta), cumulative deaths in the 0-59 age group (cyan) and cumulative deaths in the 60+ age group (black) in two years. The solid line represents the scenario of prolonging lock down for 1.5 years. The dotted line (- -) represents the scenario of ending lockdown on 8th May. Dotted line (:-) represents the scenario of ending lock down on the 8th May for the 0-59 population but prolonging it for two years for the 60+. Dotted line (---) represents the scenario of prolonging lockdown for the whole population for two years.

**Scenario 4**

Sensitivity of COVID-19 testing is still uncertain. The NHS is testing using polymerase chain reaction (PCR) [44], its sensitivity has recently been estimated to be around 83.3% [45], while earlier reports from China showed it being as low as 60-70% [46, 47]. At the same time, while computed tomography (CT) seems to have a sensitivity of 97.2% [45] it may not be feasible for universal testing. Here we assume a test with sensitivity 80%.

We describe weekly universal testing as in Scenario 1 with efficacy $\varepsilon=80\%$, (i.e. 80% of positive cases identified among the totality of positive cases tested), without lockdown ($\beta = \beta_1$).

The model’s structure is thus modified as in scenario 2 (Figure S9).

Thus, people move from $I_A$ to $I_{AI}$ at a rate $(1/3.5)(0.8) I_A$.

We additionally tested the possible impact of facemasks. The rationale for this analysis is that several studies have discussed the potential impact of facemasks on transmission of respiratory diseases, highlighting potential higher efficacy (around 50%) if worn by already infectious individuals [28, 48-52], especially in community and healthcare settings [27, 53, 54]. We here assume that notified infectious individuals wear facemasks, while the general population wears face coverings which have been estimated to be 3 times less effective than face masks [52]. Because facemasks efficacy depends on adherence, we make a conservative assumption of 30% reduction on transmission due to notified isolated cases. At the same time, we assume 10% effectiveness for face coverings, i.e. a 10% reduction in transmission from undetected symptomatic and undetected asymptomatic cases.

Facemasks efficacy of 30% translates in $e_1 = 0.7$ fold reduction in $\xi$ in the transmission caused by isolated identified cases ($I_I$, and $I_{II}$), and 10% efficacy for face coverings, i.e. a $e_2 = 0.9$ reduction on the transmission also caused by asymptomatic and symptomatic cases not yet identified ($I_A$ and $I_S$). If facemasks and face coverings are not used, $e_1 = e_2 = 1$. 

15
Thus, the new transmission rate becomes

$$\beta(t)S_i \sum_{j=1}^{3} C_{ij}[\alpha e_2(I_A(t) + f_A(t)) + e_3 \xi \alpha I_{Aij}(t) + e_2 f_{Sj}(t) + e_1 I_{Sj}(t) + e_2 I_{Sj}(t) + e_1 I_{IJ}(t)]$$

The impact of this strategy is heavily affected by the efficacy of face masks and face coverings as shown in Figure S9. Results of scenario 4 can be seen in Figure S10.

**Figure S9** Sensitivity analysis on facemasks and face coverings efficacy in a scenario where lockdown is lifted from 8th May.

**Figure S10**. Numbers of infections and deaths in London when universal testing and using facemasks, from 8th May when lockdown is ended. We compared the scenario of a prolonged lock-down with no additional control interventions (solid line), to the removed lock-down with no intervention (dotted line (-)), with weekly universal testing of the whole population from 8th May (dotted line (-)), to the addition of facemask usage by identified positive cases and face coverings by the general population, dotted line (--).

**Scenario 5**
Universal testing as in Scenario 2, with lockdown ($\beta = \beta_3$).

Facemasks as in Scenario 4.

Contact tracing affects the number of exposed individuals identified and isolated before becoming infectious. We test results when a percentage $\ell$ of infected (but not yet infectious) contacts is identified and isolated in $E_t$. Once they become infectious, they progress to isolated symptomatic or isolated asymptomatic at an average rate $k/2$. The new modelling structure is illustrated in Figure S11.
Figure S11 Model’s structure in scenario 5.

References


Béraud, G., et al., The French connection: the first large population-based contact survey in France relevant for the spread of infectious diseases. 2015. 10(7).


Kiti, M.C., et al., Quantifying age-related rates of social contact using diaries in a rural coastal population of Kenya. 2014. 9(8).


