RESEARCH ARTICLE

Estimating required ‘lockdown’ cycles before immunity to SARS-CoV-2: model-based analyses of susceptible population sizes, ‘S0’, in seven European countries, including the UK and Ireland [version 1; peer review: 1 approved with reservations, 1 not approved]

Rosalyn J. Moran1, Erik D. Fagerholm1, Maell Cullen1, Jean Daunizeau2, Mark P. Richardson3, Steven Williams1, Federico Turkheimer1, Rob Leech1, Karl J. Friston4

1Department of Neuroimaging, IOPPN, King’s College Hospital, London, London, UK
2Institut du Cerveau et de la Moelle épinière, INSERM UMRS, 1127, Paris, France
3Department of Basic and Clinical Neuroscience, IOPPN, King’s College London, London, UK
4Wellcome Trust Centre for Human Neuroimaging, University College London, London, UK

First published: 04 May 2020, 5:85
https://doi.org/10.12688/wellcomeopenres.15886.1
Latest published: 04 May 2020, 5:85
https://doi.org/10.12688/wellcomeopenres.15886.1

Abstract

Background: Following stringent social distancing measures, some European countries are beginning to report a slowed or negative rate of growth of daily case numbers testing positive for the novel coronavirus. The notion that the first wave of infection is close to its peak begs the question of whether future peaks or ‘second waves’ are likely. We sought to determine the current size of the effective (i.e. susceptible) population for seven European countries—to estimate immunity levels following this first wave.

Methods: We used Bayesian model inversion to estimate epidemic parameters from the reported case and death rates from seven countries using data from late January 2020 to April 5th 2020. Two distinct generative model types were employed: first a continuous time dynamical-systems implementation of a Susceptible-Exposed-Infectious-Recovered (SEIR) model, and second a partially observable Markov Decision Process or hidden Markov model (HMM) implementation of an SEIR model. Both models parameterise the size of the initial susceptible population (‘S0’), as well as epidemic parameters.

Results: Both models recapitulated the dynamics of transmissions and disease as given by case and death rates. Crucially, maximum a posteriori estimates of S0 for each country indicated effective population sizes of below 20% (of total population size), under both
the continuous time and HMM models. Using a Bayesian weighted average across all seven countries and both models, we estimated that 6.4% of the total population would be immune. From the two models, the maximum percentage of the effective population was estimated at 19.6% of the total population for the UK, 16.7% for Ireland, 11.4% for Italy, 12.8% for Spain, 18.8% for France, 4.7% for Germany and 12.9% for Switzerland.

**Conclusion:** Our results indicate that after the current wave, a large proportion of the total population will remain without immunity.

**Keywords**
Coronavirus, SARS-CoV-2, Covid-19, DCM, SEIR, Modelling, Susceptibility

This article is included in the Coronavirus (COVID-19) collection.
Introduction

As of early April 2020, the Coronavirus pandemic has reached different affected epicentre in Europe with its first reported cases on 24th January 2020 (Reusken et al., 2020) with cases reported shortly after in Germany, then the United Kingdom, Italy, Spain, Switzerland and in Ireland on February 29th. Subsequently outbreaks have emerged across the European continent. The daily rates of new confirmed cases of the Covid-19 virus (SARS-CoV-2) have begun to decrease in some of these countries; in particular, in Italy and Spain, with promising signs that extensive social distancing measures have been effective and that these countries have reached or are past ‘the peak’ of infections. Epidemiological models that predict the progression of populations from Susceptible (S) to Exposed (E), Infected (I) and Recovered (R) (SEIR models (Kermack & McKendrick, 1927)) can be used to investigate the properties of these peaks, given the initial susceptibility of a population. For SARS-CoV-2, no (or limited) immunity can be assumed a priori in humans and thus the majority of the entire population is deemed susceptible (Eurosurveillance Editorial Team Team, 2020).

Several studies have developed and simulated SEIR models using epidemic parameters to ‘nowcast’ and forecast transmission (Wu et al., 2020). Parameters of the model are being continuously improved and modified, such as reduced serial interval estimates (Nishiura et al., 2020; Yang et al., 2020), initially derived from observed cases in the initial outbreak in Wuhan, China (Sun et al., 2020; Wang et al., 2020a; Wang et al., 2020b). In most studies, these compartmental models are applied as dynamic generative (i.e., causal or mechanistic) models that assume a set of parameters and predict cases or clinical resources (Moghadas et al., 2020) and intervention effects (Prem et al., 2020; Wells et al., 2020). The (initial) susceptible size of a population (termed ‘S0’) is assumed to be the size of a particular city, e.g. 10 million in Wuhan (Prem et al., 2020) or—for a country—is assumed to comprise of multiples of smaller city sized outbreaks, e.g. 100k (Ferguson et al., 2020). Such models have lent important insight into the likely disease and clinical trajectories of countries as a whole, enabling planning and management for predicted numbers of cases requiring hospitalization and ventilation (Moghadas et al., 2020).

Given the lockdowns around Europe, which likely averted larger case surges (Wang et al., 2020a), we sought to investigate the current effective population size in seven European countries. Therefore, we used the SEIR model to determine the initial population size (S0) that was susceptible (i.e., would eventually become infected) at the beginning of the first wave and thus determine the levels of immunity that might exist in these countries after this wave (by assuming the susceptible population will eventually become infected and develop immunity). One approach, to perform this inverse modelling, is to apply dynamic causal modelling (Friston et al., 2007)—enabling the incorporation of prior values for parameters (e.g. the serial interval, incubation period or number of daily contacts) and prior uncertainty about these values.

Quarantine and social isolation are often explicitly accounted for in SEIR models (Feng, 2007; Ridenhour et al., 2018; Wearing et al., 2005) making them appropriate for the current Government advised social distancing. Importantly, two particular forms of this SEIR model (with social distancing) have recently been developed (Friston et al., 2020; Moghadas et al., 2020) that also account for deaths following hospitalization or treatment via ventilation within an Intensive Care Unit (ICU). Specifically, they account for a potential time lag between becoming infected and developing acute respiratory distress. This makes these models putatively ‘fit for purpose’, when using death as well as case reporting data to fit or invert the models to recover (posterior) parameter values—and estimate their uncertainty.

We aimed to apply these two models to data from seven countries: Ireland, the United Kingdom, Italy, Spain, France, Germany and Switzerland. Our goal was to estimate S0. One model, (the ‘ODE model’—see Methods and (Moghadas et al., 2020)) is based on a classical compartment model where a person in an epidemic can occupy only one compartment or ‘state’ and moves from state to state: from Susceptible to eventually (through intermediate states) either Recovery or Death. The other model—the ‘hidden Markov model’ (HMM; (Friston et al., 2020))—features several factors that change together; including where a person is located (out of the home vs. in the home, for example), as well as their infectious, testing and clinical status. We apply both models to daily case and death reports to assess whether there is convergence on estimates of the initially susceptible (i.e., effective) population sizes.

Methods

Data

Data from a repository for the 2019 Novel Coronavirus at John’s Hopkins University Center for Systems Science and Engineering were used (Dong et al., 2020). Using these date-stamped entries of daily reported cases and reported deaths, we extracted seven timeseries pairs for the countries (including all territories) of Ireland, the United Kingdom, Italy, Spain, Germany, France and Switzerland. Data records from January 22nd to April 5th, 2020 were modelled. For the ordinary differential equation (ODE) model, daily cases and daily accumulated deaths were fitted, while for the HMM model, daily cases and daily deaths were fitted, corresponding to the state equations (see (Friston et al., 2020)).

Models

ODE model. A dynamic transmission model comprising a set of 12 coupled ordinary differential equations was adapted from Moghadas et al. (2020, #4)4 The original model included 12 states for four different age categories. We simplified the model structure by collapsing across age (see Extended data (Moran et al., 2020)). The 12 states or compartments in this simplified model (flow function, Figure 1) described susceptible (S) individuals who became infected with the disease through exposure (E) to other infected individuals. Infected individuals comprised three categories, an asymptomatic or
Figure 1. Models. Left: Flow function, or population dynamics. From the susceptible state (where initially at time t= 0 S = S0), the infected population will enter one of three categories: IH (infected requiring hospitalization), InH (infected NOT requiring Hospitalization), or A (asymptomatic). From the IH state, subjects transfer either to H (+Hospital) or ICU, from which subjects transfer either to R or D. Both the InH and A states lead to Recovery (R). The observer function evaluates the dependent variable at each iteration of the integration process of the flow function. The resultant model-based data is compared against empirical John's Hopkins University data. Right: Network showing the transition between states for the HMM model. The ODE model is a 12-compartmental model—with one factor with 12 states). The HMM model is a 256-compartment model—with four factors (location, infection and clinical) each with four levels, giving $4^4 = 256$ states. The structural difference between the ODE and HMM model rests upon the allowable combination of factors that describe the state of an individual in a population. For example, in the factorial (HMM) model it is possible to die from acute respiratory distress at (e.g. a care) home. Conversely, in the ODE model, one can only die after being in hospital. Certain transitions among these states are allowable. For example, in the ‘testing’ factor, an individual could transit from untested to waiting to a positive result. Or an individual could transition from untested to untested—where they remain untested even if they are infected (see Friston et al., 2020 for a priori probability values).

ODE Model

<table>
<thead>
<tr>
<th>State</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IH</td>
<td>Infected requiring hospitalization</td>
</tr>
<tr>
<td>InH</td>
<td>Infected NOT requiring hospitalization</td>
</tr>
<tr>
<td>A</td>
<td>Asymptomatic</td>
</tr>
<tr>
<td>H</td>
<td>Hospital</td>
</tr>
<tr>
<td>ICU</td>
<td>ICU</td>
</tr>
<tr>
<td>D</td>
<td>Death</td>
</tr>
<tr>
<td>R</td>
<td>Recovery</td>
</tr>
</tbody>
</table>

Subclinical state (A), a symptomatic state who would not require hospitalization (InH) and a symptomatic state who would require hospitalization (IH). Each of these infected categories could also self-isolate—representing three more states defined by lower a priori contacts. People in states InH and A were assumed to recover, while those in states (IH) would transition to either hospitalized (H) or ICU states (ICU). From these states people would recover (R) or die (D) (Figure 1). Time constants of the mode included the incubation period, recovery period, time to self-isolate, time from symptom onset to hospitalization, time from ICU admission to death, time from non-ICU admission to death, length of stay in ICU and length of hospital stay. Parameters controlling proportions that entered branching states (e.g. proportion of all hospitalised cases admitted to ICU were also included (see Extended data for full parameter list (Moran et al., 2020)), as well as the transmission rate and contact per day either within or without self-isolation. Parameters were equipped with a priori values and optimisation was performed on log scale factors to ensure positivity (Extended data (Moran et al., 2020)) of these proportions and rate constants. To link these ODEs to the observed data we employed an observer function which assumed a variable rate of case reporting for symptomatic (without requiring hospitalization) and asymptomatic individuals. A priori we assumed that only 1% of individuals infected who were asymptomatic received tests. We assumed that 20% of symptomatic cases who do not require hospitalization receive tests, and that 100% of infected individuals who are hospitalised receive tests. The levels of 1% and 20% testing were free parameters in our model. The 100% for hospitalised tests was fixed. Finally, 100% of deaths were assumed to be recorded. Finally, we placed a priori on the initial number of individuals in each state. A priori, we assume 100 individuals in infected states. We tested two alternatives for S0:

In the ODE model (Model: ODE) we initialised S0 to 1 Million × $\theta_{S0}$ individuals, where $\theta_{S0} = 1$. This parameter would be optimised for each individual dataset and so could accommodate...
total sizes; e.g. if $\theta_{S0} = 4.9$ a posteriori then the total population of Ireland would be considered initially susceptible.

We also tested a ‘cities’-based version of the ODE model (Model: ODE_City) that might recapitulate the death and case rates for each country. For this, we altered the observer function and imposed a prior of 1 Million $\times \theta_{S0}$ individuals, where $\theta_{S0} = 1$. Then we scaled the case and death rates by the population in millions (see Extended data for equations (Moran et al. 2020)). For this if we obtained $\theta_{N0} = 1$ a posteriori then the total initial susceptible population would also correspond to the total population of Ireland, but the epidemic dynamic would comprise 4.9 distinct outbreaks.

**HMM model.** Our second model of the cases and death rates was a Dynamic Causal Model of Covid-19. See (Friston et al., 2020) for a complete description of the model. In brief, the model represents four factors describing location, infection status, test status and clinical status. Within each factor people may transition among four states probabilistically. The transitions generate predicted outcomes; for example, the number of people newly infected who have tested positive or the number of people newly infected who will remain untested. The location factor describes if an individual is at home, at work, in a critical care unit or deceased. Similar to the early states in the ODE model, the HMM has a second factor describing infection status susceptible, infected, infectious or immune, where it is assumed that there is a progression from a state of susceptibility to immunity—through a period of (pre-contagious) infection to an infectious (contagious) status. The clinical status factor comprises asymptomatic, symptomatic, acute respiratory distress syndrome or deceased. Finally, the fourth factor represents diagnostic status where an individual can be untested or awaiting the results of a test that can either be positive or negative. As with the ODE model, transitions amongst states are controlled by rate constants (inverse time constants) and non-negative probabilities. Similar to the ODE model above, we initialised (and set as priors) S0 to 1 Million $\times \theta_{S0}$ individuals, where $\theta_{S0} = 1$.

For the HMM and both ODE models (ODE model and ODE_City) to estimate the model parameters, we employed a standard (variational Laplace) Bayesian scheme to optimise parameters of corresponding DCM (spm_NLSI_GN) (Friston et al., 2007).

**Results**

The key aim of our analysis was to estimate the likely immunity after the current set of cases and deaths. To ascertain the initial susceptibility $S0$, we examined the posterior estimate from both model types and its Bayesian credible intervals. However, first we examined the evidence for each model, relative to the worst performing model. We used two ODE models, with different constructs for epidemic sizes/meta-populations. The first ODE model (ODE) assumed a prior of 1 million susceptible individuals ($S0$). The second ODE model accounted for several effective populations of size 1 million (ODE_City). The third model was the HMM model, which also assumed a prior of 1 million initial susceptible individuals. Of all three models, ODE_City was the worst performing model for all countries data (Figure 2A).

From the two better performing models, we then estimated the effective population size of $S0=S(t=0)$ as a proportion of the total population (Figure 2B). Taking a Bayesian average—across all models and countries—the estimated proportion of people that were initially susceptible at the start of this outbreak—and thus immune at the end of the outbreak—was 6.4% of the total population of each country.

The ODE model produced consistently higher estimates of $S0$ at the end of the wave than the HMM. These values suggest that after the current wave of cases, between 3 (lowest estimates for Ireland and the UK) and 12 (highest estimate for Germany) more cycles (with identical dynamics to those from Jan 22nd) would be required to bring the total population to probable herd immunity levels (we assume herd immunity of 60%; Figure 2B). We plot this fall in susceptibility state S (increase in immunity) over time, from the initial size $S0$ in Figure 2C for the ODE and HMM models separately for Ireland and the UK.

Our model inversion procedure produced fits to the data that recapitulated the rates up to April 5th for both models (Figure 3). Systematic differences in future predictions were observed however between the ODE and HMM models (though predictions were of similar orders of magnitude). For all countries the peak date and peak number of cases was higher for the ODE model. However, both models exhibited peaks for dates in the past for four countries (Italy, Spain, Germany and Switzerland). For France the models were discrepant with the ODE model predicted a peak in the future on April 20th and the HMM model estimating a peak had already occurred on April 7th.

Peaks in the future were expected for Ireland and the UK. For Ireland, the peak reported case rate predictions were estimated at April 9th for the HMM and April 23rd for the ODE. The estimate of the number of daily cases at the peak were 720 cases and 392 cases for the ODE and HMM models. For the UK, the peak case rate predictions were estimated at April 11th for the HMM and April 17th for the ODE model. The peak case rates (i.e. tested cases) were estimated at 9304 daily cases for the ODE and 5411 daily cases for the HMM models (Figure 3).

The cumulative deaths (Figure 4) evinced relatively small discrepancies between the models, with the ODE model predicting a larger cumulative death toll, of 1250 for Ireland compared with 1008 deaths in Ireland given by the HMM model. For the UK, the ODE and HMM were remarkably consistent, predicting a cumulative death toll of 49296 and 49785, respectively1 (Figure 4). In other European countries, however, the discrepancies between the model predictions were greater in

---

1These predictions fall to substantially lower levels, when empirical priors from a hierarchical or parametric empirical Bayes analysis that incorporates data from all countries are used.
Figure 2. Effective susceptible population estimates. (A) lower bound on log model evidence given by variational free energy (Log Bayes Factor) for different models across countries. This shows that the models with a prior of 1 million S0 outperform the (multiple outbreak ODE_Cities) model, with priors of initial susceptible S0 equal to the total population of the country. (B) From the two better performing models, we report the percent of the country’s total population that are initially susceptible. We also plot the Bayesian parameter average of this percent, across all countries and models (6.4%). (C) From the ODE model, we replot the fall in the susceptible population for Ireland and the UK as the dynamics of the current wave unfold. At the end of this wave large percentages of the population remain susceptible. The projections under the HHM are more pessimistic because the effective population size at the start of the first outbreak (i.e., susceptible population) is smaller (and more precise).
Figure 3. Model predictions of daily reported case rates across countries. Note both models assume that many infected individuals will not be tested or reported in the daily case rates.

Figure 4. Model predictions of cumulative deaths across countries.
some countries, such as France, Spain and Switzerland, with the HMM suggesting considerably lower cumulative deaths.

Finally, to test the assumption that low S0 proportions of the population may be indicative of a ‘next wave’ or several ‘next waves’, we estimated the initial susceptible size from the initial peak of the Spanish Flu pandemic of 1918-1919 using data collated from approximately half of the United Kingdom, i.e. a population of approximately 22 million. Using the HMM model and variational Laplace, we see fits to the data that capture the falling peak. Here, we estimated the effective or susceptible population size was $S_0 = 4.03\%$ of the total population size (Figure 5). Though dramatically different in terms of hospital care, the general picture remains – large waves may be possible after low S0.

**Discussion**

We used a variational Bayesian scheme (Friston et al., 2007) to optimise the parameters of two distinctly constructed models of viral transmission (Friston et al., 2020; Moghadas et al., 2020). We optimised the parameters of these models based on daily reported cases and daily reports of death due to Covid-19. We optimised the model from data acquired for seven European countries. Both models were able to predict (i.e. fit) the current epidemic dynamics with plausible estimated trajectories. The models differed in their exact case rate predictions but predicted commensurate figures for the deaths in the United Kingdom and Ireland. How do these estimates relate to previous predictions of Covid-19 deaths in the UK? It was predicted (Ferguson et al., 2020) that without interventions 510,000 deaths could occur in the UK due to Covid-19. This analysis (Ferguson et al., 2020) also predicted, that even with an optimal mitigation scenario, these death rates would reduce only by one half, i.e. to 255,000. Thus, the predicted death cases of our models ~50,000 in the current cycle are in line with the predictions of mitigation effects, if we assume that several more cycles are possible.

Importantly, both models predicted that we are currently nearing or past the peak of daily case rates in all seven countries. However, the estimates suggest that after this cycle more than 80% of each country’s total population in all countries studied
remain susceptible. Therefore, we assume that future cycles will occur.

The predicted S0 was higher for the ODE model relative to the HMM model. In turn, the ODE model predicted a more prolonged cycle in the current period relative to the HMM model. This speaks to a trade-off between S0 and cycle times. Assuming herd immunity requires 60% of the susceptible population to be immune (Cohen & Kupferschmidt, 2020), one may conclude that further cycles are possible. However, this is not to say that populations within current outbreak areas may not reach herd immunity after the current cycle. Yet, if this is the case (immunity is clustered in geographic or some other organisation of communities), then parts of the country—particularly those communities with high contact numbers that have not ‘been involved’ in the current cycle – may be more likely to participate in future cycles. And while it is obviously unrealistic to suppose that an additive linear effect of populations will emerge (Eubank et al., 2004; Sirakoulis et al., 2000) (i.e., identically shaped cycles), given the complexity of contacts and population movement, our analysis may offer a rough guide to cycle immunity numbers.

As with most scientific research at this time, the modelling described above was conducted with haste. In line with the sentiments of the World Health Organization’s Dr Mike Ryan “Perfection is the enemy of the good when it comes to emergency management. Speed trumps perfection. And the problem in society we have at the moment is everyone is afraid of the consequence of making a mistake. Everyone is afraid of the consequence of error, but the greatest error is not to move, the greatest error is to be paralyzed by the fear of failure.” Therefore, we are grateful to the coding repositories listed in our Underlying data, where interested researchers can reproduce or nuance our analyses.

Data availability

Underlying data

John’s Hopkins University Center for Systems Science and Engineering 2019 Novel Coronavirus data available from (data in this analysis used up to April 5th): https://github.com/CSSEGISandData/COVID-19


Archived code as at time of publication: https://doi.org/10.5281/zenodo.3766243 (Moran et al., 2020).

License: GNU Affero General Public License v3.0

Extended data


This project contains the following extended data:

- ODE FLOW FUNCTION
- ODE Observer FUNCTION
- ODE CITY Observer FUNCTION

These data are under a GNU Affero General Public License v3.0.

Acknowledgements

A previous version of this article is available from medRxiv: https://doi.org/10.1101/2020.04.10.20060426

References


Pubmed Abstract | Publisher Full Text | Free Full Text


Pubmed Abstract | Publisher Full Text | Free Full Text


Publisher Full Text


Publisher Full Text


Pubmed Abstract | Publisher Full Text | Free Full Text


Publisher Full Text


Pubmed Abstract | Publisher Full Text | Free Full Text


Pubmed Abstract | Publisher Full Text | Free Full Text


Pubmed Abstract | Publisher Full Text | Free Full Text


Pubmed Abstract | Publisher Full Text | Free Full Text


Pubmed Abstract | Publisher Full Text | Free Full Text
The paper aims to estimate the final size of susceptible individuals after the first wave of infections in seven European countries. The authors present their analysis based on two different approaches; a deterministic model, described by a system of ODE, containing 9 compartments; and a 'hidden Markov model", a 256-compartment model—with four factors (location, infection and clinical) each with four levels, giving 256 states.

The authors point out an important topic which addresses future scenarios of the second waves of COVID-19 in the studied populations. Under both methods the authors show that a pool of less than 20% of the population would be immune to the disease, and therefore, showing the risk of second waves.

Nevertheless, there are several points regarding the methodology that was not explored:

- The authors do not present sensitivity and identifiability analysis in order to assess effects of model parameters in the dynamics and possible correlations between parameters.

- The lack of data may compromise the parameters estimates of both models, once they rely solely on the daily number of cases and deaths of each country.

- Another point that was not explored is the scenario where the HMM may give the same results of the deterministic model. What are the necessary priors or the worst scenario when HMM ‘over estimate’ or is close to the estimates of the deterministic model?

- Comparison with homogeneous models are not new. Many researchers seek to see the effects of projections under the scenario of an heterogeneous population, it would be more interesting to see comparison of the HMM model and an compartmental model that include heterogeneity (see references).
The mathematical descriptions of the model must be presented at least in Supplementary materials, together with the description of parameters and graphical analysis to assess the fits of the models.

References

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?
No

If applicable, is the statistical analysis and its interpretation appropriate?
No

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

Are the conclusions drawn adequately supported by the results?
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Mathematical modeling of infectious diseases

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, however I have significant reservations, as outlined above.
Sri Venkatramanan
Biocomplexity Institute and Initiative, University of Virginia, Charlottesville, VA, USA

The authors use Bayesian model inference to estimate the effective/susceptible population for seven European countries following the first wave of the COVID-19 pandemic. They use a continuous time version and a POMDP/HMM implementation of SEIR.

While the paper asks an important question and is backed by well-formulated ODE and POMDP models, their analyses and interpretations need to be revised to address the concerns below.

Questions:
1. The authors in the abstract use effective=susceptible in Background but effective=immune in the Results. These are two starkly contrasting definitions (S vs R state of SEIR). They need to clarify what they mean by effective population, and possibly fix one of these.

2. The main results of the paper have less to do with 'lockdown cycles' but more to do with estimating how far these countries are from herd immunity. Though these are related, would be better to revise the title to match the contents of the paper.

3. Why do we need to estimate S0, the initial population that was susceptible for a novel coronavirus? Isn't this assumed to be the entire population? Need some clarification.

4. The ODE_City model is not well explained. Is this a metapopulation model? What does it mean to have 4.9 distinct outbreaks?

5. What does this sentence mean? "the estimated proportion of people that were initially susceptible at the start of this outbreak— and thus immune at the end of the outbreak—" These are two different terms as I read it S(0) and R(T).

6. How do the authors arrive at the 3 or 12 more cycles to herd immunity? Is this just by dividing 60 by the estimated immune population? Given that the recovered individuals will still be mixing, subsequent cycles of the outbreak may not have the same effective reproductive number. So it is not appropriate to assume a linear scaling (or as authors put it - identical dynamics). The authors mention it in the discussion, but this analysis seems to be central to them (reflected in the title). Would be better to refine or remove it.

7. Since the authors use Bayesian inference, I assume they get a distribution over the parameters. So it would be better to include the uncertainty bounds in model predictions (Figures 3 and 4). It seems to be present in Fig 2B (which is what I presume the red lines are).

8. The Spanish Flu analysis is not well motivated. It is evident that without further control measures or completely eliminating infections there is bound to be future waves of
outbreaks. And this example also provides a counter-example to the additive linear model for future cycles, since the 1918 pandemic had a much bigger second wave (partly due to seasonality).

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?  
Partly

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?  
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Computational epidemiology, Mathematical modeling.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.