Long-term forecasting of the COVID-19 epidemic

Dynamic Causal Modelling, UCL, UK

- **The real-time estimate of the R-number is 1.46 (credible interval from .95 to 1.98) on 18 April 2021.** The reproduction ratio will fall over the next few weeks.
- **The current estimate of the efficacy of vaccination stands at 74.0% (credible interval from 72.2 to 75.9%).** This efficacy pertains to the prevention of viral transmission (as opposed to pathogenicity); i.e., sterilising immunity.
- **Daily deaths are predicted to fall to low levels by the end of June 2021**, with a slight resurgence — to about 50 deaths per day (credible interval from 0 to 100) — in mid-May.
- **These predictions rest upon a gentle and prevalence–dependent unlocking**, with an imminent return to levels seen last summer. A final lifting of restrictions is anticipated on 7 July 2021.
- **An effective immunity threshold (of 69.6%) will be reached on 1 May 2021.** This threshold fluctuates with transmission risk (e.g., with viral evolution, aerosol transmission, et cetera) that could fall during spring.

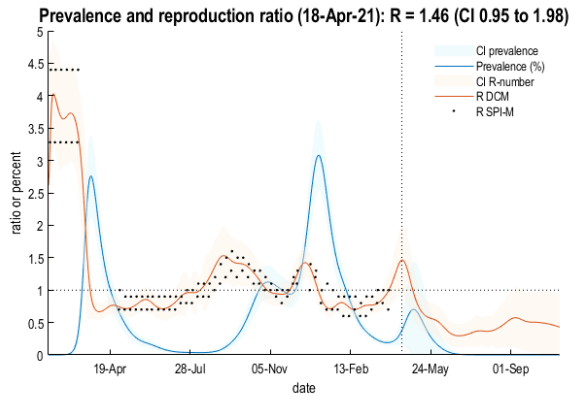
Disclaimer: *the modelling and accompanying estimates are reported in these pages for purely academic (open science) purposes. This modelling has not been commissioned. In particular, dynamic causal modelling is not commissioned by the Independent SAGE (on which Prof Friston serves as a panellist). The independent SAGE does not commit to – or engage in – any particular modelling initiative.*

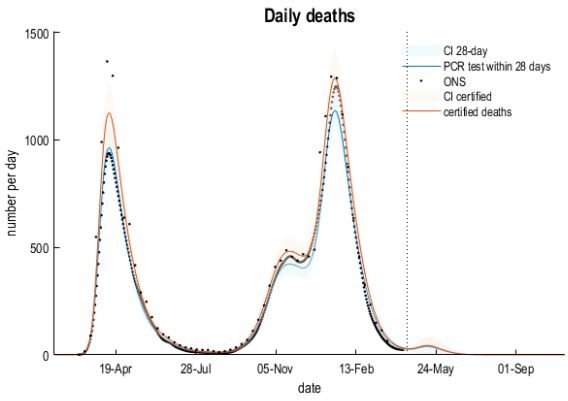
These long-term forecasts are based upon a dynamic causal model (DCM) of viral transmission and mitigated responses. This particular (age-stratified) model is equipped with a vaccination state that affords sterilising immunity (i.e., precludes transmission and clinical susceptibility). Immune efficacy is modelled as the probability of moving to an effectively vaccinated state following vaccination. For example, an efficacy of 52% means that you have a 52% chance of being immune, following your first vaccination.

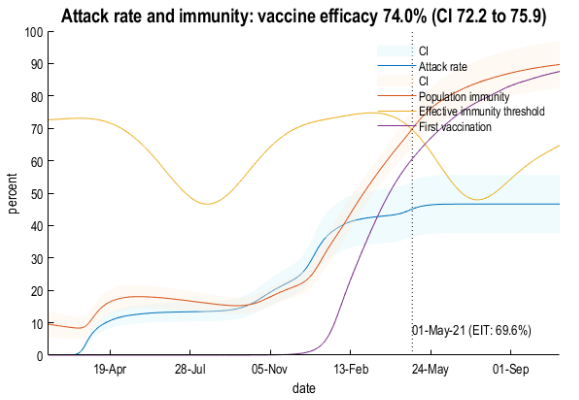
The DCM generates various data that quantify the progression of the epidemic, including **the number of reported (first) vaccinations**. These data are used to estimate the model parameters controlling contact rates, transmission risk and periods of infectiousness using standard variational procedures. Crucially, these variables are themselves time-dependent and depend upon mitigating responses, modelled as the (prevalence-dependent) probability of moving from low to high contact rate locations (e.g. from home to the workplace). The data that informs these estimates include daily positive tests, reported deaths within 28 days of a positive PCR test, certified deaths disaggregated by age and place of death, hospital admissions, contact rate proxies, such as car use and Google mobility data and so on.

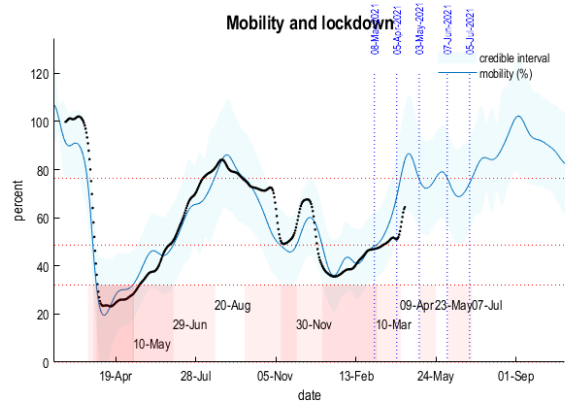
Much like long-term weather forecasts, the ensuing predictions should not be taken too seriously because there is an inherent (although quantified) uncertainty about underlying epidemiological and socio-behavioural variables. These reports will be released on a weekly basis so that people can see how the predictions change — and uncertainty resolves — as time progresses and more data are assimilated. In the summary graphs below, the lines and shaded intervals correspond to predictions and 90% credible intervals, while the black dots are (smoothed) data upon which the estimates are based. A more detailed breakdown of estimated epidemiological and testing states can be found here.

Because the DCM is, effectively, an amalgamation of a conventional epidemiological model and an agent-based behavioural (and testing) model it predicts mitigating responses to changes in community transmission. These DCM predictions are the **most likely outcomes** given our responses to date. In other words, it does not rely on any planned mitigation. This complements — and contrasts with — SPI-M projections of **reasonable worst-case scenarios** that assume a particular sequence of interventions. Generally, the most likely *predictions of mitigated responses* — i.e., what is likely to happen — are more optimistic than worst-case *projections of unmitigated* responses — i.e., what could happen.









Prevalence and reproduction ratio: this panel provides a forecast of (i) the **prevalence of infection** and (ii) the reproduction ratio or **R-number** (blue and orange lines, respectively) with their accompanying confidence intervals (shaded areas). These forecasts are based upon parameters estimated from the data up until the reporting date (the vertical line). The data include GOV.UK estimates of the R-number, which are shown for comparison with the DCM estimates. The black dots correspond to the GOV.UK (SPI-M consensus) estimates moved backwards in time by 16 days from their date of reporting.

The DCM estimate of the R-number is based upon a generative model (i.e., a real-time estimate using data assimilation), while the consensus estimates from the SPI-M are based upon retrospective (e.g., Bayesian regression) analysis of recent data.

Attack rate and immunity: this panel shows long-term forecasts of **attack rate, population or herd immunity** and the percentage of people who have been vaccinated. In addition, an estimate of the **effective immunity threshold** is provided (yellow line). The vertical line illustrates the time at which population immunity reaches the effective immunity threshold. Based upon the changes in testing, death rate and other data, one can estimate the efficacy of vaccination. The attack rate corresponds to the number of people who have been infected since the onset of the outbreak (blue line). This can be supplemented with a small proportion of the population that is estimated to have pre-existing immunity (e.g., mucosal immunity or cross immunoreactivity with other SARS viruses), shown in red. The combination can be read as the herd or population immunity.

The effective immunity threshold is based upon the effective reproduction ratio under pre-pandemic contact rates. The reproduction ratio corresponds to the product of the contact rate, transmission risk and mean infectious period. Note that the effective immunity threshold fluctuates. This reflects the fact that transmission risk changes with time. In this model, transmission risk is modelled as a seasonal fluctuation multiplied by smooth function of time (parameterised with a discrete cosine set). A fluctuating transmission risk accommodates changes in transmissibility (e.g., due to viral evolution) that is contextualised by seasonal variations in transmission (e.g., due to changes in temperature, humidity, socialising outdoors and the propensity for aerosol transmission).

Daily deaths: this panel shows fatality rates as assessed by patients who died within 28 days of positive PCR test and people who died from certified COVID-19. The former represents an underestimate of COVID-related mortality, where the degree of underestimation depends upon testing rates. The discrepancy is adequately modelled by evaluating the probability of succumbing to COVID-19 and having had a positive PCR test within 28 days.

Mobility and lockdown: a long-term forecast of locking and unlocking, based upon car use as quantified by the Department of Transport and - in this graph - Google mobility data. These measures of mobility are expressed in terms of the percentage of pre-pandemic levels. The expected mobility has been thresholded at three levels to illustrate different levels of lockdown. The dates on the lower (graded pink) bar annotate a transition from a more restrictive level of mobility to a less restrictive level. Forecasts of mobility are based upon underlying contact rates that depend upon the prevalence of infection—and are then modulated with a smooth function of time (parameterised with Fourier coefficients). The vertical dotted lines depict dates from Roadmap 3 of scenario modelling by Imperial College. Predictions of unlocking are in remarkable agreement with these tentative dates.

This dynamic causal model includes age-stratification into three groups (below the age of 25, between 25 and 65 and over 65 years of age). The contact rates within and between the three groups (for high and low contact rate locations) are estimated from the data, under mildly informative lognormal shrinkage priors. Please see the following technical report for further technical details.

Changes since last report:

- No change.

Software note: The figures in this report can be reproduced using annotated (MATLAB) code available as part of the free and open source academic software SPM (https://www.fil.ion.ucl.ac.uk/spm/), released under the terms of the GNU General Public License version 2 or later. The routines are called by a demonstration script that can be invoked by typing `>> DEM_COVID_UK` at the MATLAB prompt.

Data sources: (also available as CSV files)

https://coronavirus.data.gov.uk
https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/coronaviruscovid19infectionsurveydata
https://covid.joinzoe.com/data#levels-over-time
https://www.gov.uk/guidance/the-r-number-in-the-uk#contents
https://www.gov.uk/government/statistics/transport-use-during-the-coronavirus-covid-19-pandemic
https://www.google.com/covid19/mobility/

Peer-reviewed references [1-3] and **archival papers** [4-11]

1. Friston, K.J., et al., Second waves, social distancing, and the spread of COVID-19 across America. Wellcome Open Research, 2020. 5(103): p. 103.
2. Friston, K.J., et al., Dynamic causal modelling of COVID-19. Wellcome Open Research, 2020. 5(89): p. 89.
3. Friston, K.J., A. Costello, and D. Pillay, 'Dark matter', second waves and epidemiological modelling. BMJ Global Health, 2020. 5(12): p. e003978.
4. Moran, R.J., et al., Using the LIST model to Estimate the Effects of Contact Tracing on COVID-19 Endemic Equilibria in England and its Regions. medRxiv, 2020.
5. Friston, K.J., G. Flandin, and A. Razi Dynamic causal modelling of mitigated epidemiological outcomes. 2020. arXiv:2011.12400.
6. Daunizeau, J., et al., On the reliability of model-based predictions in the context of the current COVID epidemic event: impact of outbreak peak phase and data paucity. medRxiv, 2020.
7. Daunizeau, J., et al., Modelling lockdown-induced secondary COVID waves in France. medRxiv, 2020.
8. Gandolfi, D., et al., Dynamic causal modeling of the COVID-19 pandemic in northern Italy predicts possible scenarios for the second wave. medRxiv, 2020: p. 2020.08.20.20178798.
9. Moran, R.J., et al., 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. medRxiv, 2020: p. 2020.04.10.20060426.
10. Friston, K.J., et al., Viral mutation, contact rates and testing: a DCM study of fluctuations. medRxiv, 2021: p. 2021.01.10.21249520.
11. Friston, K.J., et al., Tracking and tracing in the UK: a dynamic causal modelling study.

- Report on DCM long-term forecasting - 2nd February 2021
- Report on DCM long-term forecasting - 6th February 2021
- Report on DCM long-term forecasting - 14th February 2021
- Report on DCM long-term forecasting - 21st February 2021
- Report on DCM long-term forecasting - 27th February 2021
- Report on DCM long-term forecasting - 7th March 2021
- Report on DCM long-term forecasting - 14th March 2021
- Report on DCM long-term forecasting - 20th March 2021
- Report on DCM long-term forecasting - 27th March 2021
- Report on DCM long-term forecasting - 4th April 2021
- Report on DCM long-term forecasting - 11th April 2021
- Report on DCM long-term forecasting - 18th April 2021