Estimating the impact of mobility patterns on COVID-19 infection rates in 11 European countries

Patrick Bryant Corresp., 1, 2 , **Arne Elofsson** 1, 2

1 Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

2 Science for Life Laboratory, Solna, Sweden

Corresponding Author: Patrick Bryant
Email address: patrick.bryant@scilifelab.se

Background

As governments across Europe have issued non-pharmaceutical interventions (NPIs) such as social distancing and school closing, the mobility patterns in these countries have changed. Most states have implemented similar NPIs at similar time points. However, it is likely different countries and populations respond differently to the NPIs and that these differences cause mobility patterns and thereby the epidemic development to change.

Methods

We build a Bayesian model that estimates the number of deaths on a given day dependent on changes in the basic reproductive number, R_0, due to differences in mobility patterns. We utilize mobility data from Google mobility reports using five different categories: retail and recreation, grocery and pharmacy, transit stations, workplace and residential. The importance of each mobility category for predicting changes in R_0 is estimated through the model.

Findings

The changes in mobility have a considerable overlap with the introduction of governmental NPIs, highlighting the importance of government action for population behavioral change. The shift in mobility in all categories shows high correlations with the death rates one month later. Reduction of movement within the grocery and pharmacy sector is estimated to account for most of the decrease in R_0.

Interpretation

Our model predicts three-week epidemic forecasts, using real-time observations of changes in mobility patterns, which can provide governments with direct feedback on the effects of their NPIs. The model predicts the changes in a majority of the countries accurately but overestimates the impact of NPIs in Sweden and Denmark and underestimates them in France and Belgium. We also note that the exponential nature of all epidemiological models based on the basic reproductive number, R_0 cause small errors to have extensive effects on the predicted outcome.

Estimating the impact of mobility patterns on COVID-19 infection rates in 11 European countries

Patrick Bryant[1,2] and Arne Elofsson[1,2]

1) Science for Life Laboratory, Solna, Sweden
2) Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

Corresponding Author: Patrick Bryant, email: patrick.bryant@scilifelab.se
Postal address: Box 1031, 171 62 Solna, Sweden

Abstract

Background

As governments across Europe have issued non-pharmaceutical interventions (NPIs) such as social distancing and school closing, the mobility patterns in these countries have changed. Most states have implemented similar NPIs at similar time points. However, it is likely different countries and populations respond differently to the NPIs and that these differences cause mobility patterns and thereby the epidemic development to change.

Methods

We build a Bayesian model that estimates the number of deaths on a given day dependent on changes in the basic reproductive number, R_0, due to differences in mobility patterns. We utilize mobility data from Google mobility reports using five different categories: retail and recreation, grocery and pharmacy, transit stations, workplace and residential. The importance of each mobility category for predicting changes in R_0 is estimated through the model.

Findings

The changes in mobility have a considerable overlap with the introduction of governmental NPIs, highlighting the importance of government action for population behavioral change. The shift in mobility in all categories shows high correlations with the death rates one month later. Reduction of movement within the grocery and pharmacy sector is estimated to account for most of the decrease in R_0.

Interpretation

Our model predicts three-week epidemic forecasts, using real-time observations of changes in mobility patterns, which can provide governments with direct feedback on the effects of their NPIs. The model predicts the changes in a majority of the countries accurately but overestimates the impact of NPIs in Sweden and Denmark and underestimates them in France and Belgium. We also note that the exponential nature of all epidemiological models based on the basic reproductive number, R_0 cause small errors to have extensive effects on the predicted outcome.

Keywords: COVID-19, infection rate, mobility pattern, non-pharmaceutical intervention, Bayesian model

Introduction

In December 2019 a new coronavirus (COVID-19) emerged in Wuhan, China. China implemented a quick strategy of suppression by imposing a lockdown in the city of Wuhan on January 23 (https://www.reuters.com/article/us-china-health-who-idUSKBN1ZM1G9, last accessed 1 May 2020), and implementing social distancing procedures nationwide, with a successful outcome [1]. Still, the virus rapidly spread across the world through our increasingly interconnected flight network, and shortly arrived in Europe. In February 2020 the number of cases started to increase quickly in some European countries. European countries introduced non-pharmaceutical interventions (NPIs) similar to those used in China to limit the spread of the virus. These NPIs include social distancing, school closures, restrict international travel and lockdown [2]. The NPIs results in behavioral changes, and these can be traced by tracking the location of mobile phones.

After an initial rapid spread in China, control measures proved very successful to stop the spread both in China[3] and in other parts of the world[4],[5]. However, there is still a risk for subsequent infections upon lifting of these restrictions[5,6]. There is, therefore, an urgent need both for understanding and tracking the effects of governmental interventions and their removals. Largescale testing could provide valuable information about the impact of interventions. However, these are expensive, sometimes inaccurate and might violate privacy rights. In contrast, the use of largescale data from anonymous tracking of mobile phones is inexpensive and readily available.

Google recently released a time-limited sharing of mobility data (https://www.google.com/covid19/mobility, last accessed 29 March 2020) from across the world as represented by summary statistics to combat COVID-19. The mobility data is measured in 6 different sectors: retail and recreation, grocery and pharmacy, parks, transit stations, workplace and residential. The effects of the government-issued NPIs can be seen through changes in these patterns.

It is likely that different countries respond in different manners to the same NPIs, why it is vital to consider the effect of NPIs country wise. Here, we show that by using real-life mobility data to model changes in the basic reproductive number, R_0, the impact of NPIs across different countries can be modelled more accurately. The mobility data utilised here have some uncertainties and lack resolution. Still, to the best of our knowledge, this data is the best openly available data source for tracking a population's movement in the eleven studied countries. Governments can, in collaboration with telephone companies, obtain much more fine-grained data, enabling them to evaluate the effect of the NPIs in more detail.

Recently, a group from Imperial College released a report [5] that estimates the effects of NPIs on R_0. Subsequently, a modified version of this report was published [7]. The report had a massive impact on how the UK government changed its intervention strategy (https://www.imperial.ac.uk/news/196477/j-ideas-neil-ferguson-tells-mps-lockdown/, last accessed 1 May 2020). A limitation of the ICL model is the assumption that each intervention has the same impact in all countries, ignoring cultural and sociological differences as well as differences in the details of the NPIs. Here, we try to overcome this by developing an extension to their model utilizing country-specific mobility data in a Bayesian framework [8], we estimate the impact of each change in mobility pattern on R_0. The resulting information provides a smooth, straightforward way for governments to analyze if NPIs are working and to what extent. We show that in a three-week forecast, our method makes a better prediction than the model from Imperial College.

Methods

Here, we introduce an MCMC model to estimate the spread of the COVID-19 infection in various countries. The ICL model strongly inspires the model, and all parameters are taken from earlier studies. For each country, we define a starting point when the total number of observed deaths has reached ten. The model is trained using data starting 30 days before this day and until 29 of March 2020. Finally, the model is used to simulate a three-week forecast from 30 March to 19 April.

Infection model

The number of cases acquired at day τ in country m, $c_{\tau,m}$ is modelled with a discrete renewal process [9,10]:

$$c_{\tau,m} = R_{\tau,m}\sum_{\tau=0}^{t-1} c_{\tau,m}\, g_{\tau-t}, \qquad \text{(i)}$$

where

$$g_{\tau-t} \sim Gamma(6.5, 0.62) \qquad \text{(ii)}$$

(Gamma distribution with a mean of 6.5 days and a standard deviation of 0.62 days) is the serial interval distribution used to model the number of cases [5,11].

g_s is discretized in steps of 1 day accordingly:

$$g_s = \int_{\tau=s-0.5}^{s+0.5} g(\tau)\,d\tau \;\; for\; s = 2,3,\ldots \; and\; g_1 = \int_{\tau=0}^{1.5} g(\tau)\,d\tau \qquad \text{(iii)}$$

The number of cases today is thus dependent on the cumulative number of cases from the previous days, weighted by the serial interval distribution, multiplied with the basic reproductive number (R_0) at day t. The discretizations, here and elsewhere, of 1 day are motivated by the intervals in reporting. Just as in the ICL model [5], we assume the starting point for the infection was 30 days before the day after each country has observed ten deaths in total. The time delay of 30 days is necessary due to the relationship between infection and death (see Death model described below). From this assumed starting point, we initialize our model with six days [1] of cases drawn from an Exponential(0.03) distribution, which are inferred in the Bayesian posterior distribution ($D_{t,m}$).

Impact on the basic reproductive number

Our model is based on the model used in the recent report [5] from Imperial College London (ICL). The ICL report tries to estimate the impact of NPIs on R_0 in the same 11 countries modelled here. The main difference between the ICL model and the current one is the modelling of the change of R_0. In the ICL model, the basic reproductive number at day t in country m ($R_{t,m}$) is estimated as a function of the NPI indicators $I_{k,t,m}$ in place at day t in country m as:

$$R_{t,m} = R_{0,m} e^{-\sum_{l=1}^{6} \alpha_k I_{k,t,m}}, \qquad \text{(iv)}$$

where I=1 when intervention k is implemented at day t in country m and α the impact of each intervention.

Here, we instead estimate $R_{t,m}$ to be a function of the relative change in mobility pattern for each country:

$$R_{t,m} = R_{0,m} e^{\alpha_1 I_{1,t,m} + \alpha_2 I_{2,t,m} + \alpha_3 I_{3,t,m} + \alpha_4 I_{4,t,m} - \alpha_5 I_{5,t,m}}, \qquad \text{(v)}$$

where $I_{1\text{-}5,t,m}$ is the relative mobility in retail and recreation, grocery and pharmacy, transit stations, workplace and residential sectors respectively at day t in country m. The residential mobility parameter has a negative sign as it is assumed that when people stay at home it lowers R_0. In our model, we assume that the impact of each relative mobility change has the same relative impact across all countries and across time. This assumption is a requirement to enable the estimation of the impact of mobility on R_0. If the mobility impacts were allowed to differ between countries, it would not be possible to discern between other country-specific factors and the effect of changes in mobility.

The parameter alpha is set to be gamma distributed with mean 0.5 and a standard deviation of 1. A narrow gamma distribution was chosen due to the assumption that the impact on R_0 is almost instantaneous, with an effect that decreases quickly with time. We did not include the data for the mobility category "Parks" as this data displayed much noise and cyclic peaks, possibly caused by varying weather (https://www.google.com/covid19/mobility, last accessed 29 March). The prior for R_0 is set to:

$$R_{0,m} \sim \text{Normal}(2.79|\kappa), \text{ with } \kappa \sim \text{Normal}(0,0.5) \qquad \text{(vi)}$$

The value of 2.79 is chosen from the median value of a recent analysis of 12 modelling studies [12], and the normal distribution from [1].

The relative mobility is modelled as the relative value change compared to a mobility baseline estimated by Google (https://www.google.com/covid19/mobility, last accessed 29 March). The baseline is the median value, for the corresponding day of the week, during the 5-week period of 2020-01-03 to 2020-02-06. For the days for which no mobility data is available, the values were set to 0. The mobility data for the forecast (and days beyond the date for the last available mobility data) was set to the same values as the last observed days. The time points for the interventions were taken from the ICL report[5], whose initial efforts were crowdsourced.

Death model

As the number of deaths in each country is likely to be the most accurate COVID-19 related data, we use this as the core of the model, being the posterior in the Bayesian simulations. The number of deaths in country m at day t is modelled as a negative binomial distribution as used in earlier models [9,13] with mean and variance accordingly:

$$D_{t,m} \sim Negative\ Binomial\ (d_{t,m}, \frac{d_{t,m}^{2}}{\psi}), \psi \sim Normal^{+}\ (0,5) \qquad \text{(vii)}$$

The expected number of deaths, $d_{t,m}$, at day t in country m is given by:

$$d_{t,m} = \sum_{\tau=0}^{t-1} c_{\tau,m} \pi_{t-\tau,m}, \qquad \text{(viii)}$$

where π_m is the infection to death distribution in the country m given by a combination of the infection to onset distribution (Gamma(5.1,0.86)) and onset to death distribution (Gamma(17.8,0.45)) (combined with mean 22.9 days and standard deviation 0.45 days) times the infection fatality rate (ifr) [5],[14],[15] :

$$\pi_{t,m} \sim ifr_m \cdot Gamma(5.1 + 17.8, 0.45) \qquad \text{(ix)}$$

$\pi_{t,m}$ is discretized in steps of 1 day accordingly:

$$\pi_{s,m} = \int_{\tau=s-0.5}^{s+0.5} \pi_m(\tau) d\tau\ for\ s\ =\ 2,3,...\ and\ \pi_{1,m} = \int_{\tau=0}^{1.5} \pi_m(\tau) d\tau \qquad \text{(x)}$$

The *ifrs* are taken from previous estimates of the population at risk is about 1% [16] and adjusted for the predicted attack rate in the age group 50-59 years of age, assuming a uniform attack rate[5,6],[14], chosen due to having the least predicted underreporting in analyses of data from the Chinese epidemic [14]. The uniform attack rate is required due to a lack of age-specific data. The number of deaths today is thus dependent on the cumulative number of cases from the previous days, weighted by the country-specific infection to death distribution.

The implications on R_0 due to relative mobility variations were estimated simultaneously for all countries in a hierarchical Bayesian framework using Markov-Chain Monte-Carlo (MCMC)[8] simulations in Stan[17]. The death data (https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide, last accessed 20200419) used in the form of the number of deaths per day is from ECDC (European Centre of Disease Control), available and updated daily. We ran the model with eight chains, using 4000 iterations (2000 warm-up), as in the earlier work [5,17]. The parameter specifics of the simulation are available in the code (see below).

MCMC Convergence

MCMC simulations are considered to converge when the Rhat statistics (a metric for comparing the variance between pooled and within-chain inferences) reach one[18]. A histogram of Rhat statistics for the modelled parameters in all simulation runs were constructed and analyzed. We also ensure that no divergent transitions were observed by setting the adapt delta in the sampler (see code).

Leave One Country Out Analysis

Since all countries are in different stages of their epidemics, different amounts of data are available for each country. To analyze how the model is influenced by different countries, we fit models using data from all countries except one using all 11 combinations[19]. We then estimate the importance of each mobility parameter in the leave-one-country-out analysis. The relative difference in each mobility parameter provides an estimate of how each country affects R_0 and thus the number of cases and deaths as well. Furthermore, the Pearson correlation coefficients for the mean R_0 across all time points are calculated for each country in the different runs when the other ten were left out (see Figure S1).

Forecast validation

To ensure the forecasts are reliable, we leave out three weeks of data (30 March - 19 April) and fit a model using data from the beginning of the epidemic up to the date for the beginning of the left-out data. We then evaluate the model with one-week intervals from the 30th of March to the 19th of April. We evaluate by the average error and the average fractional error (average error÷Σobserved deaths) during each of the three weeks. We compare our results with simulations obtained from the ICL model [5]. We should note here that the ICL model does not converge for three-week predictions using 4000 iterations (see Figure S2).

EpiEstim estimates of the basic reproductive number (R_0)

To validate our estimates of R_0, we estimate R_0 independently using case data from ECDC and the R package EpiEstim [20], based on the SIR model [21]. The serial interval used for the estimations is variable accordingly: estimate_R(country_cases, method="uncertain_si", config = make_config(list(mean_si = 7.5, std_mean_si = 2,min_mean_si = 1, max_mean_si = 8.4, std_si = 3.4, std_std_si = 1, min_std_si = 0.5, max_std_si = 4, n1 = 1000, n2 = 1000))), allowing more possible scenarios to be explored (see code, Methods section). The R_0 estimates are smoothed using one-week averages, since they are uncertain in the beginning of the epidemic when cases are few. These values are compared with those of the mobility model, only including values under 5 due to the high uncertainty of the larger values in the beginning of the epidemic. The correlations are high and the average errors are low, mainly arising in areas of large uncertainties (see figure S8).

Correlation analysis

To ensure that there is a true relationship between the daily deaths and the mobility changes, correlations between the deaths per day and the different mobility parameters were analyzed. Both the death data and the mobility data were first smoothed using one-week averages. The correlations were made by shifting the daily deaths to infer the time delay of which type of mobility affects the daily deaths. The shifts are from 0-48 days, ensuring all countries have at least ten days of data for the correlation analysis. The correlations, without shifts, between the different mobility parameters, were also analyzed (see Figure S3).

Code

The code is written in Python using the Stan package pystan (v. 2.19.1.1) for MCMC simulations. The code is freely available under the GPLv3 license. https://github.com/patrickbryant1/COVID19.github.io/tree/master/simulations/mobility

Results

Estimating the cumulative number of cases, the number of deaths per day and changes in the basic reproductive number, R_0

In Figure 1, for Italy and Sweden, and Figure S4, for all eleven modelled countries, estimates of cumulative cases, daily deaths and the basic reproductive number R_0 are shown. We simulate a three-week forecast from 30 March to 19 April using data up to 29 March from the European Centre of Disease Control (ECDC) in the form of the number of deaths per day, and relative mobility data estimated by Google (https://www.google.com/covid19/mobility, last accessed 29 March). According to the model, most countries appear to have their epidemic under control (April 19) (Table 1). The most successful nation in terms of reducing R_0 is Italy ($R_0 \approx 0.22$), and the least is Sweden ($R_0 \approx 2.01$).

From Figure S4, it can be seen that in all countries, the interventions have some positive effect, decreasing the estimated R_0 between the epidemic start and March 29. It can be noted that during the development of the epidemic, R_0 displays a wide range of values. In some countries, the mean of the estimated R_0 shows a rapid increase to values as high as 10, coupled with an increase in mobility (primarily) to grocery and pharmacies exactly when the interventions were introduced. Most posterior distributions for the mean R_0 values are centered around the prior of 2.79 (Figure 2). Notable is that Italy and Spain, which both had very rapid spread have distributions centered higher than the prior.

The estimated number of deaths for up to three weeks after the model is trained, have a good correspondence with the observed number (Figures 1, S4 and Table 2). Compared with the Imperial College London (ICL) model [5], our model displays both lower errors and less uncertainty (see Figures 3, S5 and Table S1). The average absolute errors over the 11 countries in the number of deaths are lower across all three weeks (week 1: 60 vs 159, week 2: 95 vs 472, and week 3: 88 vs 1429 for our model and the ICL model respectively).

Comparing mobility data across countries

When overlaying the implementation dates of the NPIs with the mobility data, it is clear that governmental decisions have a significant impact on the populations in the 11 modelled countries (see Figure S4). Most countries display very similar relative changes in their mobility patterns, with mobility in retail and recreation, grocery and pharmacy, transit stations and workplace decreasing, while mobility in the residential category is increasing.

Most countries have similar relative changes across the sectors (Figure S4). The ones that display smaller relative changes (Denmark, Norway and Sweden) also demonstrate more modest reductions in R_0, which is a natural consequence of our model, as it assumes that changes in R_0 are directly related to changes in mobility. The mobility patterns in Sweden display barely half of the relative changes compared with France, Spain, and Italy, and the reduction in R_0 is, therefore, smaller in Sweden.

The importance of mobility sectors for modelling changes in R_0

Analyzing the importance of each mobility parameter for predicting the reduction in R_0 ($1-e^{-alpha}$) shows that the grocery and pharmacy sector appears to be the clearest indicator for R_0 change (see Figure 4). The grocery and pharmacy sector is estimated to account for most of the reduc revision2_trackedtion of R_0, with a median reduction of 95.6 % compared to less than 10 % for the other sectors (retail and recreation 3.8 %, transit stations 3.0 %. workplace 4.0 %, residential 7.9 %).

Investigating the correlation between the deaths per day and the different mobility parameters (Figure 5), one can see that all sectors display high opposite correlations with a shift of about 20 days. These correlations are due to the time-delayed relationship between the initial spread of the disease, causing deaths occurring after the reduction in mobility, see Figure S4. The mobility changes have the highest correlations with the number of deaths 30-40 days after they occur, suggesting that the mobility affects the death rate with a time delay of 30-40 days. Roughly in agreement with the 22.9 days in our model. Since the grocery and pharmacy sector displays the most significant correlations, the model assigns most weight to that sector, although the mobility in all sectors is highly correlated with each other (Figure S3).

Model validation

The posterior distributions for the mobility parameters (see Figure S6) are almost identical in the leave-one-country-out analysis (LOO) analysis. A bimodal distribution is observed when leaving Italy out in the grocery and pharmacy sector though, emphasizing the importance of the Italian data. The variable R_0 values in the LOO analysis show Pearson correlations close to 1, with Italy and especially the United Kingdom displaying lower correlations of around 0.8 and consistently below 0.8 respectively (see Figure S1). Italy and the United Kingdom correlate badly with each other, with Pearson correlations of close to 0. 11 of 4000 iterations ended with a divergence (0.275 %) Spain was excluded. A histogram of Rhat statistics for the modelled parameters in all simulations for the main analysis is displayed in Figure S7.

To validate the R_0 estimates, we used a SIR model using EpiEstim [20] to estimate R_0 independently from case data (and not death data as in our and the ICL models). This model does not try to determine the cause of changes in R_0, but just estimates the changes from the number of reported cases. In general, the overlap of the two estimates of R_0 estimates is high, in particular at the crucial time points before and after the effects of NPI implementation (see Table S2 and Figure S8). Denmark, Norway and Spain display the most substantial differences between the estimates, differing 2.98, 1.94 and 3.48 respectively at the point before NPI implementation. The differences that do arise are mainly during the periods with considerable uncertainty in the R_0 estimates, i.e. when the number of reported cases is low. Sweden shows the most substantial error between the estimates after NPI implementation (0.98). Further, the models show very different speeds of the changes in R_0 values, EpiEstim having a much slower response than the mobility model.

Discussion

The model makes it clear that the non-pharmaceutical interventions (NPIs) introduced by governments across Europe have had substantial effects on both mobility patterns and in preventing the spread of COVID-19. By tracking the relative change in mobility in the grocery and pharmacy sector, it is possible to account for most of the reduction in the basic reproductive number, R_0, in our model. This information can, therefore, provide a useful, straightforward way for governments to analyze the effect of their NPIs.

Why the grocery and pharmacy sector has been assigned the highest importance is likely because this sector displays the strongest correlation with the daily deaths. The correlations are highest assuming a 30-40 day shift, suggesting that mobility affects the death rate with a time delay of 30-40 days, in rough agreement with our model. Since R_0 is strongly dependent on the changes in mobility, rapid changes in mobility lead to rapid changes in R_0, with drastic consequences to the estimated development of the epidemic in a country. However, changes in R_0 will not manifest in the number of deaths per day until about three weeks later (the mean value in the gamma distribution for infection to death is 22.9 days, see methods section). Therefore, a three-week forecast is provided.

The estimates have an acceptable correspondence with the observed numbers in most countries (see Figure 3 and Table 2), and compared with the ICL-model, our model displays both lower errors and less uncertainty (Figures 3, S5 and Tables 2, S1). It can also be noted that the ICL model overpredicts the number of deaths in all countries. The higher accuracy when including mobility data, further suggests the usefulness of our model.

The estimated number of cases has considerable uncertainty across all countries. One limitation of our model is that it does not take herd-immunity effects into account, which should be reached when around 60-80 % of the population is infected [22]. Still, it is unlikely that sufficiently high infection has been reached yet for this to have a significant effect. Another limitation of the model is the assumption that the impact of each relative mobility change has the same relative impact across all countries and across time. If the mobility impact were allowed to differ between countries and in time, it would not be possible to discern between other country-specific and time factors and the mobility impact. Likely both more detailed mobility data and intermixing patterns need to be considered, metrics that are not available.

The number of cases is also highly dependent on having the correct infection-fatality-rate (*ifr*). This quantity is only modelled for the age group 50-59 years and does thereby not consider the attack rates for the whole of each country's population (see methods section). If a nation managed to avoid the elderly being infected, that would lower the *ifr* [23], which could explain prediction differences to some extent.

The model validation, by a leave-one-country-out analysis, comparing with independent R_0 estimates from EpiEstim [20] and predicting a three-week forecast ensures the model's robustness. The leave-one-country-out analysis shows that the estimates are mostly affected by the data from Italy and the UK, likely due to these countries having more available data and higher death tolls early in the epidemic, making the model somewhat biased to these data in the beginning of the estimates (Figure S4). The comparison with the R_0 estimates from EpiEstim show differences that arise mainly during the periods with considerable uncertainty in the R_0 estimates, i.e. when the number of reported cases are low. The estimates also show very different speeds of the changes in R_0 values, EpiEstim having a much slower response than the mobility model (Figure S8).

The countries in the three-week forecast where the errors stand out are Denmark and Sweden, with over-predictions, and Belgium and France, which are under-predicted. We note that these two pairs of countries are close both geographically and culturally [24,25], possibly explaining the systematic differences. The differences may also be caused by differences in reporting between the countries (https://www.bloomberg.com/news/articles/2020-04-09/french-virus-deaths-jump-with-more-nursing-home-patients-counted, last accessed May 1; https://www.politico.com/news/2020/04/19/why-is-belgiums-death-toll-so-high-195778, last accessed May 1). For instance, on April 5 more than 2000 deaths were reported in France, due to sudden inclusion of potential COVID-19 attributed deaths in nursing homes occurring at earlier dates (https://www.usnews.com/news/world/articles/2020-04-02/frances-coronavirus-death-toll-jumps-to-nearly-5-400-as-nursing-homes-included, last accessed May 1). We note the sensitivity to small errors of all epidemic models using exponential measures, such as the basic reproductive number, and the significant effects these minor errors have on the predicted outcome.

Conclusions

Here, we present a model to estimate the effects of public interventions on the spread of COVID-19 that does not assume that interventions have identical results in different geographical and cultural settings. In contrast, our model uses *observational* data of mobility patterns in five environments to estimate changes in the transmission rate. Our model creates the possibility to track rapid changes in the spread, instantaneously and predict their consequences three weeks ahead in time. Therefore, our model enables governments to use anonymous real-time data to adjust their policies. We do foresee that such models will become incrementally more powerful as more detailed mobility data becomes available in the future.

Acknowledgements

We acknowledge Claudio Bassot's contribution by sharing both the Imperial College London report and Google mobility data. Without this information, this study would not be possible. We are also grateful to various colleagues and friends that contributed to the discussion. Finally, we thank the authors of the Imperial College London report for making their data and model freely available.

Declaration of interests

We declare no competing interests.

Funding statement

Financial support: Swedish Research Council for Natural Science, grant No. VR-2016-06301 and Swedish E-science Research Center. Computational resources: Swedish National Infrastructure for Computing, grant No. SNIC-2019/3-319.

References

1. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, Shaman J. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science. 2020. doi:10.1126/science.abb3221

2. European Centre of Disease Control. Coronavirus disease 2019 (COVID-19) pandemic: increased transmission in the EU/EEA and the UK –seventh update [Internet]. 2020. [cited 23 Apr 2020]. Available: https://www.ecdc.europa.eu/sites/default/files/documents/RRA-seventh-update-Outbreak-of-coronavirus-disease-COVID-19.pdf

3. Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luo W, Floyd JR, Wesolowski A, Santillana M, Zhang C, Du X, Yu H, Tatem AJ. Effect of non-pharmaceutical interventions for containing the COVID-19 outbreak in China. medRxiv (preprint). doi:10.1101/2020.03.03.20029843

4. Milne GJ, Xie S. The Effectiveness of Social Distancing in Mitigating COVID-19 Spread: a modelling analysis. medRxiv (preprint). doi:10.1101/2020.03.20.20040055

5. Flaxman S, Mishra S, Gandy A, Unwin HJT , Coupland H, Mellan TA , Zhu H, Berah T, Eaton JW, Guzman PNP, Schmit N, Cilloni L, Ainslie KEC, Baguelin M, Blake I, Boonyasiri A, Boyd O, Cattarino L, Ciavarella C, Cooper L, Cucunubá Z, Cuomo-Dannenburg G, Dighe A, Djaafara B, Dorigatti I, van Elsland S, FitzJohn R, Fu H, Gaythorpe K, Geidelberg L, Grassly N, Green W, Hallett T, Hamlet A, Hinsley W, Jeffrey B, Jorgensen D, Knock E, Laydon D, Nedjati-Gilani G, Nouvellet P, Parag K, Siveroni I, Thompson H, Verity R, Volz E, Walters C, Wang H, Wang Y, Watson O, Winskill P, Xi X, Whittaker C, Walker PGT, Ghani A, Donnelly CA, Riley S, Okell LC, Vollmer MAC, Ferguson NM and Bhatt S. Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries (Report 13). 2020 March 30 (preprint). In: Imperial College London [Internet]. [cited 23 Apr 2020]. Available: http://www.imperial.ac.uk/medicine/departments/school-public-health/infectious-disease-epidemiology/mrc-global-infectious-disease-analysis/covid-19/report-13-europe-npi-impact/

6. Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Ilaria Dorigatti, Han Fu, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Lucy C Okell, Sabine van Elsland, Hayley Thompson, Robert Verity, Erik Volz, Haowei Wang, Yuanrong Wang, Patrick GT Walker, Caroline Walters, Peter Winskill, Charles Whittaker, Christl A Donnelly, Steven Riley, Azra C Ghani. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand (Report 9). 2020 March 16 (preprint). Imperial College London. [cited 23 Apr 2020]. doi:10.25561/77482

7. Flaxman S, Mishra S, Gandy A, Unwin JH, Mellan TA, Coupland H, Whittaker Charles, Zhu H, Berah T, Eaton JW, Monod M, Imperial College COVID-19 Response Team; Ghani AC, Donnelly CA, Riley SM, Vollmer MAC, Ferguson NM, Okell LC, Bhatt S. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020. doi:10.1038/s41586-020-2405-7

8. Banerjee, S. Carlin, B. P., and Gelfand, A. E. Hierarchical Modeling and Analysis for Spatial Data. Second Edition. Boca Raton, Florida: CRC Press/Chapman & Hall;2015 ISBN-13: 978-1-4398-1917-3.

9. Fraser C. Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic. PLoS ONE. 2007. p. e758. doi:10.1371/journal.pone.0000758

10. Cauchemez S, Valleron A-J, Boëlle P-Y, Flahault A, Ferguson NM. Estimating the impact of school closure on influenza transmission from Sentinel data. Nature. 2008;452: 750–754.

11. Backer JA, Klinkenberg D, Wallinga J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. Eurosurveillance. 2020. doi:10.2807/1560-7917.es.2020.25.5.2000062

12. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. J Travel Med. 2020;27. doi:10.1093/jtm/taaa021

13. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. Nature. 2005;438: 355–359.

14. Verity, Robert; Okell, Lucy C; Dorigatti, Ilaria; Winskill, Peter; Whittaker, Charles; Imai, Natsuko; Cuomo-Dannenburg, Gina; Thompson, Hayley; Walker, Patrick G T; Fu, Han; Dighe, Amy; Griffin, Jamie T; Baguelin, Marc; Bhatia, Sangeeta; Boonyasiri, Adhiratha; Cori, Anne; Cucunubá, Zulma; FitzJohn, Rich; Gaythorpe, Katy; Green, Will; Hamlet, Arran; Hinsley, Wes; Laydon, Daniel; Nedjati-Gilani, Gemma; Riley, Steven; van Elsland, Sabine; Volz, Erik; Wang, Haowei; Wang, Yuanrong; Xi, Xiaoyue; Donnelly, Christl A; Ghani, Azra C; Ferguson, Neil M. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis. 2020. doi:10.1016/S1473-3099(20)30243-7

15. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, Azman AS, Reich NG, Lessler J. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Ann Intern Med. 2020;172: 577–582.

16. Lourenco J, Paton R, Ghafari M, Kraemer M, Thompson C, Simmonds P, Klenerman P, Gupta S. Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. medRxiv. 2020; 2020.03.24.20042291.

17. Stan. In: stan-dev.github.io [Internet]. [cited 23 Apr 2020]. Available: https://mc-stan.org

18. Brooks SP, Gelman A. General Methods for Monitoring Convergence of Iterative

Simulations. Journal of Computational and Graphical Statistics. 1998. pp. 434–455. doi:10.1080/10618600.1998.10474787

19. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media; 2013.

20. Thompson RN, Stockwin JE, van Gaalen RD, Polonsky JA, Kamvar ZN, Demarsh PA, Dahlqwist E, Li S, Miguel E, Jombart T, Lessler, J, Cauchemez S, Cori A. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics. 2019;29: 100356.

21. McKendrick AG. Applications of Mathematics to Medical Problems. Proc Edinb Math Soc. 1925;44: 98–130.

22. Kwok KO, Lai F, Wei WI, Wong SYS, Tang JWT. Herd immunity – estimating the level required to halt the COVID-19 epidemics in affected countries. Journal of Infection. 2020. doi:10.1016/j.jinf.2020.03.027

23. Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Med. 2020. doi:10.1007/s00134-020-05991-x

24. Warner-Søderholm G. Culture Matters. SAGE Open. 2012. p. 215824401247135. doi:10.1177/2158244012471350

25. Hofstede GH, Hofstede G. Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations. SAGE; 2001.

Table 1 (on next page)

Changes in R_o and mobility in the Grocery and Pharmacy sector during the epidemic.

Country	Modelled start of the epidemic	Estimated mean R_0 at epidemic start	Estimated mean R_0 at 29 March	Relative change in Groceries and pharmacies on 29 March
Austria	2020-02-22	3.11	0.36	−64%
Belgium	2020-02-18	3.24	0.51	−53%
Denmark	2020-02-21	3.02	1.36	-22%
France	2020-02-07	2.91	0.30	−72%
Germany	2020-02-15	3.08	0.56	-51%
Italy	2020-01-27	3.17	0.22	−85%
Norway	2020-02-24	2.82	0.92	-32%
Spain	2020-02-09	3.19	0.29	−76%
Sweden	2020-02-18	2.89	2.01	−10%
Switzerland	2020-02-14	2.81	0.53	-51%
United Kingdom	2020-02-12	2.82	0.61	-46%

Table 2(on next page)

Average error and average fractional error in the number of deaths for the mobility model.

Average error and average fractional error in the number of deaths for each country between the mean predicted number of deaths per day and the observed number in one, two and three week forecasts respectively. A corresponding table for the ICL model can be found in Table S2.

Three-week predictions for the number of deaths per day						
	Average error			Average fractional error		
Country	**week 1**	**week 2**	**week 3**	**week 1**	**week 2**	**Week 3**
Austria	−3	−6	−2	−2.3 %	−4.0 %	−1.7 %
Belgium	−46	−179	−186	−5.0 %	−8.7 %	−8.8 %
Denmark	0	10	28	0.4 %	10.2 %	32.2 %
France	−318	−445	−427	−6.1 %	−7.1 %	−7.8 %
Germany	−21	−7	−26	−2.2 %	−0.5 %	−1.6 %
Italy	144	201	29	2.7 %	4.9 %	0.8 %
Norway	1	1	3	1.7 %	2.2 %	6.9 %
Spain	−98	84	−8	−1.6 %	1.8 %	−0.2 %
Sweden	−3	28	180	−1.2 %	5.4 %	28.9 %
Switzerland	13	41	48	4.3 %	14.1 %	17.1 %
United Kingdom	17	−42	32	0.4 %	−0.7 %	0.5 %
Average absolute error	60	95	88	2.5 %	5.4 %	9.7 %

Figure 1

Model results for Italy and Sweden.

Model results in the form of the cumulative number of cases, deaths per day and R_0 for Italy and Sweden, are displayed on the left axes. The model results start from 30 days before ten accumulated deaths had been observed. The blue curves represent the estimations so far, while the green represents a three-week forecast (30 March-19 April). The 50 % and 95 % confidence intervals are displayed in darker and lighter shades respectively, with the mean as a solid line. The histograms represent the number of cases and deaths reported by the European Center for Disease Control (ECDC). Mobility data for the five modelled sectors represented in terms of relative change compared to baseline (observed in a five-week period of 2020-01-03 to 2020-02-06) is displayed on the right axes. The dates for the introduction of different NPIs are marked with vertical lines. As can be seen, the NPIs have very strong implications for the mobility patterns. The mobility data ranges from 2020-02-15 to 2020-03-29, after which the final levels are fixed. The graph for Rt includes a dashed horizontal line marking the value one of halted epidemic growth.

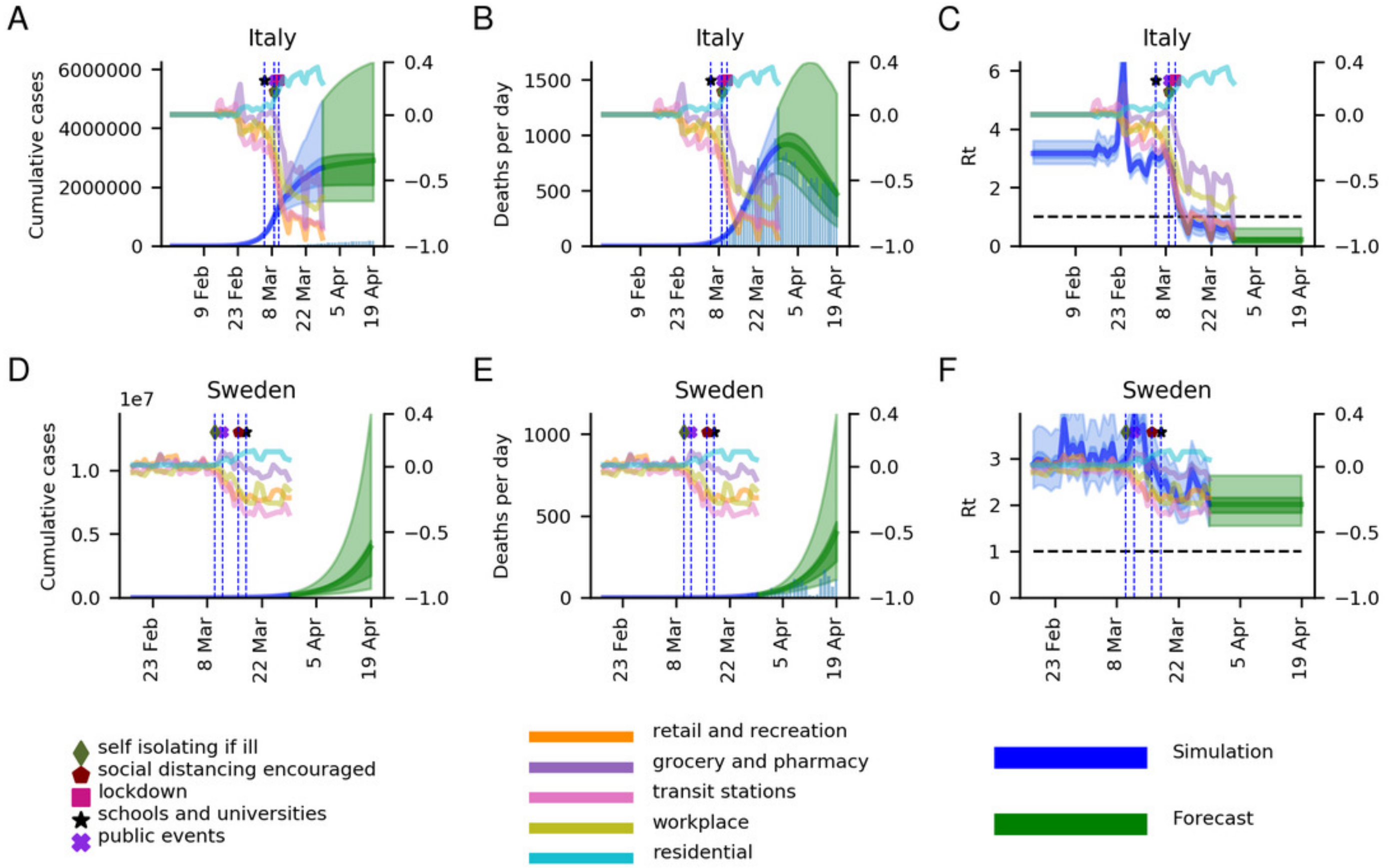
A
Italy
Cumulative cases
B
Italy
Deaths per day
C
Italy
Rt
D
Sweden
Cumulative cases
1e7
E
Sweden
Deaths per day
F
Sweden
Rt
9 Feb
23 Feb
8 Mar
22 Mar
5 Apr
19 Apr
self isolating if ill
social distancing encouraged
lockdown
schools and universities
public events
retail and recreation
grocery and pharmacy
transit stations
workplace
residential
Simulation
Forecast

Figure 2

Posterior distributions for the mean initial R0 sampled per country. The dashed line corresponds to the prior mean, set to 2.79.

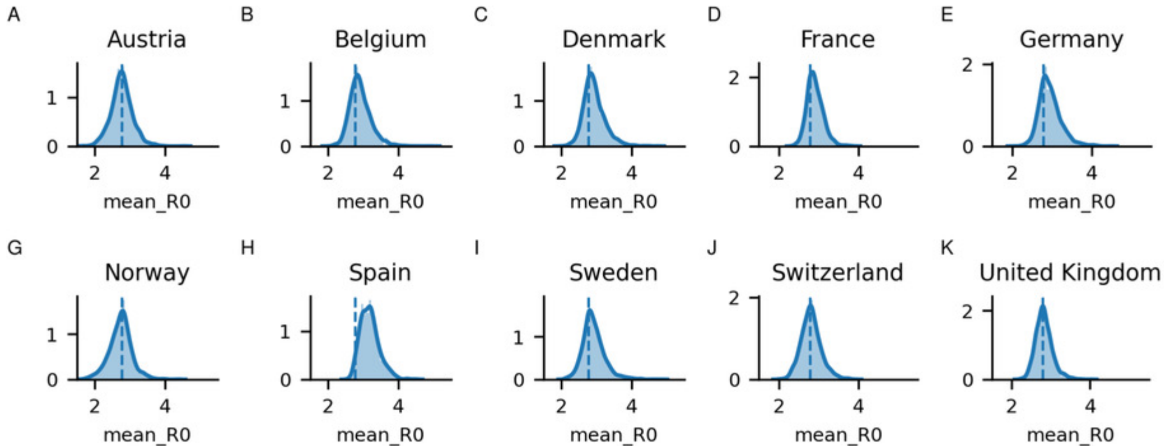

Figure 3

Three-week predictions for all countries.

Three-week predictions for all countries in the form of deaths per day for the weeks 1: (Mar 30 - April 5), week 2 (April 6 - April 12) and week 3 (April 13 - April 19). The 50 % and 95 % confidence intervals are displayed in darker and lighter shades respectively, with the mean as a solid line. The blue histogram represents the observed deaths.

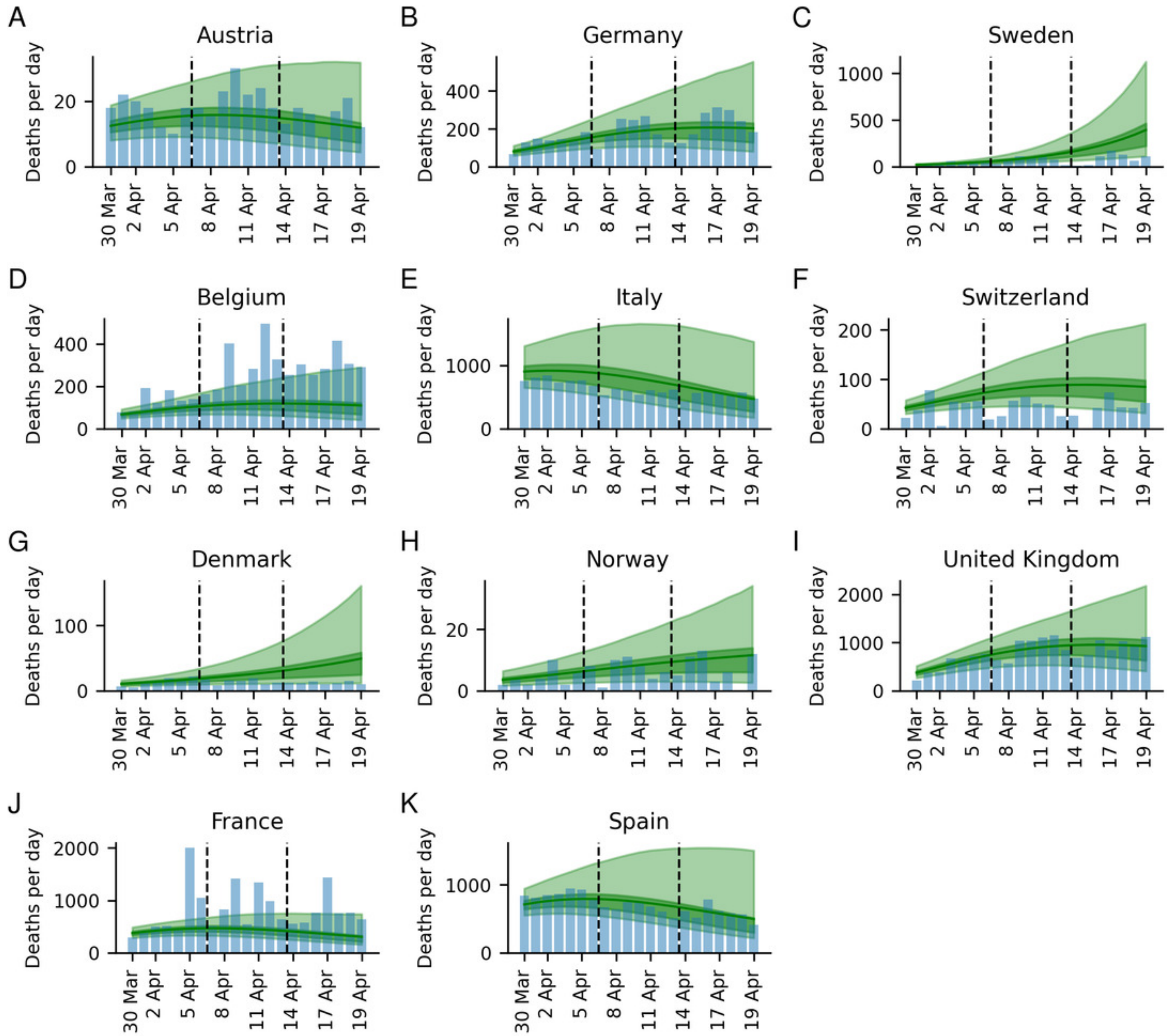
A
Austria
B
Germany
C
Sweden
D
Belgium
E
Italy
F
Switzerland
G
Denmark
H
Norway
I
United Kingdom
J
France
K
Spain
Deaths per day
30 Mar
2 Apr
5 Apr
8 Apr
11 Apr
14 Apr
17 Apr
19 Apr

Figure 4

Posterior distributions of the impact of each mobility parameter.

Posterior distributions of the impact of each mobility parameter for predicting the reduction in R0. The grocery and pharmacy sector appears to be the clearest indicator for R0 change. The median impacts are 3.8, 95.6, 3.0, 4.0 and 7.9 % for the retail and recreation, grocery and pharmacy, transit, workplace and residential sectors respectively.

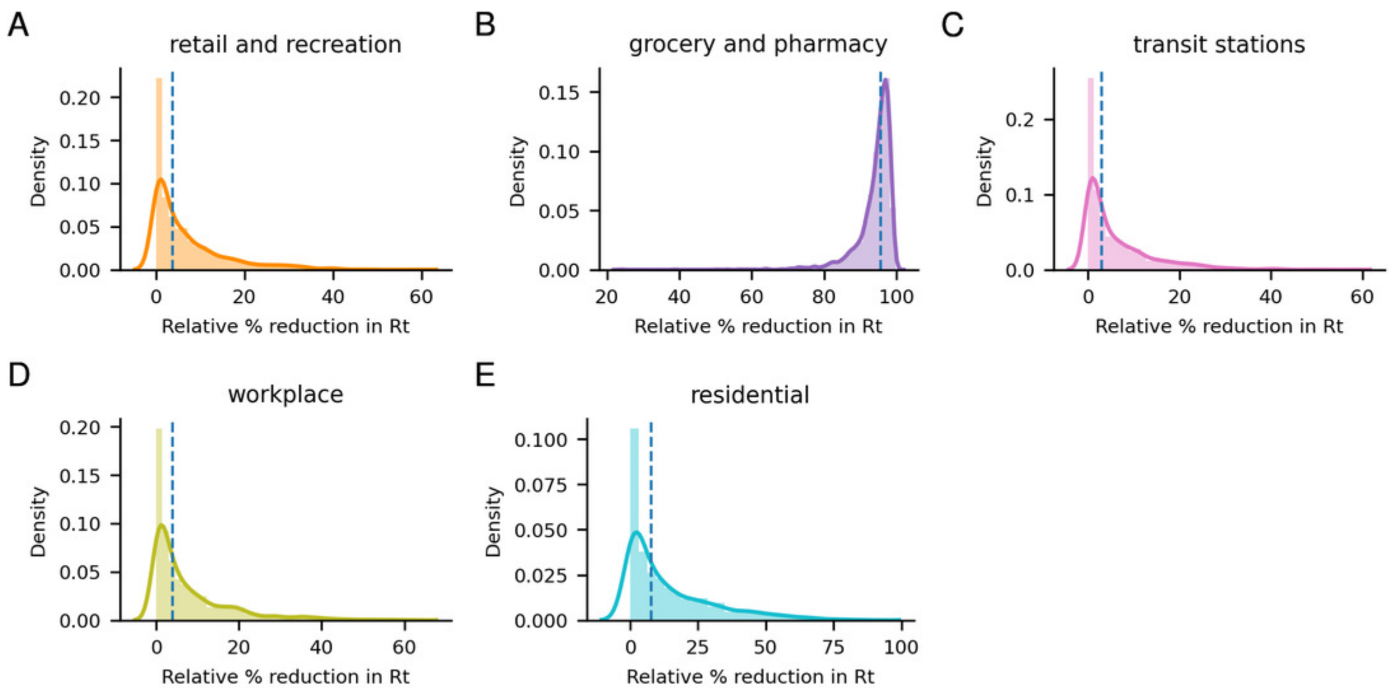

Figure 5

Correlation between daily deaths and mobility changes.

Correlation between deaths per day and mobility changes for different time delays. Each country is represented by one line. The mobility changes have the highest correlations with the deaths about 30-40 days after they occur, suggesting that mobility affects the death rate with a time delay of 30-40 days.