TITLE

Ilia Kohanovski[a], Uri Obolski[b,c], and Yoav Ram[a,*]

[a]School of Computer Science, Interdisciplinary Center Herzliya, Herzliya 4610101, Israel
[b]School of Public Health, Tel Aviv University, Tel Aviv 6997801, Israel
[c]Porter School of the Environment and Earth Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
[*]Corresponding author: yoav@yoavram.com

April 27, 2020

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Introduction

The COVID-19 pandemic has resulted in extreme non-pharmaceutical interventions (NPIs) in many affected countries. These interventions, from social distancing to lockdowns, are applied in a rapid and widespread fashion. The NPIs are designed and assessed using epidemiological models, which follow the dynamics of the viral infection to forecast the effect of different mitigation and suppression strategies on the levels of infection, hospitalization, and fatality. However, compliance to NPIs differ between interventions and people and may be impacted both by the number of domestic cases as well as by the number of cases in other regions and countries. For example, using a survey of 2,108 adults in the UK during Mar 2020, Atchison et al.[2] report that those over 70 years old were more likely to adopt social distancing than young adults (18-34 years), and that those with lower income were less likely to be able work from home and to self-isolate.

However, most epidemiological models assume the effect of NPIs on the epidemiological dynamics to begin at their officially declared date (e.g. Gatto et al.[6], Li et al.[7]). Even models that allow the effect of NPIs to by more dynamics usually assume that it increases or decreases as a function of time (e.g. Banholzer et al.[3]). Here we apply a *Susceptible-Exposed-Infected-Recovered* (SEIR) disease transmission model and *Markov Chain Monte Carlo* (MCMC) parameter estimation framework to estimate the effective start date of NPIs in several geographical regions using publicly available confirmed COVID-19 case data.

Models and Methods

All source code will be publicly available under a permissive open-source license at github.com/yoavram-lab/EffectiveNPI.

Data. We use daily confirmed case data $\mathbf{X} = (X_1, \ldots, X_T)$ from several different countries. These incidence data summarize the number of individuals X_t tested positive for SARS-CoV-2 RNA (using RT-qPCR) at each day t. Data was retrieved for X regions, see Table 1 for details and references. In regions in which there were multiple sequences of days with zero confirmed cases (e.g. France), we cropped the data to begin with the last sequence so that our analysis focuses on the first community-transmitted outbreak rather then isolated imported cases.

Region	Start date	End date	Reference
Austria	X Feb		Flaxman et al.[4]
Wuhan, China	10 Jan	8 Feb	Pei and Shaman[8]

Table 1: Reference for confirmed cases incidence data. All dates in 2020.

SEIR model. We model SARS-CoV-2 infection dynamics by following the number of susceptible S, exposed E, reported infected I_r, and unreported infected I_u individuals in a population of size N. This model distinguishes between reported and unreported infected individuals: the reported infected are those that have enough symptoms to eventually be tested and thus appear in daily case reports, to which we fit the model.

Susceptible (S) individuals become exposed due to contact with reported or unreported infected individuals (I_r or I_u) at a rate β_t or $\mu\beta_t$. The parameter $0 < \mu < 1$ represents the decreased transmission rate from unreported infected individuals, who are often subclinical or even asymptomatic. The

transmission rate $\beta_t \geq 0$ may change over time t due to behavioral changes of both susceptible and infected individuals. Exposed individuals, after an average incubation period of Z days, become reported infected with probability α_t or unreported infected with probability $(1 - \alpha_t)$. The reporting rate $0 < \alpha_t < 1$ may also change over time due to changes in human behavior. Infected individuals remain infectious for an average period of D days, after which they either recover, or becomes ill enough to be quarantined. They therefore no longer infect other individuals, and the model does not track their frequency. The model is described by the following equations:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta_t S\frac{I_p}{N} - \mu\beta_t S\frac{I_s}{N} \\
\frac{dE}{dt} &= \beta_t S\frac{I_p}{N} + \mu\beta_t S\frac{I_s}{N} - \frac{E}{Z} \\
\frac{dI_r}{dt} &= \alpha_t\frac{E}{Z} - \frac{I_r}{D} \\
\frac{dI_u}{dt} &= (1-\alpha_t)\frac{E}{Z} - \frac{I_r}{D}.
\end{aligned} \tag{1}$$

The initial numbers of exposed $E(0)$ and unreported infected $I_u(0)$ are considered model parameters, whereas the initial number of reported infected is assumed to be zero $I_r(0) = 0$, and the number of susceptible is $S(0) = N - E(0) - I_u(0)$. The vector θ of model parameters is

$$\theta = \Big(Z, D, \mu, \{\beta_t\}, \{\alpha_t\}, \{p_t\}, E(0), I_u(0)\Big). \tag{2}$$

This model is inspired by Li et al.[7] and Pei and Shaman[8], who used a similar model with multiple regions and constant transmission β and reporting rate α to infer COVID-19 dynamics in China and the continental US, respectively.

Likelihood function. The *expected* cumulative number of reported infected individuals until day t is

$$Y_t = \int_0^t \alpha_s \frac{E(s)}{Z}\, ds, \quad Y_0 = 0. \tag{3}$$

We assume that reported infected individuals are confirmed and therefore observed in the daily case report of day t with probability p_t (note that an individual can only be observed once, and that p_t may change over time, but t is a specific date rather than the time elapsed since the individual was infected). Hence, we assume that the number of confirmed cases in day t is binomially distributed,

$$X_t \sim Bin\big(n_t, p_t\big),$$

where n_t is the *realized* (rather than expected) number of reported infected individuals yet to appear in daily reports by day t. The cumulative number of confirmed cases until day t is

$$\tilde{X}_t = \sum_{i=1}^{t} X_i, \quad X_0 = 0.$$

Given $\tilde{X}_{t-1}$, we assume n_t is Poisson distributed,

$$\big(n_t \mid \tilde{X}_{t-1}\big) \sim Poi\Big(Y_t - \tilde{X}_{t-1}\Big), \quad n_1 \sim Poi(Y_1).$$

Therefore, $\big(X_t \mid \tilde{X}_{t-1}\big)$ is a binomial conditioned on a Poisson, which reduces to a Poisson with

$$\big(X_t \mid \tilde{X}_{t-1}\big) \sim Poi\Big(\big(Y_t - \tilde{X}_{t-1}\big) \cdot p_t\Big), \quad X_1 \sim Poi(Y_1 \cdot p_1). \tag{4}$$

For given vector θ of model parameters (Eq. (2)), we compute the expected cumulative number of reported infected individuals $\{Y_t\}_{t=1}^{T}$ for each day (Eq. (3)). Then, since $\tilde{X}_{t-1}$ is a function of $X_1, \ldots, X_{t-1}$, we can use Eq. (4) to write the probability to observe the confirmed case data $\mathbf{X} = (X_1, \ldots, X_T)$ as

$$\mathbb{L}(\theta \mid \mathbf{X}) = P(\mathbf{X} \mid \theta) = P(X_1 \mid \theta) P(X_2 \mid \tilde{X}_1, \theta) \cdots P(X_T \mid \tilde{X}_{T-1}, \theta). \tag{5}$$

This defines a *likelihood function* $\mathbb{L}(\theta \mid \mathbf{X})$ for the parameter vector θ given the data $\mathbf{X}$.

NPI model. To model non-pharmaceutical interventions (NPIs), we set the beginning of the NPIs to day τ and define

$$\beta_t = \begin{cases} \beta, & t < \tau \\ \beta\lambda, & t \geq \tau \end{cases}, \quad \alpha_t = \begin{cases} \alpha_1, & t < \tau \\ \alpha_2, & t \geq \tau \end{cases}, \quad p_t = \begin{cases} 1/9, & t < \tau \\ 1/6, & t \geq \tau \end{cases}, \tag{6}$$

where $0 < \lambda < 1$. The values for p_t follow Li et al.[7], who estimated the average time between infection and reporting in Wuhan, China, at 9 days before the start of NPIs (Jan 23, 2020) and 6 days after start of NPIs. The parameter τ is then added to the parameter vector θ (Eq. (2)).

Parameter estimation. To estimate the parameters θ of our model (Eq. (1)) from the data $\mathbf{X}$, we apply a Bayesian inference approach. We define the following flat priors on the model parameters $P(\theta)$:

$$\begin{aligned} Z &\sim \mathit{Uniform}(2, 5) \\ D &\sim \mathit{Uniform}(2, 5) \\ \mu &\sim \mathit{Uniform}(0.2, 1) \\ \beta &\sim \mathit{Uniform}(0.8, 1.5) \\ \lambda &\sim \mathit{Uniform}(0, 1) \\ \alpha_1, \alpha_2 &\sim \mathit{Uniform}(0.02, 1) \\ E(0) &\sim \mathit{Uniform}(0, 3000) \\ I_u(0) &\sim \mathit{Uniform}(0, 3000) \\ \tau &\sim \mathit{Uniform}(1, T-1), \end{aligned} \tag{7}$$

where T is the number of days in the data $\mathbf{X}$. Most priors follow Li et al.[7], except λ, which is used to enforce that the transmission rates are lower after the start of the NPIs ($\lambda < 1$). The likelihood function is defined in Eq. (5). The posterior distribution on the model parameters $P(\theta \mid \mathbf{X})$ is then estimated using an *affine-invariant ensemble sampler for Markov chain Monte Carlo* (MCMC) implemented in the `emcee` Python package[5].

Model selection. We perform model selection using DIC (deviance information criterion)[9],

$$\begin{aligned} DIC(\theta, \mathbf{X}) &= 2\mathbb{E}[D(\theta)] - D(\mathbb{E}[\theta]) \\ &= 2 \log \mathcal{L}(\mathbb{E}[\theta] \mid \mathbf{X}) - 4\mathbb{E}[\log \mathcal{L}(\theta \mid \mathbf{X})], \end{aligned} \tag{8}$$

where $D(\theta)$ is the Bayesian deviance, and expectations $\mathbb{E}[\cdot]$ are taken over the posterior distribution $P(\theta \mid \mathbf{X})$. We compare models by reporting their relative DIC; lower is better.

Results

Several studies have described the effects of non-pharmaceutical interventions in different regoins[4,6,7]. These studies have assumed that the parameters of the epidemiological model change at a specific date, as in Eq. (6), and set the change date τ to the official NPI date τ^*. They then fit the model once for $t < \tau^*$ and once for $t \geq \tau^*$ (see TABLE2 for a summary of official NPI dates.) For example, Li et al.[7] estimate the dynamics in China before and after τ^* at Jan 23. Thereby, they effectively estimate (β, α_1) and (λ, α_2) separately.

Here we estimate the posterior distribution of *effective* start date of the NPI, $P(\tau \mid \mathbf{X}$, as well as maximum a priori (MAP) estimates, $\hat{\tau}$, by jointly estimating $\tau, \beta, \lambda, \alpha_1, \alpha_2$ on the entire time series per region (e.g. Italy, Austria), rather than splitting the region time series at τ^*. In all examined cases the effect of an NPI is significant: the DIC of a model without NPI ($\beta_t \equiv \beta, \alpha_t \equiv \alpha, p_t \equiv p$ for all t) was higher than the DIC of a model with NPI (Eq. (6)) by at least Z. Therefore, FIGURE compares the official dates τ^* and our MAP estimates $\hat{\tau}$, with confidence intervals. It can be seen that un most regions $\hat{\tau}$ and τ^* differ significantly: that is, the effective start of NPI was either advanced or delayed compared to the official date. Do we want to report DIC of model with τ compared to model with fixed $\tau = \tau^*$? Or just that ($P(\tau \neq \tau^*) > zzz$) ? Or confidence intervals?

In the following, we describe our findings on delayed and advanced start of NPI.

Delayed effective start of NPI. We find that our MAP estimates $\hat{\tau}$ often differ significantly from the official dates tau^*. For example, in Italy, the first case officially confirmed on Feb 21, a lockdown was delayed in Northern Italy on Mar 8, with social distancing implemented in the rest of the country, and the lockdown was extended to the entire nation on Mar 11[6]. That is, the official date τ^* is either Mar 8 or 11. However, we estimate the effective date $\hat{\tau}$ at Mar 16 (the posterior probability that τ is later than Mar 11 is ($P(\tau > \tau^*)$ =???). Similarly, in Wuhan, China, lockdown was declared on Jan 23[7], but we estimate that the effective start of NPIs to be 3-4 days layer ($P(\tau > \tau^*)$ =???).

Advanced effective start of NPIs. In contrast, in some regions we estimate an effective start of NPIs $\hat{\tau}$ that is *earlier* then the official date τ^*. For example, social distancing was encouraged starting on Mar 8[4], but mass gatherings still occurred on Mar 8, including a march of 120,000 people for the International Women's Day, and a football match between Real Betis and Real Madrid (2-1) with a crowd of 50,965 in Seville. A national lockdown was only announced on Mar 14 (τ^*)[4]. Nevertheless, we estimate the effective start of NPI $\hat{\tau}$ at Mar 8 or 9, rather than Mar 14 ($P(\tau < \tau^*)$ =???).

The exception that proves the rule. We have also found a single case in which the official and effective dates match: Switzerland ordered a national lockdown on Mar 20 (τ^*), after banning public evens and closing schools on Mar 13 and 14[4]. Indeed, our MAP estimate $\hat{\tau}$ is Mar 20, and the posterior distribution shows two density peaks: a smaller one between Mar 10 and Mar 14, and a taller one between Mar 17 and Mar 22. It's also worth mentioning that Switzerland was the first to mandate self isolation of confirmed cases[4]. This seems to be

Discussion

We have estimated the effective start date of NPIs in several geographical regions using an SEIR epidemiological model and an MCMC parameter estimation framework. We find that in most of the examined regions the effective and official NPI start dates differ significantly FIGURE. We find examples of both advanced and delayed response to NPIs: for example, in Italy and Wuhan, China, the effective start of the lockdowns seems to have occurred 3-5 after the official date. This could be explained by low compliance: in Italy, it seems that a leak about the intent to lockdown Northern provinces results in people leaving those provinces[6]. However, delayed effect of NPIs could also be due to the time required by both the government and the citizens to organize for a lockdown. In contrast, in Spain and France transmission rates seem to have been reduced even before official lockdowns were imposed, possibly due to adoption of social distancing and similar behavioral adaptations in part of the population, maybe in response to domestic or international COVID-19-related reports.

As several countries (e.g. Austria, Israel) have began to relieve lockdowns and ease restrictions, we expect similar delays and advances to occur: in some countries people will begin to behave as if restrictions were eased before the official date, and in some countries people will continue to self-restrict even after restrictions are officially removed. Such delays and advances could confuse analyses and lead to wrong conclusions about the effects of NPI removals.

Conclusions. We have estimated the effective start date of NPIs and found that they often differ from the official dates. Our results emphasize the complex interaction between personal, regional, and global determinants of behavioral. Thus, our results highlight the need to further study variability in compliance and behavior over both time and space. This should be accomplished both by surveying differences in compliance within and between populations[2], and by incorporating specific behavioral models into epidemiological models[1].

Acknowledgements

This work was supported in part by the Israel Science Foundation 552/19 (YR) and XXX/XX (Alon Rosen)

References

[1] Arthur, R. F., Jones, J. H., Bonds, M. H. and Feldman, M. W. 2020, 'Complex dynamics induced by delayed adaptive behavior during outbreaks', *bioRxiv* pp. 1–23.

[2] Atchison, C. J., Bowman, L., Vrinten, C., Redd, R., Pristera, P., Eaton, J. W. and Ward, H. 2020, 'Perceptions and behavioural responses of the general public during the COVID-19 pandemic: A cross-sectional survey of UK Adults', *medRxiv* p. 2020.04.01.20050039.

[3] Banholzer, N., Weenen, E. V., Kratzwald, B. and Seeliger, A. 2020, 'The estimated impact of non-pharmaceutical interventions on documented cases of COVID-19 : A cross-country analysis', *medRxiv* .

[4] Flaxman, S., Mishra, S., Gandy, A., Unwin, J. T., Coupland, H., Mellan, T. A., Zhu, H., Berah, T., Eaton, J. W., Guzman, P. N. P., Schmit, N., Cilloni, L., Ainslie, K. E. C., 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., Gt Walker, P., Walters, C., Wang, H., Wang, Y., Watson, O., Xi, X., Winskill, P., Whittaker, C., Ghani, A., Donnelly, C. A., Riley, S., Okell, L. C., Vollmer, M. A. C., Ferguson, N. M. and Bhatt, S. 2020, 'Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries', *Imp. Coll. London* (March), 1–35.

[5] Foreman-Mackey, D., Hogg, D. W., Lang, D. and Goodman, J. 2013, 'emcee : The MCMC Hammer ', *Publ. Astron. Soc. Pacific* **125**(925), 306–312.

[6] Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R. and Rinaldo, A. 2020, 'Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures', *Proc. Natl. Acad. Sci.* p. 202004978.

[7] Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W. and Shaman, J. 2020, 'Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2)', *Science (80-.).* p. eabb3221.

[8] Pei, S. and Shaman, J. 2020, 'Initial Simulation of SARS-CoV2 Spread and Intervention Effects in the Continental US', *medRxiv* p. 2020.03.21.20040303.

[9] Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. 2002, 'Bayesian measures of model complexity and fit', *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64**(4), 583–616.