

Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave

Nicolas Banholzer¹, Eva van Weenen¹, Adrian Lison¹, Alberto Cenedese¹, Arne Seeliger¹, Bernhard Kratzwald¹, Daniel Tschernutter¹, Joan Puig Salles¹, Pierluigi Bottrighi¹, Sonja Lehtinen², Stefan Feuerriegel¹, and Werner Vach^{3,4,*}

¹Department of Management, Technology, and Economics, ETH Zurich, 8092 Zurich, Switzerland

²Department of Environmental Systems Science, ETH Zurich, 8092 Zurich, Switzerland

³Basel Academy for Quality and Research in Medicine, CH-4051 Basel, Switzerland

⁴Department of Environmental Sciences, University of Basel, CH-4055 Basel, Switzerland

* werner.vach@unibas.ch

Abstract

The novel coronavirus (SARS-CoV-2) has rapidly developed into a global epidemic. To control its spread, countries have implemented non-pharmaceutical interventions (NPIs), such as school closures, gathering bans, or even stay-at-home orders. Here we study the effectiveness of seven NPIs in reducing the number of new infections, which was inferred from the reported cases of COVID-19 using a semi-mechanistic Bayesian hierarchical model. Based on data from the first epidemic wave of $n = 20$ countries (i.e., the United States, Canada, Australia, the EU-15 countries, Norway, and Switzerland), we estimate the relative reduction in the number of new infections attributed to each NPI. Among the NPIs considered, event bans were most effective, followed by venue and school closures, whereas stay-at-home orders and work bans were least effective. With this retrospective cross-country analysis, we provide estimates regarding the effectiveness of different NPIs during the first epidemic wave.

Keywords: COVID-19, coronavirus, non-pharmaceutical interventions, government interventions, health-policy measures, Bayesian modeling

1 Introduction

The novel coronavirus (SARS-CoV-2) has developed into a global epidemic. Efforts to control the spread of SARS-CoV-2 focused on non-pharmaceutical interventions (NPIs). These represent public health-policy measures that were intended to diminish transmission rates and, to this end, aimed at reducing person-to-person contacts via so-called social distancing¹. Examples include school closures, venue closures, or stay-at-home orders.

Early studies on the population-level effects of NPIs analyzed their effectiveness mostly within a single country²⁻⁹. Thereby, NPIs were often packaged into bundles and their combined effectiveness was assessed and confirmed. By combining data from multiple countries, a couple of studies have attempted to compare the effectiveness of individual NPIs¹⁰⁻¹⁵; however, the evidence from these studies regarding which NPIs were particularly effective is still inconclusive.

Here we contribute further evidence on the combined and individual effectiveness of NPIs. Using a semi-mechanistic Bayesian hierarchical model, we estimated the effects of NPIs on the number of new infections across $n = 20$ Western countries during the first epidemic wave: the United States, Canada, Australia, the EU-15 countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom), Norway, and Switzerland. This amounts to ~ 3.3 million reported cases of coronavirus disease 2019 (COVID-19) and covers a population of ~ 0.8 billion people.

2 Methods

2.1 Data

Reported SARS-CoV-2 cases for each country between February and May 2020 were obtained from the Johns Hopkins Coronavirus Resource Center, which was developed for real-time tracking of reported cases of COVID-19 and directly aggregates cases recorded by local authorities in order to overcome time delays from alternative reporting bodies¹⁶. Hence, these numbers are supposed to account for all COVID-19 cases identified on a specific day.

Data on NPIs were collected by our research team. Their implementation dates were systematically obtained from government resources and news outlets before the NPIs were being classified into seven categories, based on definitions applicable across our sample of countries (Tbl. 1): (1) event

bans, (2) school closures, (3) venue closures (e.g., shops, bars, restaurants, and venues for other recreational activities), (4) gathering bans, (5) border closures, (6) stay-at-home orders prohibiting public movements without valid reason, and (7) work bans on non-essential business activities. Note that stay-at-home orders always also implied the presence of event bans, gathering bans and venue closures as NPIs if these had not been explicitly implemented earlier.

The implementation date of an NPI refers to the first day a measure went into action. The implementation dates were thoroughly checked to ensure correctness and consistency across countries. Overall, eight authors were involved in collecting, categorizing, and checking the data. Furthermore, local residents and/or native speakers were recruited from some countries in order to verify our encoding in cases where the interpretation of legal terms was not ambiguous or where it was difficult to distinguish, for instance, whether an NPI was enforced or recommended. Sources and details on data collection for NPIs are provided in Supplement 6.

NPI	Definition
Event ban	Cancellation of mass gatherings (i.e., 50 people or more)
School closure	Closure of schools (for primary schools)
Venue closure	Closure of venues for recreational activities and/or shops, bars, and restaurants
Border closure	Closure of national borders for individuals
Gathering ban	Prohibition of small gatherings in public or private spaces of people from different households
Stay-at-home order	Prohibition of movement without valid reason (e.g., restricting mobility except to/from work, local supermarkets, and pharmacies)
Work ban	Closure of non-essential business activities (i.e., all businesses except supermarkets, food suppliers, and pharmacies), thus prohibiting corresponding mobility

Table 1. List of non-pharmaceutical interventions (NPIs).

Our selection of Western countries defined a sample that followed a similar and comparable overall strategy in controlling the COVID-19 outbreak. On the one hand, the national strategies consisted of similar NPIs, which we can expect to work in a similar manner, despite cultural and organisational differences between countries. On the other hand, the national strategies differed in the choice, timing, and sequencing of NPIs. Taken together, the setting of this study resembles that of a natural experiment, which allows us to learn about the effects of different NPIs. However, in countries with a federal structure, the timing of NPIs may differ between regions (e.g., states or territories). In countries with such regional variation, NPI data was collected at the regional level and, similar to Hsiang et al.¹⁴, we took into account the cumulative share of the country's population that is affected by an NPI. Fig. 1 summarizes our NPI data by comparing choice, timing

and sequencing of NPIs across countries after considering regional variation within countries.

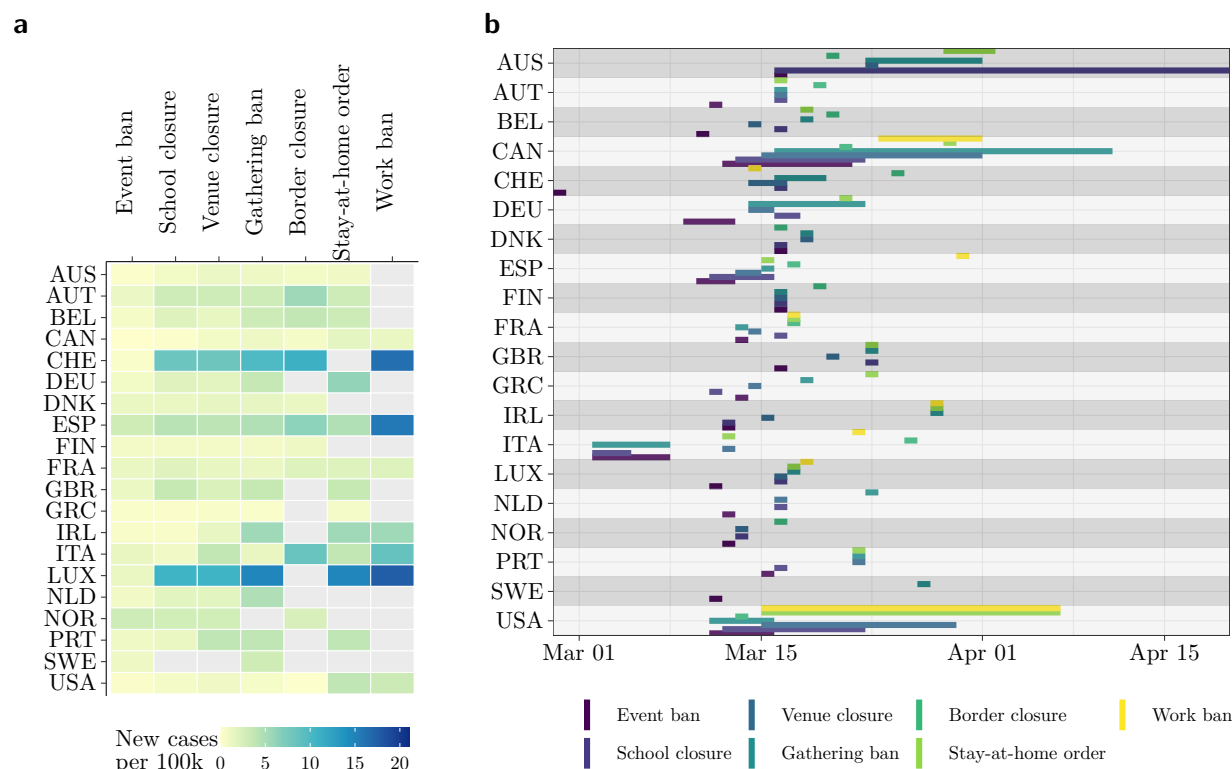


Figure 1. Timing of NPIs by country (ISO 3 country code). (a) Number of new cases per 100,000 (rolling 7-day mean) until NPIs were implemented across countries. For countries with regional variation in the implementation of NPIs, the number of new cases was averaged across regions. (b) Timeline of the implementation of NPIs. The horizontal lines show the time period in which NPIs were implemented within each country's regions. For most countries, there was no regional variation and the NPIs were implemented at one day across the entire country.

2.2 Model summary

In this section, we provide a short description of our model. A detailed description, including all modeling and prior choices is given in Supplement 1.

Fig. 2 provides a visual summary of our model structure. Our model links two unobserved quantities (i.e., the daily number of contagious subjects and the daily number of new infections) to an observed quantity (i.e., the number of reported new cases). The links consist of three components: (1) a regression type model relating the number of new infections to the number of contagious subjects, the country-specific daily transmission rate, and the presence of active measures; (2) a link between the number of new infections to the number of reported new cases; and (3) a link between the number of new infections and the number of contagious subjects.

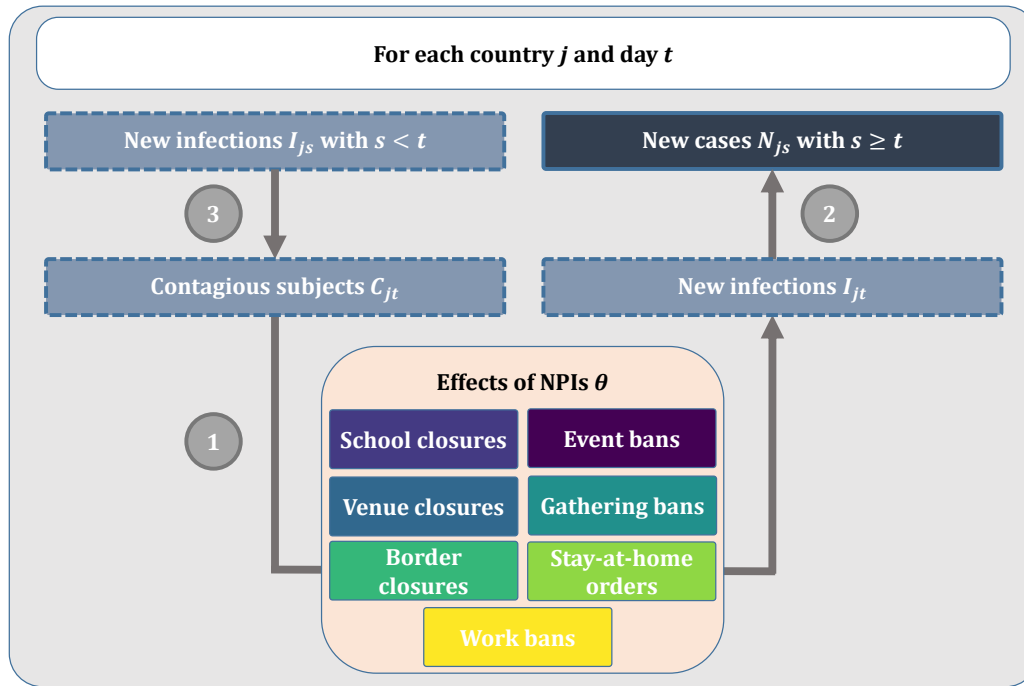


Figure 2. Visual summary of the model structure: (1) the number of new infections is modelled as a function of the number of contagious subjects, the country-specific daily transmission rate, and the reductions from active NPIs; (2) the observed number of new cases is a weighted sum of the number of new infections in the previous days; and (3) the number of contagious subjects is a weighted sum of the number of new infections in the previous days.

The fundamental part of the first component is a model for the expected number μ of new infections I_{jt} in country j at day t . In the absence of any measure, this would be $\mu^{I_{jt}} = C_{jt} \delta_j$, where C_{jt} are the number of contagious subjects and δ_j is the country-specific daily transmission rate. In the presence of NPIs, we multiply this with the reduction due to avoided infections, resulting in

$$\mu^{I_{jt}} = C_{jt} \delta_j \cdot \prod_{m=1}^M (1 - a_{mjt}), \quad (1)$$

where a_{mjt} denotes the fraction of avoided infections due to NPI $m = 1, \dots, M$ in country j at day t . In case NPI m is implemented and fully effective, a_{mjt} is set equal to the value θ_m . However, an NPI may not be fully effective in a country due to regional differences or because it may take a few days until subjects respond to the new measures. Hence, the general structure of a_{mjt} is

$$a_{mjt} = \theta_m \sum_{r=1}^{R_j} p_{rj} f(T_{mrjt}), \quad (2)$$

where p_{rj} is region's $r = 1, \dots, R_j$ proportion of the total population in country j , where T_{mrjt} denotes the number of days since the implementation of NPI m in region r of country j at time t (such that $T_{mrjt} = 1$ denotes the first day at which a reduction in the number of new infections could be expected), and where $f(T_{mrjt})$ is a time-delayed response function, which is specified such that the response to an NPI increases from zero on day $T_{mrjt} = 0$ to one on day $T_{mrjt} = 3$ (Fig. 3a), reflecting our expectation that NPIs typically require a few days until they are fully effective. The choice for the time-delayed response is varied as part of the sensitivity analysis.

The effect of an NPI when fully implemented is equal to θ_m . Within our Bayesian framework, a mixture prior was used to model the effect θ_m . The prior consists of a half-normal distribution for negative effects and a uniform distribution for positive effects (Fig. 3b). This prior allows for small increases in the number of new infections with a probability of 10%, while being uninformative about positive effects leading to reductions in the number of new infections.

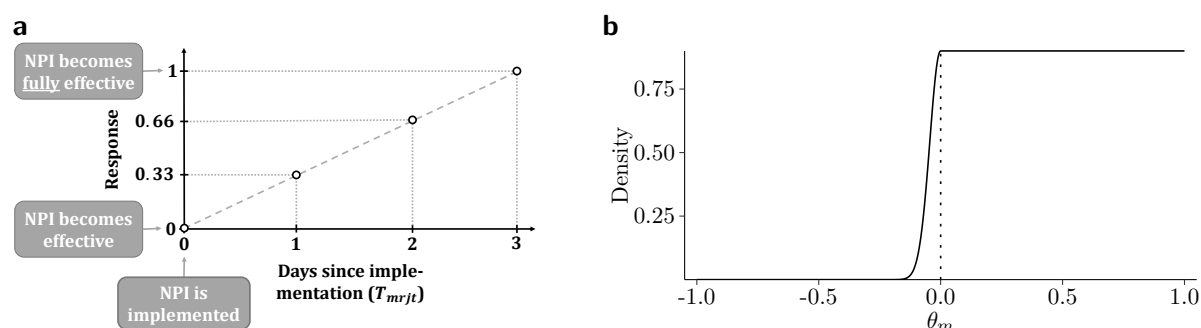


Figure 3. Modeling choices for the effects of NPIs. (a) Time-delayed response function as a first-order spline. (b) Prior for the effects of NPIs θ_m .

In the second component, the expected number of new cases is calculated as a weighted sum of the number of new infections in the previous days. The weights reflect the distribution of the time from infection to reporting, and this distribution is estimated from the data assuming a log-normal distribution. Informative priors are used for the parameters of the log-normal distribution, reflecting prior knowledge. These priors are derived by decomposing the time from infection to reporting into the incubation period and the reporting delay, thereby using a meta-analysis¹⁷ and estimates from multiple studies (in particular¹⁸) to obtain prior information about the distributions. The observed number of new cases is then modeled as a negative binomial distribution with the specified mean, allowing for overdispersion. Note that this part of the overall model was only applied after the

day when a country reached 100 reported cumulative cases (called the modeling phase), in order to avoid modeling of highly irregular case numbers in the very early stages of the epidemic when most countries still had to set up their reporting practices. At the preceding days (called non-modeling phase), only the other two components of the model were used (i.e., the number of new infections and contagious subjects).

In the third component, the number of contagious subjects is calculated as a weighted sum of the number of new infections in the previous days. The weights reflect the probability of being contagious on a specific day after being infected and can be determined from the generation time distribution. This distribution is assumed to be known and our choice is based on an estimate by a recent study using data on the exposure for both the index and secondary case¹⁹.

2.3 Parameter estimation

Choices of priors not mentioned so far are weakly informative, following general recommendations in Bayesian modeling²⁰. All model parameters are estimated with a semi-mechanistic Bayesian hierarchical model. Specifically, Markov chain Monte Carlo (MCMC) sampling is used as implemented by the Hamiltonian Monte Carlo algorithm with the No-U-Turn Sampler (NUTS) from the probabilistic programming language Stan, version 2.19.2²¹. Each model is estimated with 4 Markov chains and 2,000 iterations of which the first 1,000 iterations are discarded as part of the warm-up. Estimation power is evaluated via the ratio of the effective sample size ($\hat{n}_{\text{eff}}/N$), and convergence of the Markov chains is assessed with the Gelman-Rubin convergence diagnostic ($\hat{R}$). Further checks pertain to the detection of influential observations and correlations between the parameters of interest. If not stated otherwise, we report posterior means and credible intervals (CrIs) based on the 2.5% and 97.5% quantile of the posterior samples. Detailed estimation results and model diagnostics are provided in Supplement 3. The sensitivity of the results to the following modeling choices was investigated: start and end of the modeling phase, time-delayed response function, prior distribution for the effect of an NPI θ_m , prior choices for the time from infection to reporting and for the generation time. Furthermore, the sensitivity of the results with respect to exclusion of individual countries was assessed with a leave-one-country-out analysis. Finally, a comparison to a similar study that was recently published¹¹ is presented.

2.4 Data and code availability

We collected data from publicly available data sources (Johns Hopkins Coronavirus Resource Center¹⁶ for epidemiological data; news reports and government resources for policy measures). All the public health information that we used is documented in the main text, the extended data, and supplementary tables. A preprocessed data file together with reproducible code is available from https://github.com/nbanho/npi_effectiveness_first_wave.

3 Results

3.1 Estimated effects of NPIs

Using data from the first epidemic wave, we estimated the relative reduction in the number of new infections for each NPI (Fig. 4a). Event bans were associated with the highest reduction in the number of new infections (37%; 95% CrI 21% to 50%). The reduction was lower for venue closures (18%; 95% CrI -4% to 40%) and school closures (17%; 95% CrI -2% to 36%), followed by border closures (10%; 95% CrI -2% to 21%) and gathering bans (9%; 95% CrI -4% to 23%). stay-at-home orders (4%; 95% CrI -6% to 17%) and work bans on non-essential business activities (1%; 95% CrI -8% to 12%) appeared to be the least effective among the NPIs considered in this analysis.

The estimates for the individual effects suggest a particular strong effect of event bans. This result is further supported by analyzing the posterior ranking of the effects (Fig. 4b), which indicates that we could be at least 98% sure that event bans were among the two most effective NPIs. Conversely, we could be at least 76% sure that work bans were among the two least effective NPIs.

All NPIs together lead to an estimated relative reduction in the number of new infections by 67% (95% CrI 64% to 71%). The combined effectiveness of NPIs was also analyzed (Fig. 4c-d). Thereby, we could be at least 97% sure that at least five NPIs simultaneously lead to a reduction in the number of new infections. Regarding the magnitude of the effects, we could be at least 91% sure that three NPIs simultaneously lead to a reduction in the number of new infections of more than 10%.

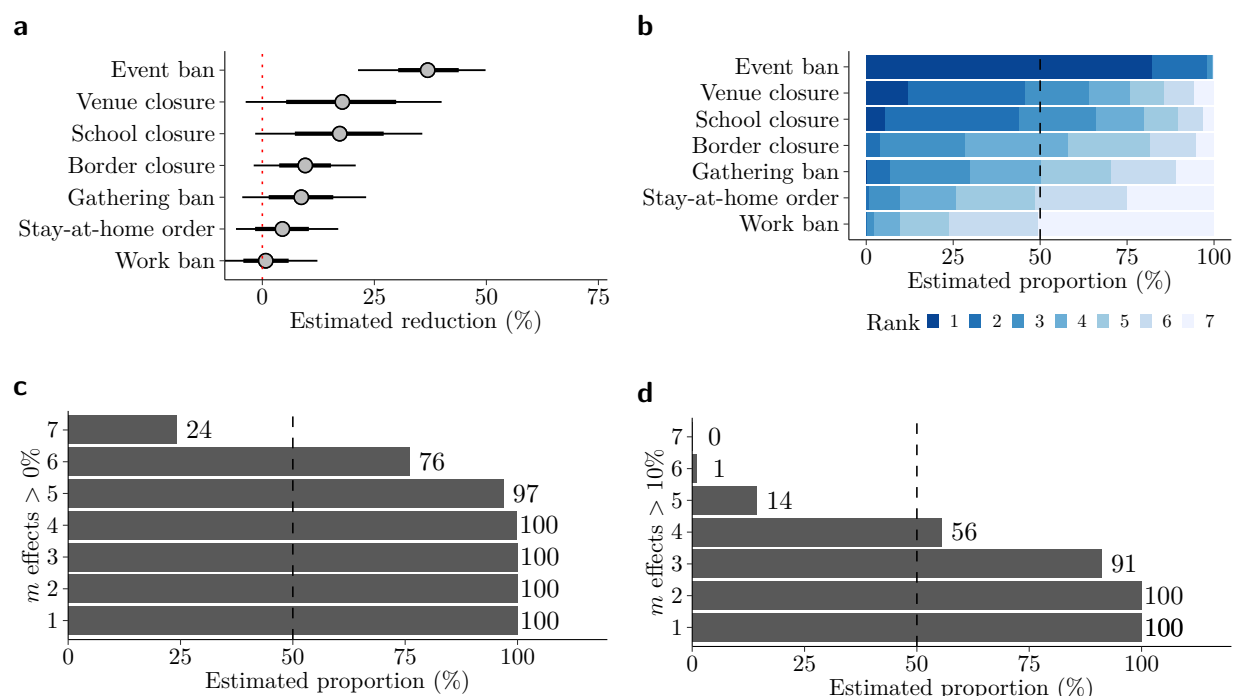


Figure 4. Estimated effects of NPIs. **(a)** Reduction in new infections (posterior mean as dots with 80% and 95% credible interval as thick and thin lines, respectively). **(b)** Ranking of the effects of NPIs from highest (1) to lowest (7) (posterior frequency distribution). **(c)** Frequency of at least m positive effects (posterior frequency distribution). **(d)** Frequency of at least m effects greater than 10% (posterior frequency distribution).

3.2 Sensitivity analysis

Sensitivity of our results was assessed with respect to varying modeling and prior choices (Supplement 4). Overall, our results were only slightly sensitive to alternative modeling and prior choices, particularly when deriving conclusions about the ranking of the effects. A leave-one-out analysis for the countries in our sample indicated some sensitivity with respect to the exclusion of Australia, Sweden, and Switzerland. This sensitivity was further investigated and subsequent analysis indicated that data from these countries are particularly informative for the effect of school closures and event bans.

3.3 Estimated number of new infections and cases over time

The model fit was assessed by comparing the expected number of new cases based on our model to the observed number of new cases (Fig. 5). Here an acceptable degree of agreement is observed for Australia, Germany, Sweden, the UK, and all other countries (figures for all other countries are

listed in Supplement 5). In addition, the posterior distribution for the number of new infections is depicted, illustrating how NPIs lead to a reduction in new infections, which later implies a reduction in new cases.

3.4 Comparison with a similar study

Recently, an analysis similar to our was published by Brauner et al.¹¹. In the following, we present a brief comparison. Brauner et al. used the same data source to generate country-specific case numbers per day, but included 30 countries only partially overlapping with our selection. The authors used similar sources to collect data on NPIs but classified some NPIs differently. The modeling approach was very similar, except for avoiding to take the number of contagious subjects explicitly into account and considering both the number of reported cases and reported deaths as outcomes. The authors included country-specific NPI effects (although not explicitly analyzed), while we took into account the share of the country's total population being affected by NPIs and a time-delayed response to their implementation. The authors used similar priors for the distribution of the time from infection to reporting and for the generation time distribution, but estimated the latter explicitly as part of fitting the model.

To facilitate a comparison of the results between the modeling approach by Brauner et al. and our approach, we reanalyzed their data using our model. The estimated effects from our model were generally smaller but the overall ranking of the effects was very similar when fitted to the same data (Tbl. 2a). In contrast to that, when comparing the results from our model on different datasets (our data and the data from Brauner et al.), there was only a similarity in the overall pattern (Tbl. 2b). For instance, we found similar strong effects for event bans and school closures as well as weak effects for stay-at-home orders, but the effect of venue closures is stronger and the effect of work bans weaker in our data. In addition, the larger set of countries in the data from Brauner et al. was associated with smaller credible intervals for the estimates.

4 Discussion

4.1 Estimated NPI effects

We performed a cross-country analysis based on $n = 20$ countries in order to assess the effectiveness of seven NPIs during the first epidemic wave. Our findings suggest that some NPIs, particularly in

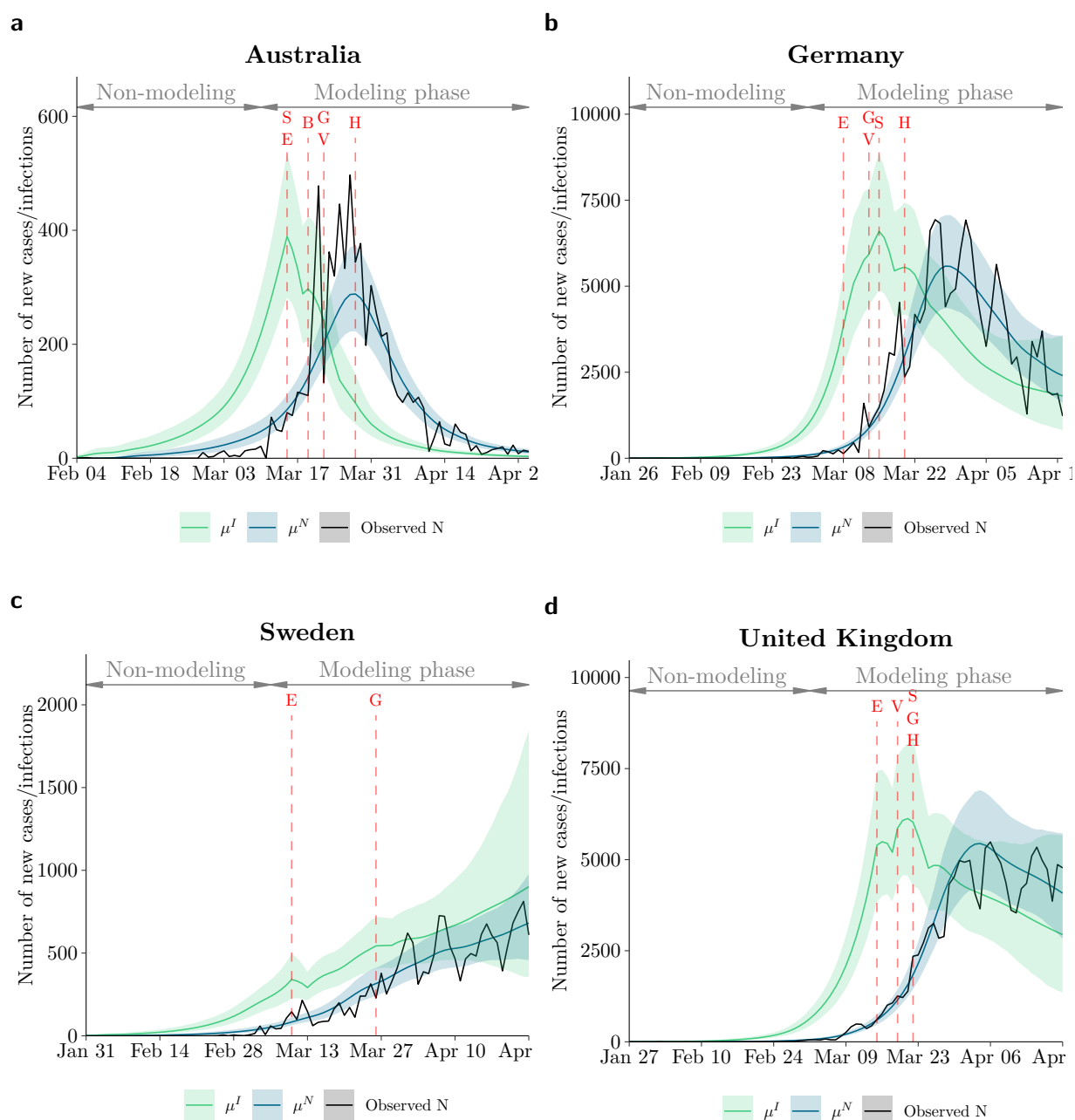


Figure 5. Expected number of new infections μ^I and new cases μ^N (posterior mean as colored lines with 95% credible interval as shaded area) and the observed number of new cases by country over time. **(a-d)** Time series for four selected countries. Red letters and lines indicate the first day an NPI was implemented within a country (S: School closures, B: Border closure, E: Event ban, G: Gathering ban, V: Venue closure, H: stay-at-home order, W: Work ban). The non-modeling phase is the time period before 100 cumulative cases were observed, which was used to seed infections in the early outbreak of the epidemic. Plots for all countries are provided in Supplement 5.

combination, lead to a strong reduction in the number of new infections. Among the NPIs considered, event bans were most effective in reducing the number of new infections across countries. Event bans

	Model and data by Brauner et al		Our model and data by Brauner et al	
	Mean (PrI)	Rank	Mean (CrI)	Rank
Gatherings <1000	23 (0 – 40)	5	17 (5 – 28)	5
Gatherings <100	34 (12 – 42)	3	24 (11 – 36)	2
Gatherings <10	42 (17 – 60)	1	32 (17 – 47)	1
Schools and universities closed	38 (16 – 54)	2	23 (12 – 34)	3
Some businesses closed	18 (– 8 – 40)	6	5 (– 5 – 17)	7
Most businesses closed	27 (– 3 – 49)	4	18 (5 – 30)	4
stay-at-home order	13 (– 5 – 31)	7	12 (5 – 18)	6

(a) Comparison of modeling.

Our NPI	NPI by Brauner et al.	Our model and data		Our model and data by Brauner et al.	
		Mean (CrI)	Rank	Mean (CrI)	Rank
Event bans	Gatherings <1000	37 (21 – 50)	1	17 (5 – 28)	2
Venue closure	Some businesses closed	18 (– 4 – 40)	2	5 (– 5 – 17)	6
School closure	Schools and universities closed	17 (– 2 – 36)	3	23 (12 – 34)	1
Border closure		10 (– 2 – 21)	4		
Gathering ban	Gatherings <100 + Gatherings <10	9 (– 4 – 23)	5	15 (1 – 29)	3
stay-at-home order	stay-at-home order	4 (– 6 – 17)	6	12 (5 – 18)	5
Work bans	Most businesses closed	1 (– 8 – 12)	7	12 (2 – 22)	4

(b) Comparison of data.

Table 2. Comparison of modeling and data with results from Brauner et al.¹¹. (a) Estimated effects by model and data from Brauner et al. (posterior mean in %, 95% prediction interval (PrI) and rank) and by our model and data from Brauner et al. (posterior mean in %, 95% credible interval and rank). Note that, in these analyses, we report cumulative effects for gathering bans and businesses closed. (b) Estimated effects by our model on our data (posterior mean in %, 95% credible interval, and rank) and by our model and data from Brauner et al. (posterior mean in %, 95% credible interval, and rank). Similar NPIs were matched but their definitions are not exactly the same. Note that in applying our model to the data by Brauner et al., we report the cumulative effect of “gatherings < 100” and “gatherings < 10” when referring to our gathering ban.

are targeted towards large gatherings of people and may thus prevent so-called “superspreader events”, which have been shown to account for a substantial fraction of the total number of infections^{22–24}. Many superspreading events originated in points-of-interest such as bars and restaurants²⁵, which aligns with our finding of a sizeable reduction from venue closures.

The effectiveness of school closures in transmission control is subject to debate^{26,27}. Although children are less susceptible to the virus than adults, it is less clear how children and adults compare regarding their infectiousness²⁸. Early findings suggested that school closures were only marginally effective in transmission control²⁹. Our results provide contrary evidence, in line with recent findings from other population-based studies^{11,13,15,30,31}. However, note that our findings relate to the closure of primary schools, which often coincided with the closure of secondary schools and universities. The study by Brauner et al. differentiated between NPIs for the closure of schools and the closure of universities, but could afterwards not disentangle the estimated effects¹¹. Thus, it is subject

to further investigation whether the closure of primary schools is less effective than the closure of secondary schools and universities²⁸.

A small effect was estimated for stay-at-home orders. This seems to contradict findings about the high effectiveness of the lockdown from Flaxman et al.¹². However, one should consider that their definition of a lockdown encompasses multiple NPIs that we differentiated (e.g., gathering bans, venue closures, and stay-at-home orders). Taken together, our estimated “lockdown effect” would therefore also be large. Finally, although our findings suggest that work bans were not effective, it should be considered that our definition of a work ban referred to a strict ban of non-essential business activities, while many countries only issued recommendations.

4.2 Methodological aspects

Flaxman et al.¹² were the first who attempted to link NPIs to observed cases or deaths using a semi-mechanistic Bayesian hierarchical model. Both the study of Brauner et al.¹¹ and our study can be seen as extensions of this approach, thereby making use in particular of data from more countries. This also implies the possibility of refined modeling.

The above studies have in common that they model the effect of the NPIs on the number of new infections. Whereas the other studies approached this by modeling explicitly the effect on the reproduction number, we directly modelled the number of new infections in relation to the number of contagious subjects. Thereby, the generation time distribution was used in a way to approximate the time from infection to becoming contagious³².

A further common property of the above studies is the use of prior information on the distribution of certain quantities that play a central role in the spread of infectious diseases. Flaxman et al. equal the generation time distribution with the serial distribution, whereas Brauner et al. and our study avoid this, considering that there can be a substantial difference between these distributions³³. Brauner et al. estimated the generation time distribution from their data on cases and deaths using an informative prior, whereas we made use of an explicit assumption based on prior knowledge. The second quantity used in both studies is the time from infection to reporting, which is assumed to be known by Flaxman et al. and estimated by Brauner et al. and our study. Similar to the other studies, we investigated the sensitivity of the estimated NPI effects with respect to the assumptions made or the priors chosen.

A specific property of our approach is to take the regional variation in the implementation of NPIs explicitly into account. We approached this by incorporating the share of the country's total population that is affected by active NPIs in our model. The other studies ignored this variation or restricted the analysis to countries with no or very little regional variation. A further specific property was to allow for a gradual increase in the response to NPIs over the first few days, whereas the other studies assumed a full response on the first day NPIs were implemented.

4.3 Insights from a comparison with a similar study

With the data on reported cases and deaths gathered by the John Hopkins University and data on NPIs gathered by government and news websites, there is a unique source for analyzing the impact of health policy measures on the course of the COVID-19 pandemic. However, there are many degrees of freedom with respect to preprocessing the available data, constructing a model, and making prior assumptions and choices. To understand the influence of these data and modeling choices on the results, it is desirable that different research groups analyse the same data in different ways. A very recent publication by Brauner et al.¹¹ gave us the possibility to perform a first check of this type. The check indicated that the choice of countries and definition of NPIs has a larger influence on the estimated effects than the detailed choices in modeling. A re-analysis of the original study by Flaxman et al. could also demonstrate sensitivity of the results with respect to specific choices of the definition of NPIs, but also found sensitivity to specific modeling choices³⁴.

4.4 Limitations

Our analysis is subject to limitations. First, our modeling assumptions do not allow for (random) variation in the effect of NPIs across countries and assume a fixed effect. Brauner et al.¹¹ demonstrated that, in principle, it is feasible to allow for country-specific variation in the effects. However, the specific parametrization chosen in our model (i.e., to take into account regional variation in the implementation of NPIs via the population share) made it challenging to incorporate random variation of the NPI effect θ_m across countries. Brauner et al observed in the sensitivity analysis of their study that assuming a fixed effect does not alter the main conclusions that they discussed regarding the effectiveness of NPIs.

Second, any approach of explaining changes in the observed number of cases solely by specific NPIs makes the implicit assumption that these changes were not the result of some other factors.

For instance, it is possible that additional measures or an increasing general awareness encouraged social distancing and hence lead to less infections. If this is the case, such effects will erroneously be assigned to the NPIs and possibly overstate their overall impact.

Third, it is challenging to distinguish between the effects of single NPIs due to their concurring introduction in many countries (Supplement Tbl. 4). This is reflected by wide credible intervals and a negative association between effects (Supplement Fig. 3.2), suggesting that the effect of one NPI may be attributed partially to another. Note that the effect of event bans could be estimated with comparably narrower credible intervals than that of venue and school closures. A reason for this could be that all three measures were implemented by most countries but event bans had on average a greater distance in implementation to other NPIs as compared to school and venue closures (Supplement Tbl. 3). Despite the difficulty in disentangling the effects of individual NPIs, we were able to demonstrate that we can be fairly confident with respect to their relative ranking.

Fourth, our analysis is limited by the type of data utilized to define the outcomes. Using the number of reported cases as outcome implies that reporting practices may have an influence on the results. In particular, definitions and reporting practices differed between countries and over time. However, since we modeled the ratio between the number of new cases and the number of contagious subjects, we can expect that these effects cancel out. The same argument applies to the challenge that we have to expect a substantial number of undetected cases, with detection rates varying across countries and over time. In addition, we modelled the number of reported cases as a random variable, thereby allowing for overdispersion. Moreover, we took into account that countries had still to develop their reporting practice in the very early phase of the epidemic by starting the modeling phase after at least 100 cumulative cases were reported in a country.

Fifth, our analysis is limited by the need to classify NPIs in a comparable manner across countries and to determine dates when NPIs were exactly implemented. There is always some subjectivity in making corresponding decisions. The comparison of our study with the one by Brauner et al. suggests that such definitions can be crucial. We provide a detailed description of our decisions (Supplement 6).

4.5 Outlook

With the model presented in this paper and that developed by Brauner et al., we may have reached a first level of maturity. This may be sufficient to justify the use of these models in analyzing the effects of NPIs. In principle, these models could also be used to study the effect of lifting NPIs. A natural next step would be to apply these models to data from the second wave, which is still ongoing. An interesting question would then be whether we can expect similar effects in the first and the second wave. It is likely that some effects may have changed, as the situations are not necessarily comparable and experience from the first wave may have helped in dealing with the second wave.

Of course, there is still room for improvement and refinement of the models. If information on new cases is also available at a regional level, regional variation in implementing NPIs can be used as an additional source of information using a two-level hierarchical approach. Furthermore, between-country and regional variation may be linked to certain characteristics which may help to understand the conditions under which an NPI is most effective. Similarly, information on cases stratified by patient characteristics may provide new insights. Also the modeling of the outcome can be improved, e.g., by taking weekday effects into account.

4.6 Conclusion

Our analysis makes a contribution to the emerging evidence about the effectiveness of different NPIs in the first epidemic wave. Event bans in particular were identified as an effective measure, whereas little evidence was found for substantial effects of stay-at-home orders and work bans. A comparison with another study indicates robustness of some conclusions, but also a dependence on the choice of data and definitions.

References

1. World Health Organization. Coronavirus diseases (COVID-19) advice for the public; 2020. Available from: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/videos>.
2. Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong:

an observational study. The Lancet Public Health. 2020 5;5(5):e279–e288. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2468266720300906>.

3. Davies NG, Kucharski AJ, Eggo RM, Gimma A, Edmunds WJ, Jombart T, et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. The Lancet Public Health. 2020 7;5(7):e375–e385. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S246826672030133X>.

4. Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. Science. 2020 7;369(6500):eabb9789. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.abb9789>.

5. Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. Proceedings of the National Academy of Sciences. 2020 5;117(19):10484–10491. Available from: <https://www.pnas.org/content/early/2020/04/22/2004978117><http://www.pnas.org/lookup/doi/10.1073/pnas.2004978117>.

6. Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT, et al. Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. The Lancet Infectious Diseases. 2020 6;20(6):678–688. Available from: [https://doi.org/10.1016/S1473-3099\(20\)30162-6](https://doi.org/10.1016/S1473-3099(20)30162-6)<https://linkinghub.elsevier.com/retrieve/pii/S1473309920301626>.

7. Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luo W, Floyd JR, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature. 2020 9;585(7825):410–413. Available from: <http://www.nature.com/articles/s41586-020-2293-x>.

8. Maier BF, Brockmann D. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. Science. 2020 5;368(6492):742–746. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.abb4557>.

9. Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. The Lancet Public Health. 2020 5;5(5):e261–e270. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2468266720300736>.

10. Banholzer N, Weenen Ev, Kratzwald B, Seeliger A, Tschernutter D, Bottrighi P, et al. Estimating the impact of non-pharmaceutical interventions on documented infections with COVID-19: A cross-country analysis. medRxiv. 2020 4;Available from: <https://www.medrxiv.org/content/10.1101/2020.04.16.20062141v3>.
11. Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenčiak T, et al. Inferring the effectiveness of government interventions against COVID-19. Science. 2020 12;p. eabd9338. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.abd9338>.
12. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020 8;584(7820):257–261. Available from: <http://www.nature.com/articles/s41586-020-2405-7>.
13. Haug N, Geyrhofer L, Londei A, Dervic E, Desvars-Larrive A, Loreto V, et al. Ranking the effectiveness of worldwide COVID-19 government interventions. Nature Human Behaviour. 2020 12;4(12):1303–1312. Available from: <http://www.nature.com/articles/s41562-020-01009-0>.
14. Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. Nature. 2020 8;584(7820):262–267. Available from: <http://www.nature.com/articles/s41586-020-2404-8>.
15. Li Y, Campbell H, Kulkarni D, Harpur A, Nundy M, Wang X, et al. The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. The Lancet Infectious Diseases. 2020 10;Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1473309920307854>.
16. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious Diseases. 2020 5;20(5):533–534. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1473309920301201>.
17. McAloon C, Collins A, Hunt K, Barber A, Byrne AW, Butler F, et al. Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. BMJ Open. 2020 8;10(8):e039652. Available from: <https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2020-039652>.

- 374 **18.** Cereda D, Tirani M, Rovida F, Demicheli V, Ajelli M, Poletti P, et al. The early phase of the
375 COVID-19 outbreak in Lombardy, Italy. arXiv. 2020 3;Available from: [http://arxiv.org/abs/](http://arxiv.org/abs/2003.09320)
376 [2003.09320](http://arxiv.org/abs/2003.09320).
- 377 **19.** Ferretti L, Ledda A, Wymant C, Zhao L, Ledda V, Abeler L, et al. The timing of COVID-19
378 transmission. medRxiv. 2020 9;Available from: [https://www.medrxiv.org/content/10.1101/](https://www.medrxiv.org/content/10.1101/2020.09.04.20188516v2)
379 [2020.09.04.20188516v2](https://www.medrxiv.org/content/10.1101/2020.09.04.20188516v2).
- 380 **20.** Stan Development Team. Prior Choice Recommendations; 2020. Available from: [https:](https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations)
381 [//github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations](https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations).
- 382 **21.** Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. Stan : A
383 Probabilistic Programming Language. Journal of Statistical Software. 2017;76(1). Available from:
384 <http://www.jstatsoft.org/v76/i01/>.
- 385 **22.** Adam D. Special report: The simulations driving the world's response to COVID-
386 19. Nature. 2020 4;580(7803):316–318. Available from: [http://www.nature.com/articles/](http://www.nature.com/articles/d41586-020-01003-6)
387 [d41586-020-01003-6](http://www.nature.com/articles/d41586-020-01003-6).
- 388 **23.** Lemieux JE, Siddle KJ, Shaw BM, Loreth C, Schaffner SF, Gladden-Young A, et al. Phylogenetic
389 analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events. Science. 2020
390 12;p. eabe3261. Available from: [https://www.sciencemag.org/lookup/doi/10.1126/science.](https://www.sciencemag.org/lookup/doi/10.1126/science.abe3261)
391 [abe3261](https://www.sciencemag.org/lookup/doi/10.1126/science.abe3261).
- 392 **24.** Wang CJ, Ng CY, Brook RH. Response to COVID-19 in Taiwan. JAMA. 2020 4;323(14):1341.
393 Available from: <https://jamanetwork.com/journals/jama/fullarticle/2762689>.
- 394 **25.** Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, et al. Mobility network models
395 of COVID-19 explain inequities and inform reopening. Nature. 2021 1;589(7840):82–87. Available
396 from: <http://www.nature.com/articles/s41586-020-2923-3>.
- 397 **26.** Couzin-Frankel J. School openings across globe suggest ways to keep coronavirus at bay, despite
398 outbreaks. Science. 2020 7;Available from: [https://www.sciencemag.org/news/2020/07/](https://www.sciencemag.org/news/2020/07/school-openings-across-globe-suggest-ways-keep-coronavirus-bay-despite-outbreaks)
399 [school-openings-across-globe-suggest-ways-keep-coronavirus-bay-despite-outbreaks](https://www.sciencemag.org/news/2020/07/school-openings-across-globe-suggest-ways-keep-coronavirus-bay-despite-outbreaks).
- 400 **27.** Levinson M, Cevik M, Lipsitch M. Reopening Primary Schools during the Pandemic. New
401 England Journal of Medicine. 2020 9;383(10):981–985. Available from: <http://www.nejm.org/>

[doi/10.1056/NEJMms2024920](https://doi.org/10.1056/NEJMms2024920).

28. Goldstein E, Lipsitch M, Cevik M. On the Effect of Age on the Transmission of SARS-CoV-2 in Households, Schools, and the Community. *The Journal of Infectious Diseases*. 2020 10; Available from: <https://academic.oup.com/jid/advance-article/doi/10.1093/infdis/jiaa691/5943164>.

29. Viner RM, Russell SJ, Croker H, Packer J, Ward J, Stansfield C, et al. School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. *The Lancet Child & Adolescent Health*. 2020 5;4(5):397–404. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S235246422030095X>.

30. Auger KA, Shah SS, Richardson T, Hartley D, Hall M, Warniment A, et al. Association Between Statewide School Closure and COVID-19 Incidence and Mortality in the US. *JAMA*. 2020 9;324(9):859. Available from: <https://jamanetwork.com/journals/jama/fullarticle/2769034>.

31. John J, Kang G. Public health during the pandemic in India. *Science*. 2020 11;370(6517):663–664. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.abe9707>.

32. Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. *American Journal of Epidemiology*. 2013 11;178(9):1505–1512. Available from: <https://academic.oup.com/aje/article-lookup/doi/10.1093/aje/kwt133>.

33. Lehtinen S, Ashcroft P, Bonhoeffer S. On the relationship between serial interval, infectiousness profile and generation time. *medRxiv*. 2020 9; Available from: <https://doi.org/10.1101/2020.09.18.20197210>.

34. Soltesz K, Gustafsson F, Timpka T, Jaldén J, Jidling C, Heimerson A, et al. The effect of interventions on COVID-19. *Nature*. 2020 12;588(7839):E26–E28. Available from: <http://www.nature.com/articles/s41586-020-3025-y>.

Acknowledgements We thank various people around the world for checking our data on non-pharmaceutical interventions.

Ethics approval Ethics approval was not required for this study.

Competing interests SF reports membership in a COVID-19 working group of the World Health Organization but without competing interest. JPS declares part-time employment at Luciole Medical outside of the submitted work. SF reports grants from the Swiss National Science Foundation outside of the submitted work. All other authors declare no competing interests.

Funding NB, EvW and SF acknowledge funding from the Swiss National Science Foundation (SNSF) as part of the Eccellenza grant 186932 on “Data-driven health management”. The funding bodies had no control over design, conduct, data, analysis, review, reporting, or interpretation of the research conducted.

Contributions NB contributed to conceptualization, data collection, data analysis, results interpretation and manuscript writing. EvW contributed to data collection and manuscript writing. AC, AS, AC, BK, DT, JPS, and PB contributed to data collection. SL contributed to manuscript writing. SF contributed to conceptualization, results interpretation and manuscript writing. WV contributed to conceptualization, data analysis, results interpretation and manuscript writing.