Impact of non-pharmaceutical interventions against COVID-19 in Europe: A quasi-experimental non-equivalent group and time-series study

Paul R Hunter[1,2], Felipe J Colón-González[3,5,6], Julii Brainard[1]*, Steven Rushton[4]

[1] Norwich Medical School, University of East Anglia, Norwich.
Paul.Hunter@uea.ac.uk, J.Brainard@uea.ac.uk
[2] Department of Environmental Health, Tshwane University of Technology, Pretoria, South Africa
[3] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Felipe.Colon@lshtm.ac.uk
[4] School of Natural and Environmental Sciences, Newcastle University Steven.Rushton@ncl.ac.uk
[5] School of Environmental Sciences, University of East Anglia, Norwich
[6] Tyndall Centre for Climate Change Research, University of East Anglia, Norwich
*corresponding author

ORCID numbers, Paul R. Hunter 0000-0002-5608-6144; Felipe J Colón-González 0000-0002-9671-3405; Julii Brainard 0000-0002-5272-7995; Steve Rushton (none)

COUNTS
Word counts: main manuscript: 5580 (without references), abstract:240
# of references: 47
# of tables: 4
# of figures: 6
Supplemental material: 2 documents. 5 sections in main document

DECLARATIONS

Funding
Professor Hunter and Dr. Brainard were funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Emergency Preparedness and Response at King's College London in partnership with Public Health England (PHE), in collaboration with the University of East Anglia. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, UEA, the Department of Health or Public Health England. Dr. Colón-González is funded by the UK Space Agency Dengue Mosquito Simulation from Satellites (D-MOSS) project.

Conflicts of Interest
The authors declare that we have no conflicts of interest.

Data: The datasets used are in the public domain with copies available from the authors.

Author contributions
PRH and JB conceived of the study. PRH and JB collected the data. PRH, FCG and SR undertook and refined analysis. PRH wrote the first draft which was revised by all authors. JB assembled revisions which all contributed to. Sharp-eyed reviewers and readers of the preprint contributed many small corrections via email.

Impact of non-pharmaceutical interventions against COVID-19 in Europe: a quasi-experimental non-equivalent group and time-series design study

ABSTRACT

Introduction

The current epidemic of COVID-19 is unparalleled in recent history as are the social distancing interventions that have led to a significant halt on the economic and social life of so many countries.

Objective

We aimed to generate empirical evidence about which social distancing measures had the most impact in reducing case counts and mortality.

Methods

We report a quasi-experimental (observational) study of the impact of various interventions for control of the outbreak. Chronological data on case numbers and deaths were taken from the daily published figures by the European Centre for Disease Control and dates of initiation of various control strategies from the Institute of Health Metrics and Evaluation website and published sources. Our complementary analyses were modelled in R using Bayesian generalised additive mixed models and in Stata using multi-level mixed effects regression models.

Results

From both sets of modelling, we found that closure of education facilities, prohibiting mass gatherings and closure of some non-essential businesses were associated with reduced incidence whereas stay at home orders and closure of additional non-essential businesses was not associated with any independent additional impact.

Conclusions

Our pertinent findings are that schools and some non-essential businesses operating "as normal" as well as allowing mass gatherings were incompatible with suppressing disease spread. Closure of all businesses and stay at home orders are less likely to be required to keep disease incidence low. Our results can help inform strategies for staying out of lockdown.

Keywords: COVID-19; control measures; stay at home; collinearity; Bayesian generalised additive mixed models

INTRODUCTION

The current pandemic of COVID-19 is unprecedented in modern history. Not only is the impact of the epidemic being measured by the number of cases and deaths, but also by its impact on overloaded health services and deleterious impacts on quality of life and near-future economic prospects. Wider society was subjected at times to an almost total stasis of social and cultural life. The benefits of social distancing was shown earliest in China, Italy and Spain that turned the tide on their country's epidemics using often severe social distancing strategies. What these examples do not do is indicate the relative importance of the different non-pharmaceutical/ social distancing interventions. Given the potentially high economic and social costs arising from stringent control measures [1-5], it has been imperative to determine which social distancing measures are most effective at controlling the pandemic. Imposition and relaxation of control measures should be informed by such knowledge. Early on in pandemic response, much policy was driven by the results of mathematical models [6]. However, there has been debate about the validity and limitations of the different models for policy making and modelling approaches that have been used [7-10]. It is also useful to assess empirical evidence of what aspects of currently applied non-pharmaceutical interventions (NPIs) have or have not been effective.

A quasi-experimental study design is an observational study where the allocation to receive the intervention (or not) is not randomly made [11, 12]. Most European states introduced a similar suite of interventions aimed at reducing contact between individuals to reduce transmission. However, the different types of intervention used and their timing varied from one country to another, and was in response to political processes in each country, rather than arose through a randomised assignment. No measure was imposed by all European countries. Where measures were put in place, they were often imposed at different points in the development of the epidemics. By late April 2020, some European countries were easing control measures so late April was a good point to take stock of intervention effects. This situation offered a unique opportunity to investigate the putative impacts of the various types of intervention, as each individual-country epidemic forms what is effectively a chrono-sequence of disease spread. The intervention strategies could then be compared as interrupted time series.

We report here analyses of trends in both reported cases and deaths across 30 European countries with rather different approaches to and timing of restrictions. We use a quasi-experimental approach to identify what affects such restrictions may have had on the control of the epidemic.

METHODS

Data

Data on new cases and deaths reported by all countries were obtained by the European Centre for Disease Control (https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide). Data up to 24th April are included. For the UK we used only pillar 1 case numbers. Pillar 1: swab testing in Public Health England laboratories and National Health Service hospitals for those with a clinical need, and health and care workers. Pillar 2 results (swab testing for health, social care and other essential workers and their households) as reported daily on https://www.gov.uk/guidance/coronavirus-covid-19-information-for-the-public#history were removed from the case numbers, as pillar 2 sampling was only introduced late in the course of the UK epidemic and inflated total case numbers relative to earlier in the UK outbreak. We also adjusted by the number of tests reported per 1 million population taken as of 16th April from worldometer (https://www.worldometers.info/coronavirus/). In order to compare time series for different countries with different dates of onset for their own epidemics we chose to define the onset as the first day after the latest time where there were two or more consecutive days with no cases reported.

The dates when (if at all) each of the various social restrictions were imposed for 30 European countries were given by the Institute of Health Metrics and Evaluation Data (IHME) (https://covid19.healthdata.org/). The six categories are “Mass gathering restrictions”, “Initial business closure”, “Educational facilities closed”, “Non-essential services closed” and “Stay at home order” and “Travel severely limited”. However, no country was listed in the dataset as having severe travel restrictions during the monitoring period so this was dropped from any further analysis. The IHME definitions of these measures are given on their website. Paraphrasing the definitions.

- Mass gathering restrictions are mandatory restrictions on private or public gatherings of any number.
- The first time that there was any mandatory closure of businesses, not necessarily all businesses. Usually such initial closures would primarily affect business such as entertainment venues, bars and restaurants.

- Where non-essential businesses are ordered to close, this usually include many more businesses than were in the first closure category. The second wave of closures likely includes general retail stores and services such as hairdressers.
- Education facilities closure includes all levels (primary, secondary and higher) education that stop direct teacher to student teaching.
- Stay at home orders affect all individuals unless travelling for essential services. They allow close contact only with people of the same household and perhaps some outdoors exercise.

In three countries (Germany, Italy and Spain) the restrictions were not implemented uniformly through the country on precisely the same dates so we took the median date for the nation; the actual variations in dates were extremely small in Italy and Spain and only somewhat diverse in Germany. See Supplemental Material 1. Among the 16 German states, 15 states imposed mass gathering restrictions within 2 days of the median date used; 9 states had initial business closures within 2 days of the median date; 15 states closed educational establishments within 2 days of the median date; 9 states closed non-essential businesses within 2 days of the median German date; all states imposed stay at home orders within 2 days of the median national date. All models adjusted for when countries started to advise or mandate their citizens to wear face masks or coverings (dates of face cover measures are listed in Supplemental Material 2). We included when countries either mandated or encouraged the wearing of face coverings or masks in public places as an independent control measure in the models. However, it was obvious that how such advisories or mandates were implemented considerably from one country to another. For example, in some countries face masks were required both outdoors and indoors in public and in others only in indoor settings. Sometimes, mask wearing was required in relatively few settings such as on public transport, other times in several settings such as on public transport, in shops and in schools. Also mask-wearing mandates where implemented were only introduced relatively late in the monitoring period, even as other control measures were being relaxed, which complicates interpretation of how much masks may have helped reduce transmission. Consequently, although we included the wearing of face coverings in the analyses, we caution against drawing any strong conclusions over their value based on these analyses alone. We also adjusted by the number of tests reported per 1 million population taken on the 17th April from WorldoMeter (www.worldometers.info/coronavirus). Ethics approval was not required because this is an analysis of data in the public domain.

Analyses

We undertook two sets of analyses. In order to ensure comparability between countries with different timing of their outbreaks we counted dates as being from the start of the growth phase of the epidemic in that country. The epidemic in each country was assumed to have commenced on the first day after the last time that no cases were reported as occurring on two consecutive days.

The first analysis was done in R using Bayesian generalised additive mixed effects models. These incorporate both fixed and random effects (ie mixed effects) to adjust for spatial dependency in disease between nation states. Random effects correspond to those for which levels are samples from a larger population, whereas fixed effects correspond to average effects for the whole population. Examples of fixed effects would be interventions such as shutting all schools and making people work at home. Other sources of variation that contribute may be more random, and associated with unmeasured features of the sampling unit (the nation state). Key here is the fact that the nation states will differ culturally and in other features such as recording methods. We have not measured the source of the variation but we know it is associated with the sampling unit (state) with which the response is recorded through time. In addition we also anticipate that there are spatial effects associated with the fact that nation states suffering from COVID-19 epidemics are geographically juxta-positioned (eg Germany abuts Austria along a land border). We expect that there will be some spatial dependency between states as the closer they are to each other the more likely it is that they have similar patterns of disease. Bayesian models are very useful as they allow us to quantify the relative contributions of fixed, random, temporal and spatial dependency in the same modelling framework.

The variance in the COVID-19 data was four orders of magnitude larger than the mean number of cases, and three orders of magnitude larger than the mean number of deaths. Consequently, models were fit using a negative binomial specification to account for potential over-dispersion in the data, and within a conditional autoregressive model (Besag-York-Mollié) framework [13] to allow for potential spatial autocorrelation and unstructured between-country variation.

Let $Y_{i,t}$ be the number of COVID-19 cases or deaths for country $i = 1, \cdots, I$ at time $t = 1, \cdots, T$
The general algebraic definition of the models is given by:

$$Y_{i,t} | \mu_{i,t}, \phi \sim NegBin(\mu_{i,t}, \phi),$$

where $Y_{i,t}$ is the number of COVID-19 cases or deaths for country $i = 1, \cdots, I$ at time $t = 1, \cdots,$ $\mu_{i,t}$ is the predicted number of COVID-19 cases or deaths for country i and time t, and $\phi > 0$ is the negative binomial dispersion parameter. A logarithmic link function of the expected number of cases or deaths was modelled as:

$$\log(\mu_{i,t}) = \alpha + \log(P_{i,d[t]}) + \delta D_{i,d[t]} + \epsilon R_{i,d[t]} + \sum_k \beta X_{i,t,k} + u_i + \nu_i,$$

where α corresponded to the intercept; $\log(P_{i,d[t]})$ denotes the logarithm of the population at risk for country i and day $d_{[t]}$ was included as an offset to adjust case counts by population. $D_{i,d[t]}$ is a linear term for the number of days since the outbreak started, with coefficient δ. $R_{i,d[t]}$ is a linear function of the number of COVID-19 tests carried out per country i at day $d_{[t]}$, with regression coefficient ϵ. X is a matrix of k intervention measures (e.g. school and business closures) with regression coefficients β. Intervention measures comprise of an index of $1, \cdots, N$ number of days following the intervention being implemented (day 1 was the day following the intervention implementation). We assumed that the imposition of each intervention led to cumulative changes in effect. Intervention measures were included in the model as a random effect to account for potential non-linearities in the exposure-response relationship. A random effect adjustment was appropriate because the observation data (case counts) were samples from a larger population (due to limited testing to confirm symptomatic cases and possible asymptomatic cases). Unknown confounding factors with spatial dependency that represent, for example, human mobility, were incorporated using spatially correlated (i.e. structured) random effects (u_i) and independent, identical and normal distributed (i.e. unstructured) random effects (ν_i) for each country i. Spatial random effects were specified using a Besag-York-Mollie model to account for spatial dependencies and unstructured variation between countries [14] . Goodness of fit was evaluated using the Deviance Information Criterion (DIC). Models were fitted in R version 3.6.1 using the INLA package.

The second analysis was a multi-level mixed effects regression analysis in STATA v 16.1. We used a mixed effects negative binomial regression model with cases or deaths on a specific day as the outcome variable, country population as the exposure variable, country as a mixed effect, and days from start of the epidemic as a fixed effect. Fixed effect was appropriate for days elapsed because we were looking for possible effect of NPI relevant to a fixed start point and over the entire population. All main interventions were included as categorical variables with the week number included as a linear variable after the start of the intervention. Monitoring by week number was appropriate with regard to case counts, given that incubation period tends to be about 5 days [15-

18], and a small lag between symptom onset and obtaining test results is likely: thus, total days elapsed from exposure to changes in recorded case counts has tended to be about 7 days. A lag from symptom onset to hospitalisation of about 7 days [19, 20], and a similar subsequent lag (about 7 days) from hospitalisation to death are reported in COVID-19 literature [19-21]. Figure 1 indicates these key likely onset of intervention impact on an exemplar epidemic curve. For simplicity and brevity we report only on the results for the 7 day categorisation in this manuscript. However, in view of the variation in incubation period and the possibility that this might have interfered with the parameter estimates, we repeated Analysis 2 for three alternative response time periods (post-intervention) as sensitivity analyses. These alternative response periods were 4 days, 10 days and 14 days. The resulting incident risk ratios (between our preferred response period of 7 days and alternatives) could then be compared for possible trend differences. In further sensitivity and collinearity checks, we dropped each of the main predictor variables (intervention timings) from the final equation and noted if the regression parameter and standard errors of remaining predictor variables changed dramatically or if the coefficients reversed trend (eg., went from suggesting increase to suggesting decrease).

We also checked for collinearity between the predictor variables by calculating the variance inflation factors (VIF) for the predictors and by calculating the condition number using the coldiag2 command in STATA. A VIF of < 10 suggests that model predictors do not have multi-collinearity problems. VIF values > 10.0 need to be considered with regard to other model diagnostics, such as condition index and eigenvalues. A condition number > 15 with any variance proportions above 0.9, or if eigenvalues were < 0.01 could suggest collinearity that undermines confidence in coefficient estimates, according to guidance in Chatterjee and Hadi 2015 [22] and Regorz 2020 [23]. In addition, as sensitivity analysis, within analysis 2 we reran the model dropping each predictor variable in turn to determine whether or not the regression parameters and their standard errors were changed substantially.

RESULTS

Table 1 gives the estimated date of the start of the epidemic in each country and when each of the five intervention types were implemented, according to the IHME website. "Mass gathering restrictions", "initial business closure", "educational facilities closed", "non-essential services closed" and "stay at home order" were respectively implemented by 29, 28, 29, 23 and 19 countries. Italy was the first country to enter the epidemic on 22nd February and Lithuania the last on 14th march. By our criteria, half of all countries had their epidemic start on or before 27th February.

Table 1. Timing of estimated start of each country's main epidemic and the introduction of social distancing measure across 30 European countries (all dates in 2020).

Country	Start of main epidemic	Mass gathering restrictions	Initial business closure	Educational facilities closed	Non-essential services closed	Stay at home order	Face covering encouraged or compulsory
Austria	26/02	10/03	16/03	16/03	16/03	16/03	06/04
Belgium	02/03	13/03	13/03	14/03	18/03	18/03	NA
Bulgaria	12/03	13/03	13/03	13/03	13/03	17/03	30/03
Croatia	11/03	09/03	19/03	16/03	19/03	17/03	NA
Cyprus	10/03	24/03	24/03	13/03	24/03	24/03	NA
Czech Rep	02/03	10/03	10/03	10/03	14/03	16/03	18/03
Denmark	27/02	18/03	18/03	16/03	NA	NA	NA
Estonia	11/03	13/03	13/03	16/03	NA	NA	05/04
Finland	27/02	12/03	18/03	18/03	04/04	NA	NA
France	26/02	04/03	14/03	12/03	14/03	16/03	05/04
Germany	26/02	22/03	17/03	16/03	23/03	22/03	01/04
Greece	05/03	08/03	12/03	11/03	22/03	23/03	NA
Hungary	05/03	12/03	12/03	16/03	16/03	28/03	NA
Ireland	04/03	12/03	15/03	12/03	24/03	27/03	NA
Italy	22/02	11/03	11/02	05/03	11/03	11/03	06/04
Latvia	08/03	13/03	NA	12/03	NA	NA	NA
Lithuania	14/03	15/03	14/03	16/03	15/03	15/03	01/04
Luxembourg	07/03	13/03	18/03	16/03	18/03	NA	20/04
Malta	08/03	NA	17/03	13/03	23/03	NA	NA
Netherlands	28/02	10/03	21/03	15/03	NA	NA	NA
Norway	27/02	12/03	12/03	12/03	NA	NA	05/04
Poland	07/03	10/03	31/03	12/03	NA	24/03	NA
Portugal	03/03	19/03	16/03	16/03	19/03	19/03	16/04
Romania	04/03	06/03	21/03	11/03	21/03	23/03	NA
Slovakia	07/03	12/03	16/03	12/03	16/03	NA	14/03
Slovenia	05/03	12/03	15/03	16/03	15/03	20/03	29/03
Spain	25/02	15/03	15/03	14/03	15/03	15/03	13/04
Sweden	27/02	11/03	NA	NA	NA	NA	NA
Switzer'd	26/02	28/02	16/03	13/03	16/03	NA	NA
UK	28/02	23/03	20/03	23/03	24/03	23/03	NA

Note: NA = not applicable, this control was not implemented.

Analysis 1

Model metrics are presented in Table 2. The dispersion parameter evaluates whether the model is able to cope with potential dispersion in the data. When the value is close to 1 (as it is here) the model is shown to do well at accounting for dispersion.

Table 2. Model metrics

Model	Deviance Information Criterion	Watanabe-Akaike Information Criterion	Conditional predictive ordinate	Dispersion
Cases	18009.4	18012.6	-9006.6	1.01
Deaths	8032.4	8035.9	-4018.4	0.89

Notes: The Watanabe-Akaike Information Criterion is described by Watanabe 2010 [24] and was developed to specifically help identify best model fit in Bayesian models. Smaller W-AIC values mean better fit compared to alternative model specifications. The conditional predictive ordinate is a Bayesian diagnostic [25] that detects surprising observations.

The exposure-response relationships estimated by the models are presented in Figures 2 (cases) and Figure 3 (deaths). The *X* axis represents the days since the intervention started and the *Y* axis indicates the logarithm of the risk ratio. It can be observed that mass gathering restrictions have a negative effect on the number of cases with fewer cases occurring as the number of days since intervention started increases. A similar effect is observed for the initial closure of business and the closure of education facilities with less cases occurring as the number of days since the intervention increases. The closure of non-essential business does not appear to have a significant effect on the number of COVID-19 cases. This is evident as the estimated relationship and its 95% credible interval stay close to zero on the Y axis. Surprisingly, stay-home measures showed a positive association with cases. This suggests that as the number of lock-down days increased, so did the number of cases. Negative associations with deaths (Figure 3) were estimated for mass gatherings, initial business closure and the closure of educational facilities; while a non-significant effect was estimated for non-essential business closure. The stay-home measures showed an inverted U quadratic effect with an initial rise of deaths up to day 20 of the intervention followed by decrease. These results suggest that stay at home orders may not be required to ensure outbreak control and reduce outbreak harms, provided all the other control measures are implemented. Of course, if stay at home measures are implemented then all the other measures such as business closures, banning mass gatherings and school closures would also follow.

The patterns seen in Figures 2 and 3 fit with the understood disease incubation, development and concurrent ascertainment processes. The median incubation period is understood to be 4-7 days

[15-17], while case ascertainment tended to require an elapse of 2-10 more days [26]. For severe cases (those who are hospitalised), 8-14 days post symptom onset tends to coincide with start of a 5-7 day long period of peak disease severity [20]. As a result, we expect no intervention should be cited as affecting case counts in under about 7 days, and no intervention is likely to strongly reduce counts of death in less than 2-3 weeks.

For cases and deaths, mask wearing mandates/advisories seem to have either initial effects which were negative (case) or neutral (deaths), followed by rises (in cases or deaths). The overall effect is quite small, which we confirmed with further sensitivity analysis show below. The additional benefit of mask-wearing advisories/mandates to the other outbreak control measures seemed to be small and inconsistent. However, for the reasons discussed above we hesitate to interpret these results as certain effects of face cover/mask mandates/advisories.

Figures 4 and 5 shows the association between actual cases and deaths in each country, expressed as 7 day rolling means, and the numbers predicted by the models on cases and deaths. Although for many countries there is a reasonable correlation between the two this is not the case for all countries and particularly countries with smaller populations. The model outputs especially did not fit Sweden which had much lower numbers of cases and deaths than predicted. This could be explained by partial implementation of controls and unmandated behavioural change in the population. We acknowledge that, at least for some countries, our model could not capture all the temporally changing variables influencing the spread of the disease.

Figure 6 shows the maps of the posterior mean for the country-specific relative risks of (A) COVID-19 cases, and (C) COVID-19 deaths. These country-specific risks enable comparisons of individual countries to case/death incidence in whole study area, having accounted for the effects of all other covariates in the model. Figures 6-A and 6-C indicate whether the cases or deaths were relatively higher or lower in each country relative to full-region incidence (cases or deaths per 100,000). Posterior means in the top two categories (in shades of orange) indicate especially excess country-specific risk relative to cases/deaths in the whole region. Posterior means lower than 1.0 (dark blue) indicate a lower relative risk than that of the whole region. Maps 6-B and 6-D show the country-specific posterior probability (range 0-1) of observing a relative risk larger than one (compared to case/death incidence in all 30 countries). The proportion of spatial variance explained by the models is 16% for the case-specific model, and 15% for the death-specific model. These values (15-16%) are not high, indicating that the spatial components of the models are not highly explanatory of the variability in cases/deaths.

Analysis 2

For confirmation and comparison, the analysis was repeated using a multilevel mixed effects model with results shown in Table 3. The conclusions of this analysis were broadly the same as for the hierarchical probabilistic models described above. The coefficients for these models assess the independent contributions of the interventions to the outcomes whilst holding the others at their mean (as we would expect from a multivariate linear model). The Incident Risk Ratios (IRR) are shown in Table 3 with 95% confidence intervals, for each of deaths and cases, for each period (each week) after the intervention started. Larger IRR values suggest greater effects; a value of 1 implies no effect, values above 1.0 suggest increase in cases/deaths, while values below 1 imply decrease. For time periods 1-7 and 8 -14 the IRR values were above 1, indicating a positive association between cases/death and the intervention variable. For periods starting 15 days onwards the IRR was generally below 1 suggesting a negative association between the outcome and the intervention. This patterns probably reflects the time lag between exposure, latency and disease detection, so that impact interventions only 'kick-in' after what is effectively a lag period of 14 days. Closing schools, banning mass gatherings and initial business closures most reduced cases and deaths. Other measures had smaller and less consistent effects. In addition, we looked at the impact of removing each intervention or all interventions on the model log likelihoods (Table 4). The biggest impact came from removing educational closures from the model. The next biggest change came from removal of stay-home orders, but this intervention was associated with a smaller decline in epidemic risk (deaths). We note that removing mask-wearing as a control measure had a moderate effect on case counts but very minor effect in mortality outcomes; this difference may reflect the relatively late imposition of mask-wearing mandates/advisories.

Table 3. Results of mixed effects negative binomial model of effect of each intervention on case numbers and deaths

		Cases			Deaths		
Intervention	Timing	IRR	L95%CI	U95%CI	IRR	L95%CI	U95%CI
Mass gathering restrictions	Before	1	--	--	1	--	--
	1-7 d after	1.32	1.10	1.57	0.76	0.55	1.03
	8-14 d after	1.13	0.88	1.43	0.58	0.41	0.84

Intervention	Timing	Cases IRR	L95%CI	U95%CI	Deaths IRR	L95%CI	U95%CI
	15-21 d after	0.99	0.73	1.34	0.59	0.38	0.92
	22-28 d after	0.80	0.56	1.15	0.56	0.33	0.93
	29-35 d after	0.74	0.48	1.13	0.50	0.28	0.91
	36+ d after	0.66	0.40	1.09	0.49	0.25	0.98
Initial business closures	Before	1	--	--	1	--	--
	1-7 d after	1.18	0.96	1.46	1.07	0.80	1.43
	8-14 d after	0.87	0.66	1.15	1.07	0.75	1.54
	15-21 d after	0.69	0.49	0.96	0.72	0.47	1.11
	22-28 d after	0.61	0.41	0.91	0.50	0.29	0.83
	29-35 d after	0.47	0.29	0.76	0.42	0.22	0.77
	36+ d after	0.32	0.18	0.56	0.37	0.18	0.77
Educational facilities closed	Before	1	--	--	1	--	--
	1-7 d after	1.47	1.22	1.79	2.51	1.89	3.34
	8-14 d after	1.38	1.05	1.80	3.14	2.14	4.62
	15-21 d after	0.95	0.67	1.33	2.76	1.74	4.36
	22-28 d after	0.52	0.35	0.78	2.02	1.19	3.43
	29-35 d after	0.26	0.16	0.42	1.10	0.60	2.01
	36+ d after	0.14	0.08	0.25	0.55	0.28	1.10
Non-essential services closed	Before	1	--	--	1	--	--
	1-7 d after	1.14	0.92	1.41	1.40	1.03	1.90
	8-14 d after	1.15	0.90	1.47	1.41	1.00	1.97
	15-21 d after	1.02	0.78	1.33	1.42	0.99	2.03
	22-28 d after	0.83	0.60	1.13	1.44	0.95	2.17
	29-35 d after	0.76	0.52	1.10	1.04	0.65	1.68
	36+ d after	0.76	0.46	1.26	0.77	0.42	1.39
Stay at home order/advisory	Before	1	--	--	1	--	--
	1-7 d after	1.19	0.97	1.47	1.30	0.96	1.76
	8-14 d after	1.95	1.56	2.44	2.01	1.45	2.77
	15-21 d after	2.28	1.79	2.90	2.23	1.58	3.14
	22-28 d after	2.55	1.94	3.35	1.99	1.36	2.89
	29-35 d after	2.49	1.78	3.48	1.84	1.19	2.83
	36+ d after	2.39	1.49	3.84	1.21	0.70	2.10
Mask order/advisories	Before	1	--	--	1	--	--
	1-7 d after	0.66	0.55	0.79	0.91	0.75	1.11
	8-14 d after	0.53	0.43	0.65	0.89	0.71	1.12
	15-21 d after	0.52	0.40	0.67	0.97	0.73	1.29
	22-28 d after	0.68	0.48	0.98	1.40	0.91	2.15
	29-35 d after	1.15	0.70	1.87	1.36	0.72	2.55
	36+ d after	1.06	0.56	2.01	1.45	0.60	3.54
Days from epidemic start	per day	1.14	1.12	1.15	1.17	1.15	1.19

		Cases			Deaths		
Intervention	Timing	IRR	L95%CI	U95%CI	IRR	L95%CI	U95%CI
Tests per 1000 population as of 16 Apr		1.06	1.04	1.07	1.02	0.99	0.06
Random effects							
Country (Variance)		0.26	0.15	0.46	1.19	0.70	2.03

Note: IRR = Incident Risk Ratio. The IRR is generated by exponentiating the results of the model raw outputs which were generated in a default log scale.

Table 4. Log likelihood of each model for full model compared with models excluding each of the interventions and all interventions

	Model	Log likelihood	Change
	Full model (Cases)	-9081	
Excluded	Mass gathering restrictions	-9096	-15
	Initial business closures	-9097	-16
	Educational facilities closed	-9157	-76
	Non-essential services closed	-9085	-4
	Stay at home advisory	-9112	-31
	Face coverings	-9109	-28
	All interventions	-9617	-536
	Full model (Deaths)	-4096	
Excluded	Mass gathering restrictions	-4101	-5
	Initial business closures	-4109	-13
	Educational facilities closed	-4163	-66
	Non-essential services closed	-4104	-8
	Stay at home advisory	-4113	-17
	Face coverings	-4100	-4
	All interventions	-4569	-472

Collinearity and sensitivity analyses

Regression diagnostics for the alternative specifications of response time periods in Analysis 2 (4, 10 or 14 rather than 7 days) are shown in Supplementary Material 5 with visual comparisons available in Supplementary Material 6. There was little difference in the overall rate of decline in risk ratio with increased time since intervention regardless of time unit used. There were noticeable outliers in a few model IRR values at the longest time periods (above 40 or 50 days) when data contributions tended to be from just one or two countries (see Supplemental Material 5 and 6).

The VIF values for the predictor variables in Analysis 1 were all less that 10 (mean VIF 5.7) except for initial business closures which gave a VIF of 10.4 (Supplementary Material 3). Collinearity diagnostics for Analysis 2 were almost identical, in that the VIF only just exceed the 10.0 threshold and only for the initial business closures variable (Supplementary Material 4). The condition index exceeded 15.0 in the 9th dimension and suggested some collinearity between initial and non-essential business closure parameters. However, corresponding variance proportions in all dimensions for each control measure were well below 0.9. The smallest eigenvalue (Supplementary Material 4) was 0.059, which is above the suggested threshold of 0.01. These tests as a group indicate that collinearity between predictor variables did not harmfully bias the apparent separate contributions of each disease control measure (as indicated by coefficient central estimates) in our models. In addition, the standard errors of the predictors in both models were relatively small and in the sensitivity and collinearity checks, dropping each of the main predictor variables from the final equation of analysis 2 did not strongly change the coefficients and standard errors of remaining predictor variables. We conclude that there was some collinearity in our models, notably between the business closure variables, but that this was not enough to affect our conclusions.

Discussion

Our analyses confirm that the imposition of non-pharmaceutical control measures have been effective in controlling epidemics in each country. However, we were unable to demonstrate a strong impact from every intervention. Closure of educational facilities, banning mass gatherings and early closure of some but not necessarily all commercial businesses were all associated with reduction of the spread of infection. Widespread closure of all non-essential businesses and stay at home orders seem not to have had much additional value. Other analyses of actual intervention impositions and subsequent case/death counts also have found that school closures were especially effective control measures for reducing spread of COVID-19 [27-30]. However, it is vital that we caveat this finding (about closing educational establishments) by noting that it relates to closing schools that operated ‘as normal’ rather than when they operated with COVID-secure policies. We also do not attempt here to discuss what the best COVID mitigation measures might be within schools.

It seems likely that many possible combinations of social distancing measures can be effective. The apparent effects of the measures as described here may be biased by the measures themselves tending to have a sequence in common among all countries. Measures imposed later may seem less effective simply because of the order in which they happened (additional benefits were small after other measures were put in place). Other analysts have drawn this same conclusion about coronavirus NPIs [30]. Our analyses indicated that school closures and stopping mass gatherings were most effective, but we acknowledge that these measures were among the earliest taken in Europe; the data didn’t allow us to see what marginal gains might have been achieved if school closures had been the last of all measures taken. Also, different measures reinforced and enabled each other: for instance, there was little incentive to leave home if schools and businesses are already closed and weather was inclement (as it often is in early spring in Europe, when most national lockdowns started). Business and school closures usually preceded stay at home measures in Europe, so it may not have been possible for data on stay at home orders to be linked to large additional effects. This potential ordering problem is at least somewhat mitigated for by our use of individual lag measures (in timing) from when each intervention was effected. It is also worth noting that outside of institutional and crowded settings, there is evidence that much if not most COVID-19 transmission was within households in this period [31]; stay at home orders intensify contact within households which would be expected to increase household transmission. It could be therefore not surprising that stay at home measures on their own are not very effective outbreak control measures and may not generate large additional benefits.

There has been uncertainty about how beneficial the closing of educational establishments can be on coronavirus respiratory disease transmission [30, 32-37], especially given that children often have mild or no symptoms [38]. We cannot resolve the lack of consensus in these lines of evidence about how likely children are to pass SARS-COV-2 to adults. Emergences of novel and seemingly more infectious variants [39] of the virus may complicate attempts to understand transmission patterns between children and to adults using historical data, as well as understanding relatively effectiveness of specific non-pharmaceutical interventions. Our study similarly does not identify which level of school closure has the most benefit whether it is primary, junior, senior school or even higher education, though more recent evidence tends to point towards schooling between 11 and

19 as being more likely to drive transmission than education for younger children [35]. Note that our own results are based on total closure rather than schools operating with at least partial social distancing. The impacts of partial school closures or social distancing controls within open schools need to be evaluated separately.

After closing educational establishments, the next greatest impact on the epidemiology of the European COVID-19 controls was from banning mass gatherings (which could be of any size), both public and private gatherings. A 2018 review of the impact of mass gatherings on outbreaks of respiratory infectious disease [40] found that most evidence was linked to the Islamic Hajj pilgrimage, where most infections were respiratory, mainly rhinovirus, human coronaviruses and influenza A virus. The evidence for respiratory disease outbreaks arising from other mass gatherings such as music festivals or sporting events is less established, but not absent. Several outbreaks of respiratory infectious disease have been linked to large festivals [40, 41]. For instance, during the 2009 influenza season pandemic influenza A(H1N1)pdm09 outbreaks were recorded at three of Europe's six largest music festivals, while some 40% of pandemic flu cases that season in Serbia were linked with the Exit music festival. Analyses by other investigators using different approaches than ours on COVID-19 NPIs also tend to find that banning large gatherings can be especially effective for reducing disease transmission [30].

The types of business closures are interesting. We established that there was weak collinearity between the two types of business closures in the models. However, the stronger association was with the initial business closures. Given that those initial closures were mostly directed at business where people congregate and have a purpose of facilitating socialising (i.e. the hospitality industry), this would suggest that control measures among these businesses are where the most impact may be had. Although outbreaks of food poisoning are frequently linked with venues where food is consumed, outbreaks of respiratory infections are much more rarely so. One exception was an outbreak of SARS at a restaurant where live palm civets were caged close to customer seating [42]. The link with COVID-19 is probably much less about food and beverage consumption, and simply about time people spend in close proximity to each other.

Similar to other authors who have tried to assess relative importance of possible NPIs in controlling COVID-19 and not found strong benefits for face-cover usage [43], we hesitate to interpret our findings on mask-wearing as definitive. Mask-advisories have not been implemented in isolation, and were often implemented relatively late in the sequence of NPIs in the group of European countries that we studied. Mask-interventions were also implemented unevenly (as advisories or mandates) and usually only in limited settings. Our separate evidence review [44] found that mask-wearing to stop respiratory disease transmission is likely to be only be modestly effective, but we agree that when it comes to a pandemic situation, small protective measures may have cumulative important benefits [45].

Limitations

Although our study suggests that closures of educational interventions and banning mass gatherings are the most important measures, this is caveated with several observations. Many interventions were implemented in different ways and at different points in the local epidemic. We relied on published and observed data which may have suffered from problems of under-ascertainment; the true effect of specific interventions may depend on true community prevalence that was not measured accurately enough. We did not undertake a systematic sensitivity analysis (excluding just one country per model, for instance) or adjustments in categorisations. It is likely that there will be serial dependency in the data as the level of disease at one time point is dependent (inevitably) on prior states of disease in the nation state, but we did not attempt to measure serial dependency in our models which might have further informed relative NPI efficacy. For example, in accordance with the IHME assignment, we treated Sweden as a country without school closures because schools for persons under 16 stayed open, although upper secondary and tertiary education facilities were actually shut in Sweden from late March 2020 [46]. Given recent evidence that secondary (age 11 to 19) rather than junior schools that may play an important role in transmission of COVID, the educational closures in Sweden may explain in part the divergence from our predictions in that country [35]. Our models cannot allow for differences in school building construction materials or ventilation rates between countries that might influence transmissibility. The findings in support of school closures to contain the virus can truly only refer to schools when schools operate 'as normal' and not with COVID mitigation practices in place. The exact timing of restrictions being introduced varied over time in Italy, Spain and between individual federal states in Germany [47]. Which types of work places could stay open varied; while the acceptable reasons for being outdoors also varied between countries. Stay at home orders in some countries was an advisory but not enforced whilst elsewhere stay-home orders were enforced by police with penalties. In some countries, children could go outside and outdoor exercise was permitted whilst in others either or both might be banned. In some countries, severe travel restrictions were a separate intervention whilst in others travel bans were a consequence of a stay at home order and could not be identified separately.

Because of this variety in how interventions were implemented and described, the results for the potential of stay at home advisories especially may be under-estimated. All models are simplifications of the complex nature of reality; our modelling was unable capture many subtle variations in how control measures were implemented. We acknowledge that lack of direct observation of these variations may have biased our results.

Conclusion

Relaxing stay-at-home orders and allowing reopening of non-essential businesses appeared to be the lowest risk measures to relax as part of plans to carefully lift COVID-19 lockdown measures. There is still even now relatively little unclear empirical evidence on the relative value of different interventions. And yet, the reasons to implement minimal control measures are compelling, given the social and economic harm linked to tight control measures. Hence, whilst we need to be cautious about using preliminary results, public health officials will have to use evidence as it emerges rather than expect to wait for a final full view to decide what might be (was) the best control strategy. Careful monitoring of how relaxation of each control measure affects transmissibility of COVID-19 is required and will help to minimise the inevitably imperfect results.

REFERENCES

1. Jacobs LA. Rights and quarantine during the SARS global health crisis: differentiated legal consciousness in Hong Kong, Shanghai, and Toronto. Law & Society Review. 2007;41(3):511-52.
2. Johnson HC, Gossner CM, Colzani E, Kinsman J, Alexakis L, Beauté J, et al. Potential scenarios for the progression of a COVID-19 epidemic in the European Union and the European Economic Area, March 2020. Eurosurveillance. 2020;25(9):2000202.
3. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: Rapid review of the evidence. The Lancet. 2020;395:912-20.
4. Office for Budget Responsibility. Coronavirus reference scenario 2020 [updated 7 April. Available from: https://obr.uk/coronavirus-reference-scenario/.
5. Massaro E, Ganin A, Perra N, Linkov I, Vespignani A. Resilience management during large-scale epidemic outbreaks. Scientific Reports. 2018;8(1):1-9.
6. Adam D. Special report: The simulations driving the world's response to COVID-19. Nature. 2020;580(7803):316-8.
7. Sridhar D, Majumder MS. Modelling the pandemic. British Medical Journal 2020;369:m1567.
8. Borenstein S, Johnson CK. Modeling coronavirus: ‘Uncertainty is the only certainty’. Associated Press,. 2020 7 April. https://apnews.com/88866498ff5c908e5f28f7b5b5e5b695.
9. Andrews C. Predicting the pandemic: mathematical modelling tackles Covid-19. Engineering & Technology. 2020 10 April.

10. McCoy D. Faith in coronavirus modelling is no substitute for sound political judgment. The Guardian. 2020 10 April. https://www.theguardian.com/commentisfree/2020/apr/10/modelling-pandemic-politicians-decisions-science.
11. Rosenbaum PR. How to see more in observational studies: Some new quasi-experimental devices. Annual Review of Statistics and Its Application. 2015;2:21-48.
12. Harris AD, McGregor JC, Perencevich EN, Furuno JP, Zhu J, Peterson DE, et al. The use and interpretation of quasi-experimental studies in medical informatics. Journal of the American Medical Informatics Association. 2006;13(1):16-23.
13. Besag J, York J, Mollié A. Bayesian image restoration, with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics. 1991;43(1):1-20.
14. Obaromi D. Spatial Modelling of Some Conditional Autoregressive Priors in A Disease Mapping Model: the Bayesian Approach. Biomedical Journal of Scientific & Technical Research. 2019;14(3).
15. Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. The Lancet Infectious Diseases. 2020;20(8):911-9.
16. 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;25(5):2000062.
17. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Annals of Internal Medicine. 2020;172(9):577-82.
18. Elias C, Sekri A, Leblanc P, Cucherat M, Vanhems P. Incubation period of COVID-19: A meta-analysis. International Journal of Infectious Diseases. 2021.
19. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 16,749 hospitalised UK patients with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol. medRxiv. 2020.
20. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet. 2020;395(10229):1054-62.
21. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet infectious diseases. 2020.
22. Chatterjee S, Hadi AS. Regression analysis by example. 4th ed: John Wiley & Sons; 2015 Feb 25.
23. Regorz A. How to interpret a Collinearity Diagnostics table in SPSS 2020 [updated 18 Jan 2020. Available from: http://www.regorz-statistik.de/en/collinearity_diagnostics_table_SPSS.html.
24. Watanabe S. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. Journal of Machine Learning Research. 2010;11(Dec):3571-94.
25. Pettit L. The conditional predictive ordinate for the normal distribution. Journal of the Royal Statistical Society: Series B (Methodological). 1990;52(1):175-84.
26. Wright O. People told to wait ten days for coronavirus test results. The Times. 2020 May 7. https://www.thetimes.co.uk/article/people-told-to-wait-ten-days-for-coronavirus-test-results-7dpjb0t96.
27. Pan A, Liu L, Wang C, Guo H, Hao X, Wang Q, et al. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. Journal of the American Medical Association. 2020;323(19):1-9.
28. Zhang J, Litvinova M, Liang Y, Wang Y, Wang W, Zhao S, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. Science. 2020;368:1481-6.

29. Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. The Lancet Infectious Diseases. 2020;20(7):793-802.
30. 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;4(12):1303-12.
31. Koh WC, Naing L, Rosledzana MA, Alikhan MF, Chaw L, Griffith M, et al. What do we know about SARS-CoV-2 transmission? A systematic review and meta-analysis of the secondary attack rate, serial interval, and asymptomatic infection. PLoS One. 2020;15(10):e0240205.
32. 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;4(5):397-404.
33. Cowling BJ, Lau EH, Lam CL, Cheng CK, Kovar J, Chan KH, et al. Effects of school closures, 2008 winter influenza season, Hong Kong. Emerging Infectious Diseases. 2008;14(10):1660.
34. Hens N, Ayele GM, Goeyvaerts N, Aerts M, Mossong J, Edmunds JW, et al. Estimating the impact of school closure on social mixing behaviour and the transmission of close contact infections in eight European countries. BMC Infectious Diseases. 2009;9(1):187.
35. Flasche S, Edmunds WJ. The role of schools and school-aged children in SARS-CoV-2 transmission. The Lancet Infectious Diseases. 2020;S1473-3099(20):30927-0.
36. Folkhälsomyndigheten: Public Health Agency of Sweden. COVID-19 in schoolchildren: A comparison between Finland and Sweden 2020 [Available from: www.folkhalsomyndigheten.se/publicerat-material/.
37. von Bismarck-Osten C, Borusyak K, Schönberg U. The role of schools in transmission of the SARS-CoV-2 virus: Quasi-experimental evidence from Germany: Ruhr Economic Papers; 2020 [Available from: https://www.rwi-essen.de/media/content/pages/publikationen/ruhr-economic-papers/rep_20_882.pdf.
38. Shen K, Yang Y, Wang T, Zhao D, Jiang Y, Jin R, et al. Diagnosis, treatment, and prevention of 2019 novel coronavirus infection in children: experts' consensus statement. World Journal of Pediatrics. 2020;16:1-9.
39. Public Health England. Investigation of novel SARS-COV-2 variant: Variant of Concern 202012/01. 2020 Dec 21.
40. Hoang V-T, Gautret P. Infectious diseases and mass gatherings. Current Infectious Disease Reports. 2018;20(11):44.
41. Botelho-Nevers E, Gautret P. Outbreaks associated to large open air festivals, including music festivals, 1980 to 2012. Eurosurveillance. 2013;18(11):20426.
42. Wang M, Yan M, Xu H, Liang W, Kan B, Zheng B, et al. SARS-CoV infection in a restaurant from palm civet. Emerging Infectious Diseases. 2005;11(12):1860.
43. 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;10.1126.
44. Brainard J, Jones NR, Lake IR, Hooper L, Hunter PR. Community use of face masks and similar barriers to prevent respiratory illness such as COVID-19: a rapid scoping review. Eurosurveillance. 2020;25(49):2000725.
45. Cowling BJ, Leung GM. Face masks and COVID-19: don't let perfect be the enemy of good. Eurosurveillance. 2020;25(49):2001998.
46. Johnson S. Swedish parents fret as schools stay open amid European virus shutdown. Reuters. 27 March 2020. https://uk.reuters.com/article/uk-health-coronavirus-sweden-education/swedish-parents-fret-as-schools-stay-open-amid-european-virus-shutdown-idUKKBN21E2NH.
47. Anonymous. How are Germany's 16 states dealing with the coronavirus pandemic? The Local DE. 19 March 2020. https://www.thelocal.de/20200319/how-are-germanys-16-states-dealing-with-coronavirus-pandemic.

Figure 1. Exemplar timeline of possible NPI impositions and potential epidemic response

Figure 2. Incidence Rate Ratios (cases) following implementation of country level non-pharmaceutical control measure and daily reported COVID 19 case numbers in 30 European countries.

Figure 3. Incidence Rate Ratios (deaths) following implementation of country level non-pharmaceutical control measure and daily reported deaths from COVID-19 in 30 European countries.

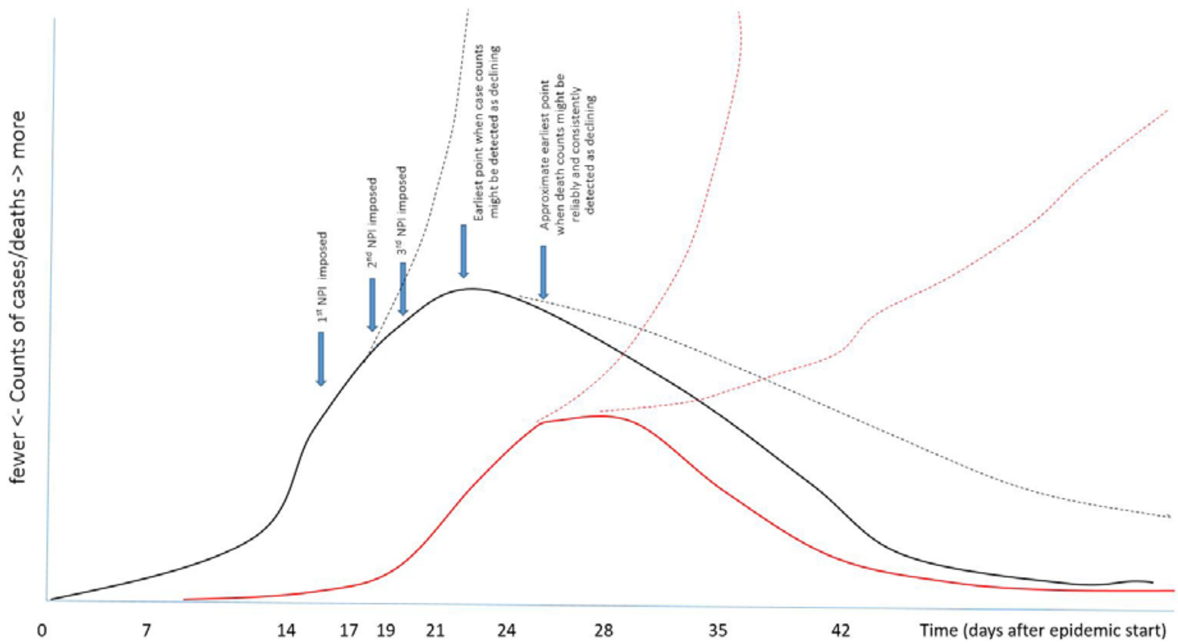
Figure 4. Comparison of predicted daily reports of case numbers of COVID-19 with seven day rolling average actual numbers across 30 European countries.

Figure 5. Comparison of predicted daily numbers of reports of deaths COVID-19 with seven day rolling average actual numbers across 30 European countries.

Figure 6. Posterior mean of the country-specific risk ratio of COVID-19 A) cases and C) deaths; and posterior probability of exceeding one COVID-19 B) case or D) death.

Figure 1. Exemplar timeline of possible NPI impositions and potential epidemic response

Figure 1. Exemplar timeline of possible NPI impositions **and** potential epidemic response



Notes: Black lines: cases; red lines: deaths. Solid lines: actual data, dashed lines: trajectories in absence of NPIs imposed. Exemplar timeline shows how epidemic measures (counts of cases or deaths) might respond, in case of 3 NPIs imposed, 2 of which were quite effective and one NPI that is less effective. In this exemplar, 1st NPI imposed is especially effective, 2nd NPI is somewhat effective (and thus there are two alternative trajectories for both cases and deaths), while the 3rd NPI makes no additional reduction to the epidemic. "Earliest point" refers to the first hypothesised time point in the exemplar when the epidemic might be expected to deviate from trajectory (ies) if the NPIs hadn't been imposed.

Figure 2. Incidence Rate Ratios (cases) following implementation of country level non-pharmaceutical control measure and daily reported COVID 19 case numbers in 30 European countries.

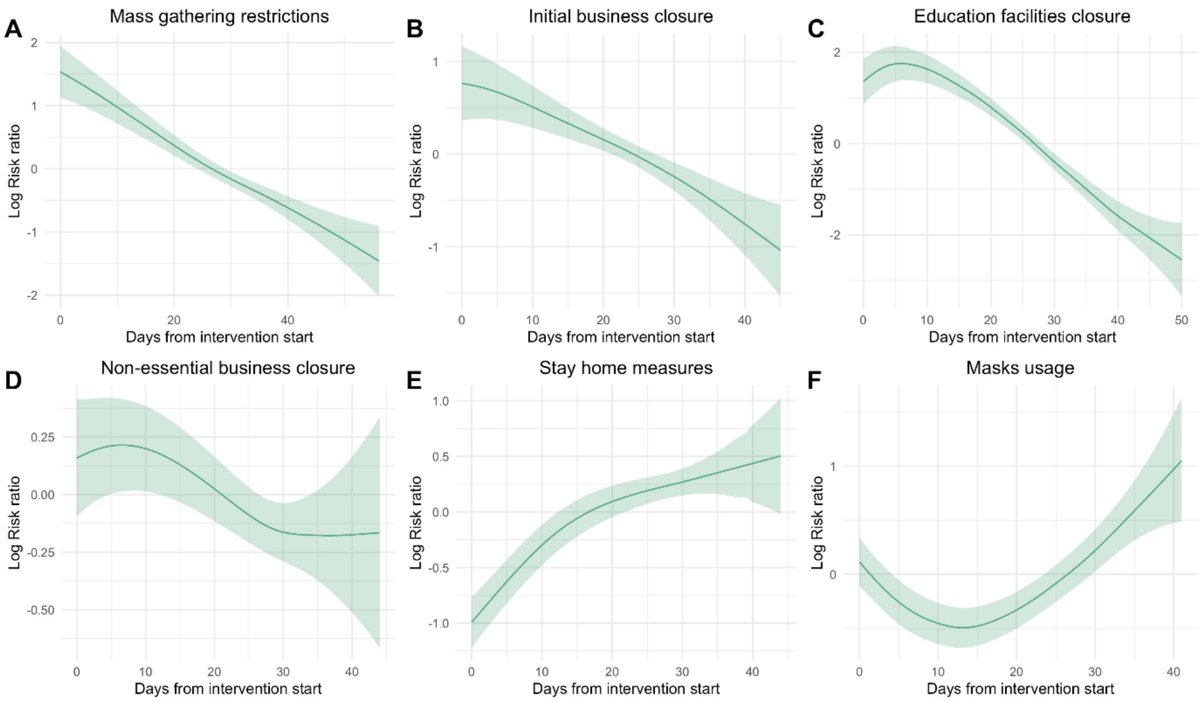


Figure 2 Notes: central line is posterior mean of the exposure-response relationship; shading is 95% confidence interval.

Figure 3. Incidence Rate Ratios (deaths) following implementation of country level non-pharmaceutical control measure and daily reported deaths from COVID-19 in 30 European countries.

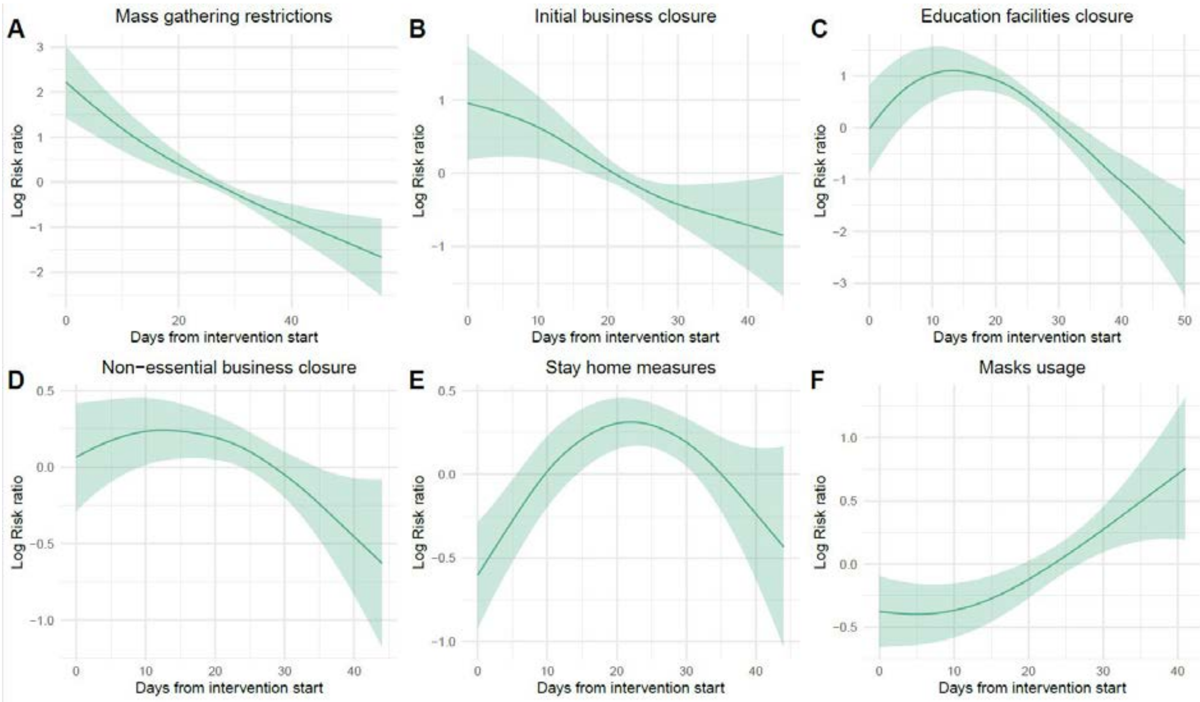


Figure 3 Notes: central line is posterior mean of the exposure-response relationship; shading is 95% confidence interval.

Figure 4. Comparison of predicted daily reports of case numbers of COVID-19 with seven day rolling average actual numbers across 30 European countries.

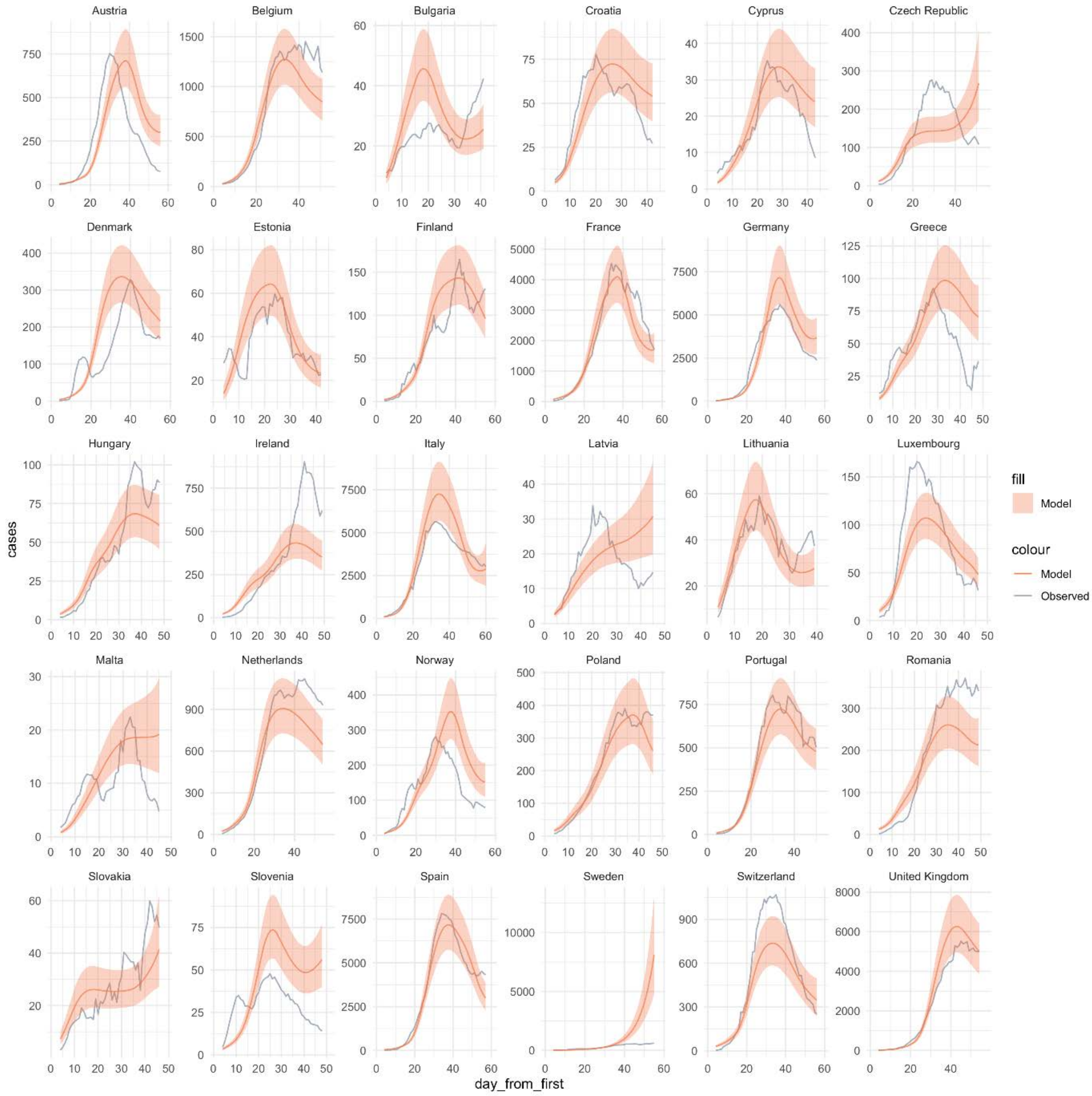


Figure 4 Notes: central line is posterior mean of the predictions made by the models (for individual countries over time); shading is 95% confidence interval.

Figure 5. Comparison of predicted daily numbers of reports of deaths COVID-19 with seven day rolling average actual numbers across 30 European countries.

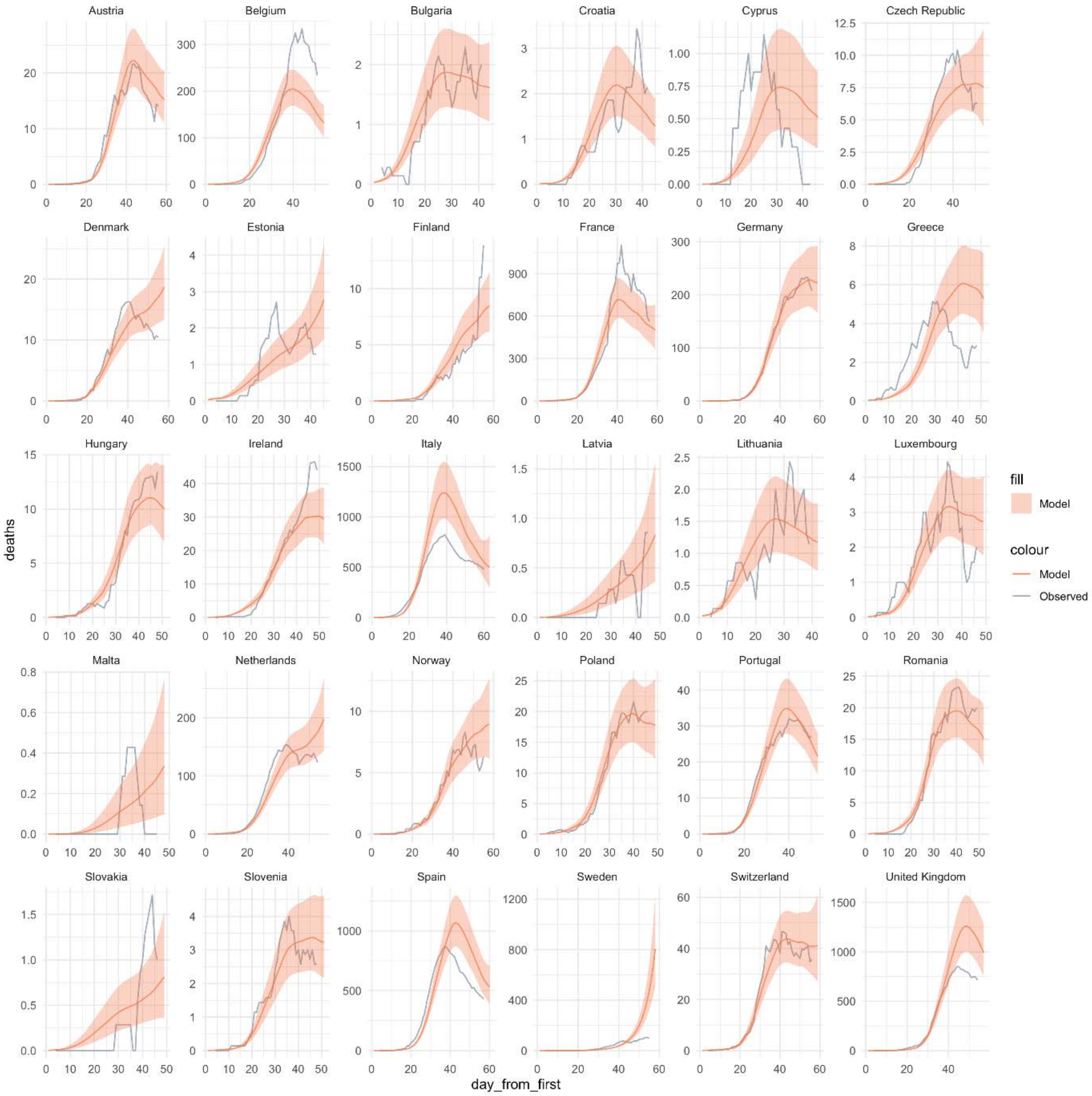
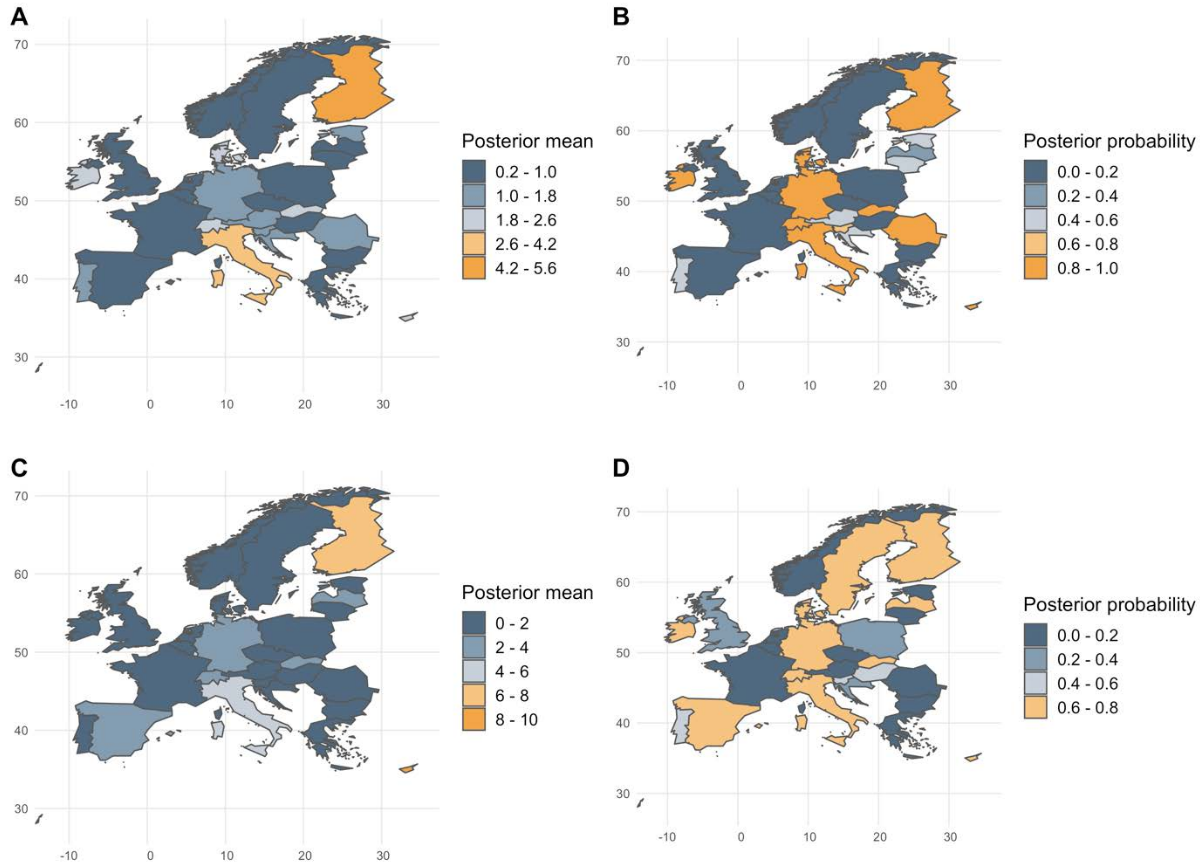


Figure 5 Notes: central line is posterior mean of the predictions made by the models (for individual countries over time); shading is 95% confidence interval.

Figure 6. Posterior mean of the country-specific risk ratio of COVID-19 A) cases and C) deaths; and posterior probability of exceeding one COVID-19 B) case or D) death.



Note: Figure 6 shows the maps of the posterior mean for the country-specific relative risks of (A) COVID-19 cases, and (C) COVID-19 deaths compared to the whole of the study area after having accounted for the effects of all other covariates in the model. Figure 6 also shows the country-specific posterior probability of exceeding (B) one case or (D) one death (per 100,000 persons after adjusting for covariates).

This supplementary material is hosted by *Eurosurveillance* as supporting information alongside the article "Impact of non-pharmaceutical interventions against COVID-19 in Europe: a quasi-experimental non-equivalent group and time-series design study ", on behalf of the authors, who remain responsible for the accuracy and appropriateness of the content. The same standards for ethics, copyright, attributions and permissions as for the article apply. Supplements are not edited by *Eurosurveillance* and the journal is not responsible for the maintenance of any links or email addresses provided therein.

SUPPLEMENTARY MATERIAL 1. Dates that restrictions were imposed in states/regions of Italy, Spain, and Germany

Dates in March when restrictions were imposed: eg., 15 = 15 March,

Spain	massgath	initbuscl	educclosed	nonessent	SAHO
Andalucia	15	15	14	15	15
Aragon	15	15	14	15	15
Asturias	15	15	14	15	15
Islas B	15	15	14	15	15
Basque C	15	15	14	15	15
Canaries	15	15	14	15	15
Cantabria	15	15	14	15	15
C&Leon	15	15	14	15	15
C-LaM	15	15	14	15	15
Catalonia	11	12	14	15	15
Ceuta	15	15	14	15	15
C Madrid	15	13	11	13	15
Extremadura	15	15	14	15	15
Galicia	15	15	14	15	15
La Rioja	15	15	14	15	15
Melilla	15	15	14	15	15
Murcia	15	15	14	15	15
Navarre	15	15	14	15	15
C Valencian	15	15	14	15	15
medians	15	15	14	15	15

SUPPLEMENTARY MATERIAL 1. Dates that restrictions were imposed in states/regions of Italy, Spain, and Germany (continued)

Dates in March when restrictions were imposed: eg., 11 = 11 March, -7 = 7 days before 1 March, or 23 Feb.

Italy	massgath	initbuscl	educclosed	nonessent	SAHO
Abruzzo	11	11	5	11	11
Basilicata	11	11	5	11	11
Calabria	11	11	5	11	11
Campania	11	11	5	11	11
Emilia-R	7	7	1	11	11
F-V Giulia	11	11	5	11	11
Lazio	11	11	5	11	11
Liguria	11	11	5	11	11
Lombardia	-7	-7	1	8	8
Marche	7	7	5	11	11
Molise	11	11	5	11	11
Piemonte	7	7	5	11	11
di Bolzano	11	11	5	11	11
di Trento	11	11	5	11	11
Puglia	11	11	5	11	11
Sardegna	11	11	5	11	11
Sicilia	11	11	5	11	11
Toscana	11	11	5	11	11
Umbria	11	11	5	11	11
V d'Aosta	11	11	5	11	11
Veneto	-7	-7	1	11	11
medians	11	11	5	11	11

SUPPLEMENTARY MATERIAL 1. Dates that restrictions were imposed in states/regions of Italy, Spain, and Germany (continued)

Dates in March when restrictions were imposed: eg., 22 = 22 March, -3 = 3 days before 1 March, or 27 Feb.

	massgath	initbuscl	educclosed	nonessent	SAHO
Germany	21	21	17	21	21
	21	17	16	21	21
	23	14	23	23	23
	23	17	18	17	17
	17	20	16	20	22
	22	15	16		22
	22	15	16		22
	23	27	16	27	23
	23	18	16	18	23
	23	-3	16	23	23
	22	23	16	23	22
	21	15	16		21
	22	24	16	24	22
	23	23	23	23	23
	24	14	16	24	24
	22	15	17		22
medians	22	17	16	23	22

SUPPLEMENTARY MATERIAL 2: Orders or recommendations to wear facemasks or face coverings, by sovereign territory in Europe

Austria 6 April 2020 (compulsory)

Compulsory in shops and most commercial premises from 6 April 2020, soon widened to public transport and shops that were due to re-open on 14 April.

https://uk.reuters.com/article/us-health-coronavirus-austria/austria-to-make-basic-face-masks-compulsory-in-supermarkets-idUKKBN21H16A

https://uk.reuters.com/article/uk-health-coronavirus-austria-masks/austrian-supermarkets-hand-out-face-masks-before-they-become-compulsory-idUKKBN21J5XP

=====================

Bulgaria 30 March 2020 (mix recommended/compulsory)

Strong recommendation from 30 March; made compulsory in all public places from 12 April; extended to 13 May.

https://sofiaglobe.com/2020/03/30/covid-19-bulgaria-makes-wearing-a-protective-mask-in-public-places-compulsory/

https://www.bnr.bg/en/post/101257255/bulgaria-introduces-mandatory-wearing-of-masks-in-public-from-april-12-until-april-26-inclusive

https://www.novinite.com/articles/204264/The+Mandatory+Wearing+of+Protective+Masks+is+Extended+until+May+13

===========================

Czechia 18 March 2020 (compulsory)
Mandatory in all public spaces and many work places. Order extended until end of June 2020.

https://www.praguemorning.cz/face-masks-now-mandatory-in-all-prague-shops-and-offices/

https://news.expats.cz/weekly-czech-news/prymula-face-masks-to-remain-mandatory-in-czech-republic-until-end-of-june/

====================
Estonia 5 April 2020 (highly recommended) statement by PM

https://news.err.ee/1073236/prime-minister-we-are-unfortunately-still-in-coronavirus-deepening-phase

========================

France 5 April (encouraged not required)

https://www.thelocal.fr/20200406/mask-or-no-mask-what-is-the-official-coronavirus-advice-in-france

https://www.france24.com/en/20200405-coronavirus-abrupt-reversal-on-mask-policy-in-france-and-the-us-raises-new-questions
=======================

Germany 1 April 2020 (recommendation that became compulsory)

Variable rules for when must be worn vary by state and sometimes by city, introduced dates also variable. Nationally mandated from 27 April on public transport & also in most shops before then. RKI formally endorsed mask-wearing on 1 April.

https://www.thelocal.de/20200402/latest-face-masks-in-public-could-help-to-reduce-spread-of-coronavirus-says-germanys-robert-koch-institute

https://muscateer.om/en/news/europe-updategermany-new-face-mask-rules-in-idZ2trbg==

=======================

Italy 6 April mandatory in some regions, some settings by 6 April, endorsed by national govt previously but uneven uptake.

Lombardy, Tuscany 6 April mandatory anywhere outdoors

https://www.ansa.it/english/news/2020/04/06/coronavirus-lombardy-makes-face-masks-compulsory_a852ffdb-a0dd-4c55-a725-e852c5a2fc43.html

https://uk.reuters.com/article/us-health-coronavirus-italy-masks/scramble-for-masks-as-italian-region-orders-coronavirus-cover-up-idUKKBN21O1Y0

https://www.thelocal.it/20200406/coronavirus-where-should-you-wear-a-face-mask-in-italy

========================
Lithuania 1 April (recommendation)

https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-use-face-masks-community.pdf

====================

Luxembourg 20 April 2020 (**compulsory in some situations** , where can't keep 2m apart)

https://uk.reuters.com/article/us-health-coronavirus-luxembourg/luxembourg-enforces-use-of-masks-as-lockdown-eases-idUKKBN2221W3

=====================
Norway 5 April (encouraged)

https://www.newsinenglish.no/2020/04/05/officials-change-their-minds-about-masks/

https://www.fhi.no/en/op/novel-coronavirus-facts-advice/facts-and-general-advice/hand-hygiene-cough-etiquette-face-masks-cleaning-and-laundry/

======================

Poland 16 April (mandatory, most public places)
https://www.aljazeera.com/news/2020/04/countries-wearing-face-masks-compulsory-200423094510867.html

=====================

Slovakia 14 March (recommendation, followed by requirement from about 1 April)

https://balkaninsight.com/2020/04/09/slovak-news-crews-hailed-for-covid-19-coverage/

https://www.npr.org/sections/coronavirus-live-updates/2020/04/01/825180019/in-big-adjustment-some-european-countries-push-for-residents-to-wear-masks

https://balkaninsight.com/2020/04/09/slovak-news-crews-hailed-for-covid-19-coverage/

https://www.ecdc.europa.eu/sites/default/files/documents/COVID-19-use-face-masks-community.pdf

=========================

Slovenia 29 March (**mandatory** in many places)

https://english.sta.si/2746850/slovenia-sticking-to-use-of-masks-in-indoor-public-places

=========================

Spain 13 April (**recommended** and sometimes freely given out, ie train stations, but not compulsory)

https://www.rtve.es/noticias/20200413/como-colocar-retirar-desechar-mascarillas-higienicas-para-evitar-contagio-coronavirus/2011879.shtml

https://www.thelocal.es/20200424/what-are-the-rules-for-wearing-a-protective-mask-in-spain

===================================

Countries with no predominant government recommendation or compulsion in place (as of noon 29.4)
Belgium, Croatia, Republic of Cyprus (south), Denmark, Finland, Greece, Hungary, Ireland, Latvia, Malta, Netherlands, Portugal (likely soon), Romania, Sweden, Switzerland, UK

Netherlands Masks are not a substitute for 1.5 m
https://www.dutchnews.nl/news/2020/04/dutch-stay-firm-on-face-masks-but-they-may-have-an-exit-strategy-role/

Hungary 27 April mandatory, commuters & shoppers but in Budapest only
https://www.themayor.eu/en/budapest-makes-masks-mandatory-for-shoppers-and-commuters

UK
Scotland's first minister recommended (did not mandate) that face coverings (not surgical grade masks) should be worn in all enclosed public spaces, from 28 April 2020. Because Scotland comprises 5.5 million (just 8.2% of the total UK population of approximately 67.9 million) we still (as of 29.4.20) treated the entirety of the UK as a country without an official endorsement of face coverings in our modelling.

https://www.gov.scot/publications/coronavirus-covid-19-public-use-of-face-coverings/

SUPPLEMENTARY MATERIAL 3: Variance Inflation Factors (VIF) for Analysis 1, run as linear models

Variance inflation factors	
mass_gathering_restrictions	4.915550
initial_business_closure	10.011556
education_facilities	9.490315
non_essential_services	6.253501
stay_home	3.514925
masks	1.432351
tests per million population as of 16 April 2020	1.295814

SUPPLEMENTARY MATERIAL 4: Collinearity Diagnostics

Variance inflation factors for Analysis 2 model fit as linear regression

Variable	VIF	1/VIF
Initial business closures	10.44	0.095815
Education closures	9.87	0.101302
Mass gatherings banned	6.52	0.153397
Non essential business closures	6.27	0.159552
Days elapsed into main epidemic	6.15	0.162476
Stay at home order	3.52	0.284366
Mask advisory/mandate imposed	1.47	0.680742
Tests per 1 million/population by 16 April	1.31	0.764803
Mean VIF	5.69	

SUPPLEMENTARY MATERIAL 4: Collinearity Diagnostics (continued)
Analysis 2 Model when fit as linear regression, yielded condition indices and variance-decomposition proportions

Dimension ->	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Condition indices ->	**1**	**2.61**	**3.56**	**5.08**	**6.04**	**9.37**	**11.44**	**14.71**	**15.9**
Model Parameter	**Variance proportions**								
Constant	0	0.03	0.01	0.01	0.31	0.21	0.14	0.26	0.03
Mass gathering restriction	0	0	0	0.06	0	0.02	0.49	0.13	0.3
Initial business closure	0	0	0	0	0.03	0.01	0.27	0.04	**0.65**
Education facilities closed	0	0	0	0.01	0.02	0.04	0.1	0.33	0.5
Non essential services closed	0	0	0	0.05	0.02	0.61	0.15	0.01	**0.15**
Stay at home order	0	0.02	0.04	0.2	0.23	0.48	0	0.01	0.03
Masks advisory or mandated	0	0.17	0.72	0.01	0.02	0.04	0	0.02	0.01
Days elapsed from start main epidemic	0	0	0	0.01	0	0	0	0.78	0.2
Tests/1 mln population to 16 April	0	0.09	0.07	0.2	0.15	0.26	0.19	0	0.03

SUPPLEMENTARY MATERIAL 4: Collinearity Diagnostics (continued)

<u>Principal components/correlation</u>

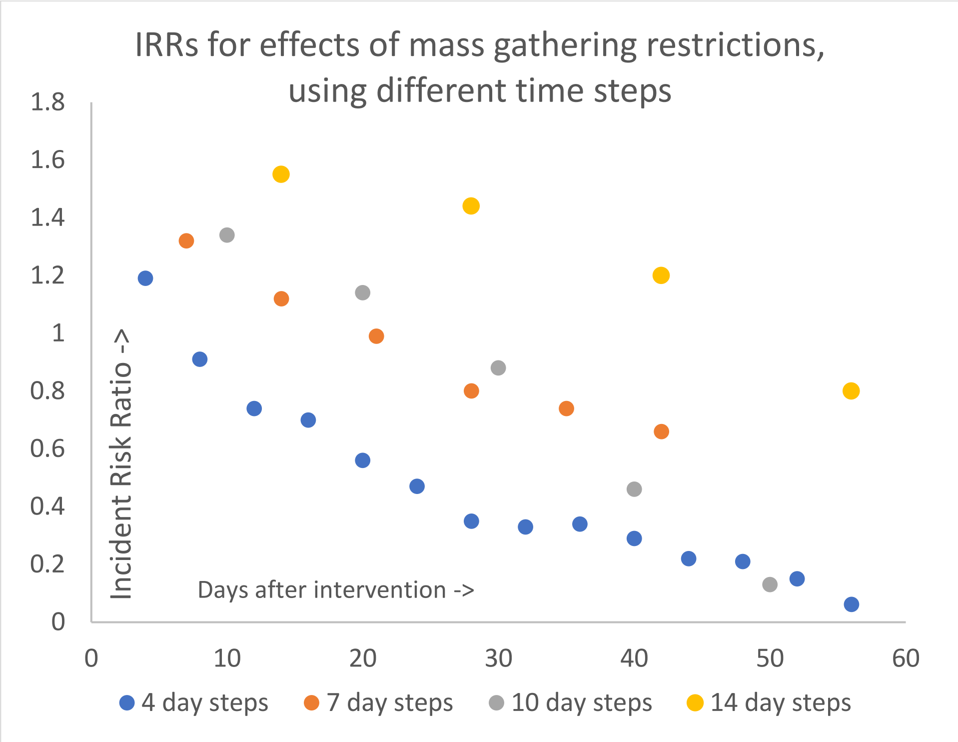
Number of observations = 1588

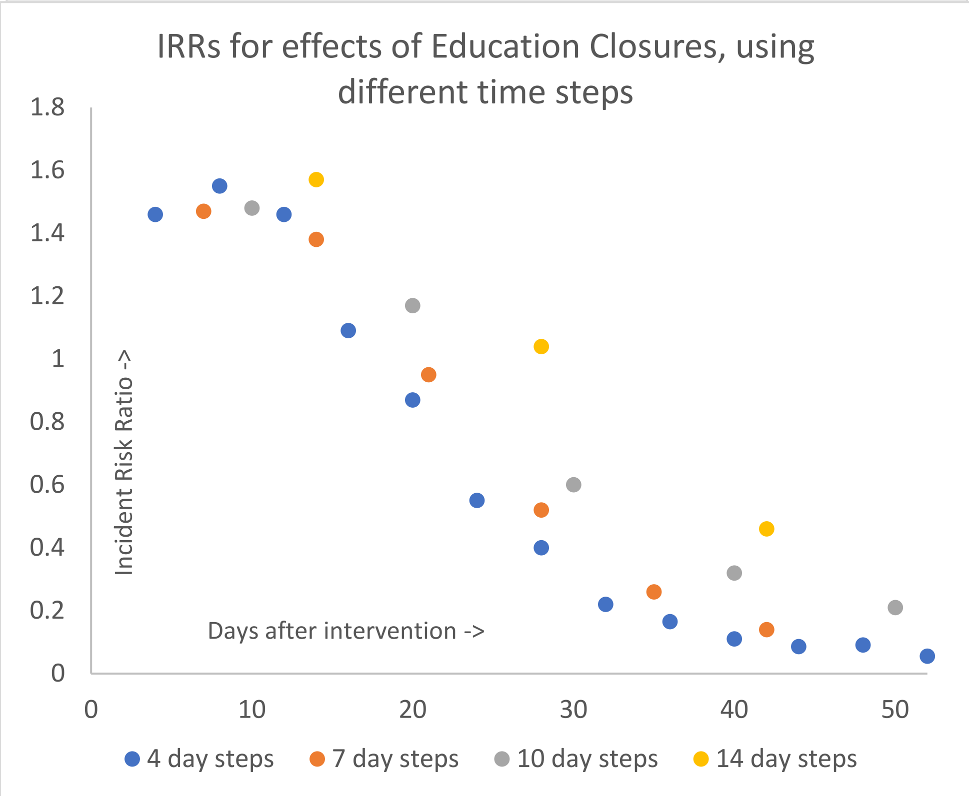
Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	5.2667	4.21318	0.6583	0.6583
Comp2	1.05353	0.302265	0.1317	0.79
Comp3	0.751261	0.26581	0.0939	0.8839
Comp4	0.485451	0.3267	0.0607	0.9446
Comp5	0.158751	0.041833	0.0198	0.9645
Comp6	0.116918	0.008604	0.0146	0.9791
Comp7	0.108315	0.049241	0.0135	0.9926
Comp8	0.059073	.	0.0074	1

SUPPLEMENTARY MATERIAL 5: Please see separate document for full regression model specifications and outputs. **Alternative results using different time step units for epidemic response, Analysis 2.** Tested variations are 4 days, 7 days, 10 days or 14 day units. The main manuscript describes the results when using 7 day time response periods, as there was little difference in the overall trends of a decline in risk ratios with time since interventions were imposed. Please see Supplementary Material 6 (below) for further detail showing between model comparisons.

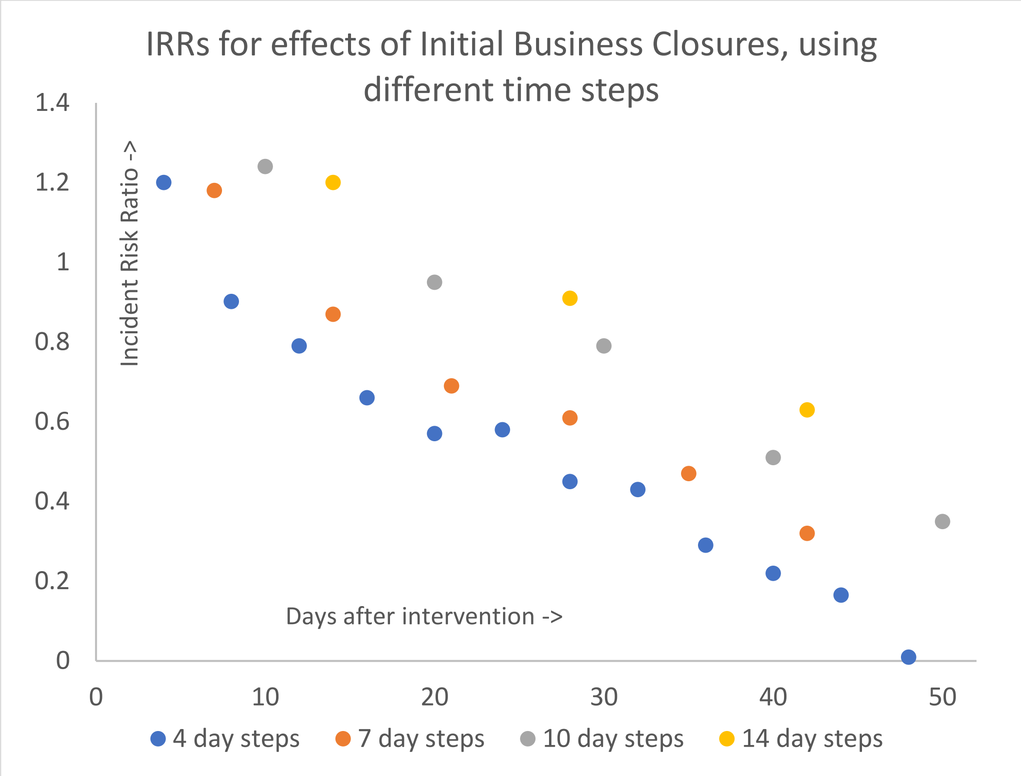
SUPPLEMENTARY MATERIAL 6 (following pages): Alternative results using different time step units for epidemic response, Analysis 2 : Comparisons of Incident Risk Ratios (IRRs) when models were generated using different time step response units: 4, 7, 10 or 14 days.

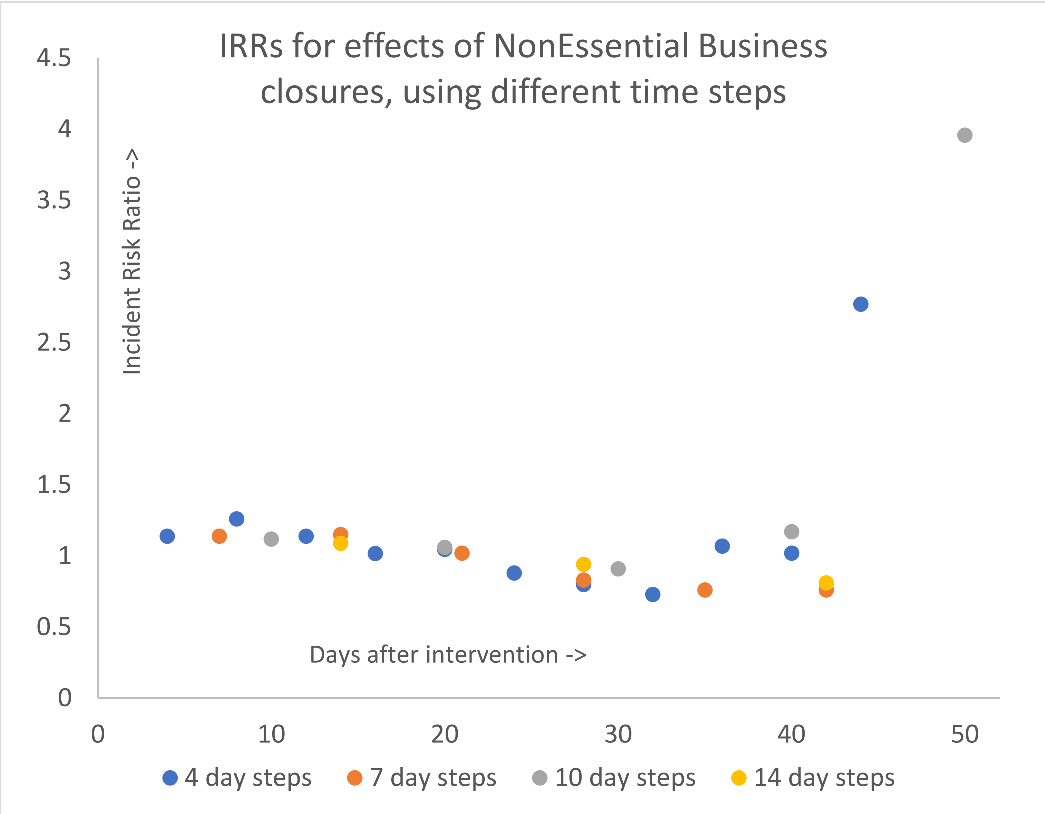
6A: IRRS for CASES



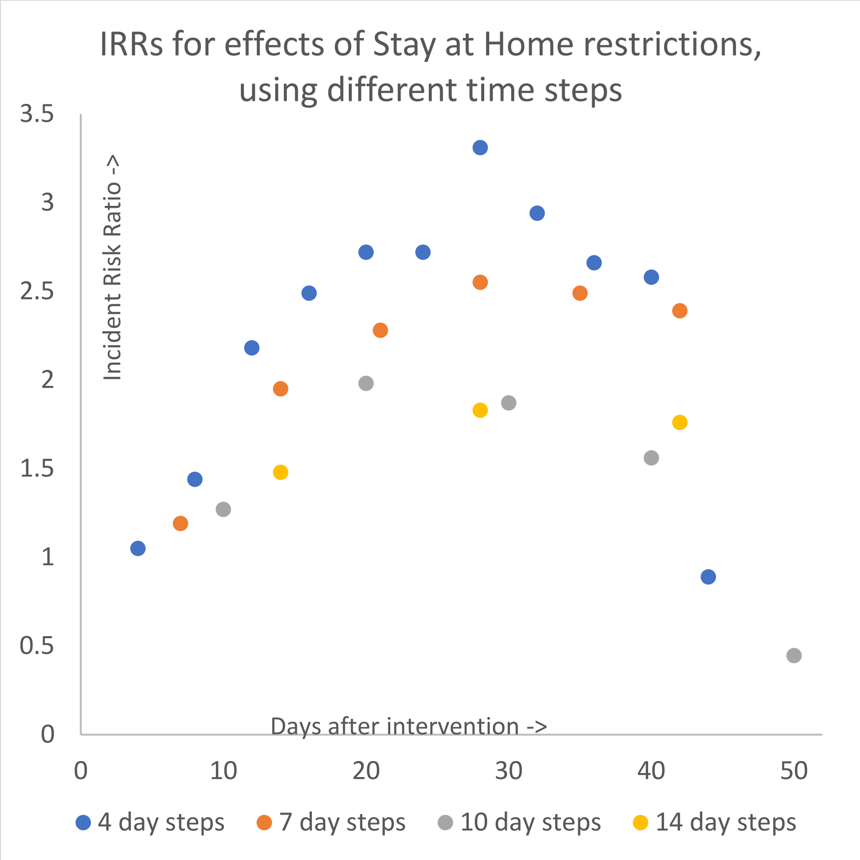


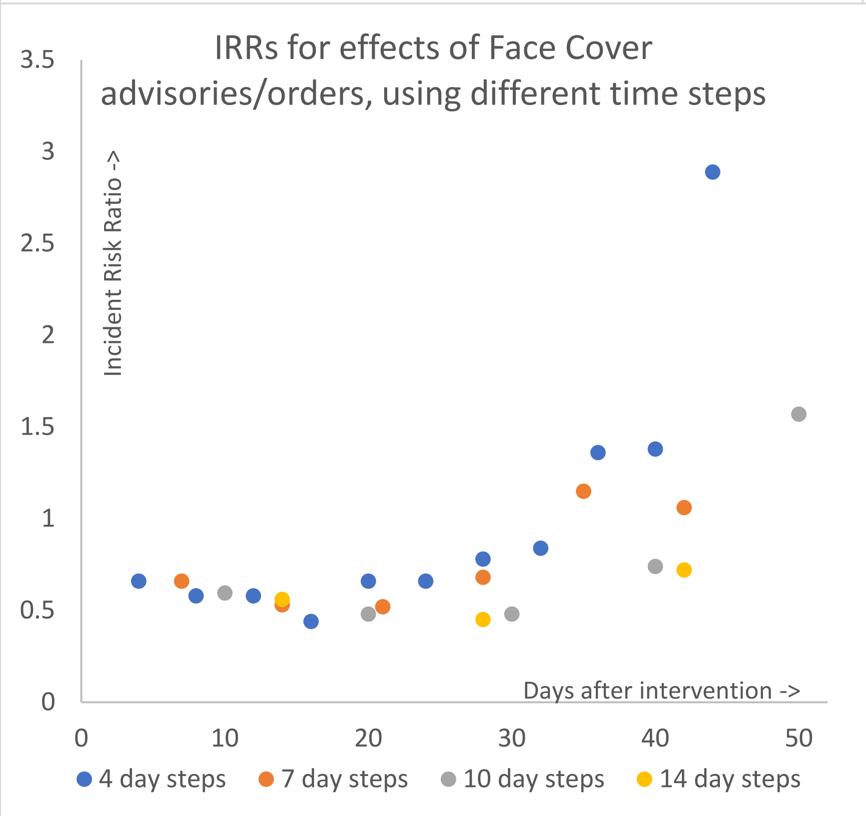
6A: IRRS for CASES (continued)



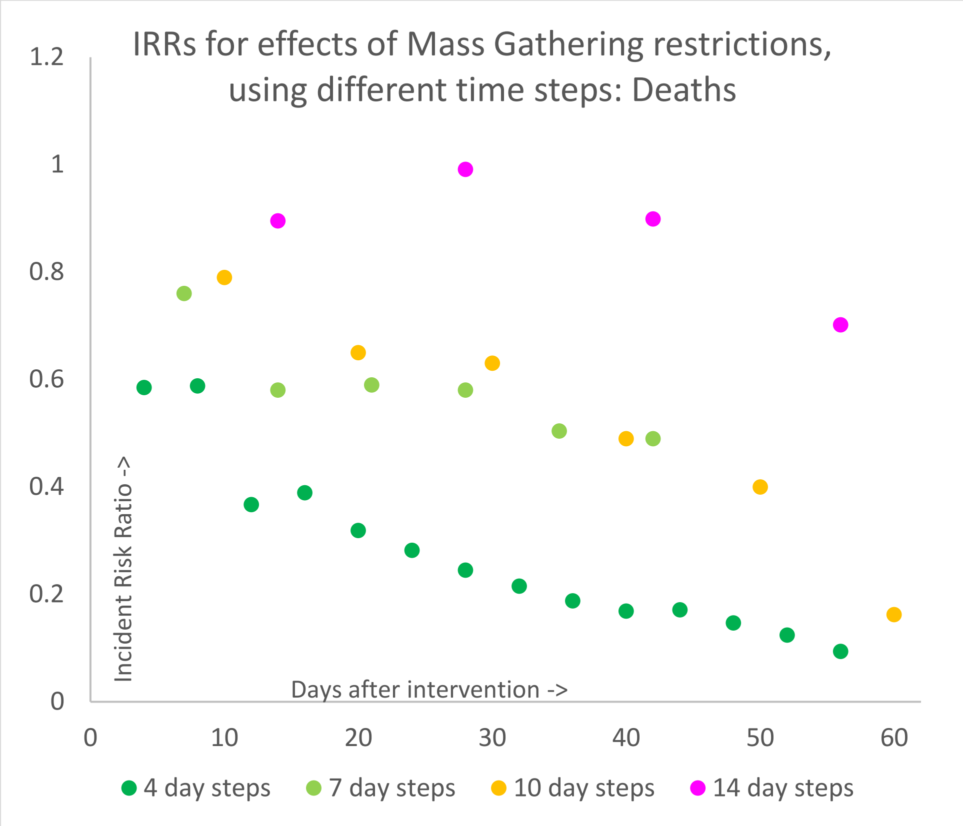


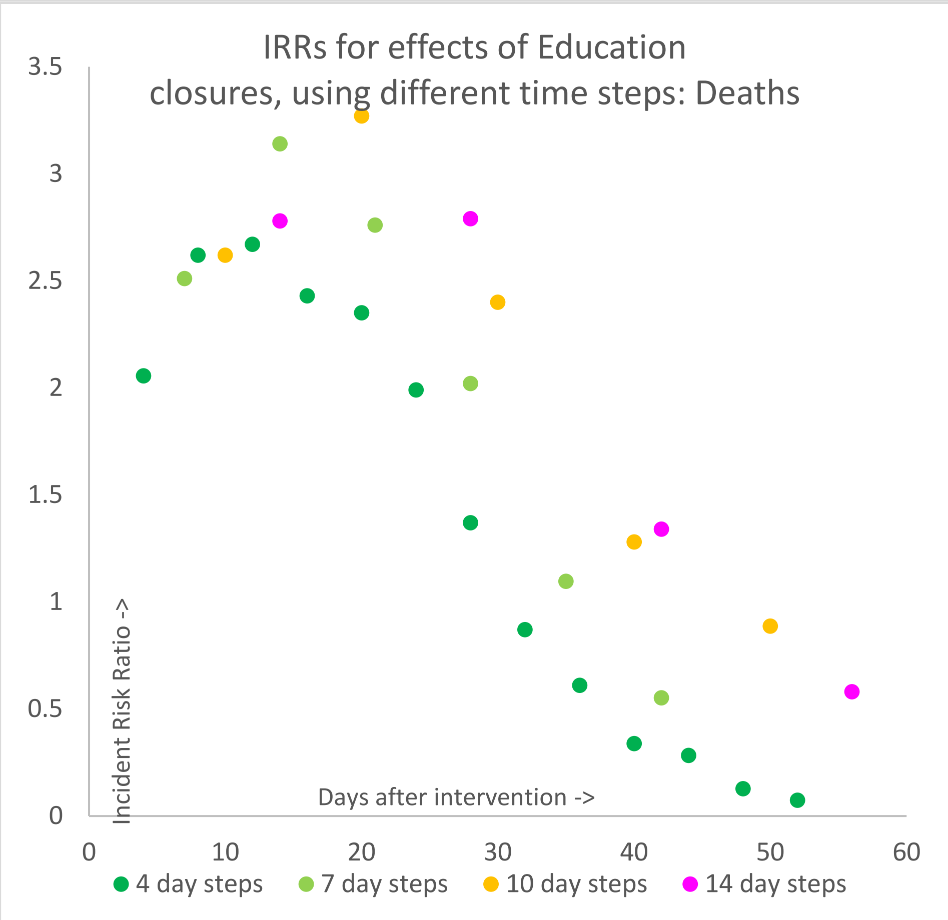
6A: IRRS for CASES (continued)



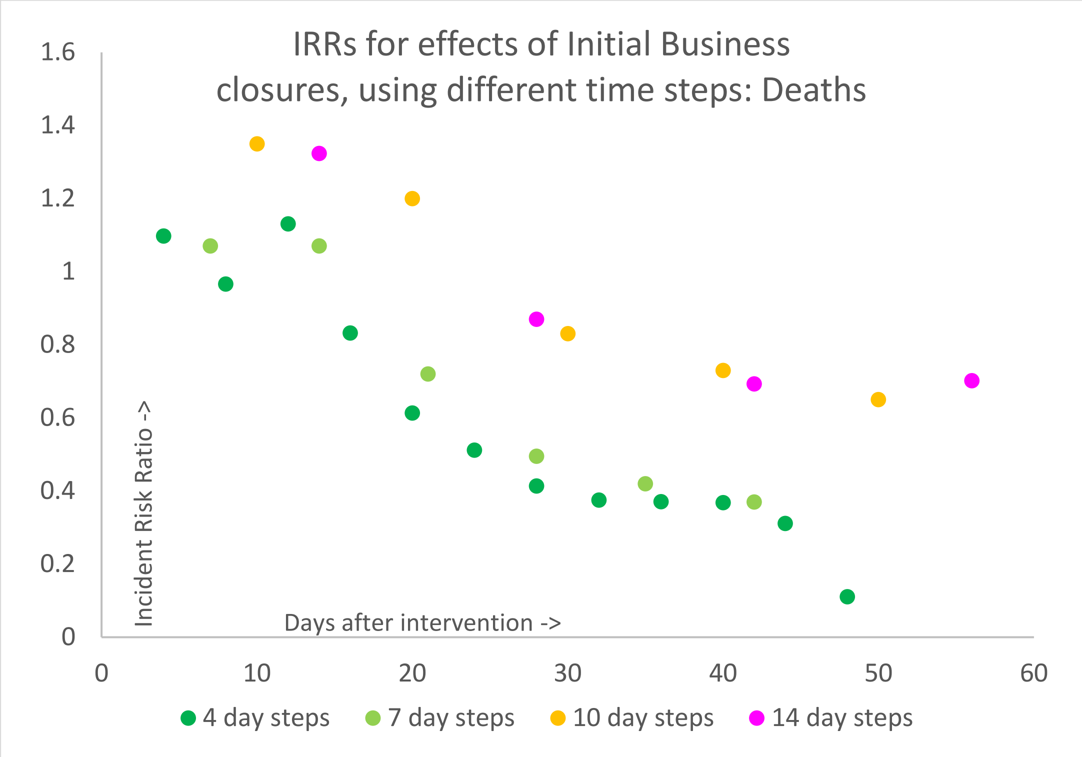


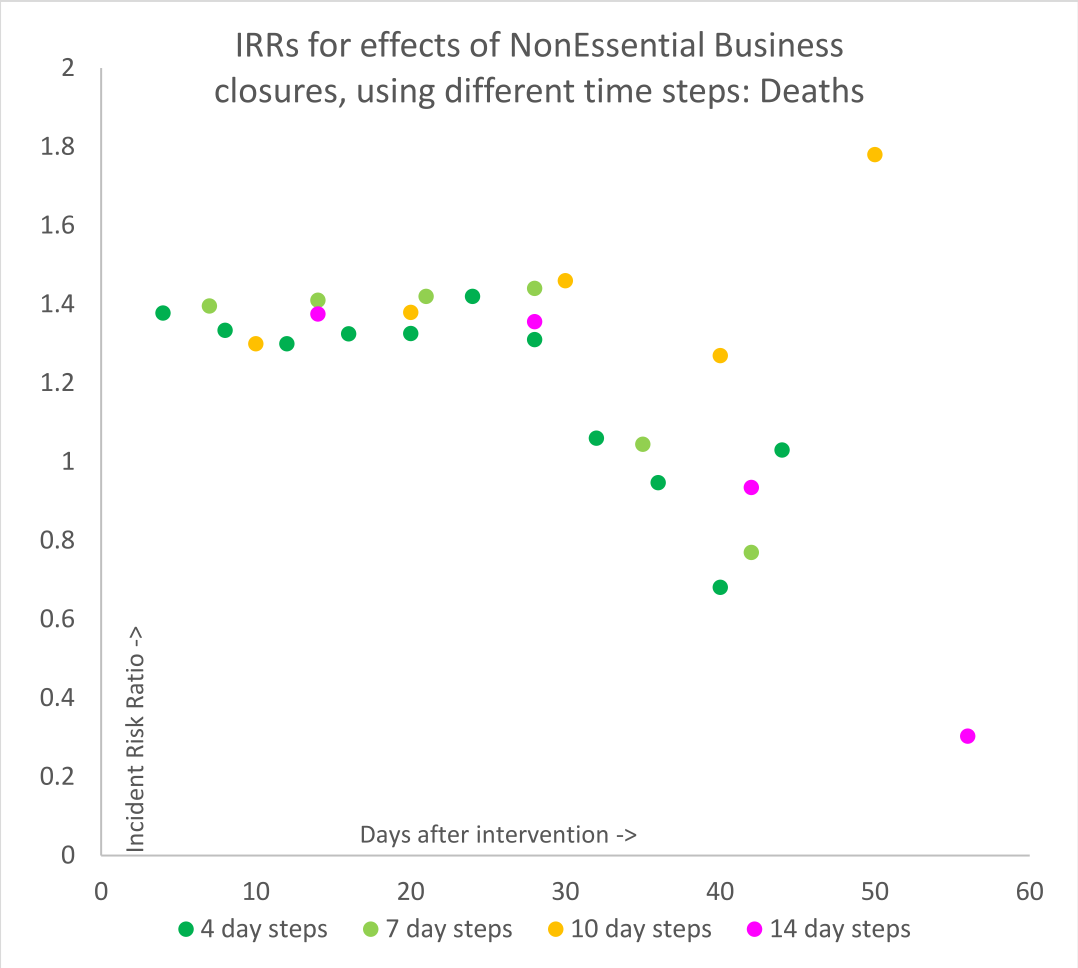
6B: IRRS for DEATHS





6B: IRRS for DEATHS (continued)





6B: IRRS for DEATHS (continued)