Liu et al. BMC Medicine (2021) 19:40
https://doi.org/10.1186/s12916-020-01872-8

BMC Medicine

RESEARCH ARTICLE Open Access

The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories

Yang Liu[1†], Christian Morgenstern[2†], James Kelly[2], Rachel Lowe[1,3], CMMID COVID-19 Working Group and Mark Jit[1*]

Abstract

Background: Non-pharmaceutical interventions (NPIs) are used to reduce transmission of SARS coronavirus 2 (SARS-CoV-2) that causes coronavirus disease 2019 (COVID-19). However, empirical evidence of the effectiveness of specific NPIs has been inconsistent. We assessed the effectiveness of NPIs around internal containment and closure, international travel restrictions, economic measures, and health system actions on SARS-CoV-2 transmission in 130 countries and territories.

Methods: We used panel (longitudinal) regression to estimate the effectiveness of 13 categories of NPIs in reducing SARS-CoV-2 transmission using data from January to June 2020. First, we examined the temporal association between NPIs using hierarchical cluster analyses. We then regressed the time-varying reproduction number (R_t) of COVID-19 against different NPIs. We examined different model specifications to account for the temporal lag between NPIs and changes in R_t, levels of NPI intensity, time-varying changes in NPI effect, and variable selection criteria. Results were interpreted taking into account both the range of model specifications and temporal clustering of NPIs.

Results: There was strong evidence for an association between two NPIs (school closure, internal movement restrictions) and reduced R_t. Another three NPIs (workplace closure, income support, and debt/contract relief) had strong evidence of effectiveness when ignoring their level of intensity, while two NPIs (public events cancellation, restriction on gatherings) had strong evidence of their effectiveness only when evaluating their implementation at maximum capacity (e.g. restrictions on 1000+ people gathering were not effective, restrictions on < 10 people gathering were). Evidence about the effectiveness of the remaining NPIs (stay-at-home requirements, public information campaigns, public transport closure, international travel controls, testing, contact tracing) was inconsistent and inconclusive. We found temporal clustering between many of the NPIs. Effect sizes varied depending on whether or not we included data after peak NPI intensity.

(Continued on next page)

* Correspondence: mark.jit@lshtm.ac.uk
†Yang Liu and Christian Morgenstern contributed equally to this work.
[1]Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, London, UK
Full list of author information is available at the end of the article

© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)
Conclusion: Understanding the impact that specific NPIs have had on SARS-CoV-2 transmission is complicated by temporal clustering, time-dependent variation in effects, and differences in NPI intensity. However, the effectiveness of school closure and internal movement restrictions appears robust across different model specifications, with some evidence that other NPIs may also be effective under particular conditions. This provides empirical evidence for the potential effectiveness of many, although not all, actions policy-makers are taking to respond to the COVID-19 pandemic.

Keywords: Non-pharmaceutical interventions, Policy evaluation, COVID-19, SARS-CoV-2, Public health intervention, Pandemic, Quantitative, Health impact assessment, Longitudinal analysis

Background

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The virus is easily transmissible between humans, with a basic reproduction number around 2–4 depending on the setting [1, 2]. To date, no vaccine or highly effective pharmaceutical treatment exists against COVID-19. Countries have used a range of non-pharmaceutical interventions (NPIs) such as testing suspected cases followed by isolation of confirmed cases and quarantine of their contacts, physical distancing measures such as schools and workplaces closures, income support for households affected by COVID-19 and associated interventions, and domestic and international travel restrictions [3]. These interventions aim to prevent infection introduction, contain outbreaks, and reduce peak epidemic size so that healthcare systems do not become overwhelmed. However, these interventions come at a cost. Testing and contact tracing require laboratory and public health resources to be successful at scale, government subsidies affect national budgets, while physical distancing disrupts economic activities and daily life [4]. Hence, the psychological, social, and economic cost of interventions needs to be balanced against their potential effectiveness in reducing SARS-CoV-2 spread.

Modelling studies suggest that travel restrictions [5, 6], contact tracing and quarantine [7, 8], and physical distancing [9, 10] may delay SARS-CoV-2 spread, based on assumptions about how they may change transmission between individuals in populations. However, the effectiveness of such interventions depends on factors such as societal compliance (e.g. the extent to which people reduce their daily contacts following government restrictions) that are difficult to prospectively measure. Empirical evidence about the effectiveness of specific policy interventions has been limited (see Additional file 1: Table S8 for a review) [11–37]. While several countries have seen disease incidence peak and fall [38], ascribing changes in transmission to particular interventions is difficult since countries tend to impose combinations of policy changes at different levels of stringency in close temporal sequence.

Several global databases of COVID-19-related policy interventions have been compiled [39]. Here, we used the regularly updated Oxford COVID-19 Government Response Tracker (OxCGRT) [3] and conducted panel analysis to understand the association between policy interventions and time-varying reproduction numbers (R_t), a measure of the rate of transmission of an infectious disease in a population. We also explore whether this relationship is modulated by definitions of policy interventions, temporal lags, and population characteristics in different countries.

Methods

Data on NPIs and R_t

Data on COVID-19-related NPI intensity from 1 January to 22 June 2020 was extracted on 5 July 2020 from version 5 of OxCGRT, based on the codebook version 2.2 (22 April 2020) [3]. This version contains publicly available information from 178 countries and territories on 18 NPI categories. We further divided these countries and territories into seven regions according to the World Bank classification [40]. Note that these 18 NPI categories are broad, so many specific policy interventions (e.g. facial covering mandates) are not independently coded in the database. See Additional file 1: Table S1 for further metadata.

From this database, we removed (i) "miscellaneous" policies as they contained no data at the time of our data extraction; (ii) "giving international support" and "investment in vaccines" policies as they did not on face validity have a causal pathway to influence local SARS-CoV-2 transmission within the timescale of the analysis; (iii) "fiscal measures" and "emergency investment in healthcare" policies as both the start and the duration of their effect is often unclear (e.g. the announcement of an investment may be implemented weeks later; funding that is allocated may be spent over a long time); and (iv) data after 22 June 2020 because > 10% of countries and territories have missing data after this date (see Additional file 1: Figure S1) [3]. Missing data fields on or before 22 June 2020 were imputed by (a) carrying forward or backwards the next or last non-missing observation when

missingness occurred at the two tails of the time-series or (b) linearly interpolating using non-missing observations when missingness does not occur at the two tails of the time series. We divided the remaining 13 policy interventions into four policy groups roughly consistent with the original database (Table 1).

Most NPIs in the database are measured on ordinal scales that capture intensity (e.g. 0 = no contact tracing; 1 = limited contact tracing; 2 = comprehensive contact tracing). Since the intervals between categories are not necessarily equally spaced, we converted NPI history into binary variables under two scenarios: (i) *any effort scenario*: all zero records were converted to 0, and non-zero records were converted to 1, and (ii) *maximum effort scenario*: all non-maximum records were converted to 0, and all records at maximum levels were converted to 1 (see Table 1).

Transmission of SARS-CoV-2 is routinely measured using the time-varying reproduction number (R_t), a metric which represents the mean number of secondary cases that arise from one index case. We used the median R_t estimates available through EpiForecasts [https://epiforecasts.io/], a publicly available repository of R_t estimates for many countries. The estimation process is based on reported incidence while accounting for a range of uncertainties surrounding the incubation period, the delays between symptom onset and reporting [41]. The underlying method has been detailed in Cori et al. [42]. In short, the transmission rate of an infectious disease is approximated by the ratio between new infections at time t and the infectious individuals at time $t - w$ where w is the associated time window. In EpiForecasts, a weekly time window is used. This measure is expected to fall when effective NPIs reduce the rate of SARS-CoV-2 transmission. Since the effects of some NPIs may take time to become evident, we explored a range of temporal lag effects between NPI implementation and R_t changes.

Between 1 January and 22 June 2020, data on NPIs and R_t are simultaneously available for 130 countries and territories, all of which are used in the panel analysis described below.

Understanding the temporal patterns

The effect of an NPI on R_t may vary over time as a result of the evolving epidemic dynamics (e.g. decreasing number of susceptibles) or time-varying factors such as public compliance (e.g. the proportion of shoppers wearing facial coverings after government mandates). To examine this effect, we split up the time series of NPIs and R_t values into two parts: before and after peak NPI intensity. This was a sensitivity analysis to examine the robustness of NPIs' effectiveness in reducing COVID-19 transmission over time.

We used OxCGRT's stringency index (SI), a combined metric of several behaviour-related NPI measures, to determine peak NPI intensity. We then fitted a Gaussian generalised additive model (GAM) with cubic splines, using SI as the response variable and date as the sole explanatory variable for each World Bank region (i.e. the predicted regional SI is informed by all stringency index time-series within it). The peak of the predicted SI splines for each region was then examined to derive an average peak across all the regions. We then constructed two time-series: (i) the *full* time series and (ii) the *truncated* time series up to the time of peak SI.

We examined temporal clustering among different NPIs to identify potential structural confounding. If two effective NPIs are temporally clustered, one may be removed due to multicollinearity, which should not by any means be interpreted as that NPI being ineffective. Similarly, if one effective and one ineffective NPI are temporally clustered, the statistical association between the effective NPI and reductions in R_t may create a statistical artefact whereby the ineffective NPI may also appear to be associated with reductions in R_t. Either way, the existence of temporal clustering could cause misinterpretation of the regression results unless it is accounted for.

To investigate the temporal clustering patterns, we conducted hierarchical cluster analysis using Ward's

Table 1 Thirteen types of NPIs from OxCGRT, their general categorisations, and the coding schema used in our analysis to quantify their intensity

NPI groups	Specific NPIs	Coding schema
Internal containment and closure	School closure; workplace closure; cancellation of public events; limits on gathering sizes; closure of public transport; stay-at-home requirement; internal movement restriction	*Any effort scenario*: NPIs are binary variables, considered "present" as long as any (non-zero) effort is made. *Maximum effort scenario*: NPIs are binary variables, considered "present" only if the maximum effort is made. For example, an intervention X has levels 0–3. A record at level 2 is converted to 1 under *any effort* and 0 under *maximum effort* scenarios.
International travel restrictions	International movement restriction	
Economic policies	Income support; debt/contract relief for households	
Health systems policies	Public information campaign; testing policy; contact tracing	

method [43], which minimises within-cluster variance, under the *any effort scenario* and the *maximum effort scenario*. The inputs of the hierarchical clustering process were 13 vectors (one for each NPI under consideration), with each vector element corresponding the NPI status aligned by a unique time and location. Euclidean distance was used as the distance function between each pair of NPIs, using all available data (i.e. the full time-series for each NPI). We chose this method to compare the entire time-series of the NPIs, without having to select time-series summary metrics (e.g. the timing when an NPI was implemented) a priori. We then used multi-scale bootstrapping ($n = 10{,}000$) to test the statistical significance of the identified clusters, defined using approximate unbiased p values less than 0.05 [44]. The complete implementation of this method can be found in the GitHub repository at [https://github.com/yangclaraliu/COVID19_NPIs_vs_Rt].

Panel analyses

We used panel (or longitudinal) regression to study the association between NPI intensity and R_t, treating the time-series of NPI intensity and R_t in each country as observations of an individual in a panel. We used a linear fixed effects model:

$$R_{it} = \alpha_i + \sum \beta X_{it} + u_{it}$$

where R_{it} is the time-varying reproduction number of location i at time t, α_i is a location-specific intercept (assumed to remain constant over the timescale of the analysis), βX_{it} represents the 13 NPIs and their corresponding coefficients, and u_{it} is the error term. The decision to use a fixed-effects model with individual intercept (as opposed to a random-effects model) was based on the results of the Durbin-Wu-Hausman test [45]. In other words, there is insufficient evidence to support a random effect model based on global data, and the effects of each NPI on R_t can be characterised by fixed estimators.

We investigated the appropriate temporal lag between NPI intensity and R_t. To do this, we calculated the deviance (natural logarithm of the sum of squared residuals divided by the number of data points) assuming errors are normally distributed for temporal lags of 1 to 21 days. Smaller deviances indicate temporal lags that provide better model fits. A temporal lag of k days regresses on R_t a particular day against NPIs implemented k days before (i.e. $X_{i(t-k)}$). This analysis was carried out at both the regional and global levels. Data from North America and South Asia were excluded from region-specific temporal lag analyses due to small sample sizes.

Stepwise backwards variable selection based on Akaike and Bayesian Information Criterion (AIC or BIC) was then used to choose the most parsimonious model. Beginning with the full model (13 independent variables, one for each NPI), independent variables were removed one at a time sequentially. We also validate our results using univariable analyses and a forward variable selection algorithm.

Statistical interpretation

For both the *any effort* and the *maximum effort scenarios*, we examined a range of model specifications including (i) different variable selection criteria: AIC and BIC, (ii) different temporal lags between the timing of NPIs and changes in R_t (selected based on deviance from the analysis of temporal patterns), and (iii) different time series lengths: one ending on 22 June 2020 and the other truncated to 13 April 2020, when NPI intensity peaked (on average). We then defined categories of "evidence strength" behind each association according to Table 2. For example, if an NPI has significantly negative effects on R_t in all but one model set-up (i.e. one of model selection criteria, temporal lags, and time-series length mentioned above), that NPI is considered to have moderate strength evidence, as long as no other NPI in the same temporal cluster has significantly positive effects on R_t. Allocating each NPI to an evidence category was done independently by two authors (YL and MJ), with differences resolved by discussion.

Software

All analyses were conducted using R version 4.0.0 [46], with packages "plm" and "pvclust" [47, 48]. Code is available at https://github.com/yangclaraliu/COVID19_NPIs_vs_Rt.

Results

Trends in NPI intensity

Temporal trends in COVID-19-related NPI intensity measured using the OxCGRT SI are relatively consistent across regions (Fig. 1). Following the initial imposition of NPIs in China, almost all regions experienced an initial increase in policy stringency in early February 2020. The East Asia and Pacific region had the highest SI up to mid-March, but by April had the lowest SI. From March, other regions registered rapid increases in their stringency indices. The stringency index peaked in mid-April for all regions, and so 13 April 2020 was used as the time of peak NPI intensity (see Additional file 1: Table S2). All regions and nearly all countries had a higher stringency index in June compared to February.

Figure 2 shows how the intensity of specific NPI groups varies in each region relative to the time of peak intensity. Under both *any* and *maximum effort scenarios*, "Health System Policies" was the first NPI group to increase across all regions. It was also the most commonly

Table 2 Expert interpretation of evidence from the statistical associations of each NPI with reductions in R_t

Evidence strength	Effect estimates	Temporal cluster
Strong	Selected and significant with intended effect signs (i.e. negative) regardless of model specifications (i.e. variable selection criteria, temporal lags, and time-series lengths).	Not in a temporal cluster with any NPI with significantly positive effect estimates.
Moderate	Selected and significant with intended effect sign (i.e. negative) in two of three model specification dimensions (i.e. variable selection criteria, temporal lags, and time-series lengths), and non-selected or non-significant in the remaining dimension.*	
Weak	Not strong or moderate	

*For the moderate category, all the model specifications that were non-significant or non-selected were examined to see if they had a value in common across one of the three criteria, e.g. all of them had a lag of 10 days. Significance was assessed using $\alpha = 0.05$

used NPI group. This was followed by "Internal Containment and Closures" and then "Economic Policies", although "International Travel Restrictions" sometimes came before "Internal Containment and Closures". NPI intensity increased only as the first case was detected in each region, except for sub-Saharan Africa where many countries took action before the first detected case. While the stringency index has decreased across all regions (Fig. 1), the decreasing trends were not apparent in most NPI groups apart from International Travel Restrictions (Fig. 2).

Hierarchical cluster analysis shows that, given the *any effort scenario*, all the NPIs are contained in two significant temporal clusters (Fig. 3). These temporal clusters align well with the broad categorisations defined in the OxCGRT, i.e. countries tend to start implementing the same categories of NPI simultaneously. However, under the *maximum effort scenario*, there are three significant temporal clusters and several NPIs are not in any cluster, i.e. countries reach their maximum level of intensity for NPIs at very different times. Specific cluster assignments are also presented in Additional file 1: Table S2–3. Direct visual representation of the association between dates when pairs of NPIs were implemented and lifted can be found in Additional file 1: Figure S4–5.

Panel analyses

We examined the goodness-of-fit (based on deviance) of the panel regression model in all scenarios both at the regional and global level to identify the most appropriate temporal lag (see Additional file 1: Figure S4–5) [3]. For both the full and *truncated* time series (ending on 22 June and 13 April 2020, respectively), we identified temporal lags to be longest in East Asia and Pacific (between 5 and 10 days), followed by Europe and Central Asia (approximately 5 days), and the shortest in Latin America and the Caribbean (below 5 days) (see Additional file 1: Figure S4–5) [3]. The results from the Middle East and North Africa and sub-Saharan were not consistent

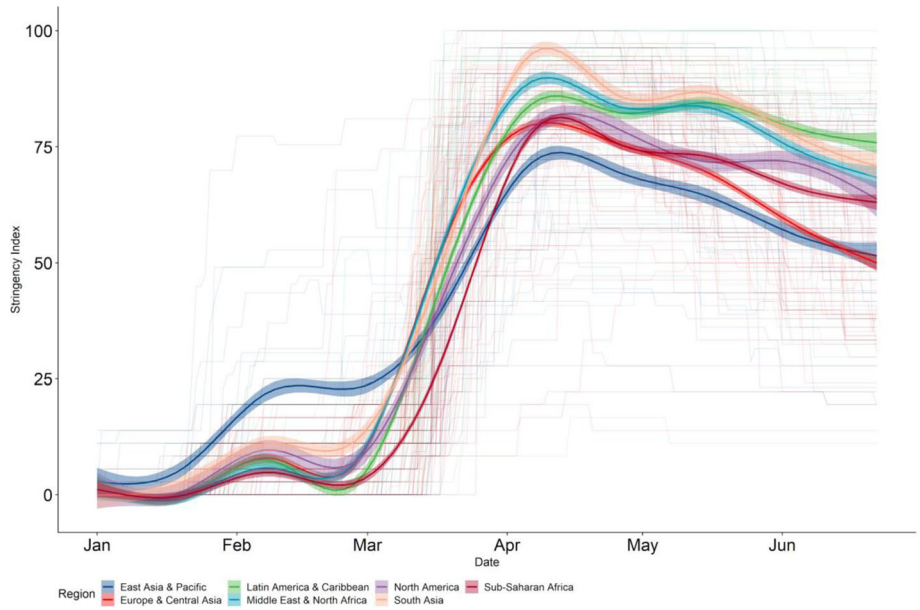


Fig. 1 Temporal changes in NPI stringency index (range = 0-100) by region. Countries with available data are assigned corresponding geographical regions based on the World Bank classification.

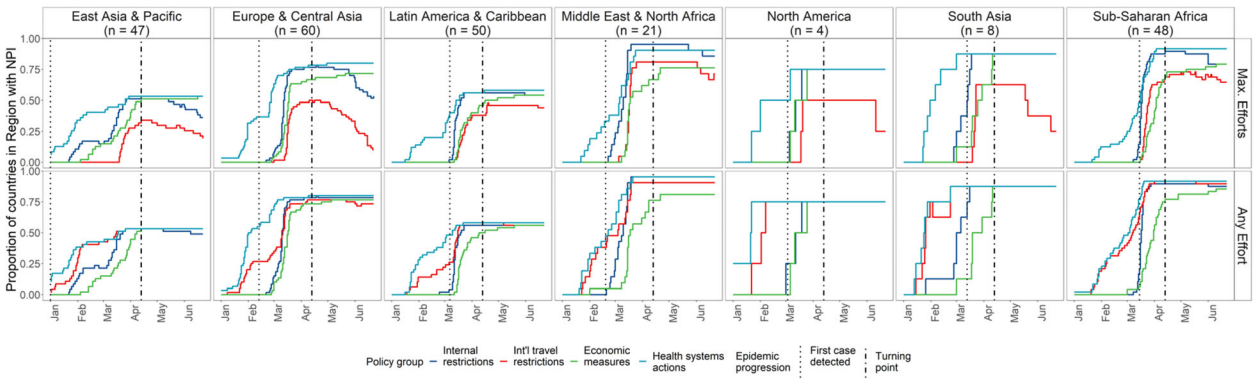


Fig. 2 The proportions of countries implementing NPIs in each group by region. Colours indicate different NPI groups: dark blue—internal containment and closure, red—international travel restrictions, green—economic measures, and light blue—health system policies; *n* values represent the total number of countries in each World Bank region; the *y*-axis shows the proportion of countries implementing NPIs per category, e.g. one means every country in the region is implementing some measures from the NPI group. The turning point of 13 April is shown as a dotted-and-dashed vertical line.

between scenarios, and there was no clear indication of the most appropriate temporal lag when countries were all combined in a global analysis. Due to the observed heterogeneity in the temporal lags, we examined three different lag values (1, 5, and 10 days) in the regression analyses for both *full* and *truncated* time-series.

The NPIs in the models selected based on AIC and BIC are shown in Fig. 4. Although we present the backward variable selection process in the main text under the assumption that all NPIs may explain variation in R_t, the final models were unaltered when a forward variable selection algorithm was used. Effect validation based on univariable panel analyses can be found in Additional file 1: Figure S6–7.

Under the *any effort scenario*, the most consistently excluded NPIs were contact tracing, restrictions on gatherings, and international travel restrictions. Public information campaigns and testing policies were excluded using the *truncated* but not the *full* time-series; stay-at-home requirements were excluded using the *full* but not the *truncated* time-series. Under the *maximum effort scenario*, the most consistently excluded NPIs were contact tracing, international travel restrictions, and closure of public transportation. Public information campaigns were excluded by one model using the *truncated* time series but were always included by models using the *full* time-series. NPIs may be excluded from models either because (i) they do not affect R_t or (ii) their effects were

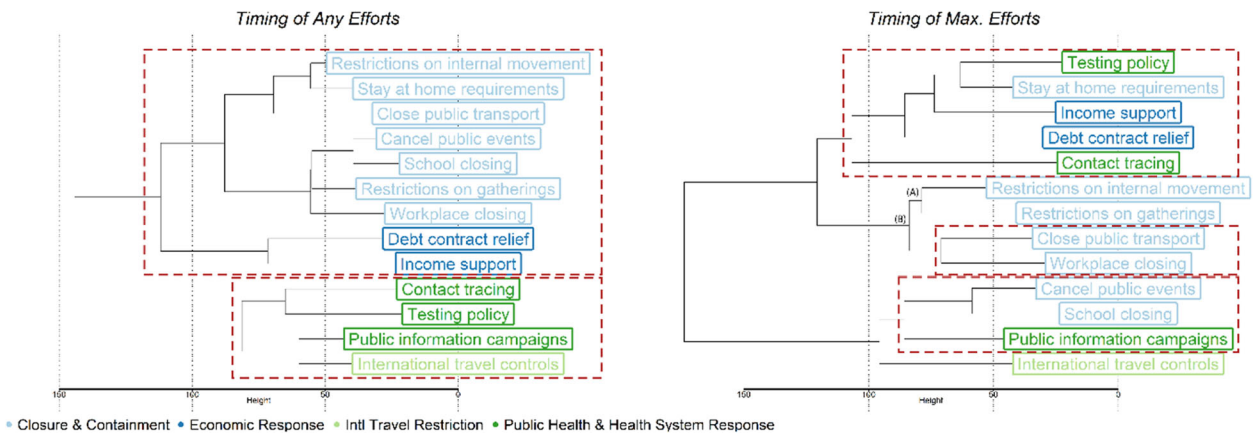


Fig. 3 Dendrogram from the hierarchical cluster analysis of NPI time-series by scenario. The height of the node connecting two NPIs on the dendrogram represents the degree of similarity between their time-series. For example, under the *Maximum Effort Scenario*, the time-series of "restrictions on internal movement" is more similar to that of "restrictions on gatherings" (linked at point A), compared to that of "close public transport" (linked at point B). The hierarchical clustering analysis relies on Ward's method and Euclidean distances. The colour of the text boxes corresponds to the group each NPI is in; red dashed boxes indicate statistically significantly temporal clusters identified through bootstrapping.

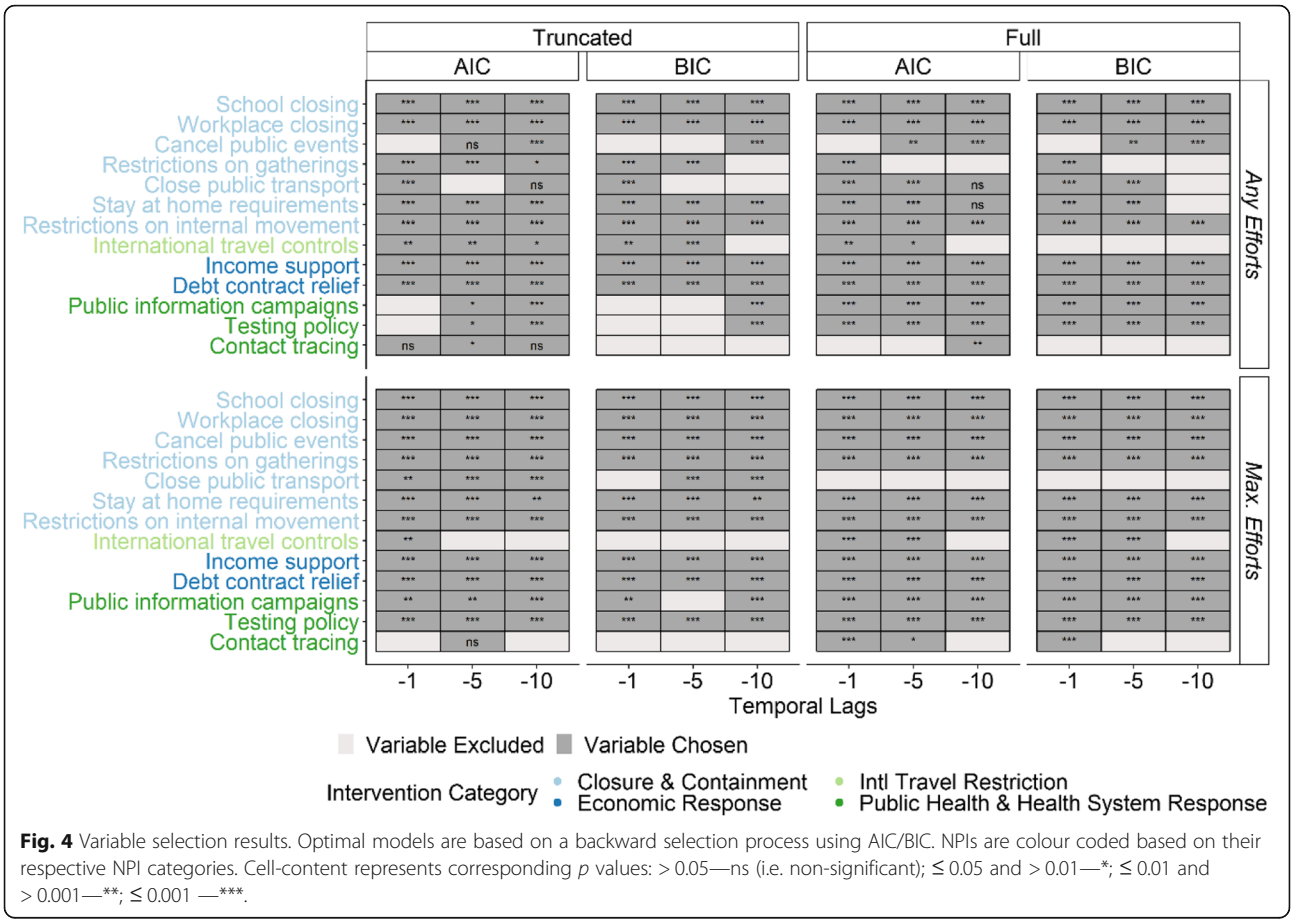


Fig. 4 Variable selection results. Optimal models are based on a backward selection process using AIC/BIC. NPIs are colour coded based on their respective NPI categories. Cell-content represents corresponding *p* values: > 0.05—ns (i.e. non-significant); ≤ 0.05 and > 0.01—*; ≤ 0.01 and > 0.001—**; ≤ 0.001 —***.

fully captured by other NPIs in the same temporal clusters, and thus, they were removed by the variable selection process.

Effect size estimates for the selected models in Fig. 4 are shown in Fig. 5. A few NPIs have statistically significantly positive effects: "closure of public transport", "stay-at-home requirements", and "contact tracing". These results indicate that the NPIs are associated with increased R_t. While it is not inconceivable that some NPIs may be transiently associated with increased

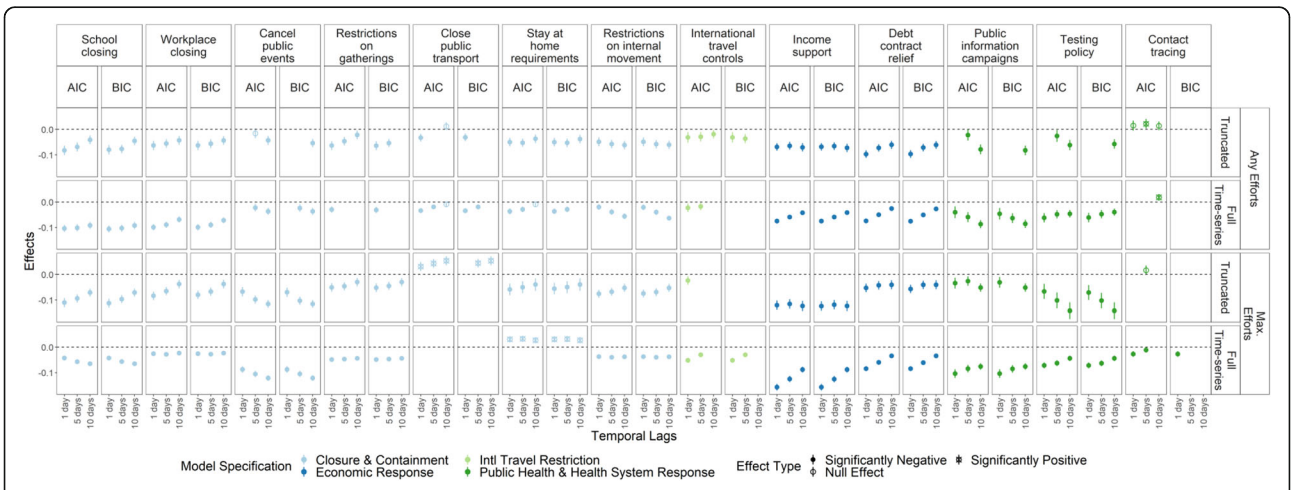


Fig. 5 Effect sizes for each NPI from the selected models. Points and lines indicate mean and 95% confidence intervals.

R_t (e.g. increased testing efforts may be associated with increased R_t because they result in better detection of COVID-19 cases), variables with positive effects are likely capturing residual non-random errors for other NPIs in the same temporal cluster. Hence, they are likely biasing effect size estimates of other temporal cluster members, likely away from the null hypotheses. Hence, NPIs that are temporally related need to be interpreted within the context of the respective clusters rather than as individual measures (see also Fig. 3).

We calculated the mean absolute errors (MAE) of the final models for each country. We found that the largest residuals (i.e. worst fits) were observed in sub-Saharan Africa (highest in Zambia and Zimbabwe), East Asia and Pacific (highest in Mongolia and China), and the Middle East North Africa regions (highest in Palestine and Djibouti); the lowest residuals (i.e. best fits) were observed in sub-Saharan Africa (lowest in Namibia and Mauritania) and Latin America and the Caribbean (lowest in Colombia and El Salvador). More details can be found in Additional file 1: Table S5–6.

Interpretation

Of the 13 NPIs in the OxCGRT, we found strong evidence for the association between two of them (school closure, internal movement restrictions) and R_t, under both *any effort* and *maximum effort* scenarios. Another three NPIs (workplace closure, income support, and debt/contract relief) had strong evidence for an association under the *any effort* scenario only, meaning that the reductions in R_t were associated with the initiation of these interventions, with no evidence of greater effect as they were intensified. There was strong evidence for two other NPIs (public events cancellation, restriction on gathering) under the *maximum effort* scenario only, meaning that a reduction in R_t was only evident when the NPIs reached their maximum intensity.

In some cases, the sequential order in which NPIs are implemented may make it more or less likely that particular NPIs capture the effects of other NPIs or they may have interactive effects on each other (e.g. one NPI may boost or reduce the effect of a subsequent NPI). For example, the back-sampling process used in Abbott et al. [41] may attribute true R_t reduction to NPIs occurring later. Thus, we verify NPIs rated as being supported by strong statistical evidence by checking their sequential ordering in COVID-19 response strategies. Most of these NPIs were not implemented particularly early or late in the sequence of NPIs (Additional file 1: Figure S8–9). Complete school closure and mandatory public events cancellations are moderately left-skewed, indicating that they tend to occur first. Some (non-maximum) levels of income support and debt/contract relief are right-skewed, making it possible that their observed effects are statistical artefacts or are dependent on the imposition of earlier NPIs.

Evidence for the other NPIs was mixed. Stay-at-home requirements had moderate evidence under the *any effort* scenario, while public information campaigns had moderate evidence under the *maximum effort* scenario. Among all NPIs, some (non-maximum) levels of stay-at-home requirements tended to occur later in the overall COVID-19 response strategy. The remaining four (public transport closure, international travel controls, testing, contact tracing) had only weak evidence for an association with R_t. Detailed interpretation of the statistics, through which these conclusions were reached, is presented in Additional file 1: Table S7. Similar methods were applied to the original raw data, without converting to *any* or *maximum effort scenarios*. However, as no statistical conclusion can be reasonably drawn, we only show the results in Additional file 1: Figure S10–11.

We observed variability in effect estimates due to differences in time-series and temporal lags used (Fig. 5). For example, the effect sizes of internal movement restrictions are smaller using the truncated time-series compared to using the full time-series. This suggests that general adherence to movement restrictions may have decreased over time. However, this variability may also be explained by the fact that full-time series include more observations. The effect sizes of public events cancellation are higher for longer temporal lags—indicating their impact on R_t may be delayed. These hypotheses need further validation using empirical evidence.

Discussion

Our study used panel regression to examine the temporal association between NPIs that countries introduced in response to the COVID-19 pandemic, and its rate of transmission in populations, represented by R_t. We explored how the association is modified according to the following model specifications: (i) level of NPI intensity (i.e. *any* vs *maximum scenarios*), (ii) model selection criteria (i.e. AIC vs BIC), (iii) varying lag effects, and (iv) different time-series lengths (i.e. truncated vs. full time-series).

We found the strength of evidence behind an association between NPIs and R_t depended on these model specifications. Only two NPIs (school closure, internal movement restrictions) showed unequivocal evidence of being associated with a decrease in R_t regardless of the assumptions made. Whether schools should stay closed has attracted debate. Keeping schools closed could potentially hurt children's educational development and general well-being. Resuming schools, on the other hand, may increase COVID-19 transmission risks for both students and teachers. Our findings are consistent with much existing literature—although school closures cannot single-

handedly suppress an outbreak, they are generally effective in terms of reducing transmission [49, 50].

We found evidence that internal movement restrictions reduced R_t, but no evidence of a similar effect for international travel restrictions. The latter is consistent with Russell et al., which shows international movement restrictions have a limited impact on the epidemic dynamics of COVID-19 for most countries [51]. This difference may be explained by the types of movement interrupted—internal movement restrictions interrupt trips of all lengths whereas international movement restrictions only disrupt longer trips, which are much less common. Additionally, internal and international movement restrictions were likely used in different epidemic contexts—internal movement restrictions tend to be used more often to prevent outbreaks from escalating whereas international travel restrictions make more sense in preventing infection introduction [52]. The latter effect is not well represented in our data since R_t can only be estimated in settings with existing COVID-19 outbreaks (i.e. after introduction).

There are differences in the strength and direction of the effects of some NPIs (such as public transport closure and stay-at-home requirements) depending on whether the whole time series of data was used, or only data up to the date of peak NPI stringency (13 April 2020). This may indicate that these NPIs might have different effects at the start of the pandemic compared to later on, so when the NPIs were removed (likely after the peak), R_t did not return to its original level before the introduction of the NPIs.

The best-fitting models also support a considerable delay between NPIs and their effect on transmission. This delay is about a week on average but differs widely between regions. It could reflect delays between policies being put in place and actual behaviour change. It could also reflect delays in reporting, although these are explicitly accounted for in the R_t estimation in EpiForecasts—the same onset-to-delay distribution is applied in all countries [41] and hence may not reflect differences between settings. Delays of up to 3 weeks between policy changes and changes in reported cases have been documented [53].

We were not able to find evidence that supports the effectiveness of contact tracing and testing policies. This may be because both contact tracing and testing policies could lead to more cases being reported, as well as interrupting onward transmission, so the overall effect is the combination of these two opposing effects. While calculating the R_t, EpiForecasts does not explicitly account for changes in reporting rate [41]. Another potential explanation is the way NPIs are reported in the OxCGRT, which largely relies on publicly available data sources, such as news articles. Contact tracing and test policies are both well-known public health intervention tools and have minimal impacts on the lives of those who are not potentially infected. Thus, they may be less likely to receive media coverage, compared to more disruptive NPIs such as workplace closures.

We focused our discussion on the direction and relative magnitude of the estimated effect of different NPIs, within the context of their temporal clusters during the on-going COVID-19 pandemic. The actual values of NPI-specific effect sizes, which were found to be greater for "School Closures" and "Workplace Closures" under the *any effort scenario* and for "Cancellation of Public Events" and "Income Support" under the *maximum effort scenario*, should be interpreted with caution. Given the statistical approach and the ecological design of the study, these numerical values are difficult to interpret due to structural confounding. For example, when a temporal cluster was effective in reducing R_t, we were not able to confidently attribute the effects to particular NPIs within the cluster. As the pandemic progresses, data on more diverse NPI implementation profiles and outcomes may become available, enabling more precise determination of effect sizes.

Many other papers have explored the impact of physical distancing measures on SARS-CoV-2 transmission. Prospective mechanistic transmission models have explicitly modelled contacts relevant to viral transmission between individuals in different subgroups (e.g. ages), as well as the impact that NPIs may have on these contacts. Such studies mainly use data from a single location only such as Wuhan [9], Hong Kong [54], the USA [55], and the UK [50]. They suggest that physical distancing interventions can have a large impact on transmission. While the impact of income-related interventions has been less well studied, country reports suggest that they often play an important role in ensuring adherence to distancing measures [56].

Another group of studies have used empirical data to retrospectively examine whether NPIs have been effective in reducing transmission, using either statistical methods or mechanistic epidemiological models. Many such studies look at single interventions such as travel restrictions [25] or "lockdowns" [22, 27]. Therefore, they are less useful to policy-makers wanting to establish which of a basket of NPIs are most effective.

Only a small number have looked at multiple interventions across multiple countries (see Additional File 1: Table S8 for a review). These relate NPIs from databases to proxies of transmission such as R_t estimated from cases and/or deaths [18, 21, 37], or the rate of change in cases directly [13, 16, 57]. Our work demonstrates the major challenges that all such studies (including ours) face—NPI introductions are highly correlated in time, so it is difficult to independently identify the effect of each

NPI due to structural confounding. A few studies partially account for this using techniques such as examining whether the number of NPIs that had already been implemented affects the impact of subsequent NPIs [11] or excluding statistically non-significant variables after all NPIs are included initially [12].

Our study extends previous work to address this problem in several ways. Firstly, we use data across a larger number of countries and territories and longer time series (January–June 2020), enhancing the power to detect independent effects even when there is partial collinearity. Second, instead of assuming that all NPIs tested have an effect like previous work, we conduct variable selection to identify only those NPIs that are retained in parsimonious models. Third, we conduct cluster analysis to explicitly identify temporal correlations, and use this in our interpretation of the strength of evidence behind each intervention. Fourth, we have conducted sensitivity analyses across a range of model specifications around the variable selection criteria, temporal lag between NPIs and change in transmission, temporal truncation, and the way NPI intensity is coded.

Nonetheless, our study also has several limitations. First, besides the information bias in the NPIs database discussed above, the coding scheme may also introduce potential bias. For example, NPIs coded as "comprehensive contact tracing for all identified cases" may have different implications in different countries. Effectiveness of contact tracing in places like Singapore [58] may be masked by seemingly similar but realistically non-comparable contact tracing programmes elsewhere. Second, compared to daily incidence, R_t estimates are much more suitable for cross-country comparisons and thus are used as the metric for COVID-19 transmission in this study. However, these estimates are based on a series of assumptions that may not always be appropriate. For example, the underlying methods assume constant case ascertainment rates over the 12-week time window (March–June 2020) over which our analysis takes place. Consequently, declines in R_t over time may have been obscured by improvements in case ascertainment, leading to some effective NPIs appearing ineffective in our analyses. We have partially adjusted for this by giving weight in our interpretation only to NPIs whose effect direction is robust to changes in the time-series length. Another limitation is that our model also does not propagate uncertainty around R_t estimates. Third, although we examined a wide range of NPIs, we did not include any potential interactions in the current model. Such interaction is a possibility, e.g. more people may comply with workplace closures when receiving income support. Future research should look into these relationships. Last but not the least, although OxCGRT is one of the most comprehensive databases of COVID-19-related NPIs to our knowledge, it does not capture individual behaviour such as face-covering use in public spaces. Thus, we were not able to assess the effectiveness of such measures in reducing COVID-19. Such behavioural measures may prove crucial to controlling COVID-19 epidemics, so analyses of datasets that capture adherence to these measures (e.g. survey of public behaviours [59]) may yield important insights in the future.

Conclusions

In conclusion, evidence from a panel of 130 countries and territories provides evidence about the effectiveness of school closure and internal movement restrictions in reducing SARS-CoV-2 transmission. Despite the inherent limitations of observational and ecological data, our study provides the broadest empirical evaluation on the relative effectiveness of NPI in reducing COVID-19 transmission, while addressing the issue of structural confounding due to temporal clustering.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-020-01872-8.

Additional file 1. Table S1: Metadata on Policy Code in the Oxford COVID-19 Government Response Tracker. **Table S2:** Peak timing of stringency indices by region. **Table S3:** Results of hierarchical clustering of time-series using the any effort scenario. **Table S4:** Results of hierarchical clustering of time-series using the maximum effort scenario. **Table S5:** Lowest performance model fit by country. **Table S6:** Highest performance model fit by country. **Table S7:** Statistical interpretation worksheets. **Table S8:** Review of existing literature. **Figure S1:** The number of countries and regions with available data in the Oxford COVID-19 Government Response Tracker. **Figure S2:** The pair-wise scatter plot of NPI timing under the Any Effort Scenario. **Figure S3:** The pair-wise scatter plot of NPI timing under the Maximum Effort Scenario. **Figure S4:** Deviance from panel analyses using different temporal lags between effective reproduction number and policy interventions based on full time-series. **Figure S5:** Deviance from panel analyses using different temporal lags between effective reproduction number and policy interventions based on the truncated time-series. **Figure S6:** Univariable panel analyses – effect sizes. **Figure S7:** Univariable panel analyses – p-values. **Figure S8:** The sequential order of different NPIs under any effort scenario. **Figure S9:** The sequential order of different NPIs under maximum effort scenario. **Figure S10:** Hierarchical cluster analysis of NPIs time-series using the multilevel scenario. **Figure S11:** Effect sizes for each NPI from the selected models based on multilevel scenario.

Acknowledgements

Funding information for the Centre for Mathematical Modelling of Infectious Disease COVID-19 Working Group: James Munday (Wellcome Trust: 210758/Z/18/Z); Hamish Gibbs (UK DHSC/UK Aid/NIHR: ITCRZ 03010); Carl A B Pearson (BMGF: NTD Modelling Consortium OPP1184344, DFID/Wellcome Trust: 221303/Z/20/Z); Kiesha Prem (BMGF: INV-003174, European Commission: 101003688); Quentin J Leclerc (UK MRC: LID DTP MR/N013638/1); Sophie R Meakin (Wellcome Trust: 210758/Z/18/Z); W John Edmunds (European Commission: 101003688, UK MRC: MC_PC_19065, NIHR: PR-OD-1017-20002); Christopher I Jarvis (Global Challenges Research Fund: ES/P010873/1); Amy Gimma (Global Challenges Research Fund: ES/P010873/1, UK MRC: MC_PC_19065); Sebastian Funk (Wellcome Trust: 210758/Z/18/Z); Matthew Quaife (ERC Starting Grant: #757699, BMGF: INV-001754); Timothy W Russell (Wellcome Trust: 206250/Z/17/Z); Jon C Emery (ERC Starting Grant: #757699); Sam Abbott (Wellcome Trust: 210758/Z/18/Z); Joel Hellewell (Wellcome Trust: 210758/Z/18/Z); Rein M G J Houben (ERC Starting Grant: #757699); Kathleen

O'Reilly (BMGF: OPP1191821); Georgia R Gore-Langton (UK MRC: LID DTP MR/N013638/1); Adam J Kucharski (Wellcome Trust: 206250/Z/17/Z); Megan Auzenbergs (BMGF: OPP1191821); Billy J Quilty (NIHR: 16/137/109, NIHR: 16/136/46); Thibaut Jombart (Global Challenges Research Fund: ES/P010873/1, UK Public Health Rapid Support Team, NIHR: Health Protection Research Unit for Modelling Methodology HPRU-2012-10096, UK MRC: MC_PC_19065); Alicia Rosello (NIHR: PR-OD-1017-20002); Oliver Brady (Wellcome Trust: 206471/Z/17/Z); Kevin van Zandvoort (Elrha R2HC/UK DFID/Wellcome Trust/NIHR, DFID/Wellcome Trust: Epidemic Preparedness Coronavirus research programme 221303/Z/20/Z); James W Rudge (DTRA: HDTRA1-18-1-0051); Akira Endo (Nakajima Foundation, Alan Turing Institute); Kaja Abbas (BMGF: OPP1157270); Fiona Yueqian Sun (NIHR: 16/137/109); Simon R Procter (BMGF: OPP1180644); Samuel Clifford (Wellcome Trust: 208812/Z/17/Z, UK MRC: MC_PC_19065); Nicholas G. Davies (NIHR: Health Protection Research Unit for Immunisation NIHR200929, UK MRC: MC_PC_19065); Charlie Diamond (NIHR: 16/137/109); Rosanna C Barnard (European Commission: 101003688); Rosalind M Eggo (HDR UK: MR/S003975/1, UK MRC: MC_PC_19065); Emily S Nightingale (BMGF: OPP1183986); David Simons (BBSRC LIDP: BB/M009513/1); Katharine Sherratt (Wellcome Trust: 210758/Z/18/Z); Graham Medley (BMGF: NTD Modelling Consortium OPP1184344); Gwenan M Knight (UK MRC: MR/P014658/1); Stefan Flasche (Wellcome Trust: 208812/Z/17/Z); Nikos I Bosse (Wellcome Trust: 210758/Z/18/Z); Petra Klepac (Royal Society: RP\EA\180004, European Commission: 101003688).

Authors' contributions

YL, CM, JK, and MJ conceived the idea and developed the model; YL visualised the results; YL and MJ drafted the manuscript; all co-authors interpreted the results, reviewed the manuscript, provided edits, and agreed on the final version. The authors read and approved the final manuscript.

Funding

CM and JK's contribution to this work were supported by the Royal Society's Rapid Assistance in Modelling the Pandemic (RAMP) scheme.
We thank Richard Pebody (WHO European Region Office, Copenhagen), Katelijn Vandemaele (WHO, Geneva), and Helen Johnson (European Centre for Disease Prevention and Control, Stockholm) for helpful comments.
YL and MJ are funded by the National Institute of Health Research (UK) (16/137/109), the Bill & Melinda Gates Foundation (INV-003174), and the European Commission project Epipose (101003688). This research was partly funded by the National Institute for Health Research (NIHR) (16/137/109) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care. This research is partly funded by the Bill & Melinda Gates Foundation (INV-003174). The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Bill & Melinda Gates Foundation. YL is also supported by the UK Medical Research Council (MC_PC_19065). CM and JK are employed by IPM Informed Portfolio Management. IPM is appreciative of the contributions of its employees above and beyond the scope of their work, their dedication to the community, and the world as a whole. The views expressed herein are made in a personal capacity and are not those necessarily made, sponsored, affiliated, or endorsed by IPM. RL was supported by a Royal Society Dorothy Hodgkin Fellow.

Availability of data and materials

This study relies entirely on data that are either publically available or from published literature. Code used can be found at [https://github.com/yangclaraliu/COVID19_NPIs_vs_Rt].

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Prof. Mark Jit is a member of the editorial board at BMC Medicine.

Author details

[1]Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, London, UK. [2]IPM Informed Portfolio Management, London, UK. [3]Centre on Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London, UK.

Received: 8 September 2020 Accepted: 25 November 2020
Published online: 05 February 2021

References

1. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect Dis. 2020;20(5):553–8.
2. Riou J, Althaus CL. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. Eurosurveillance. 2020;25(4):2000058.
3. Hale T, Angrist N, Kira B, Petherick A, Phillips T, Webster S. Variation in government responses to COVID-19. BSG-WP-2020/032. Version 5.0. 2020. Available from: https://www.bsg.ox.ac.uk/sites/default/files/2020-05/BSG-WP-2020-032-v5.0_0.pdf. [cited 2020 May 11].
4. Barrot J-N, Grassi B, Sauvagnat J. Sectoral effects of social distancing. Rochester, NY: Social Science Research Network; 2020. Report No.: ID 3569446. Available from: https://papers.ssrn.com/abstract=3569446. [cited 2020 Jun 7].
5. Quilty BJ, Diamond C, Liu Y, Gibbs H, Russell TW, Jarvis CI, et al. The effect of inter-city travel restrictions on geographical spread of COVID-19: evidence from Wuhan, China. medRxiv. 2020;21:2020.04.16.20067504. https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-020-01712-9.
6. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science. 2020;368(6489):395–400.
7. Hellewell J, Abbott S, Gimma A, Bosse NI, Jarvis CI, Russell TW, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Glob Health. 2020;8(4):e488–96.
8. Firth JA, Hellewell J, Klepac P, Kissler S, Kucharski AJ, Spurgin LG. Using a real-world network to model localized COVID-19 control strategies. Nat Med. 2020;26:1616–22. https://doi.org/10.1038/s41591-020-1036-8.
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. Lancet Public Health. 2020,0(0). Available from: http://www.ncbi.nlm.nih.gov/pubmed/32220655. Accessed 25 Mar 2020.
10. 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. Lancet Infect Dis. 2020;20(6):678–88.
11. 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 2020.07.06.20147199. https://www.nature.com/articles/s41562-020-01009-0.
12. Chen X, Qiu Z. Scenario analysis of non-pharmaceutical interventions on global COVID-19 transmissions. arXiv:200404529 [physics, q-bio, stat]. 2020; Available from: http://arxiv.org/abs/2004.04529. [cited 2020 Aug 10].
13. 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;584:262–7. https://doi.org/10.1038/s41586-020-2404-8.
14. Amer F, Hammoud S, Farran B, Boncz I, Endrei D. Assessment of Countries' Preparedness and Lockdown Effectiveness in Fighting COVID-19. Disaster Medicine and Public Health Preparedness. 2020;1–8.
15. Alfano V, Ercolano S. The Efficacy of Lockdown Against COVID-19: A Cross-Country Panel Analysis. Appl Health Econ Health Policy. 2020;18(4):509–17
16. Banholzer N, Weenen E van, Kratzwald B, Seeliger A, Tschernutter D, Bottrighi P, et al. Impact of non-pharmaceutical interventions on documented cases of COVID-19. medRxiv. 2020;2020.04.16.20062141.
17. Bellali H, Chtioui N, Chahed M. Factors associated with country-variation in COVID-19 morbidity and mortality worldwide: an observational geographic study. medRxiv [Internet]. [cited 2020 Aug 13]; Available from: https://www.medrxiv.org/content/10.1101/2020.05.27.20114280v1.
18. Brauner JM, Mindermann S, Sharma M, Stephenson AB, Gavenčiak T, Johnston D, et al. The effectiveness and perceived burden of nonpharmaceutical interventions against COVID-19 transmission: a modelling study with 41 countries. Science. 2020;eabd9338.

19. Chowell G, Rothenberg R, Roosa K, Tariq A, Hyman JM, Luo R. Sub-epidemic model forecasts for COVID-19 pandemic spread in the USA and European hotspots, February-May 2020. medRxiv. 2020;2020.07.03.20146159
20. Rey SK, Rahman MdM, Shibly KH, Siddiqi UR, Howlader A. Epidemic Trend Analysis of SARS-CoV-2 in SAARC Countries Using Modified SIR (M-SIR) Predictive Model. medRxiv [Internet]. [cited 2020 Aug 13]; Available from: https://www.medrxiv.org/content/10.1101/2020.06.29.20142513v1.
21. 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;8:1–8.
22. Ghosal S, Bhattacharyya R, Majumder M. Impact of complete lockdown on total infection and death rates: a hierarchical cluster analysis. Diabetes Metab Syndr. 2020;14(4):707–11.
23. Jüni P, Rothenbühler M, Bobos P, Thorpe KE, Costa BR da, Fisman DN, et al. Impact of climate and public health interventions on the COVID-19 pandemic: a prospective cohort study. CMAJ. 2020;192(21):E566–73.
24. Karnakov P, Arampatzis G, Kičić I, Wermelinger F, Wälchli D, Papadimitriou C, et al. Data driven inference of the reproduction number (R0) for COVID-19 before and after interventions for 51 European countries. medRxiv. 2020; 2020.05.21.20109314
25. Linka K, Peirlinck M, Costabal FS, Kuhl E. Outbreak dynamics of COVID-19 in Europe and the effect of travel restrictions. Comput Methods Biomechanics Biomed Eng. 2020;0(0):1–8.
26. Liu X. A Simple, SIR-like but Individual-Based I-i AIR Model: Application in Comparison of COVID-19 in New York City and Wuhan. Nature Human Behaviour. 2020;2020.05.28.20115121. https://www.sciencedirect.com/science/article/pii/S2211379720321288.
27. Lonergan M, Chalmers JD. Estimates of the ongoing need for social distancing and control measures post-"lockdown" from trajectories of COVID-19 cases and mortality. Eu Respir J 2020;56(1). Available from: https://erj.ersjournals.com/content/56/1/2001483. [cited 2020 Aug 10].
28. López L, Rodó X. The end of social confinement and COVID-19 re-emergence risk. Nature Human Behaviour. 2020;4(7):746–55
29. McGrail DJ, Dai J, McAndrews KM, Kalluri R. Enacting national social distancing policies corresponds with dramatic reduction in COVID19 infection rates. Nature Human Behaviour. 2020;2020.04.23.20077271. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0236619#:~:text=Globally%2C%20we%20find%20that%20social,a%20two%20week%20time%20period.
30. Mishra PK, Mishra S. A deductive approach to modeling the spread of COVID-19. medRxiv. 2020;2020.03.26.20044651.
31. Osherovich VA, Fainberg J, Osherovich LZ. Double power law for COVID-19: prediction of new cases and death rates in Italy and Spain. medRxiv. 2020; 2020.05.07.20094714.
32. Petr K, Georgios A, Fabian W, Daniel W, Costas P, Petros K. Data-driven inference of the reproduction number for COVID-19 before and after interventions for 51 European countries. Swiss Medical Weekly [Internet]. [cited 2020 Aug 13]; Available from: https://smw.ch/article/doi/smw.2020.20313.
33. Ghosal S, Sinha B, Sengupta S, Majumder M. Frequency of testing for COVID 19 infection and the presence of higher number of available beds per country predict outcomes with the infection, not the GDP of the country - A descriptive statistical analysis. medRxiv. 2020;2020.04.01.20047373
34. Tobías A. Evaluation of the lockdowns for the SARS-CoV-2 epidemic in Italy and Spain after one month follow up. Science of The Total Environment. 2020;725:138539.
35. Wang Q, Xie S, Wang Y, Zeng D. Survival-Convolution Models for Predicting COVID-19 Cases and Assessing Effects of Mitigation Strategies. medRxiv [Internet]. [cited 2020 Aug 13]; Available from: https://www.medrxiv.org/content/10.1101/2020.04.16.20067306v2.
36. Yang P, Qi J, Zhang S, Wang X, Bi G, Yang Y, et al. Feasibility Study of Mitigation and Suppression Intervention Strategies for Controlling COVID-19 Outbreaks in London and Wuhan. Nature Human Behaviour. 2020;2020.04.01.20043794. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0236857
37. 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. Lancet Infect Dis. 2020;0(0). Available from: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30785-4/abstract. [cited 2020 Nov 4].
38. CMMID epiforecasts Core Team. [EpiForecast] Covid-19: Global summary. EpiForecast: COVID-19. Available from: https://epiforecasts.io/covid/posts/global/. [cited 2020 May 11].
39. World Health Organization. Tracking public health and social measures: a global dataset. 2020. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/phsm. [cited 2020 Jun 7].
40. World Bank. World Bank Country and Lending Groups – World Bank Data Help Desk. 2020. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups. [cited 2020 Jun 14].
41. Abbott S, Hellewell J, Thompson RN, Sherratt K, Gibbs HP, Bosse NI, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. Wellcome Open Res. 2020;5:112. https://doi.org/10.12688/wellcomeopenres.16006.1.
42. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am J Epidemiol. 2013;178(9):1505–12.
43. Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? J Classif. 2014;31(3): 274–95.
44. Shimodaira H. Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. Ann Stat. 2004;32(6):2616–41.
45. Hausman JA. Specification tests in econometrics. Econometrica. 1978;46(6): 1251–71.
46. R Core Team. R: a language and environment for statistical computing: Vienna R Foundation for statistical Computing; 2020. Available from: http://www.R-project.org/. Accessed 25 Apr 2020.
47. Croissant Y, Millo G. Panel data econometrics in R: the plm package. J Stat Software. 27(2):1–43.
48. Suzuki R, Terada Y, Shimodaira H. pvclust: hierarchical clustering with P-values via multiscale bootstrap resampling. 2019. Available from: https://CRAN.R-project.org/package=pvclust.
49. Jackson ML, Hart GR, McCulloch DJ, Adler A, Brandstetter E, Fay K, et al. Effects of weather-related social distancing on city-scale transmission of respiratory viruses. medRxiv. 2020;3:2020.03.02.20027599.
50. 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. Lancet Public Health 2020;0(0). Available from: https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(20)30133-X/abstract. [cited 2020 Jun 24].
51. Russell TW, Wu JT, Clifford S, Edmunds WJ, Kucharski AJ, Jit M; Centre for the Mathematical Modelling of Infectious Diseases COVID-19 working group. Effect of internationally imported cases on internal spread of COVID-19: a mathematical modelling study. Lancet Public Health. 2020:S2468-2667(20)30263-2. https://doi.org/10.1016/S2468-2667(20)30263-2. Epub ahead of print. PMID: 33301722
52. Clifford S, Quilty BJ, Russell TW, Liu Y, Chan Y-WD, Pearson CAB, et al. Strategies to reduce the risk of SARS-CoV-2 re-introduction from international travellers. medRxiv. Available from: https://www.medrxiv.org/content/10.1101/2020.07.24.20161281v2. [cited 2020 Aug 10].
53. Stockdale JE, Doig R, Min J, Mulberry N, Wang L, Elliott LT, et al. Long time frames to detect the impact of changing COVID-19 control measures. medRxiv. 2020; 2020.06.14.20131177.
54. Wu P, Tsang TK, Wong JY, Ng TW, Ho F, Gao H, et al. Suppressing COVID-19 transmission in Hong Kong: an observational study of the first four months. Research Square. 2020.
55. Bertozzi AL, Franco E, Mohler G, Short MB, Sledge D. The challenges of modeling and forecasting the spread of COVID-19. PNAS. 2020;117(29):16732–8.
56. Armario C. Colombia's Medellin emerges as surprise COVID-19 pioneer. AP NEWS. 2020; Available from: https://apnews.com/b3f8860343323d0daeef72191b669baf. [cited 2020 Aug 13].
57. Islam N, Sharp SJ, Chowell G, Shabnam S, Kawachi I, Lacey B, et al. Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. BMJ. 2020;370. Available from: https://www.bmj.com/content/370/bmj.m2743. [cited 2020 Aug 11].
58. Vaswani K. Coronavirus: The detectives racing to contain the virus in Singapore. BBC News. 2020; Available from: https://www.bbc.com/news/world-asia-51866102. [cited 2020 Jun 15].
59. Imperial College London. COVID-19 Behaviour Tracker. Available from: http://www.coviddatahub.com/. [cited 2020 Nov 4].

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.