

Understanding the effectiveness of government interventions in Europe's second wave of COVID-19

Mrinank Sharma^{*,1,2,3}, Sören Mindermann^{*,4}, Charlie Rogers-Smith⁵, Gavin Leech⁶, Benedict Snodin³, Janvi Ahuja^{7,3}, Jonas B. Sandbrink⁷, Joshua Teperowski Monrad^{8,9,3}, George Altman¹⁰, Gurpreet Dhaliwal^{11,12}, Lukas Finnveden³, Alexander John Norman¹³, Sebastian B. Oehm^{14,15}, Julia Fabienne Sandkühler¹⁶, Thomas Mellan¹⁷, Jan Kulveit³, Leonid Chindelevitch¹⁷, Seth Flaxman¹⁸, Yarin Gal⁴, Swapnil Mishra^{+,17,19}, Jan Markus Brauner^{*,+,@,4,3}, Samir Bhatt^{+,@,20,17,19}

*Equal contribution to first authorship

+Equal contribution to senior authorship

@Correspondence: jan.brauner@cs.ox.ac.uk, s.bhatt@imperial.ac.uk

¹Department of Statistics, University of Oxford, UK

²Department of Engineering Science, University of Oxford, UK

³Future of Humanity Institute, University of Oxford, UK

⁴Oxford Applied and Theoretical Machine Learning (OATML) Group, Department of Computer Science, University of Oxford, UK

⁵OATML Group (work done while at OATML as an external collaborator), Department of Computer Science, University of Oxford, UK

⁶Department of Computer Science, University of Bristol, UK

⁷Medical Sciences Division, University of Oxford, UK

⁸Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, UK

⁹Department of Health Policy, London School of Economics and Political Science, UK

¹⁰Manchester University NHS Foundation Trust, UK

¹¹The Francis Crick Institute, London, UK

¹²School of Life Sciences, University of Warwick, Coventry, UK

¹³Mathematical, Physical and Life Sciences (MPLS) Doctoral Training Centre, University of Oxford, UK.

¹⁴Medical Research Council Laboratory of Molecular Biology, Cambridge, UK

¹⁵University of Cambridge, UK

¹⁶Independent researcher, Essen, Germany

¹⁷Medical Research Council (MRC) Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, UK

¹⁸Department of Mathematics, Imperial College London, UK

¹⁹Abdul Latif Jameel Institute for Disease and Emergency Analytics (J-IDEA), School of Public Health, Imperial College London, UK

²⁰Section of Epidemiology, Department of Public Health, University of Copenhagen, Denmark

Abstract

As European governments face resurging waves of COVID-19, non-pharmaceutical interventions (NPIs) continue to be the primary tool for infection control. However, updated estimates of their relative effectiveness have been absent for Europe's second wave, largely due to a lack of collated data that considers the increased subnational variation and diversity of NPIs. We collect the largest dataset of NPI implementation dates in Europe, spanning 114 subnational areas in 7 countries, with a systematic categorisation of interventions tailored to the second wave. Using a hierarchical Bayesian transmission model, we estimate the effectiveness of 17 NPIs from local case and death data. We manually validate the data, address limitations in modelling from previous studies, and extensively test the robustness of our estimates. The combined effect of all NPIs was smaller relative to estimates from the first half of 2020, indicating the strong influence of safety measures and individual protective behaviours--such as distancing--that persisted after the first wave. Closing specific businesses was highly effective. Gathering restrictions were highly effective but only for the strictest limits. We find smaller effects for closing educational institutions compared to the first wave, suggesting that safer operation of schools was possible with a set of stringent safety measures including testing and tracing, preventing mixing, and smaller classes. These results underscore that effectiveness estimates from the early stage of an epidemic are measured relative to pre-pandemic behaviour. Updated estimates are required to inform policy in an ongoing pandemic.

Introduction

The first wave of the novel coronavirus, SARS-CoV-2, resulted in marked excess mortality across most Western European countries. Most of these countries implemented a suite of unprecedented non-pharmaceutical interventions (NPIs) including the closure of businesses, schools, and bans on gatherings (1–4). Although partial control was achieved in the summer months, a second wave (5, 6) of the epidemic followed the reopening of European societies around September 2020. With vaccines set to reach only a minority of the global population in 2021 (7) and current vaccine roll-out delays in Europe, NPIs remain the primary tool for infection control in the short term. Furthermore, new variants of concern with mutations that contribute to higher transmissibility, severity, or antigenic escape (8–10) further increase the need for targeted methods to control infections while reducing the social and economic consequences of broad lockdowns.

The relative effectiveness of NPIs in the first wave of the pandemic has been determined by relating the implemented NPIs to the course of the pandemic in different countries (1–3, 11–14). However, NPI effectiveness estimates from the second wave are expected to be substantially more representative of future waves and thus more relevant to government decision making, for several reasons. The statistical models used to estimate NPI effectiveness rely on comparing epidemiological observations in the presence and absence of NPIs. However, disease transmission in the absence of a given NPI depends strongly on behaviour. Contact patterns changed with the arrival of the first wave, as individuals adopted voluntary practices to reduce infection including physical distancing, avoiding crowded places, mask-wearing, and working from home (15), while organizations adopted ongoing safety measures such as testing and separating cohorts in schools (16). Recent polling suggests that behavioural patterns have comparatively stabilized since the summer of 2020 (15), indicating that as long as these behavioural changes stay in place, effect sizes in the second wave are more relevant to potential further waves.

Furthermore, many NPIs in the first wave started within a short period in March 2020, which made it challenging to disentangle the effects of individual NPIs (1, 2). European governments have also adopted a more granular selection of interventions in the second wave, such as

time-staggered closure of different business types. This stratification was absent in the first wave. Moreover, NPIs were implemented subnationally, which greatly increases the number of data points available for this study. Finally, in the first wave, the data quality was affected by pervasive undercounting and sampling bias in the case data, and the variable time lag between infection and death (17) reduced the statistical precision of mortality data (1). Taken together, the data from the second wave is predicted to give a much more detailed, robust, and contemporary picture of the impact of specific interventions.

Here, we provide the first effect estimates for individual interventions during Europe's second wave. Countries typically implemented several NPIs concurrently, or in grouped tiers (18–20), making their individual effects non-identifiable. A multinational dataset is required to overcome this limitation since different countries implemented different sets of NPIs at different times. We also require subnational data since modelling on a national scale would obscure local heterogeneity from transmission timings (21, 22), socioeconomic heterogeneity, and differing policy decisions. Ignoring this subnational variation can result in ecological fallacies (23) and biased effect estimates. Infection heterogeneity preceding the second wave in the UK is a salient example, as there was a strong north/south geographic divide meaning that analysis on a national scale would have obscured increases in transmission due to aggregation. Modelling efforts for the second wave thus require a novel NPI dataset, since the utility of existing trackers is limited by a lack of granular subnational data, inconsistencies (24), missing entries (25), and broad intervention definitions ill-suited for the second wave.

We introduce a bespoke and systematic categorisation of interventions and estimate effect sizes across a randomized sample of 114 regions in 7 European countries (Austria, the Czech Republic, England, Germany, Italy, the Netherlands, Switzerland). We manually collected the intervention data with several validation procedures to ensure high quality and consistency. Using a semi-mechanistic Bayesian transmission model with a latent stochastic process, we estimate effects for all NPIs jointly from case and death data. Since estimates of NPI effectiveness can be highly sensitive to modelling decisions (14, 26), we evaluate robustness to changes in the data, model, epidemiological assumptions, and potential unobserved confounding factors.

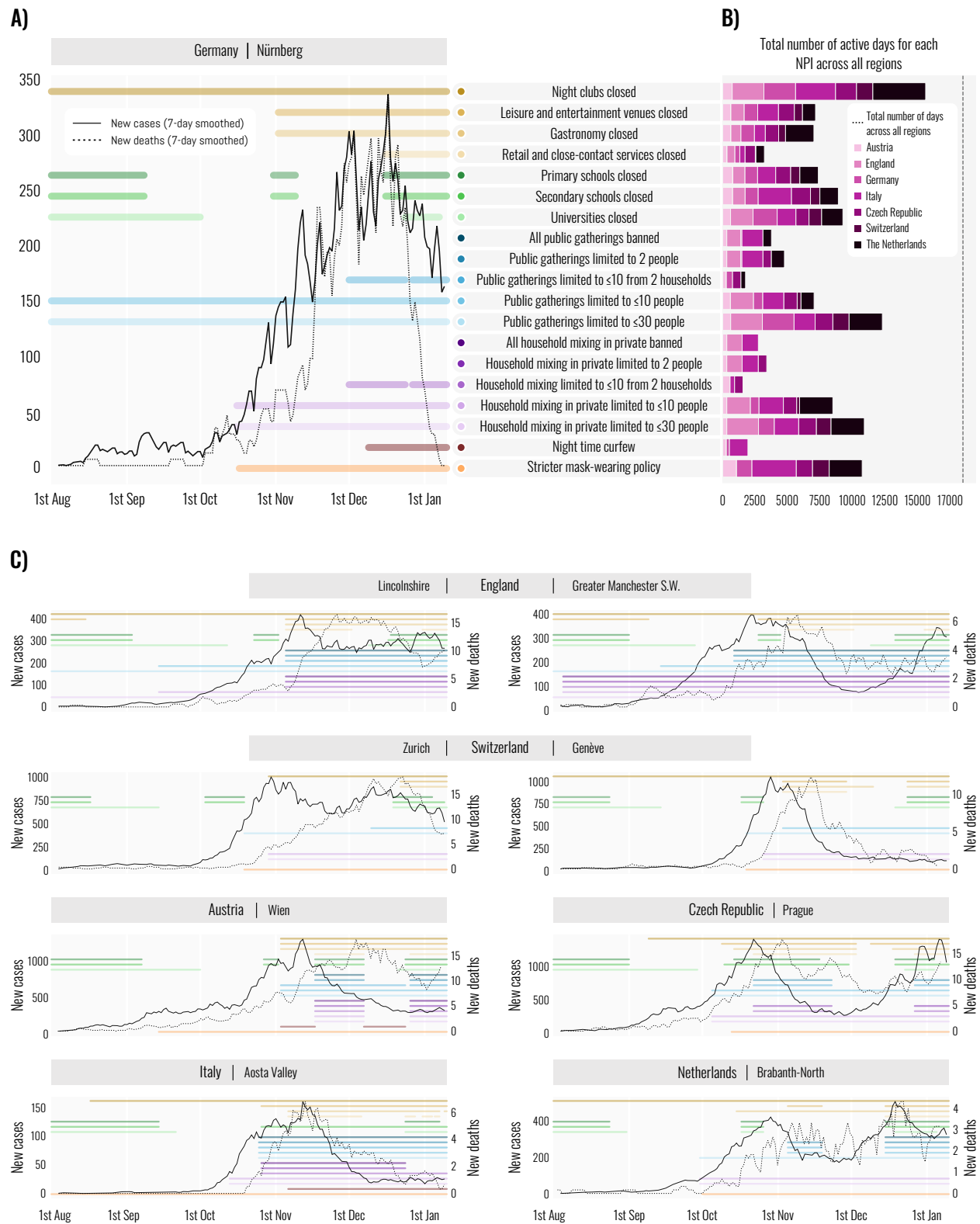


Fig. 1. Dataset. (A) Cases, deaths and implementation dates of nonpharmaceutical interventions in an example region (Nürnberg, Germany). Colored lines indicate the dates that each intervention was active. Colors represent different interventions. (B) The total number of days that each intervention was used in our dataset, aggregated across regions but separated by country. The dashed vertical line indicates the total-number of region-days in our dataset. (C) Additional example timelines showing the cases, deaths, and interventions in 6 regions. Comparing two regions within England (Lincolnshire and Greater Manchester S.W.) and within Switzerland (Zürich and Gèneve) reveals significant subnational variation, both in the interventions used and in the evolution of the epidemic.

Table 1. Main dataset characteristics. In total, we collated more than 5,500 intervention entries through a bespoke systematic categorisation.

Countries	7
Regions of analysis	114
Period	1 August 2020 - 9 January 2021*
Days across all regions	19,000
NPI entries in the dataset	> 5,500
Data validation (manual)	Semi-independent double entry; interviews with local epidemiologists; validation against external sources; cross-country consistency checks

*We ended the period of analysis before 9 January 2021 for English regions depending on their prevalence of a new variant of concern (see *Methods*).

Differences between Europe's first and second wave

A key challenge for distinguishing the effects of specific interventions is that governments often implement several NPIs simultaneously (Fig. 1). During the first wave, interventions were implemented within a short time window: for a given intervention and region, on average 83% of the other interventions in that region started in the same 10 day period (2). In the second wave, NPI implementation was spaced out (Fig. 1), with only 23% of interventions starting in the same 10 day period. The increased temporal spacing combined with an even larger and subnational dataset (9.2x more NPI implementations than the largest study that focused on Europe (2)) allow us to disentangle the effects of 17 individual NPIs. For most NPIs, we were able to robustly isolate their individual effects because their collinearity was limited and our dataset large. Specifically, for each pair of NPIs that we are able to disentangle, we observed one without the other for 6969 region-days on average (with a minimum of 635 region-days for limiting household mixing in private to ≤ 10 attendees and to ≤ 30 attendees). However, we only show the combined effect of indoor and outdoor gathering bans (of various stringencies) since these comprise all 6 NPI pairs that score lowest on the aforementioned metric. Our estimates are robust to changes in data and model parameters (see *Robustness of estimates*), in contrast to smaller datasets on first-wave data (1, 26), indicating (27) that the data are sufficient to overcome collinearity.

Using this modelling approach, we estimate intervention effect sizes and express these as percentage reductions in the (instantaneous) reproduction number R_t (Fig. 2). The effect sizes in the second wave were considerably smaller than those estimated for the first half of 2020. All NPIs included in the study together reduced R_t by 66% [95% CI: 61%-69%], compared to median reductions of 77%-82% in the first wave (1, 2). The difference between the waves is even more pronounced if we consider the effectiveness of the most stringent set of NPIs actually implemented in each region, rather than the (hypothetical) combined effectiveness of all NPIs included in the study. The most stringent set of NPIs implemented in each region reduced R_t by an average of 56% [95% CI: 40%-64%], compared to 76%-82% in the first wave, even though NPIs in the second wave were often similarly strict or stricter (1, 2). Finally, R_t was reduced from an average maximum of 1.7 [95% CI: 1.4-2.4] to a minimum of 0.7 [95% CI: 0.5-0.8] across

regions in the second wave, compared to an average maximum of 3.3-3.8 and minimum of 0.7-0.8 in the first wave (1, 2).

The most important explanation for these differences between the first and second wave is likely a combination of persistent behavioural change, such as avoiding close contact or increased hygiene, and the widespread adoption of safety protocols, such as distancing rules and adequate ventilation. These changes likely made various areas of public life safer and thereby reduced the effect of strict bans or closures. This hypothesis is corroborated by the stability of behavioural polls after the first wave (28). Several other factors seem less important but may still have contributed to the variation. First, a build-up of population immunity likely does not explain the reduction in NPI effectiveness: attack rates were low in our period of analysis (29) and the in- or exclusion of population immunity did not change the estimated effect sizes. Second, reduced adherence to NPIs in the second wave may have played a role, although adherence seems much more relevant for some NPIs (e.g. restrictions on private gatherings) than for others (e.g. closure of businesses or educational institutions). Finally, the ascertainment rate of cases increased (30, 31) throughout the first wave but was far more stable in the second wave. However, this difference is expected to increase the estimated effectiveness of NPIs in the second wave compared to the first, the opposite of what we found.

If all NPIs together reduced the original value of R_t to 18-23% in the first wave and to 35% in the second wave, this suggests an overall reduction in R_t of 34-49% due to behaviour changes, safety protocols, and potentially other effects. The result underscores the importance of viewing the effectiveness of an NPI relative to the counterfactual safety measures and behaviour in the absence of the NPI. Since first-wave NPI effectiveness was measured relative to pre-pandemic behaviour, it is likely inadequate to inform policy during an ongoing pandemic. While behaviour will continue to change dynamically, the increased stability seen in behavioural polls (15) means that second wave NPI effectiveness estimates are expected to generalise better to future waves.

A detailed and subnational assessment of interventions in Europe's second wave

We find that business closures were particularly effective, with a combined effect of reducing R_t by 35% [95% CI: 29-41%] (Fig. 2A). Closures of gastronomy (restaurants, pubs, and cafes) had a large effect on transmission with an estimated reduction in R_t of 12% [95% CI: 8-17%], an estimate that is broadly in line with the increases estimated to have occurred as a consequence of the UK's "eat out to help out" scheme (32) in August. We find a similar substantial effect for closing night clubs [12%, 95% CI: 8-17%], which were predominantly shut earlier than other businesses; this effect size may reflect early second wave superspreading (33). The combined effect of closing retail and close contact services (such as hairdressers and beauty salons) is also considerable [12%, 95% CI: 7-18%]. Assuming that much of this effect is due to the more common type of business – namely, retail – it underscores the potential risks of brief but very numerous indoor contacts (34). Closing leisure and entertainment venues such as zoos, museums, and theatres had a small effect [3%, -1-10%].

As a broad intervention, we found that banning all gatherings, including 1-on-1 meetings, had a large effect: a 26% [95% CI: 18-32%] reduction in R_t . By recording the number of persons and households allowed to meet, we were able to test the impact of various thresholds. We found no evidence for diminishing returns in the number of persons allowed to meet; indeed, only strict thresholds had a statistically significant effect. This is consistent with previous studies on the English tier 1 and tier 2 restrictions, which limited gatherings to 6 people and had a small total impact in contrast to the more stringent tier 3 (18, 19, 35). The small effect associated with more lenient person limits (10 or higher) contrasts with estimates from the first wave (2). The difference could be due to voluntary protective behaviours such as avoiding crowds and distancing, but also due to limited adherence to rules on private mixing (36). These factors may also explain why all gathering bans combined were somewhat less effective compared to business closures. Defining a "lockdown" policy as a ban on all gatherings and closure of all non-essential businesses, we estimate a total reduction in R_t of 52% [95% CI: 47-56%].

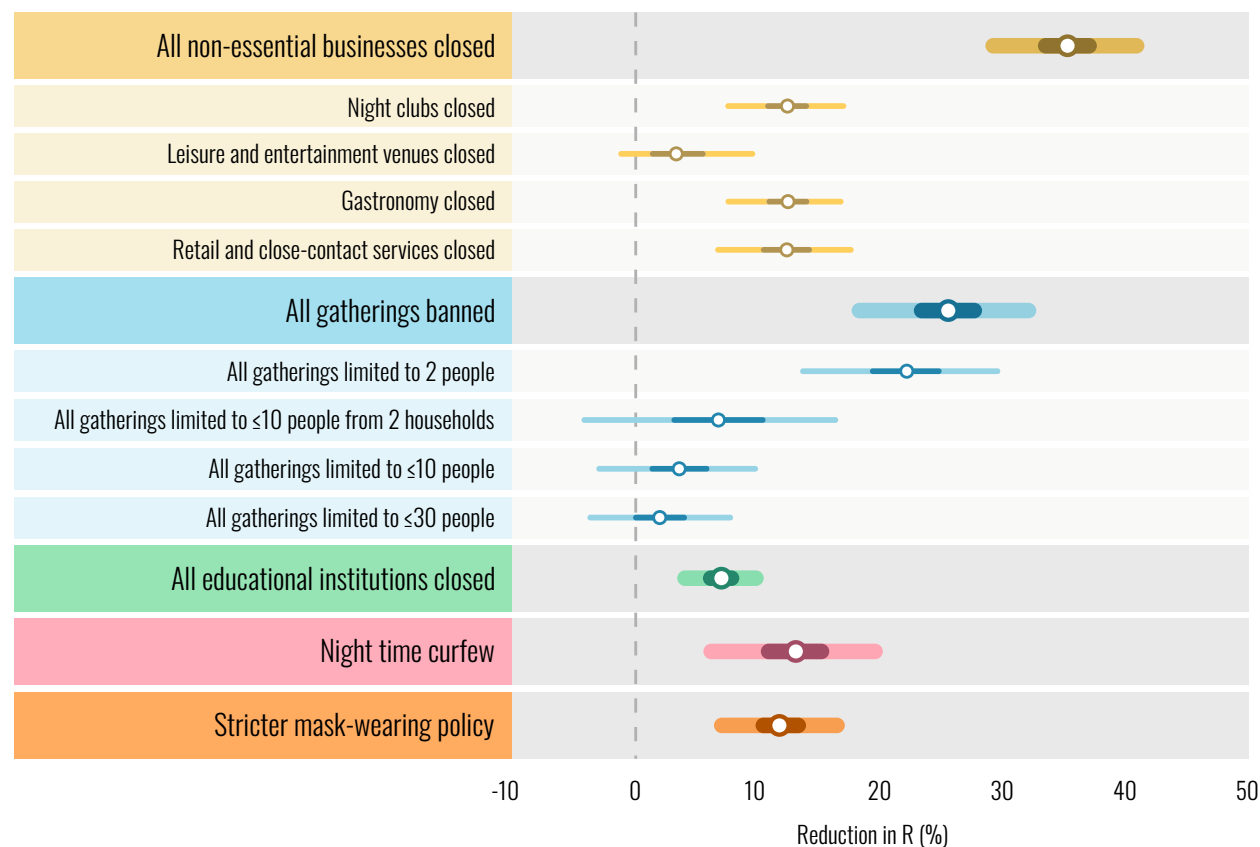
Most countries adopted different limits for public gatherings and household mixing in private at various times. We are therefore able to disaggregate the effect of these gathering types and examine their relative effects (Fig. 2B). While the total effect of banning all private mixing exceeds that of banning public gatherings, private mixing restriction was only effective at a strict threshold of 2 people allowed to meet. As discussed above, this could be due to a combination of low adherence and individuals voluntarily avoiding crowds.

Observational studies of the first wave consistently found that closing educational institutions was among the most effective NPIs (2, 11–13, 37). In strong contrast, we discover that their effect was small in the second wave [7%, 95% CI: 4–10%]. We conjecture that a combination of safety measures, behaviour changes, and epidemiological factors (38) in the education sector prevented large undetected clusters which may have developed in the first half of 2020 (39–42). Schools in Europe's second wave operated under safety measures that some other organizations lacked, including symptom screening, asymptomatic testing, contact tracing, sanitizing, ventilation, distancing, reducing group sizes, and preventing the mixing of groups (39, 43). However, measures can vary by country (16) and detailed further assessments are required. Without sufficient measures, opening schools could lead to a resurgence (44).

We documented student presence separately for universities (or higher education) and schools (both primary and secondary) by recording their local term times, holidays, and closure dates in all 114 regions, and identifying regions without universities. However, we refrain from quantifying the relative effects of schools and universities here, and instead only report their combined effect since it was more robust than the individual effects in one sensitivity analysis that was designed to adjust for undetected infections (45) in schools (Supplementary text 1.1.3). The small combined effect suggests that safety measures and behaviour changes significantly reduced transmission associated both with schools and with universities.

Finally, we find that a stricter mask-wearing policy (mandatory in most or all shared/public spaces) and a night time curfew had moderate, but statistically significant effects [12%, 95% CI: 7–17%] and [13%, 6–20%]). However, due to the broad nature of these interventions, they are also likely to interact with other active NPIs. In contrast, the other NPIs affect largely distinct areas of social activity and therefore are not expected to mutually interact to a great extent.

A)



B)

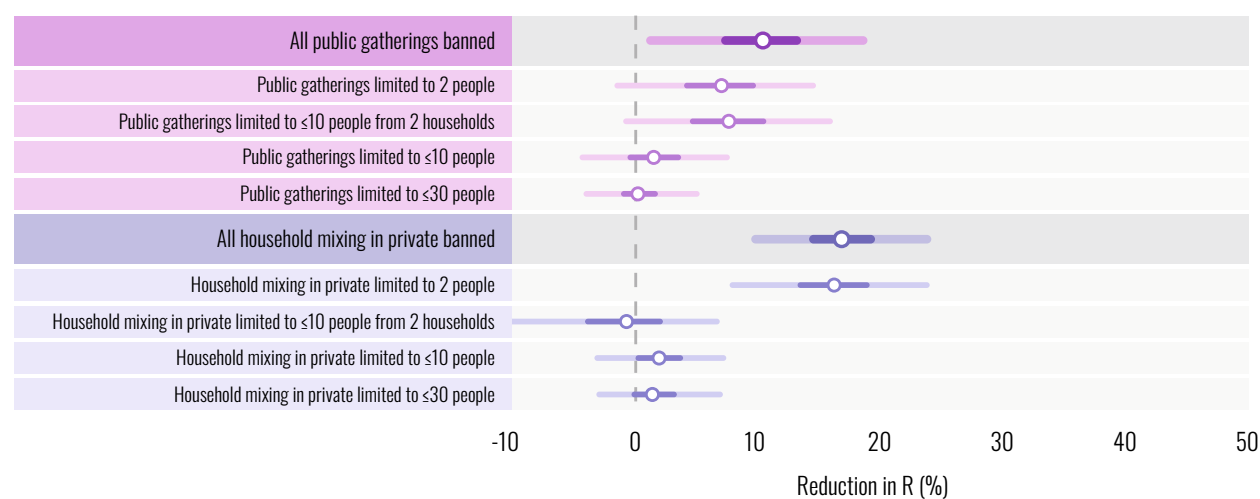


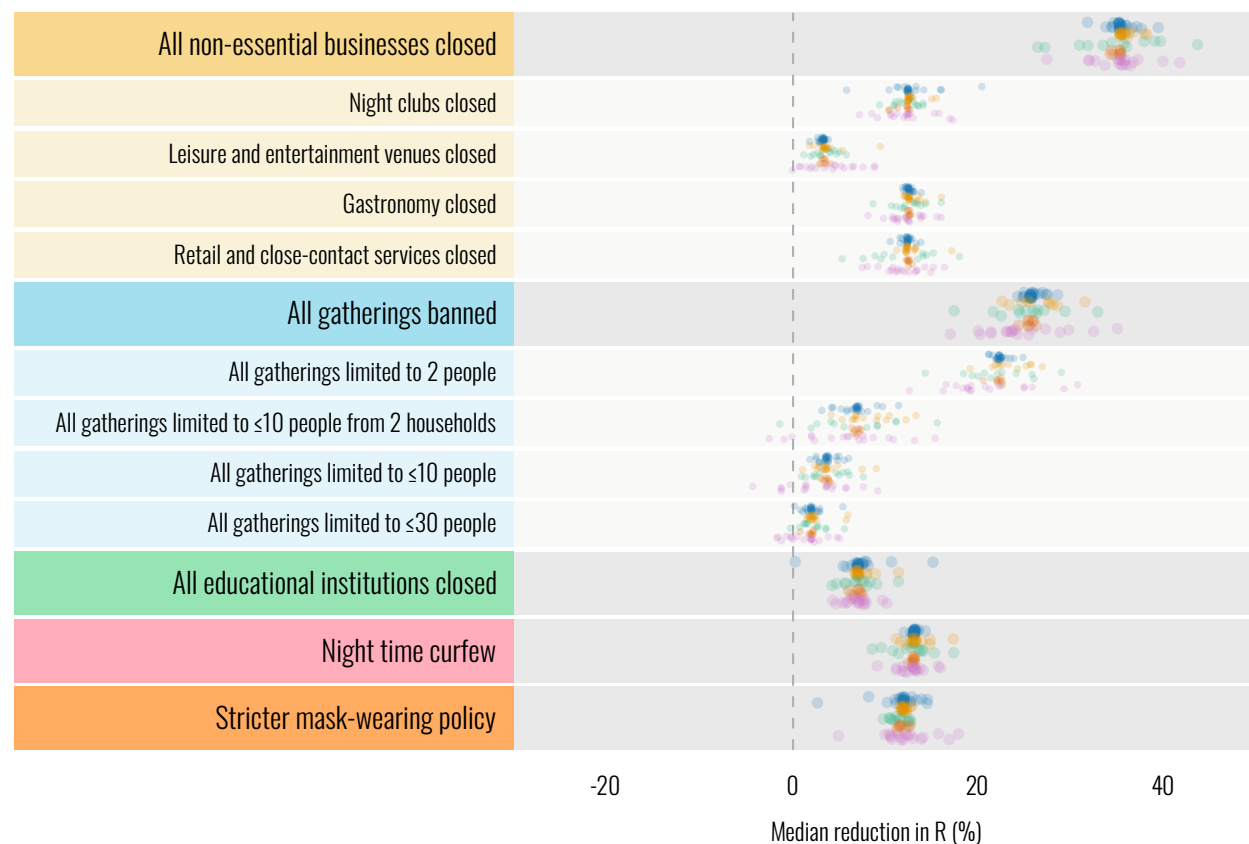
Fig. 2. Intervention effectiveness under default model settings. Posterior percentage reductions in R_t . Markers indicate posterior median estimates. Lines indicate the 50% and 95% posterior credible intervals. A negative 1% reduction refers to a 1% increase in R_t . **(A)** Effectiveness of the main interventions included in our study. Intervention names preceded by “All” show the combined effect of multiple interventions. For example, “All gatherings banned” shows the combined effect of banning all public gatherings and all household mixing in private. **(B)** Individual effectiveness estimates for gathering types, separated into public gatherings and household mixing in private.

Robustness of estimates

The usefulness of estimates of intervention effectiveness hinges on their robustness, and sensitivity analysis can help diagnose when the model is misspecified or the data too collinear (27, 46). Fig. 3 shows how the median estimates of effect size from Fig. 2 vary across 17 types of sensitivity analyses, as we modify the priors and structure of the model, change the distributions of epidemiological delays, and randomly vary the set of regions and other aspects of the data. Each analysis is shown in Supplementary text section 1.1. Since we cannot model all possible factors that affect transmission, we also investigate the stability of effects sizes to unobserved effects that influence R_t , acting as possible confounders. These include unrecorded NPIs and changes to ascertainment and fatality rates. See Supplementary Table S1 for details on the experiments in Fig. 3 and Supplementary text sections 1.2-1.6 for additional validation experiments including simulations on confounding, posterior predictive checks (47), and a single-model meta-analysis across regions.

Median effects vary across a total of 86 experimental conditions in Fig. 3. However, a robust broad picture emerges in which some NPIs outperform others across all experiments. Although our results are robust to varying strengths and types of unobserved factors, the true strength of unobserved confounding is unknown and our study is therefore subject to the limitations of observational approaches.

A)



B)

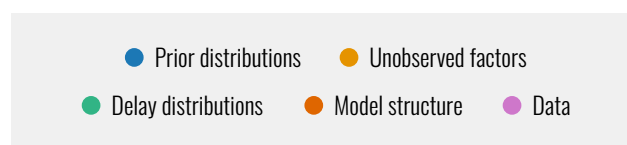
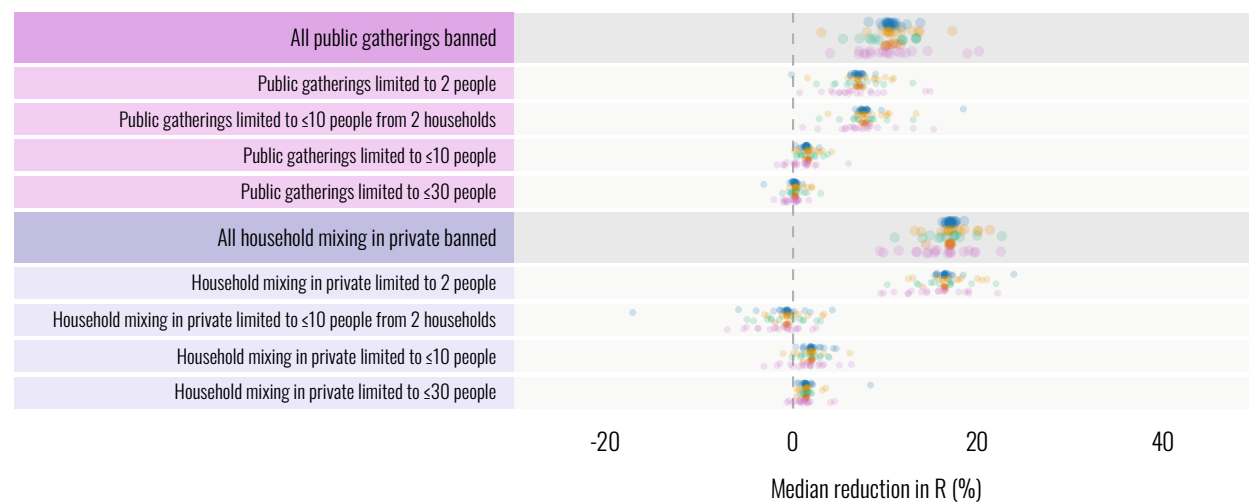


Fig. 3. Robustness of median intervention effectiveness estimates across 86 experimental conditions. Each dot represents the posterior median intervention effectiveness in a different experimental condition. Dot color indicates five categories of sensitivity analyses. Each category contains several sensitivity analyses (17 in total). Each sensitivity analysis contains several experimental conditions (86 in total). Supplementary Table S1 lists all sensitivity analyses by category. **(A)** Robustness of effectiveness estimates of the main interventions included in our study. Intervention names preceded by “All” show the combined effect of multiple interventions. For example, “All gatherings banned” shows the combined effect of banning all public gatherings and all household mixing in private. **(B)** Robustness of the individual effectiveness estimates for separately banning public gatherings or household mixing in private.

Conclusion

European governments are presently deciding which interventions to keep and which to remove. These decisions demand effectiveness estimates directly from the second wave since estimates from the first wave are measured relative to pre-pandemic contact patterns, and therefore overestimate NPI effectiveness in an ongoing pandemic.

As different areas of social life reopen, effectiveness estimates can inform which areas require additional safety measures to reopen safely. In case of a resurgence—for example, due to variants of concern that are more transmissible or evade protection from previously acquired immunity—our results provide a starting point to effectively control infections while minimizing social and economic harms.

Using a European NPI dataset of increased scale and granularity enabled us to provide the first effectiveness estimates for individual NPIs in Europe’s second wave. However, decision-makers using these results should account for several factors. Discontinuing safety measures and voluntary protective behaviours among the still-susceptible population would change intervention effectiveness, perhaps even increasing it to first-wave levels. Mass vaccination may increase the relative importance of interventions such as school closures that mainly affect the unvaccinated population. Moreover, it is still unknown whether variants of concern differentially affect certain population groups. Finally, while our model allows for spatiotemporal variation in behaviour and adherence, no single effectiveness estimate can apply

to all regions. Therefore, our estimates should be combined with expert judgment to adjust them to local and contemporary circumstances

Methods

Data

Dataset overview

We collected a custom NPI dataset for this modelling study, as existing datasets do not provide sufficient geographical resolution to model the second wave. Further advantages of our dataset are NPI definitions tailored towards the second wave and high data quality through extensive validation.

To create this dataset, we collected chronological data on NPIs that were in place between 1 August 2020 and 9 January 2021 in administrative regions, districts, and local areas of 7 European countries. The resulting dataset contains over 5,500 entries on various NPIs in 114 regions of analysis (Supplementary Table S4). Every entry includes the NPI start date and end date, quotes and comments, and one or more sources from websites of governments and universities, legal documents, and/or media reports. Daily case and death data were obtained from government websites (Supplementary Table S3). The data is available upon request.

We now describe how we selected the countries, regions of analysis within each country, and NPI definitions.

We first identified 7 European countries for which public data on daily reported cases and deaths were available at the same geographical resolution at which the country implemented NPIs (Austria, the Czech Republic, England, Germany, Italy, the Netherlands, Switzerland).

To gather initial information about the transmission-reducing NPIs used in these countries, we conducted an exploratory data collection and interviewed local epidemiologists from the countries. Based on these data, we created NPI definitions that faithfully represent the interventions that were implemented in these countries. We focused on clear-cut, major interventions that were implemented in many countries and we only recorded mandatory

restrictions, not recommendations¹. We also accounted for closures that are not due to NPIs, such as vacation and term times in schools and universities, as we surmised that these effectively function as NPIs.

The exploratory data also informed the appropriate level of geographical granularity for the NPI data collection. In each country, we set our regions of analysis to correspond to the highest possible level of administrative division for which NPI implementations were identical throughout each region. The chosen administrative divisions were (Supplementary Table S3):

- states in Austria,
- administrative regions in the Czech Republic,
- NUTS 3 statistical regions in England,
- districts in Germany,
- administrative regions in Italy²,
- safety regions in the Netherlands,
- and cantons in Switzerland.

For Austria, the Czech Republic, Italy, and the Netherlands, it was feasible to collect data from the whole country (9, 14, 21, and 25 regions of analysis). From each other country, we took a stratified random sample of 15 regions of analysis³. The sample was stratified by the regions' number of COVID-deaths in the first wave, to ensure a sufficiently diverse sample and reduce the variance of our NPI effect estimator. To increase statistical power and reduce the influence of imported infections, we did not include regions with fewer than 2000 reported cases during the analysis period.

¹ Some regions do not have the legislative capacity to regulate gatherings in private homes. As an exception, we therefore also recorded prohibitions against private gatherings that were not legally binding. This was applicable in Italy, Austria, and the German state Northrhine-Westphalia. In each case, we confirmed with locals that these restrictions were followed by the vast majority of the population.

² With the exception of the Trentino-Alto Adige region, which was split into the autonomous provinces Trentino and Alto-Adige.

³ Each of the 16 German states had different regulations for its districts. To reduce the work required for data collection, we sampled the 15 districts only from the four largest states (Northrhine-Westphalia, Bavaria, Baden-Württemberg, Lower Saxony). These four states make up 60% of the population.

Data collection

To ensure high data quality, the NPI data was collected with semi-independent double entry and several validation steps. Each country was collected by two authors of this paper, who were provided with a detailed description of the NPIs. The researchers manually researched all dates by using internet searches and screening (local) government press releases, ordinances, and legislation. There was no automatic component in the data gathering process.

In the first round of data entry, the researchers initially collected the timeline of national NPI implementations. The researchers then compared their national timeline to the Oxford COVID-19 Government Response Tracker dataset (48) and, if there were any conflicts, visited all primary sources to resolve them. The data for each region of analysis were then entered by one of the two researchers, drawing on the national timeline and additional research. Several countries operated a tier or traffic light system that governed NPI implementation in subnational administrative divisions. For these countries, the researchers did not blindly enter the NPIs prescribed by the tier or traffic light system, but additionally consulted local government websites and media reports to investigate if the NPIs prescribed by the national system were, in fact, implemented in a region of analysis.

In the second round of data entry, every entry was independently entered again by another researcher. This researcher had access to the sources found in the first round as well as the associated quotes and comments, but not to the NPI data entered in the first round. This semi-independent double entry is similar to the validation used for parts of CoronaNet (24).

Finally, data from the two rounds of entry were compared and all conflicts resolved by discussion and by visiting primary sources. A researcher then manually compared the data from all countries to ensure the consistent application of NPI definitions across countries. We also validated the data against further external sources (e.g. (49) for Italy or (50) for England), contacted local epidemiologists when in doubt, and implemented a range of automated plausibility checks.

Throughout the data collection process, the researchers discussed edge cases and judgement calls on a shared online workspace, to ensure consistency across countries. As expected, the

validation process removed various sources of error and inconsistency. Supplementary text section 4 contains detailed explanations on coding decisions and judgement calls.

The total time spent on manual data collection, not including the design of the process, was 950 hours, with 185 hours on the national timelines, 470 hours on collecting the regions of analysis, and 290 hours on the validation steps.

Data pre-processing

To prevent bias, we excluded⁴ all observations in a region of analysis after the date when the variant of concern (VOC) B.1.1.7 first made up >10 % of infections in that region. Specifically, we excluded cases 5 days after >10 % of all infections were due to the VOC and deaths 11 days after this value was reached. We chose these values to ensure that on the last included day, more than 80% of the reported cases and deaths were generated **before** the VOC exceeded >10 % of all infections, according to our delay distributions. This only affected regions in England, usually towards the end of November (9).

The last day recorded in the intervention data set is 9th January 2021. Therefore, we included cases up to 5 days after this date and deaths up to 11 days after this date, as they are predominantly generated by infections before the 9th of January (see above).

Furthermore, to prevent influence from infections generated before the start of the analysis period, we excluded cases in the first 8 days (until 8th August) and deaths in the first 25 days (until 25 August). These values were chosen such that 80% of the cases and deaths recorded on the first observed day were generated in our window of analysis (including seeded infections), according to our delay distributions.

Supplementary text section 3.1 explains how we created the NPI features used in the modelling from the raw data. The final NPIs used for modelling are described in Table 2.

⁴ The observations were excluded by *masking* the observed cases and deaths, equivalent to removing these days from the likelihood/log-probability term. This ensures that these observations do not provide evidence, in the Bayesian sense, about our estimated NPI effects.

Table 2. NPI definitions.

NPI	Definition
Primary schools closed	Most or all primary schools (ages 5/6 to 10/11) have moved all teaching online or have closed (including for school holidays).
Secondary schools closed	Most or all secondary schools (ages 10/11 to 17/18) have moved all teaching online or have closed (including for school holidays).
Universities closed	Most or all higher education institutions are on (summer) term-break, (Christmas) vacation, or have sent students away from the university town (e.g., by closing university accommodation). As a result, a large fraction of students will have left their term-time accommodation to live at their home addresses. We did not count online teaching as a university closure if students were still expected to be present in the university town because (i) this still allows (likely considerable) transmission from students mixing outside of teaching events, and (ii) universities usually moved various components of their schedule online throughout the analysis period in a gradual manner. Some of the regions of analysis did not contain universities. For these, we counted universities as closed throughout the period of analysis.
Night clubs closed	Most or all nightclubs, discos, and other late night venues are closed.
Gastronomy closed	Most or all gastronomy establishments/venues (restaurants, pubs, cafes) are closed or limited to take-away.
Leisure and entertainment venues closed	A large fraction of leisure and entertainment venues are closed. Common examples include theaters, cinemas, concert halls, museums, gyms, dance studios, indoor skating rinks, bowling alleys, public baths, indoor play areas, escape games, casinos, billiard rooms, zoos, amusement parks.
Retail and close contact services closed	All non-essential retail shops are closed. Only those retail shops designated as essential may open; common examples are supermarkets, pharmacies, and gas stations. Additionally, all non-essential services that require close contact between customers and service providers are closed. This includes beauticians, nail salons, massage parlours, and - in all countries but Italy - hairdressers, but not medical services.
Night time curfew	Individuals must stay indoors during evenings/nights. There are exemptions for limited reasons, such as emergencies or caregiving.
Stricter mask-wearing	Mask-wearing is required in most or all shared/public spaces outside the home (inside and outside) where other people are present or where

policy	social distancing is not possible.
	Already before implementing this policy, all countries in our dataset had some less strict policies in place that required mask-wearing only in select public spaces (see Supplementary text section <i>feature_creation</i>). The estimated effectiveness of this NPI thus shows the additional benefit of the stricter policy.
Public gatherings limited to ≤ 30 , ≤ 10 , 2 people, or banned.	Gatherings in public spaces are limited to a certain number of people. The limits of 30 and 6 include all regulations with at least that level of strictness. For example, a ban on public gatherings of more than 15 people would be classified as “public gatherings limited to ≤ 30 people”.
Household mixing in private is limited to ≤ 30 , ≤ 10 , 2 people, or banned.	Gatherings of individuals in private spaces are limited to a certain number of people. See the row above for additional explanations.

Model

Please refer to the supplement for the model description.

Acknowledgements

We thank Fabian Valka for advice on the Austrian COVID response, Ilaria Dorigatti for advice on Italy, Natalie Claire Ceperley for advice on Switzerland, Veronika Nyvltova for help with Czech NPI data, Paul Hunter for sharing his UK tier data, Toby Philips for advice on NPIs implemented in the second wave. We thank Muhammed Razzak for his comments on the manuscript.

M. Sharma was supported by the EPSRC Centre for Doctoral Training in Autonomous Intelligent Machines and Systems (EP/S024050/1) and a grant from the EA Funds programme. S. Mindermann's funding for graduate studies was from Oxford University and DeepMind. C. Rogers-Smith was supported by a grant from Open Philanthropy. A.J. Norman was supported by the U.K. BBSRC and Open Philanthropy. J. Ahuja was supported by Open Philanthropy. J. T. Monrad was supported by the Augustinus Foundation, the Knud Højgaard Foundation, the William Demant Foundation, the Kai Lange and Gunhild Kai Lange Foundation, and the Aage and Johanne Louis-Hansen Foundation. G. Leech was supported by the UKRI Centre for Doctoral Training in Interactive Artificial Intelligence (EP/S022937/1). S.B. Oehm was supported by the Boehringer Ingelheim Fonds. L. Chindelevitch and S. Bhatt acknowledge funding from the MRC Centre for Global Infectious Disease Analysis (MR/R015600/1), jointly funded by the U.K. Medical Research Council (MRC) and the U.K. Foreign, Commonwealth and Development Office (FCDO), under the MRC/FCDO Concordat agreement; are part of the EDCTP2 program supported by the European Union; and acknowledge funding by Community Jameel. S. Flaxman acknowledges the EPSRC (EP/V002910/1) and the Imperial College COVID-19 Research Fund. J.M. Brauner was supported by the EPSRC Centre for Doctoral Training in Autonomous Intelligent Machines and Systems (EP/S024050/1) and by Cancer Research UK. S. Bhatt acknowledges The UK Research and Innovation (MR/V038109/1), the Academy of Medical Sciences Springboard Award (SBF004/1080), The MRC (MR/R015600/1), The BMGF (OPP1197730), Imperial College Healthcare NHS Trust- BRC Funding (RDA02), The Novo Nordisk Young Investigator Award (NNF20OC0059309) and The NIHR Health Protection Research Unit in Modelling Methodology. S. Bhatt thanks Microsoft AI for Health and Amazon AWS for computational credit

Author contributions

S. Mindermann, M. Sharma, J.M. Brauner, S. Mishra, S. Bhatt, Y. Gal, L. Chindelevitch, S. Flaxman conceived the research.

J.M. Brauner, S. Mindermann, M. Sharma, S. Bhatt, S. Mishra, B. Snodin, G. Altman, J. Ahuja, L. Finnveden, J.B. Sandbrink, G. Dhaliwal, J.T. Monrad, A.J. Norman, J. Kulveit, G. Leech, J.F. Sandkühler designed and conducted the NPI data collection.

M. Sharma, S.Mindermann, S. Mishra, J.M. Brauner, S.Bhatt, Y. Gal, G. Leech, S. Flaxman, L. Chindelevitch designed the model and modelling experiments.

M. Sharma, S. Mishra, C. Rogers-Smith, G. Leech, L.Finnveden performed and analyzed the modelling experiments.

S. Mindermann, S. Mishra, J.M. Brauner, S. Bhatt, J.T. Monrad, J. Kulveit did the literature review.

S. Mindermann, S. Bhatt, J.M. Brauner, M. Sharma, S. Mishra, J.T. Monrad, S.B. Oehm, G. Leech, A.J. Norman, C. Rogers-Smith, B. Snodin, G. Dhaliwal, G. Altman, J. Ahuja, J.B. Sandbrink wrote the manuscript.

All authors read, gave input on, and approved the final manuscript.

Competing interests

No conflicts of interests.

References

1. S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, M. Monod, Imperial College COVID-19 Response Team, A. C. Ghani, C. A. Donnelly, S. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, S. Bhatt, Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*. **584**, 257–261 (2020).
2. J. M. Brauner, S. Mindermann, M. Sharma, D. Johnston, J. Salvatier, T. Gavenčiak, A. B. Stephenson, G. Leech, G. Altman, V. Mikulik, A. J. Norman, J. T. Monrad, T. Besiroglu, H. Ge, M. A. Hartwick, Y. W. Teh, L. Chindelevitch, Y. Gal, J. Kulveit, Inferring the effectiveness of government interventions against COVID-19. *Science* (2020), doi:10.1126/science.abd9338.
3. S. Hsiang, D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L. Y. Huang, A. Hultgren, E. Krasovich, P. Lau, J. Lee, E. Rolf, J. Tseng, T. Wu, The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*. **584**, 262–267 (2020).
4. H. Salje, C. Tran Kiem, N. Lefrancq, N. Courtejoie, P. Bosetti, J. Paireau, A. Andronico, N. Hozé, J. Richet, C.-L. Dubost, Y. Le Strat, J. Lessler, D. Levy-Bruhl, A. Fontanet, L. Opatowski, P.-Y. Boelle, S. Cauchemez, Estimating the burden of SARS-CoV-2 in France. *Science*. **369**, 208–211 (2020).
5. M.-K. Looi, Covid-19: Is a second wave hitting Europe? *BMJ*. **371**, m4113 (2020).
6. S. Griffin, Covid-19: Second wave death rate is doubling fortnightly but is lower and slower than in March. *BMJ*. **371**, m4092 (2020).
7. T. K. Burki, Challenges in the rollout of COVID-19 vaccines worldwide. *Lancet Respir Med* (2021), doi:10.1016/S2213-2600(21)00129-6.
8. K. R. McCarthy, L. J. Rennick, S. Nambulli, L. R. Robinson-McCarthy, W. G. Bain, G. Haidar, W. P. Duprex, Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science*. **371**, 1139–1142 (2021).
9. E. Volz, S. Mishra, M. Chand, J. C. Barrett, R. Johnson, L. Geidelberg, W. R. Hinsley, D. J. Laydon, G. Dabrera, Á. O'Toole, R. Amato, M. Ragonnet-Cronin, I. Harrison, B. Jackson, C. V. Ariani, O. Boyd, N. J. Loman, J. T. McCrone, S. Gonçalves, D. Jorgensen, R. Myers, V. Hill, D. K. Jackson, K. Gaythorpe, N. Groves, J. Sillitoe, D. P. Kwiatkowski, S. Flaxman, O. Ratmann, S. Bhatt, S. Hopkins, A. Gandy, A. Rambaut, N. M. Ferguson, The COVID-19 Genomics UK (COG-UK) consortium, Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* (2021), doi:10.1038/s41586-021-03470-x.
10. E. C. Sabino, L. F. Buss, M. P. S. Carvalho, C. A. Prete Jr, M. A. E. Crispim, N. A. Fraiji, R. H. M. Pereira, K. V. Parag, P. da Silva Peixoto, M. U. G. Kraemer, M. K. Oikawa, T. Salomon, Z. M. Cucunuba, M. C. Castro, A. A. de Souza Santos, V. H. Nascimento, H. S. Pereira, N. M. Ferguson, O. G. Pybus, A. Kucharski, M. P. Busch, C. Dye, N. R. Faria, Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet*. **397**, 452–455 (2021).
11. N. Banholzer, E. van Weenen, A. Lison, A. Cenedese, A. Seeliger, B. Kratzwald, D. Tschernutter, J. P. Salles, P. Bottrighi, S. Lehtinen, S. Feuerriegel, W. Vach, Estimating the effects of non-pharmaceutical interventions on the number of new infections with COVID-19 during the first epidemic wave. *bioRxiv* (2021), doi:10.1101/2021.01.15.21249884.
12. Y. Liu, C. Morgenstern, J. Kelly, R. Lowe, CMMID COVID-19 Working Group, M. Jit, The impact of

- non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. *BMC Medicine*. **19** (2021), doi:10.1186/s12916-020-01872-8.
13. N. Haug, L. Geyrhofer, A. Londei, E. Dervic, A. Desvars-Larrive, V. Loreto, B. Pinior, S. Thurner, P. Klimek, Ranking the effectiveness of worldwide COVID-19 government interventions. *Nat Hum Behav*. **4**, 1303–1312 (2020).
14. M. Sharma, S. Mindermann, J. M. Brauner, G. Leech, A. B. Stephenson, T. Gavenčiak, J. Kulveit, Y. W. Teh, L. Chindelevitch, Y. Gal, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., 2020), 33, pp. 12175–12186.
15. YouGov, Personal measures taken to avoid COVID-19 (2020), (available at <https://yougov.co.uk/topics/international/articles-reports/2020/03/17/personal-measures-taken-a-void-covid-19>).
16. European Centre for Disease Prevention and Control, COVID-19 in children and the role of school settings in transmission - first update (2020).
17. R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-dannenburg, H. Thompson, P. G. T. Walker, H. Fu, A. Dighe, T. Jamie, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, D. Laydon, G. Nedjati-, Estimates of the severity of {COVID}-19 disease. *Lancet Infect. Dis.* **in press** (2020), doi:10.1101/2020.03.09.20033357.
18. D. J. Laydon, S. Mishra, W. R. Hinsley, P. Samartsidis, S. Flaxman, A. Gandy, N. M. Ferguson, S. Bhatt, Impact of the Tier system on SARS-CoV-2 transmission in the UK between the first and second national lockdowns. *medRxiv* (2021) (available at <https://www.medrxiv.org/content/10.1101/2021.02.23.21252277v1>).
19. N. G. Davies, R. C. Barnard, C. I. Jarvis, T. W. Russell, M. G. Semple, M. Jit, W. J. Edmunds, Centre for Mathematical Modelling of Infectious Diseases COVID-19 Working Group, ISARIC4C investigators, Association of tiered restrictions and a second lockdown with COVID-19 deaths and hospital admissions in England: a modelling study. *Lancet Infect. Dis.* (2020), doi:10.1016/S1473-3099(20)30984-1.
20. M. Manica, G. Guzzetta, F. Riccardo, A. Valenti, P. Poletti, V. Marziano, F. Trentini, X. Andrianou, A. M. Urdiales, M. Del Manso, Others, Impact of tiered restrictions on human activities and the epidemiology of the second wave of COVID-19 in Italy. *medRxiv* (2021) (available at <https://www.medrxiv.org/content/10.1101/2021.01.10.21249532v2.full-text>).
21. M. A. C. Vollmer, S. Mishra, H. J. T. Unwin, A. Gandy, T. A. Mellan, V. Bradley, H. Zhu, H. Coupland, I. Hawryluk, M. Hutchinson, O. Ratmann, M. Monod, P. Walker, C. Whittaker, L. Cattarino, C. Ciavarella, L. Cilloni, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, N. Brazeau, G. Charles, L. V. Cooper, Z. Cucunuba, G. Cuomo-Dannenburg, A. Dighe, B. Djaafara, J. Eaton, S. L. van Elsland, R. FitzJohn, K. Fraser, K. Gaythorpe, W. Green, S. Hayes, N. Imai, B. Jeffrey, E. Knock, D. Laydon, J. Lees, T. Mangal, A. Mousa, G. Nedjati-Gilani, P. Nouvellet, D. Olivera, K. V. Parag, M. Pickles, H. A. Thompson, R. Verity, C. Walters, H. Wang, Y. Wang, O. J. Watson, L. Whittles, X. Xi, A. A. M. Ghani, S. Riley, L. C. Okell, C. A. Donnelly, N. M. Ferguson, I. Dorigatti, S. Flaxman, S. Bhatt, A sub-national analysis of the rate of transmission of COVID-19 in Italy. *medRxiv* (2020), doi:10.1101/2020.05.05.20089359.
22. H. J. T. Unwin, S. Mishra, V. C. Bradley, A. Gandy, T. A. Mellan, H. Coupland, J. Ish-Horowicz, M. A. C. Vollmer, C. Whittaker, S. L. Filippi, X. Xi, M. Monod, O. Ratmann, M. Hutchinson, F. V. Harrison, H. Zhu, I. Hawryluk, P. Milton, K. E. C. Ainslie, M. Baguelin, A. Boonyasiri, N. F. Brazeau, L. Cattarino, Z. Cucunuba, G. Cuomo-Dannenburg, I. Dorigatti, O. D. Eales, J. W. Eaton, S. L. van Elsland, R. G. FitzJohn, K. A. M. Gaythorpe, W. Green, W. Hinsley, B. Jeffrey, E. Knock, D. J. Laydon, J. Lees, G.

- Nedjati-Gilani, P. Nouvellet, L. Okell, K. V. Parag, I. Siveroni, H. A. Thompson, P. Walker, C. E. Walters, O. J. Watson, L. K. Whittles, A. C. Ghani, N. M. Ferguson, S. Riley, C. A. Donnelly, S. Bhatt, S. Flaxman, State-level tracking of COVID-19 in the United States. *Nat. Commun.* **11** (2020), doi:10.1038/s41467-020-19652-6.
23. A. A. Schuessler, Ecological inference. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10578–10581 (1999).
24. C. Cheng, J. Barceló, A. S. Hartnett, R. Kubinec, L. Messerschmidt, COVID-19 Government Response Event Dataset (CoronaNet v.1.0). *Nat Hum Behav.* **4**, 756–768 (2020).
25. *Oxford COVID-19 Government Response Tracker (OxCGRT)* (<https://github.com/OxCGRT/covid-policy-tracker>).
26. K. Soltesz, F. Gustafsson, T. Timpka, J. Jalden, C. Jidling, A. Heimerson, T. B. Schon, A. Spreco, J. Ekberg, O. Dahlstrom, Others, On the sensitivity of non-pharmaceutical intervention models for SARS-CoV-2 spread estimation. *medRxiv* (2020), doi:10.1101/2020.06.10.20127324.
27. C. Winship, B. Western, Multicollinearity and Model Misspecification. *Sociol. Sci.* **3**, 627–649 (2016).
28. J. Twyman, Getting It Right: YouGov and Online Survey Research in Britain. *Journal of Elections, Public Opinion and Parties.* **18** (2008), pp. 343–354.
29. H. Ward, C. J. Atchison, M. Whitaker, K. E. C. Ainslie, J. Elliot, L. C. Okell, R. Redd, D. Ashby, C. A. Donnelly, W. Barclay, Others, Antibody prevalence for SARS-CoV-2 in England following first peak of the pandemic: REACT2 study in 100,000 adults. *medRxiv* (2020), doi:10.1101/2020.08.12.20173690.
30. J. Hasell, E. Mathieu, D. Beltekian, B. Macdonald, C. Giattino, E. Ortiz-Ospina, M. Roser, H. Ritchie, A cross-country database of COVID-19 testing. *Sci Data.* **7**, 345 (2020).
31. S. Mishra, J. Scott, H. Zhu, N. M. Ferguson, S. Bhatt, S. Flaxman, A. Gandy, A COVID-19 model for local authorities of the United Kingdom. *medRxiv* (2020), doi:10.1101/2020.11.24.20236661.
32. T. Fetzter, *Subsidizing the Spread of Covid19: Evidence from the UK's Eat-out To-help-out Scheme* (University of Warwick, Department of Economics, 2020).
33. F. Wong, J. J. Collins, Evidence that coronavirus superspreading is fat-tailed. *Proceedings of the National Academy of Sciences.* **117** (2020), pp. 29416–29418.
34. K. S. Kwon, J. I. Park, Y. J. Park, D. M. Jung, K. W. Ryu, J. H. Lee, Evidence of Long-Distance Droplet Transmission of SARS-CoV-2 by Direct Air Flow in a Restaurant in Korea. *J. Korean Med. Sci.* **35**, e415 (2020).
35. P. R. Hunter, J. S. Brainard, A. Grant, The effectiveness of the three tier system of local restrictions for control of COVID-19. *medRxiv* (2020), doi:10.1101/2020.11.22.20236422.
36. C. Atchison, L. R. Bowman, C. Vrinten, R. Redd, P. Pristerà, J. Eaton, H. Ward, Early perceptions and behavioural responses during the COVID-19 pandemic: a cross-sectional survey of UK adults. *BMJ Open.* **11**, e043577 (2021).
37. Y. Li, H. Campbell, D. Kulkarni, A. Harpur, M. Nundy, X. Wang, H. Nair, Usher Network for COVID-19 Evidence Reviews (UNCOVER) group, 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.* **21**, 193–202 (2021).
38. L. E. Smith, H. W. W. Potts, R. Amlôt, N. T. Fear, S. Michie, G. James Rubin, Adherence to the test, trace and isolate system: results from a time series of 21 nationally representative surveys in the UK

- (the COVID-19 Rapid Survey of Adherence to Interventions and Responses [CORSAIR] study) (2020), doi:10.1101/2020.09.15.20191957.
39. E. Otte Im Kampe, A.-S. Lehfeld, S. Buda, U. Buchholz, W. Haas, Surveillance of COVID-19 school outbreaks, Germany, March to August 2020. *Euro Surveill.* **25** (2020), doi:10.2807/1560-7917.ES.2020.25.38.2001645.
 40. J. P. Torres, C. Piñera, V. De La Maza, A. J. Lagomarcino, D. Simian, B. Torres, C. Urquidi, M. T. Valenzuela, M. O’Ryan, Severe Acute Respiratory Syndrome Coronavirus 2 Antibody Prevalence in Blood in a Large School Community Subject to a Coronavirus Disease 2019 Outbreak: A Cross-sectional Study. *Clin. Infect. Dis.* **955** (2020).
 41. C. Stein-Zamir, N. Abramson, H. Shoob, E. Libal, M. Bitan, T. Cardash, R. Cayam, I. Miskin, A large COVID-19 outbreak in a high school 10 days after schools’ reopening, Israel, May 2020. *Euro Surveill.* **25** (2020), doi:10.2807/1560-7917.ES.2020.25.29.2001352.
 42. L. Tondeur, Y. Madec, C. Besombes, M. N. Jolly, M.-N. Ungeheuer, I. Cailleau, L. Kuhmel, S. Temmam, C. Huon, K.-Y. Chen, Others, Cluster of COVID-19 in northern France: A retrospective closed cohort study. *medRxiv*, doi:10.1101/2020.04.18.20071134.
 43. A. Fontanet, R. Grant, M. Greve-Isdahl, D. Sridhar, Covid-19: Keeping schools as safe as possible. *BMJ.* **372**, n524 (2021).
 44. J. D. Munday, C. I. Jarvis, A. Gimma, K. L. M. Wong, K. van Zandvoort, S. Funk, W. J. Edmunds, C. C.-19 W. Group, Others, Estimating the impact of reopening schools on the reproduction number of SARS-CoV-2 in England, using weekly contact survey data. *medRxiv* (2021), doi:10.1101/2021.03.06.21252964.
 45. Y. Dong, X. Mo, Y. Hu, X. Qi, F. Jiang, Z. Jiang, S. Tong, Epidemiology of COVID-19 Among Children in China. *Pediatrics.* **145** (2020), doi:10.1542/peds.2020-0702.
 46. C. F. Dormann, J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell, S. Lautenbach, Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography.* **36**, 27–46 (2013).
 47. George E. P. Box, Sampling and Bayes’ Inference in Scientific Modelling and Robustness. *J. R. Stat. Soc. Ser. A.* **143**, 383–430 (1980).
 48. T. Hale, A. Petherick, T. Phillips, S. Webster, B. Kira, Oxford COVID-19 Government Response Tracker, Blavatnik School of Government. *Working Paper* (2020).
 49. M. Manica, G. Guzzetta, F. Riccardo, A. Valenti, P. Poletti, V. Marziano, F. Trentini, X. Andrianou, A. M. Urdiales, M. Del Manso, Others, Effectiveness of regional restrictions in reducing SARS-CoV-2 transmission during the second wave of COVID-19, Italy. *medRxiv* (2021), doi:10.1101/2021.01.10.21249532.
 50. P. R. Hunter, J. Brainard, A. Grant, The effectiveness of the three-tier system of local restrictions for control of COVID-19. *medRxiv* (2020), doi:10.1101/2020.11.22.20236422.