DOI: 10.1377/hlthaff.2020.01409
HEALTH AFFAIRS 40,
NO. 1 (2021): 70–81

By Ravindra Prasan Rannan-Eliya, Nilmini Wijemunige, J. R. N. A. Gunawardana, Sarasi N. Amarasinghe, Ishwari Sivagnanam, Sachini Fonseka, Yasodhara Kapuge, and Chathurani P. Sigera

Increased Intensity Of PCR Testing Reduced COVID-19 Transmission Within Countries During The First Pandemic Wave

Ravindra Prasan Rannan-Eliya (ravi@ihp.lk) is executive director of the Institute for Health Policy, in Colombo, Sri Lanka.

Nilmini Wijemunige is a research associate at the Institute for Health Policy.

J. R. N. A. Gunawardana is a research assistant at the Institute for Health Policy.

Sarasi N. Amarasinghe is a research associate at the Institute for Health Policy.

Ishwari Sivagnanam is a research associate at the Institute for Health Policy.

Sachini Fonseka is a research assistant at the Institute for Health Policy.

Yasodhara Kapuge is a clinical researcher at the Institute for Health Policy.

Chathurani P. Sigera is a research officer at the Institute for Health Policy.

ABSTRACT Experts agree that reverse transcription–polymerase chain reaction (PCR) testing is critical in controlling coronavirus disease 2019 (COVID-19), but decision makers disagree on how much testing is optimal. Controlling for interventions and ecological factors, we used linear regression to quantify testing's impact on COVID-19's average reproduction number, which represents transmissibility, in 173 countries and territories (which account for 99 percent of the world's COVID-19 cases) during March–June 2020. Among interventions, PCR testing had the greatest influence: a tenfold increase in the ratio of tests to new cases reported reduced the average reproduction number by 9 percent across a range of testing levels. Our results imply that mobility reductions (for example, shelter-in-place orders) were less effective in developing countries than in developed countries. Our results help explain how some nations achieved near-elimination of COVID-19 and the failure of lockdowns to slow COVID-19 in others. Our findings suggest that the testing benchmarks used by the World Health Organization and other entities are insufficient for COVID-19 control. Increased testing and isolation may represent the most effective, least costly alternative in terms of money, economic growth, and human life for controlling COVID-19.

No medication has proved effective in slowing transmission of the novel coronavirus (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19). Pending the widespread availability of vaccines, governments have relied on nonpharmaceutical interventions, including physical distancing, travel restrictions, and hygiene measures. After China sequenced and shared the viral genome, detection of active infection using reverse transcription–polymerase chain reaction (PCR) testing has been part of the response, extending beyond clinical care and epidemiological tracking. Country experience and simulations show that testing, tracing, and isolation can reduce transmission,[1,2] as substantial asymptomatic transmission occurs in COVID-19. The World Health Organization (WHO) has urged countries to "test, test, test"[3] and has suggested a rate of ten negative to one positive test results as an indicator of adequacy.[4]

Despite this, decision makers disagree on what constitutes adequate testing.[5,6] The legacy of pandemic influenza planning, which focused on reducing morbidity and mortality and never envisaged testing for controlling spread, may contribute to this disagreement, and most research focuses on other nonpharmaceutical interventions.[7–10] In a PubMed search we found

only thirty quantitative analyses of the impact of testing. Almost all involved modeling and simulation, and none quantified real-world impacts.[11] Research problems include difficulties of isolating impacts when multiple nonpharmaceutical interventions are implemented simultaneously; increases in per capita testing rates with cases, which makes them a poor indicator of testing strategy; lack of a global testing database; failure to control for confounding factors; and nonrepresentative geographical samples.[12]

Our study addressed these research problems by compiling data on numerous factors to quantify the association of PCR testing with COVID-19 spread during the initial pandemic wave—when some countries, such as China and New Zealand, achieved near-elimination—using a study design that robustly managed data gaps to maximize sample size and covariates.

Study Data And Methods

To model impacts on COVID-19 transmission, we adapted methods from previous epidemiological studies that assessed impact of interventions, such as school closures, on COVID-19 and other respiratory viruses.[13,14] Specifically, we estimated a cross-sectional, linear regression model of transmission intensity against the average intensities of interventions and other factors.

The online eAppendices provide a full description and justification of all of our data sources and methods.[15]

DATA Our unit of observation was all countries or territories with daily cases reported in the COVID-19 data repository published by the Center for Systems Science and Engineering at Johns Hopkins University, in Baltimore, Maryland.[16] We included subnational jurisdictions with substantial autonomy over health policies and borders, such as Macau, and territories, such as the US state of Hawaii, that are not part of the contiguous United States. We defined the observation period for each as starting from the date of peak incidence in March, typically March 28–31, and ending June 15, 2020.

We obtained data on daily PCR test numbers from multiple online sources, including the Our World in Data data repository,[17] other online collections, official communications, and news reports. We enumerated tests performed instead of persons tested, as few countries report persons tested. We interpolated data gaps and corrected inconsistencies, giving preference to official sources and Our World in Data, and we excluded data such as those from Venezuela, which substantially included antibody tests, as these do little to stop transmission.

We obtained daily data on eleven categorical COVID-19 policy indicators from the Oxford COVID-19 Government Response Tracker (OxCGRT).[18] These tracked containment and closure interventions, along with public information, testing, and contact tracing policies.

To track changes in individual movement, which may be voluntary or in response to government nonpharmaceutical interventions, such as lockdowns, we obtained mobility metrics from Google and Facebook[19,20] that were derived from locational data sent by mobile devices. As these were highly collinear, we combined them into two composite measures that proxy population-level mobility: increases in time spent at home and reductions in time spent in nonresidential locations (eAppendix 1).[15]

We found and collated data on other relevant factors, giving preferences to data sets with wide coverage, from the United Nations or other official agencies, and from research groups whose sources were well documented. We filled many remaining gaps by searching online sources, including government sites and news media. eAppendix 1 describes all variables and sources.[15]

For some interventions and factors, we obtained daily data. These included national school closures, the percentage of people wearing masks or face coverings, mask mandates, temperature, specific humidity, and relative humidity. For other factors, we obtained estimates for the most recent available year. These included population exposure to air pollution by particulate matter (particulate matter with diameter less than 2.5 micrometers, written as $PM_{2.5}$), physical inactivity, geographical latitude, distances to the major pandemic epicenters during March (Wuhan, China; Korea; Italy; and Iran), having a policy for universal Bacillus Calmette-Guérin (BCG) vaccination for tuberculosis, tuberculosis incidence, the number of severe acute respiratory syndrome (SARS) cases to proxy SARS experience, an index of democratic development, and two measures of health system capacity to manage infectious disease threats that have been developed by defense analysts: the Infectious Disease Vulnerability Index (IDVI)[21] and the Global Health Security Index (GHSI).[22]

As general controls, we sourced a range of socioeconomic and health system measures from the World Bank's World Development Indicators[23] and other sources, including gross domestic product (GDP) per capita, life expectancy, hospital beds, and poverty rates. As initial high rates might constrain countries from increasing testing faster than incidence, we added as controls the peak March incidence rate and the date on which the fiftieth COVID-19 case was reported.

OUTCOME OF INTEREST The time-varying effec-

tive reproduction number (R_{eff}) quantifies the transmissibility of a virus at any given time. It represents the average number of secondary infections generated by one infected person. In the absence of interventions or behavioral changes, the SARS-CoV-2 virus is highly infectious with a reproduction number in the range of 2 to 4, indicating that each case will typically infect that many other people.[1,2] Interventions that slow transmission reduce the reproduction number; if it falls below 1, incidence declines and eventually the virus will disappear. For our analysis, we consider the impact of interventions on the average level of the reproduction number during the study period. This is the average transmissibility of the virus, which is independent of the average number of cases.

Focusing on average transmissibility has two advantages. First, alternatives used in other studies, such as cumulative cases or incidence rates, are compatible with different averages of transmissibility, as they also depend on how transmissibility varies over time. For example, a country that employs an intervention for one month that reduces transmissibility below 1 and then abandons it in the next month will experience a U-shaped outbreak with fewer cumulative cases and a lower average incidence rate than another country that starts with the same initial incidence rate but employs the same intervention only during the second month, leading to an inverted-U-shaped outbreak. From our perspective, the intervention had equal effectiveness in both, but this will be evident only from the average reproduction number, which will be identical, and not from the raw case or incidence numbers. Second, within a linear regression model when the reproduction number is logged, the covariate coefficients can be directly interpreted as their percentage effect on the reproduction number. This is more meaningful from a policy perspective, as the key epidemiological goal in controlling an epidemic is to find a mix of interventions that reduce the reproduction number below one to achieve control.

We estimated the average reproduction number in each territory by adapting an approach[14] that approximates it by assuming that spread is exponential and that negligible numbers of people have immunity[24] (eAppendix 2).[15] Specifically, we computed the constant reproduction number that would have been required to change the incidence rate at the start in March to that on June 15. By assuming that the increase in incidence is exponential, we could obtain the average daily percentage increase in incidence. From this, we derived the average transmissibility or reproduction number by making an assumption about the average number of days it takes one person to infect the next (the generation interval). Incidence rates were derived from the center-weighted, seven-day moving average of new cases, and we followed the EpiForecasts group's assumption of a generation interval of 3.6 days.[25] External validation of our method confirmed no bias and close consistency with EpiForecasts estimates.

Our findings indicate that there is no single optimal level of testing.

INTERVENTIONS AND OTHER COVARIATES We quantified PCR testing intensity using the test-to-case ratio (TCR), defined as the ratio of tests to new cases reported. This controls for increases in test numbers with cases and increases in detected cases with testing, but it is also inflated by multiple tests on the same person. It aligns with the WHO benchmark[4] and is comparable to others using test positivity rates, such as those adopted by the Trump administration.[26] The TCR proxies overall intensity, which depends on policies for testing international arrivals, contact tracing and testing of contacts, repeat testing in detection and clinical care, and symptomatic thresholds for community testing, but it does not reflect inefficiencies, such as reporting delays, which reduce impact on transmission. In the absence of global data on these details, the TCR was the most comparable metric.

We quantified exposure variables, which include interventions, ecological factors, and other controls, by their mean daily value during an exposure period that lagged the observation period by seven days. This was based on estimates of an incubation period of two to twelve days[27] and a delay between symptom onset and case reporting of two to seven days.[28]

ANALYSIS AND MODELS Many variables had missing values, with the frequency ranging from zero for most general indicators and 5 percent for TCR to 34 percent for Google mobility variables. Half of the territories lacked data for at least one variable. As values were not missing completely at random, analyzing only territories with complete data would result in a small sample, biased results, and reduced precision. To overcome this, we used multiple imputation to impute missing values, estimating all models with 300 multiply imputed data sets. We also

log-transformed TCR, per capita GDP, the peak March incidence rate, and other variables to ensure normality as desirable in multiple imputation. Full details are in eAppendix 3.[15]

Of 221 countries and territories in our data set, we included in analysis only 173 that reported more than 100 COVID-19 cases (for a listing, see eAppendix 1),[15] as small case numbers make incidence estimates unreliable. These accounted for 98 percent of the world's population and 99 percent of COVID-19 cases (data not shown).

We predefined two lists of covariates: exposures to be retained on the basis of theory and prior evidence, and exposures without strong evidence. The former were TCR, mobility changes, school closures, and mask usage. The latter were BCG policy, tuberculosis incidence, latitude, temperature, specific humidity, IDVI and GHSI, and per capita GDP and life expectancy as general controls. Other covariates formed a third category. We included the OxCGRT policy indicators in this third category, as they were highly correlated with the mobility measures, which reflect actual behavior, and added little explanatory power.

We evaluated covariates for model inclusion using stepwise backward selection. In model 1 we retained the lists of exposures to be retained on the basis of theory and prior evidence and exposures without strong evidence, and we eliminated others based on model fit (adjusted R^2) and coefficient meaningfulness and significance at univariate p value < 0.20. In model 2 we prioritized parsimony to avoid overfitting and forced inclusion of the list of exposures to be retained on the basis of theory and prior evidence only.

The Institutional Review Board of the Institute for Health Policy, in Colombo, Sri Lanka, assessed the study as not requiring full ethics review, as all data were anonymized, aggregate, and publicly available. Analyses were performed in Stata 14.2, with some data processing done using R. The statistical significance level was set at 5 percent, and statistical tests were two-tailed.

LIMITATIONS Our analysis has several limitations. First, as an observational study, it cannot support causal inferences, and relationships remain associations. Second, reported cases understate actual incidence, and testing increases detection. Although using ratios of incidence rates removes country differences, it cannot eliminate temporal changes, but as countries with higher TCRs increased these levels faster (see eAppendix 5, supplementary exhibit S5),[15] any bias is likely downward. In addition, our data cannot differentiate imported from local cases, and our estimates of transmission are biased upward when imported cases predominate. Third, our analysis cannot account for heterogeneity within large countries—for example, with regard to temperature—which reduces precision. Fourth, we cannot account for factors that our data do not capture, such as physical distancing or differences in isolation strategies, although our extensive covariates mitigate this. Fifth, the exposure lag might not adequately control for endogeneity arising from countries intensifying interventions in response to increased cases. Finally, time series analysis would be better but was not possible because of data limitations.

Study Results

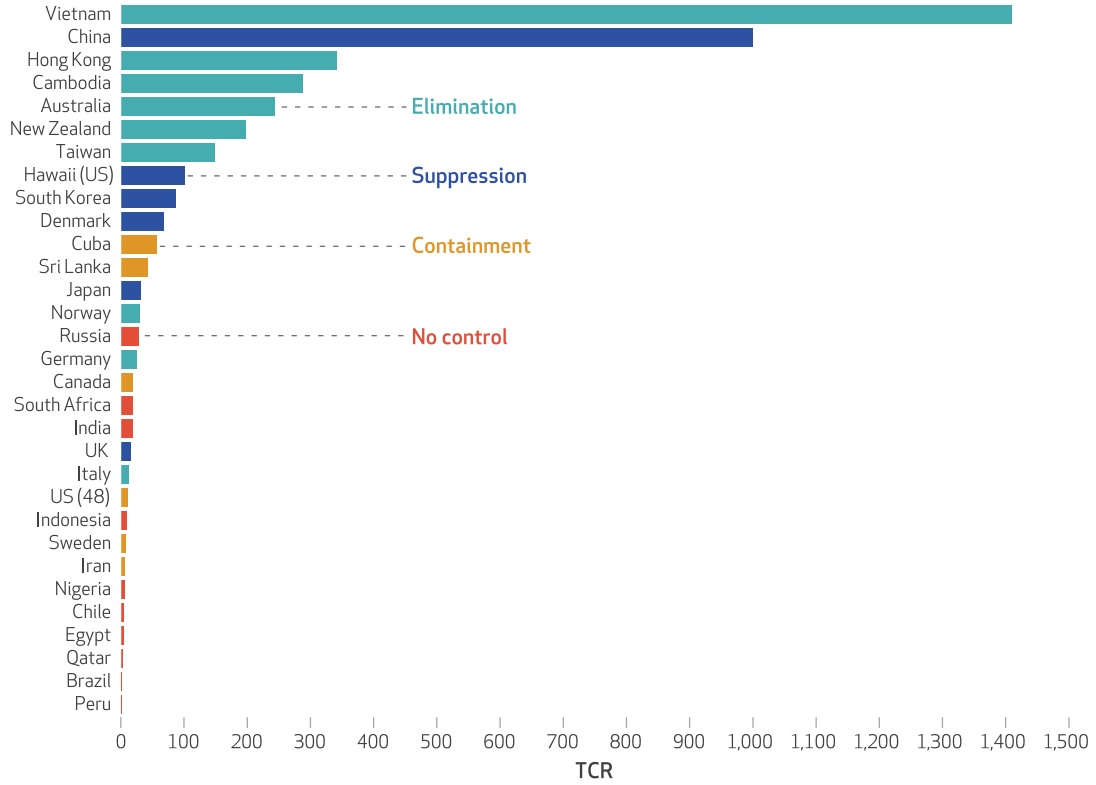
DESCRIPTIVE INFORMATION During March–June 2020 COVID-19 transmissibility fell globally, with the median reproduction number declining from more than 2.5 in early March to fluctuate above 1.0 during April–June (eAppendix 5, supplementary exhibit S1).[15] The falls in transmissibility varied by region. The World Bank's East Asia and Pacific region, which groups East Asian and Pacific Island countries, Australia, and New Zealand, was most successful in controlling COVID-19 spread, with the average reproduction number falling to 0.9 and several territories coming close to eliminating the virus. In contrast, in several regions and countries, such as South Asia and Brazil, the average reproduction number remained above 1.1 with no slowdown in spread.

Among interventions to control spread, testing intensity showed the largest variation, with the test-to-case ratio varying from 1 to more than 1,400 across countries (exhibit 1). Greater testing intensity was also associated with lower incidence, cumulative cases, and deaths (eAppendix 5, supplementary exhibits S5–S6).[15] Among other interventions (eAppendix 5, supplementary exhibits S2–S4, S8),[15] there was much uniformity in the imposition of school closures and little variation by income level in mobility reductions and lockdown measures. Mask usage varied considerably across countries, with the country average being greatest in Latin America, and the country averages for lockdown measures and mobility reductions were least in the East Asia and Pacific region.

IMPACT OF TESTING, INTERVENTIONS, AND OTHER FACTORS Our models regress COVID-19 transmissibility (represented by the natural logarithm of the average reproduction number) against the average intensity of the included interventions and factors. The estimated coefficients convey the size of impact of each factor, and by exponentiating these coefficients, we obtain their percentage effect on the reproduction number. The models also provide us with 95 per-

EXHIBIT 1

COVID-19 test-to-case ratios (TCRs) in selected countries and territories during the study period



SOURCE Authors' estimates based on data obtained from multiple sources as described in eAppendix 1 (see note 15 in text), with approximate estimate for China based on limited public data. **NOTES** The study period varies by country or territory and is from the date of peak incidence in March to June 15, 2020 (for details, see eAppendix 1). Countries and territories are categorized by their relative success in controlling COVID-19 spread, as reflected in their average reproduction number (R). The "elimination" category represents those where levels were compatible with elimination, or $R < 0.90$, the "suppression" category represents those with $0.90 \leq R < 0.95$, the "containment" category represents those with $0.95 \leq R < 1.05$, and the "no control" category represents those with $R \geq 1.05$. US (48) indicates the contiguous 48 states.

cent confidence intervals for the effect sizes.

Our final model (model 2) fits the data well, explaining 81 percent (adjusted $R^2 = 0.81$) of the variation in average COVID-19 transmissibility across countries (full details on these estimates are in eAppendix 5, supplementary exhibit S10).[15]

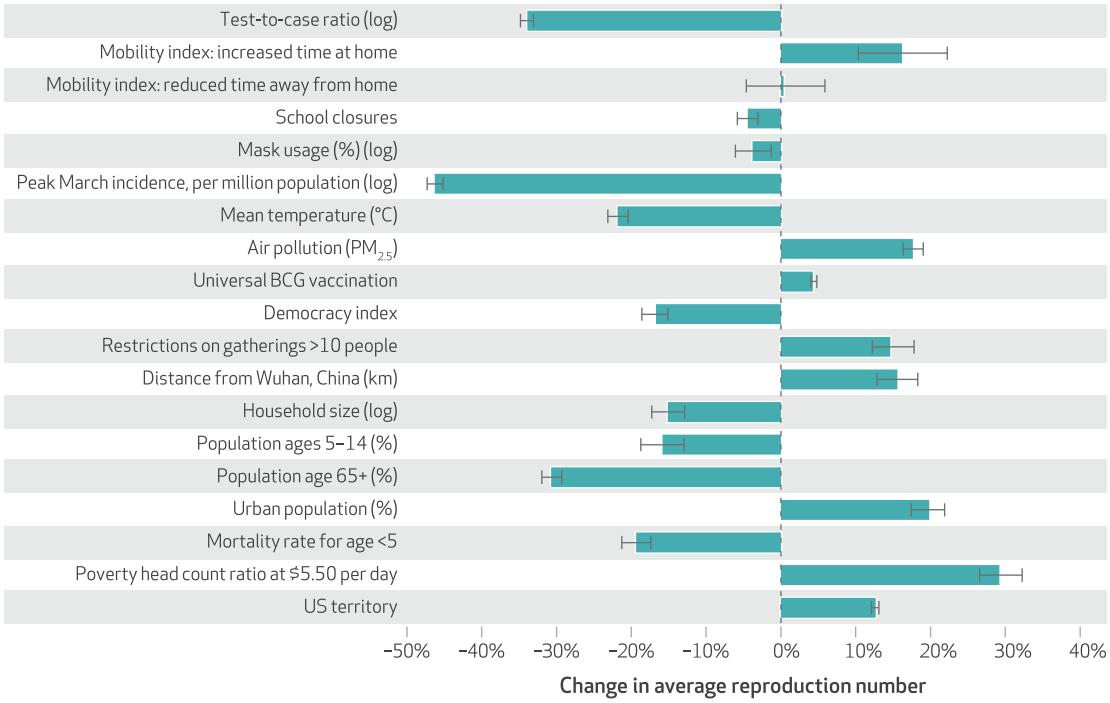
Of all intervention measures, testing intensity was the most influential and was highly significant ($p < 10^{-16}$). Its effect is logarithmic, so a tenfold increase in the TCR would reduce the average reproduction number by 8.6 percent (95% CI: 6.8, 10.3), and a hundredfold increase would reduce it by 16.4 percent (95% CI: 13.1, 19.6; computed from the model 2 estimates reported in eAppendix 5, supplementary exhibit S10).[15] Because TCR levels varied so much between countries, this translates into the largest relative impact of all intervention measures. This is shown in exhibit 2, which illustrates the relative impacts of key interventions and factors. In contrast to the effect estimates discussed here and shown in eAppendix 5, supplementary exhibit S10,[15] exhibit 2 uses the corresponding standardized coefficients, which allow direct comparison of their relative effects.

None of the other intervention measures was statistically significant ($p > 0.05$), although school closures and mask use were associated with reductions in transmissibility. Increased time spent at home was associated with increased transmissibility, although it was not statistically significant ($p = 0.15$) (data not shown). This was expected, as epidemiologists assume that increased time at home leads to more transmission within households.[29] However, there was no reduction in transmissibility associated with reduced time spent in nonresidential locations, implying that the mobility changes usually associated with lockdowns increased overall transmission globally, although none of these effects was statistically significant.

Several of the ecological factors were associated with substantial and statistically significant

EXHIBIT 2

Estimated relative effects of selected interventions and factors on COVID-19 transmissibility



SOURCE Authors' calculations of effects in sampled countries and territories during the study period, using estimates from the final model (model 2), as described in the text. **NOTES** The figure displays the standardized coefficients for each factor. These coefficients represent the percentage effect on transmissibility, represented by the average reproduction number (R), of a 1-standard-deviation increase in each factor, which allows direct comparison of their relative effect sizes. The whiskers indicate the twenty-fifth to seventy-fifth percentiles or interquartile range (IQR) of the multiple imputation (MI) estimates of these standardized effects. Confidence intervals are not displayed, as they cannot be reliably derived for standardized effects using MI. Because having universal Bacillus Calmette-Guérin (BCG) vaccination or being a US territory (defined here to include both the US states and external US territories, such as Guam) are binary outcomes for which standardized effects are not meaningful, the bars for these categories represent their unadjusted effects and not the standardized effects, with the whiskers indicating the IQR of the MI estimates to provide some comparability with the other variables. Both the unadjusted estimates with confidence intervals and the standardized estimates with IQRs are in eAppendix 5, supplementary exhibits S10 and S11 (see note 15 in text). The original coefficient for the test-to-case ratio was significant ($p < 10^{-16}$), as was that of gathering restrictions ($p < 0.02$), but for the other interventions, none was significant at the 0.10 level. $PM_{2.5}$ is a measure of fine particulate matter (air pollution).

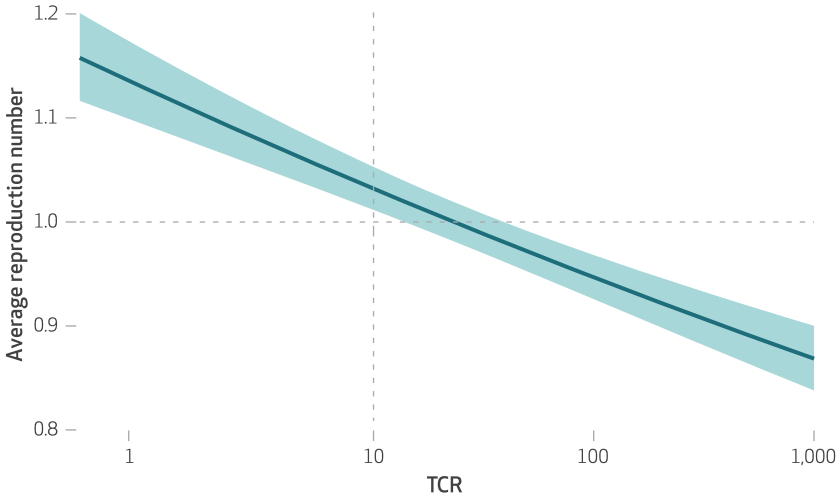
protective effects. A 1ºC increase in temperature and a 1 percent increase in the share of the population that was elderly (age sixty-five or older) reduced transmissibility by 0.4 percent (95% CI: 0.2, 0.5; $p < 10^{-4}$) and 0.7 percent (95% CI: 0.3, 1.0; $p < 10^{-4}$), respectively. Increases in the share of the population that was ages 5–14 were also protective, but this was not statistically significant ($p = 0.15$). In contrast, air pollution, urbanization, and poverty were associated with statistically significant increases in transmission. A 1 micrometer per cubic meter increase in fine particulate air pollution ($PM_{2.5}$), a 1 percent increase in the urban share of the population, and a 1 percent increase in the percentage of the population living below the $5.50 international poverty line were associated with 0.1 percent (95% CI: 0.0, 0.2; $p = 0.002$), 0.1 percent (95% CI: 0.0, 0.2; $p = 0.003$), and 0.1 percent (95% CI: 0.0, 0.2; $p = 0.02$) increases in transmissibility, respectively.

In contrast, physical inactivity, specific and relative humidity, latitude, and the various measures of health systems capacity, such as health spending, hospital beds, and the IDVI and GHSI indices were not associated with any statistically significant effects ($p > 0.20$) during model building (results not shown). Our general controls of life expectancy and per capita GDP also exhibited no relationship with transmission (joint F-test: $p = 0.70$). All were dropped from the final model.

Two results were unexpected. The March peak incidence rate was highly influential, with a tenfold increase associated with a 7.4 percent (95% CI: 5.6, 9.1; $p < 10^{-6}$) reduction in transmissibility during the subsequent three months, and universal BCG vaccination was associated with a 4.2 percent increase (95% CI: 1.0, 7.7; $p = 0.01$).

EXHIBIT 3

Simulated effects of changes in testing intensity on COVID-19 transmissibility



SOURCE Authors' analysis using estimates from the final model (model 2), developed as described in the text. **NOTES** The solid line indicates predicted levels of transmissibility, as represented by the average reproduction number, at different levels of testing intensity (test-to-case ratio, or TCR) when using model 2 results to simulate outcomes in a hypothetical country with median characteristics and levels of other interventions. The shaded region indicates the 95% confidence interval. Estimates are displayed using a logarithmic scale.

In addition, increases in the days that public gatherings of more than ten people were banned ($p = 0.015$) and increased distance from the Wuhan epicenter ($p = 0.02$) were significantly associated with increases in transmissibility. Being a US territory (defined here to include both the US states and external US territories) was also significantly associated with a 12.7 percent (95% CI: 6.5, 19.2; $p < 10^{-4}$) increase in transmissibility.

SIMULATIONS We explored our results using counterfactual simulations (for full details, see eAppendix 5).[15] In a hypothetical country with median levels of other interventions and characteristics, increasing testing intensity to TCR levels of 60 and above would have reduced the average reproduction number significantly below 1 (exhibit 3). Simulations also showed that increased testing might have reduced the reproduction number close to or below 1 in many countries or territories, including Peru, Chile, and Indonesia, where lockdowns failed to achieve this (exhibit 4). Other simulations indicated that better performance in the East Asia and Pacific region was driven primarily by greater testing and that other regions might have done much better with a similar mix of testing, masks, and mobility restrictions (eAppendix 5, supplementary exhibit S9).[15]

ROBUSTNESS CHECKS We undertook a range of robustness checks to assess possible limitations (eAppendix 4).[15] These checks evaluated changes in parameter assumptions and samples, including the observation period, the exposure lag, and the case threshold for including territories. Throughout, our results remained robust, meaningful, and statistically significant. We evaluated our choice of multiple imputation by reestimating models using only complete cases (countries and territories with no missing data), which indicated that a complete case analysis would have produced biased overestimates of testing's impact. As linear regression assumes linear relationships, we also reestimated our results by specifying the logged TCR with a restricted cubic spline to allow it to have a nonlinear relationship, which showed that our estimated effect holds across TCR levels of 1–1,000 (eAppendix 5, supplementary exhibit S7).[15]

We also investigated our failure to detect net benefits from mobility reductions associated with lockdown measures; we were able to detect net benefits, although not statistically significant ones, only in the World Bank's Europe and Central Asia region. We also confirmed a beneficial impact in the group of eleven European countries where Seth Flaxman and colleagues[8] previously reported that lockdown reduced transmission, indicating that our study is picking up real geographical differences in lockdown impacts (eAppendix 4).[15]

Discussion

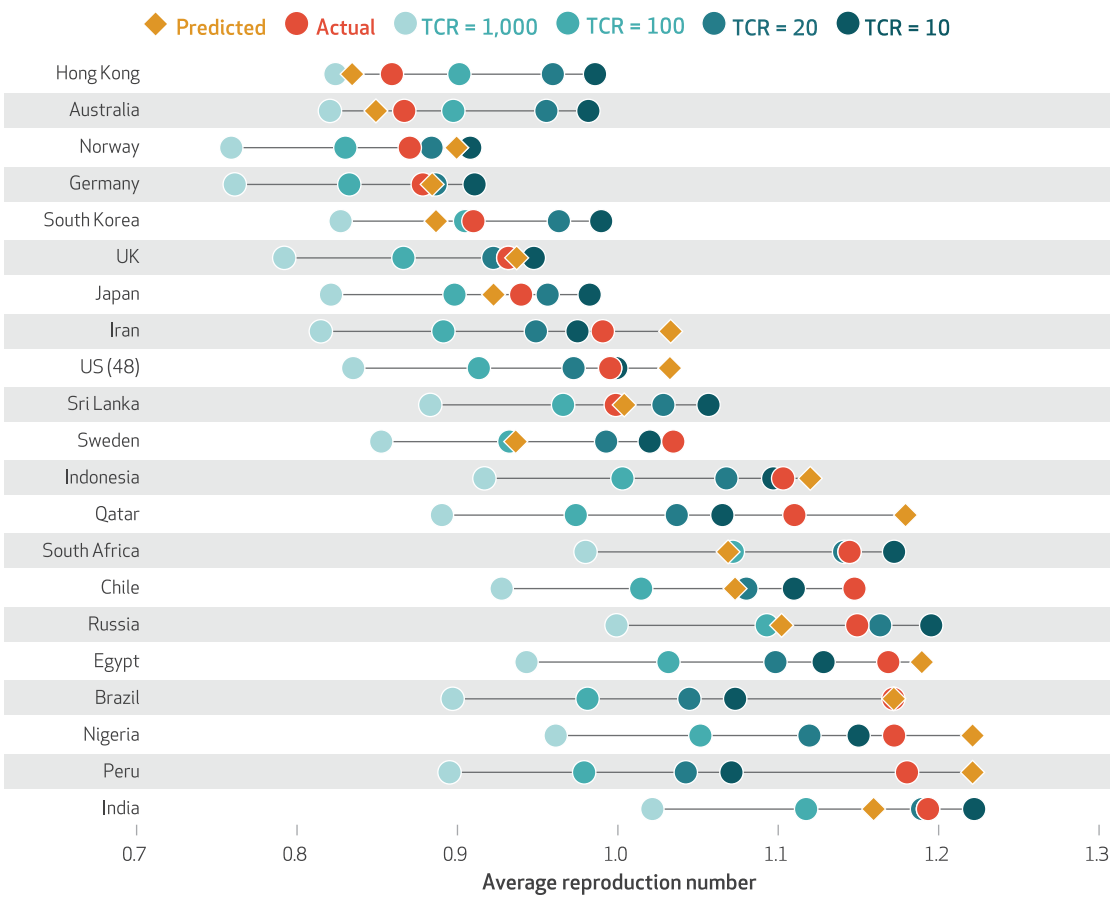
Unlike previous studies, our analysis explicitly quantifies testing for COVID-19. Its strengths are its comprehensive global sample, its accounting for both interventions and ecological factors, and extensive controls that minimize selection and omitted variable bias. These may explain why our results differ from those of previous studies.

Although our finding of a strong association between testing intensity and transmissibility cannot prove causality, the robustness of the relationship across countries and test-to-case ratio levels is consistent with a known mechanism. Around half or more of COVID-19 transmission is caused by people who are asymptomatic or who have only minor symptoms; only increases in PCR testing make it possible to increase detection and isolation of infectious cases and then increase the numbers of their potentially infectious contacts who are isolated. This remains the only known approach that blocks person-to-person transmission sufficiently to stop the epidemic.[1]

We lacked systematic data on who actually is tested in different countries and what happens after testing, so we did not adjust for this. However, as exhibit 5 illustrates, differences in test-

EXHIBIT 4

Predicted effects of changes in testing intensity on COVID-19 transmissibility in selected countries and territories



SOURCE Authors' analysis using estimates from the final model (model 2), developed as described in the text. **NOTES** Colored symbols indicate the predicted levels of COVID-19 transmissibility, as represented by the average reproduction number, at different levels of testing intensity (test-to-case ratio, or TCR) during the study period. The red circles represent the estimated actual average reproduction number during the study period; the teal circles represent the predicted changes from that value at TCR levels of 10, 20, 100, and 1,000. The diamonds indicate the average reproduction number as predicted by model 2, for reference. US (48) indicates the contiguous 48 states.

ing rates between countries are associated with significant differences in how testing is targeted and acted on and its likely impact on transmission. Where testing intensity was low, as in the US (TCR = 11) and the UK (TCR = 15), testing mostly diagnosed and isolated the most symptomatic cases; this use of testing has only a limited impact on transmission. Only with greater testing can countries screen wider circles of asymptomatic case contacts and lower the symptomatic threshold for testing individuals without obvious exposure. At the most intensive testing levels (TCR = 100–1,500), countries were either actively encouraging anyone with respiratory symptoms or fever to get tested or routinely testing such patients in addition to testing international arrivals and isolating and testing all case contacts. Such countries tolerated very low positivity rates, to increase detection of cases in the community. To the extent that such differences explain differences in testing rates, they provide the link between higher testing rates and more effective control of COVID-19 transmission.

Our findings on the effects of temperature, air pollution ($PM_{2.5}$), and age structure confirm the results of previous studies (see eAppendix 1).[15] The role of air pollution may warrant more attention, as it may explain higher transmission levels in countries such as India and Nepal, although we note that we used mostly 2018 pollution estimates, which overestimate 2020 levels in most countries. However, our inability to detect additional effects of humidity and latitude suggests that they do not have substantial effects independent of temperature.

Our results strongly reject previous findings that BCG vaccination is protective. We speculate that earlier studies failed to adequately control

EXHIBIT 5

COVID-19 testing practices in selected countries and territories during the study period, by average reproduction number

Country/territory	TCR	Symptomatic threshold for testing of individuals without exposure history	Contact tracing	International arrivals
Elimination ($R < 0.90$)				
Vietnam	1,410	A/FORS testing available and allowed	Extensive contact tracing: close contacts quarantined and tested	All tested, mandatory 14dQI
Hong Kong	342	FORS patients routinely tested in all public clinics and some private clinics	Extensive contact tracing: close contacts isolated, tested if symptomatic	All tested, mandatory 14dI
Australia	244	A/FORS encouraged to test: high levels of voluntary uptake	Extensive contact tracing: close contacts isolated and tested	Only if symptomatic during mandatory 14dQ, but shift toward universal testing, varying by state
Taiwan	148	A/FORS encouraged to test	Extensive contact tracing: close contacts isolated and tested if symptomatic	Only if symptomatic during mandatory 14dQ
Germany	25	A/FORS encouraged to test: uptake low to modest	Extensive contact tracing: close contacts quarantined and tested	Mandatory 14dI
Suppression ($0.90 \leq R < 0.95$)				
Hawaii (US)	102	Modest access to testing for A/FORS through private initiatives	Extensive contact tracing: close contacts isolated, but not tested	Mandatory 14dQI, but no universal testing
UK	15	Symptomatic cases not encouraged to test; limited community access	Contact tracing limited or suspended: close contacts not tested	No testing, no quarantine
Containment ($0.95 \leq R < 1.05$)				
Sri Lanka	42	A/FORS not encouraged to test with limited testing in practice	Extensive contract tracing: close contacts isolated and tested	All tested, mandatory 14dQI
US (48)	11	Capacity constraints typically limited community access to all but the most severely symptomatic; long delays common in receiving test results	Contact tracing decentralized with uneven and limited capacity: guidelines recommend testing of close contacts, but only partially implemented	No testing or isolation/ quarantine, except in a few states
Sweden	7	A/FORS not encouraged to test: limited community access	Contact tracing limited or largely suspended: close contacts not tested	No testing, no quarantine
No control ($R \geq 1.05$)				
India	18	Varying criteria for testing by state: in practice, restrictions, cost, availability, and social factors result in low uptake	National policy to trace and isolate close contacts, but implementation varies, with many places abandoning efforts: partial coverage and testing not the norm	Symptomatic arrivals tested, but isolation and quarantine policies vary by state from 7 to 14 days
Indonesia	9	Access limited and not encouraged; private testing expensive	National policy to trace, isolate, and test close contacts: in practice, limited capacity and coverage	Arrivals to self-isolate, but no routine testing (after June, tested if not tested 1 week before arrival)
Peru	0.8	Limited community access and polymerase chain reaction testing	Contact tracing limited: close contacts not tested	Mandatory 14dQ, but no testing

SOURCE Authors' assessment based on review of news media reports, official communications, and research literature. **NOTES** Details describe the typical situation during the study period only. Countries and territories are categorized by level of control achieved according to their average reproduction number (R). US (48) indicates the contiguous 48 states. TCR is test-to-case ratio. A/FORS denotes individuals who are asymptomatic or have fever or respiratory symptoms. 14dQI is 14-day quarantine/ isolation.

for confounding factors. Our finding that economic and health capacities have no relationship with transmission suggests that national resources are not usually a constraint, and instead it is the strategies that countries choose that matter. Surprisingly, the Global Health Security Index, which claims to assess countries' capability to prevent and mitigate epidemics and pandemics, and the Infectious Disease Vulnerability Index, which assesses countries' vulnerability to transnational infectious disease outbreaks, exhibited no relationship with COVID-19 transmissibility, indicating that they are poor measures of capacity and vulnerability and that other country characteristics need to be looked at.

The findings that locations closer to Wuhan and with worse COVID-19 outbreaks in March did significantly better in subsequent control of transmission are intriguing. We speculate that in places that were closer to or confronted the initial epidemic earlier, this engendered more fear and forceful reactions by governments and societies, which our data do not capture, than in places that had time to habituate to the threat.

The estimated effects of our two mobility measures are consistent with previous research in that they confirm an increase in transmission associated with time spent at home, but their overall effect, which increases transmission except in Europe, was not. Although neither effect was statistically significant, this may explain an anomaly in current knowledge. Studies have found that lockdowns and mobility reductions slowed COVID-19 transmission in Europe and North America,[8–10] but no empirical analysis with adequate controls has demonstrated net benefits at the global level. Problematically, many countries elsewhere, such as India, Indonesia, Peru, and Chile, failed to slow the epidemic with stringent lockdowns. In addition, a recent analysis of OxCGRT data that evaluated the global impact of nonpharmaceutical interventions found that the evidence on impact of several lockdown-related interventions was inconsistent and inconclusive.[12] We offer three linked explanations, noting that household transmission accounts for a substantial part of transmission in most countries. First, in developing regions, where personal living space is less, there may be larger increases in transmission at home during lockdowns. Second, outside developed regions, the necessities of subsistence may make it more difficult for people to remain in their homes. Third, home confinement was only effective in achieving epidemic control in Wuhan when residents were tested and quarantined if positive to prevent them infecting other household members.[30] In many countries with ineffective lockdowns, the mechanisms of testing, quarantine, and household support may be inadequate to obtain substantial benefits. Given the economic costs, better understanding of the performance of lockdowns in developing countries should be a global research priority.

Policy Implications

Our findings indicate that there is no single optimal level of testing. At any level, increases in testing further reduce transmission (exhibit 3). When incidence is high and uncontrolled, all measures, including testing, might need to be intensified to achieve control and to make widespread testing and tracing feasible. When the virus is close to elimination or the reproduction number is substantially below 1, increases in testing could be traded for relaxing other interventions, such as school and work closures, mask wearing, and social distancing, as aptly demonstrated by the sustained return to normalcy in countries with intensive testing, such as China, New Zealand, and Vietnam.

At the same time, almost all countries that reduced transmission to levels compatible with elimination were testing at TCR levels of 100 and above (exhibit 1). This implies that most benchmarks suggested by the WHO, the US government, and other agencies are inadequate. Given the effectiveness of whatever other interventions they were doing, most countries would likely have needed a TCR of at least 100 to achieve epidemic control. eAppendix 5, supplementary exhibit S6,[15] which shows that the bulk of all deaths occurred in countries that tested at the WHO benchmark level or lower, underlines this.

Our results imply that in strategies to "flatten the curve," which originated in the CDC's pandemic influenza planning,[7] critical care capacity is the wrong threshold to target for COVID-19. At high incidence rates, even the wealthiest nations, such as the US, UK, and Qatar, cannot expand testing and tracing fast enough (or they give up altogether) to achieve epidemic control. Early and continuous aggressive testing to keep incidence within the capacity to test, trace, and isolate may be the best way to flatten the curve.

Conclusion

We provide empirical evidence that testing intensity was the common factor explaining the success of countries or territories that achieved near-elimination, such as China, Cambodia, and New Zealand, and the most important predictor of performance elsewhere. Given the costs and uncertainties associated with other non-

pharmaceutical interventions, a strategy that relies much more on increased testing and isolation deserves serious consideration and resource allocation outside the East Asia and Pacific region. It is likely to be less costly in terms of money, economic growth, and human life. ■

The authors thank Saroj Jayasinghe, Rob Condon, Owen O'Donnell, Delan Devakumar, David Scheerer, and two anonymous reviewers for their comments and helpful observations and Anabella Pinton for valuable research assistance. This study was not supported by any external funder. The Institute for Health Policy expresses its appreciation to members of the research team who accepted reductions in compensation during their work on the study at a time of financial stringency for the institute. An unedited version of this article was published online December 2, 2020, as a Fast Track Ahead Of Print article. That version is available in the online appendix.

NOTES

1 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.

2 Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. Lancet Public Health. 2020;5(5):e279–88.

3 World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19—16 March 2020 [Internet]. Geneva: WHO; 2020 Mar 16 [cited 2020 Dec 2]. Available from: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---16-march-2020

4 World Health Organization. COVID-19—virtual press conference—30 March 2020 [Internet]. Geneva: WHO; 2020 Mar 30 [cited 2020 Dec 2]. Available from: https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-30mar2020.pdf?sfvrsn=6b68bc4a_2

5 Graham D, Oré D. Mexico's coronavirus death surge puts testing regime under the microscope. Reuters [serial on the Internet]. 2020 Jun 16 [cited 2020 Dec 2]. Available from: https://www.reuters.com/article/us-health-coronavirus-mexico-tests/mexicos-coronavirus-death-surge-puts-testing-regime-under-the-microscope-idUSKBN23N392

6 Hogan L. Fighting alone. Washington Post [serial on the Internet]. 2020 Jul 16 [cited 2020 Dec 2]. Available from: https://www.washingtonpost.com/outlook/2020/07/16/larry-hogan-trump-coronavirus/

7 Centers for Disease Control and Prevention. Interim pre-pandemic planning guidance: community strategy for pandemic influenza mitigation in the United States: early, targeted, layered use of non-pharmaceutical interventions [Internet]. Atlanta (GA): CDC; 2007 Feb [cited 2020 Dec 2]. Available from: https://stacks.cdc.gov/view/cdc/11425

8 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;584(7820):257–61.

9 Nouvellet P, Bhatia S, Cori A, Ainslie KEC, Baguelin M, Bhatt S, et al. Reduction in mobility and COVID-19 transmission [Internet]. London: Imperial College London; 2020 Jun 8 [cited 2020 Dec 2]. (Report 26). Available from: https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-06-08-COVID19-Report-26.pdf

10 Courtemanche C, Garuccio J, Le A, Pinkston J, Yelowitz A. Strong social distancing measures in the United States reduced the COVID-19 growth rate. Health Aff (Millwood). 2020; 39(7):1237–46.

11 We searched PubMed on August 29, 2020, for studies evaluating the impact of testing on COVID-19 spread, using keywords including "COVID-19," "2019-nCoV," "SARS-CoV-2," "transmission," "intervention," and "testing," and we identified other studies not reported in PubMed. This yielded 3,179 published studies and preprints, of which 30 reported quantitative analyses of the impact of testing (with or without tracing and isolation).

12 Liu Y, Morgenstern C, Kelly J, Lowe R, Jit M. The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. medRxiv [serial on the Internet]. 2020 Aug 12 [cited 2020 Dec 2]. Available from: https://www.medrxiv.org/content/10.1101/2020.08.11.20172643v1

13 Cowling BJ, Lau EHY, Lam CLH, Cheng CKY, Kovar J, Chan KH, et al. Effects of school closures, 2008 winter influenza season, Hong Kong. Emerg Infect Dis. 2008; 14(10):1660–2.

14 Huang KE, Lipsitch M, Shaman J, Goldstein E. The US 2009 A(H1N1) influenza epidemic: quantifying the impact of school openings on the reproductive number. Epidemiology. 2014;25(2):203–6.

15 To access the appendix, click on the Details tab of the article online.

16 Johns Hopkins University. COVID-19 data repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [Internet]. Baltimore (MD): Johns Hopkins University; 2020 [cited 2020 Dec 2]. Available from: https://github.com/CSSEGISandData/COVID-19

17 Ritchie H, Ortiz-Ospina E, Beltekian D, Mathieu E, Hasell J, Macdonald B, et al. Coronavirus (COVID-19) testing [Internet]. Cambridge (MA): Our World in Data; 2020 [last updated 2020 Dec 2; cited 2020 Dec 2]. Available from: https://ourworldindata.org/coronavirus-testing

18 Hale T, Angrist N, Kira B, Petherick A, Phillips T, Webster S. Variation in government responses to COVID-19 [Internet]. Oxford: University of Oxford, Blavatnik School of Government; 2020 May [cited 2020 Dec 2]. (BSG Working Paper Series). Available from: https://www.bsg.ox.ac.uk/sites/default/files/2020-05/BSG-WP-2020-032-v6.0.pdf

19 Google. COVID-19 Community mobility reports [database on the Internet]. 2020 Nov 29 [cited 2020 Dec 2]. Available from: https://www.google.com/covid19/mobility/

20 Facebook Data for Good. Movement range maps [serial on the Internet]. 2020 [cited 2020 Dec 2]. Available from: https://dataforgood.fb.com/tools/movement-range-maps/

21 Moore M, Gelfeld B, Okunogbe AT, Paul C. Identifying future disease hot spots: Infectious Disease Vulnerability Index [Internet]. Santa Monica (CA): RAND Corporation; 2016 [cited 2020 Dec 2]. Available from: https://www.rand.org/pubs/research_reports/RR1605.html

22 Nuclear Threat Initiative, Johns Hopkins Bloomberg School of Public Health. Global Health Security Index: building collective action and accountability [Internet]. Baltimore (MD): Johns Hopkins Bloomberg School of Public Health; 2019 Oct [cited 2020 Dec 2]. Available from: https://www.ghsindex.org/wp-content/uploads/2020/04/2019-

Global-Health-Security-Index.pdf
23 World Bank. Indicators [Internet]. Washington (DC): World Bank Group; 2020 [cited 2020 Dec 2]. Available from: https://data.worldbank.org/indicator/
24 Various serological surveys indicate that even in most affected places, with the exception of New York, less than 5 percent of the population would have been infected before June 2020.
25 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 [version 1; peer review: awaiting peer review]. Wellcome Open Research [serial on the Internet]. 2020 Jun 1 [cited 2020 Dec 2]. Available from: https://wellcomeopenresearch.org/articles/5-112
26 White House. Opening up America again [Internet]. Washington (DC): White House; 2020 [cited 2020 Dec 2]. Available from: https://www.whitehouse.gov/wp-content/uploads/2020/04/Guidelines-for-Opening-Up-America-Again.pdf
27 Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann Intern Med. 2020;172(9):577–82.
28 Glöckner S, Krause G, Höhle M. Now-casting the COVID-19 epidemic: the use case of Japan, March 2020. medRxiv [serial on the Internet]. 2020 Mar 23 [cited 2020 Dec 2]. Available from: https://www.medrxiv.org/content/10.1101/2020.03.18.20037473v1
29 Ferguson N, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand [Internet]. London: Imperial College London; 2020 [cited 2020 Jun 10]. (Report 9). Available from: https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-16-COVID19-Report-9.pdf
30 Hao X, Cheng S, Wu D, Wu T, Lin X, Wang C. Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. Nature. 2020;584(7821):420–4.