medRxiv preprint doi: https://doi.org/10.1101/2020.09.26.20201814; this version posted September 28, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Evolution of COVID-19 cases in selected low- and middle-income countries: past the herd immunity peak?

Axel S. Lexmond[1], Carlijn J.A. Nouwen† and J. Paul Callan*†*

All authors contributed equally to this work.
[1] Department of Engineering, University of Pretoria, South Africa
** Corresponding authors: nouwencarlijn@gmail.com; paul_callan@post.harvard.edu*
† These authors did this work in their personal capacities.

25 September 2020

We have studied the evolution of COVID-19 in 12 low- and middle-income countries in which reported cases have peaked and declined rapidly in the past 2-3 months. In most of these countries the declines happened while control measures were consistent or even relaxing, and without signs of significant increases in cases that might indicate second waves. For the 12 countries we studied, the hypothesis that these countries have reached herd immunity warrants serious consideration. The Reed-Frost model, perhaps the simplest description for the evolution of cases in an epidemic, with only a few constant parameters, fits the observed case data remarkably well, and yields parameter values that are reasonable. The best-fitting curves suggest that the effective basic reproduction number in these countries ranged between 1.5 and 2.0, indicating that the curve was "flattened" in some countries but not "suppressed" by pushing the reproduction number below 1. The results suggest that between 51 and 80% of the population in these countries have been infected, and that between 0.05% and 2.50% of cases have been detected – values which are consistent with findings from serological and T-cell immunity studies. The infection rates, combined with data and estimates for deaths from COVID-19, allow us to estimate overall infection fatality rates for three of the countries. The values are lower than expected from reported infection fatality rates by age, based on data from several high-income countries, and the countries' populations by age. COVID-19 may have a lower mortality risk in these three countries (to differing degrees in each country) than in high-income countries, due to differences in immune-response, prior exposure to coronaviruses, disease characteristics or other factors. We find that the herd immunity hypothesis would not have fit the evolution of reported cases in several European countries, even just after the initial peaks – and subsequent resurgences of cases obviously prove that those countries have infection rates well below herd immunity levels. Our hypothesis that the 12 countries we studied have reached herd immunity should now be tested further, through serological and T-cell-immunity studies.

Introduction

Figure 1, produced by researchers at Imperial College London[1], illustrates how case numbers are expected to evolve during an epidemic under different conditions. The red curve shows expected cases for an uncontrolled outbreak. This curve has three main features: (1) initial exponential growth in new cases, (2) a single peak as the population reaches herd immunity, (3) an exponential decline in new cases. The green curve shows expected cases when governments and people take measures to control disease spread, but those measures are not sufficient to reduce the effective basic reproduction number (R_{0_e})[2]

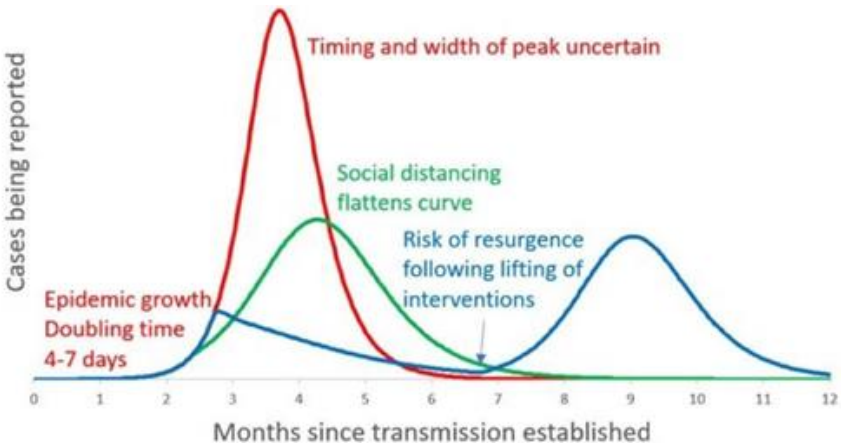
[1] Source: https://www.imperial.ac.uk/news/195976/individual-response-covid-19-important-government-action/ .
[2] Three distinct reproduction numbers characterise the disease in different situations. First, the basic reproduction number, R_0, is the average number of new infections caused by each current infected individual, that would happen in the absence of any disease control measures by governments or individuals, and before there is widespread immunity. Even this number is not a universal characteristic, because it can vary by location, even in the absence of control measures, due to differences in typical numbers of social contacts in different places. Second, we define the effective basic reproduction number, R_{0_e}, as the adjusted basic reproduction number after control measures are taken, but without significant levels of immunity. This number can, therefore, change from time to time, as governments change policies and people change their practices. Third, the effective reproduction

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.26.20201814; this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

below 1. The green curve displays largely the same features as the red – disease spread is only halted by herd immunity – but the curve is "flattened", with cases spread out more over time and fewer cases at the peak. The blue curve shows the evolution of cases when containment measures are sufficient to bring R_{0_e} below 1. The curve is "crushed"; the initial exponential growth in case numbers is halted and cases decline (at a slower rate than for the red or green curves). In this case, because most people are not immune to the disease, it is possible for the disease to return if containment measures later allow R_{0_e} to increase above 1, as shown in the blue curve in the Figure 1.

Figure 1: Typical outbreak curves in different scenarios.[1]



Patterns in reported cases

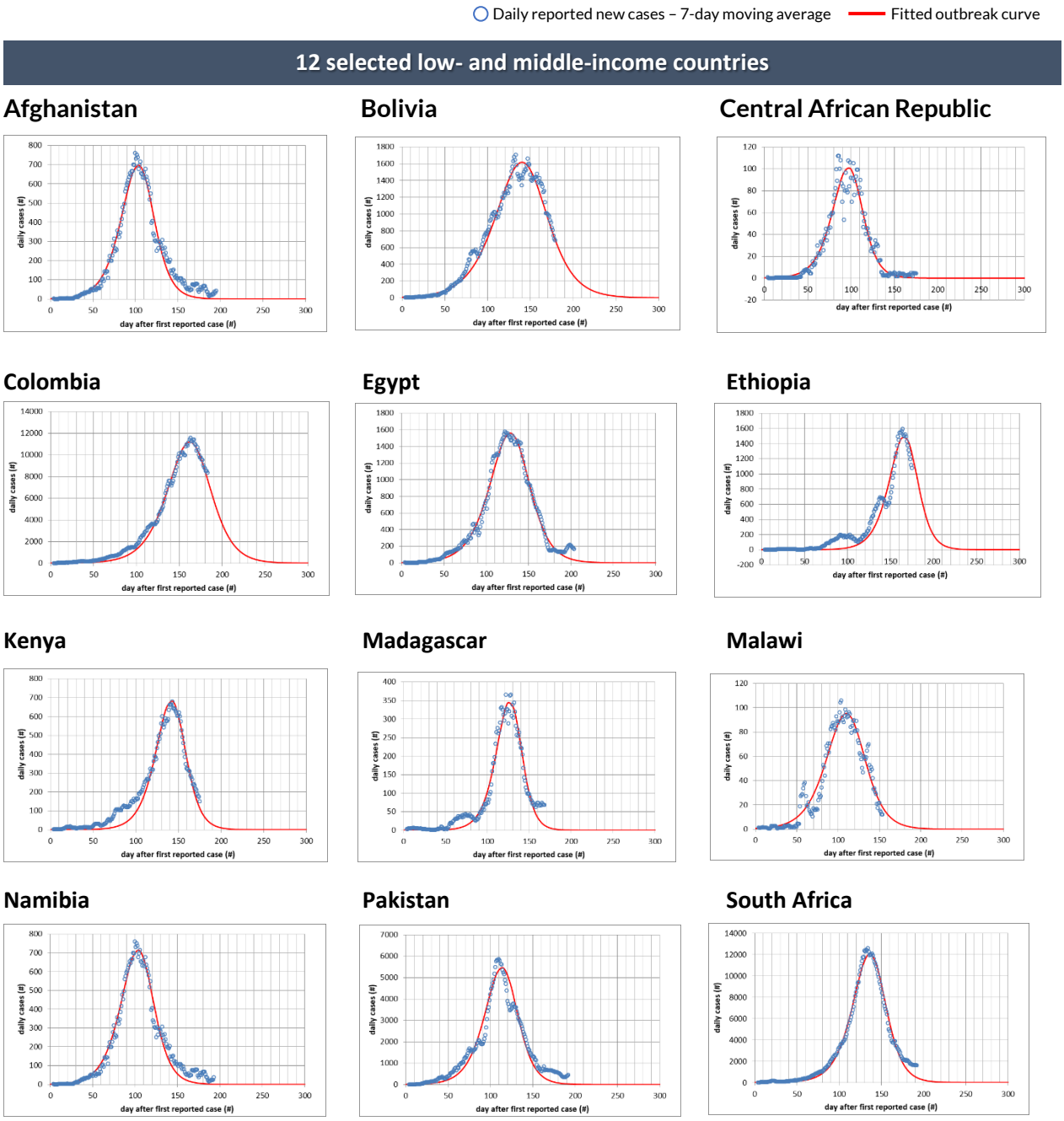
Reported cases in several low- and middle-income countries (LMICs) have evolved in a manner that is very similar to the red and the green curves of Figure 1. We show the 7-day rolling average of reported new cases for 12 such countries in Figure 2. Other countries show similar patterns; we have chosen to study a subset with the clearest similarities to expected outbreak curves for which R_{0_e} remains above 1. In all these countries, the reported cases have (1) grown exponentially, (2) reached a single clear peak and (3) declined exponentially. Regulations were most stringent, and compliance was greatest, in most of these countries, after the designation of the global pandemic in March and have relaxed to varying degrees in recent months – but cases continued to decline. None of these countries has reported a significant increase in new cases after the peak that would indicate a second wave (although cases in some countries have only recently passed the peak). Together, these observations point to a hypothesis that the outbreaks in these countries have reached herd immunity, and that the recently observed declines in new cases are because many people have already been infected and are immune – at least temporarily.

However, the numbers of cases in other countries – including most high-income countries (HICs) but also some LMICs – show patterns that are much different. Figure 3 shows 7-day rolling averages of reported cases for 6 such comparison countries. In these countries, cases have evolved in a manner that is similar to the first part of the blue curve of Figure 1. There have been peaks in numbers of reported cases, yet the decline is often longer and slower than the red or green curves would suggest. In some countries, there have been resurgences in cases, indicating that the initial suppression was not due to high levels of immunity.

number, R_e, is the actual average number of new infections caused by each current infected individual, which decreases as the number of people with immunity increases (and also depends on disease control policies and practices).

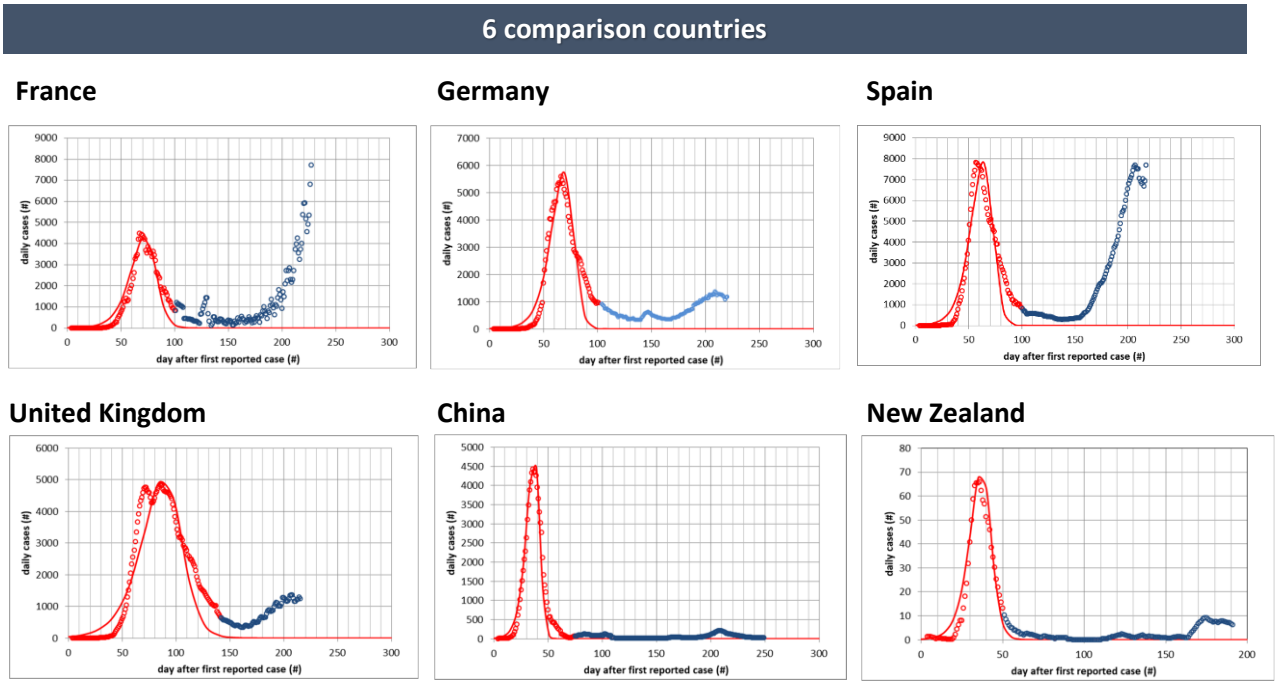
medRxiv preprint doi: https://doi.org/10.1101/2020.09.26.20201814; this version posted September 28, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 2: Daily reported new cases of COVID-19 (dots, based on a 7-day rolling average), for selected low- and middle-income countries, with best-fit curves from simple disease spread model (line).



medRxiv preprint doi: https://doi.org/10.1101/2020.09.26.20201814; this version posted September 28, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 3: Daily reported new cases of COVID-19 (based on a 7-day rolling average, the dots) for 6 comparison countries, with best-fit curves for first peak in these countries (the red line). Note that the red dots are datapoints that are included in the fits and the blue dots represent datapoints which are not used in fitting.



Fit with disease outbreak model and estimation of outbreak parameters

We test our hypothesis of herd immunity in these 12 countries by using a simple disease outbreak model and fitting to the reported cases. The outbreak case curves are those described by a linearised Reed-Frost model[3], the textbook deterministic mathematical model for an epidemic. The curves produced by this model depend on just two parameters, namely the effective basic reproduction number (R_{0_e}) and the generation time (t_g), which is the average time from infection of one person and when that person infects other people. Expected reported cases are calculated by scaling the Reed-Frost model's results by a detection rate (p)[4]. The parameters that produce the best-fit curve for reported cases in each country are determined partly analytically from the observed data and partly from least squares regression. Our model parameters (R_{0_e}, t_g and p) are constant over time – a beneficial assumption in that it avoids having too many free parameters, which might lead to good fits even if the model incorrectly describes the disease dynamics. Furthermore, in most of the countries studied, reported cases and reported deaths have followed similar trends (with changes in deaths lagging the corresponding changes in cases), even for countries with very low absolute numbers of reported cases and deaths – which suggests that the shapes of the curves likely reflect trends in actual cases and deaths, and that detection rates for both do not vary wildly over time.

The best-fit curves are shown with red lines in Figure 2, for each of the 12 LMICs studied, and the corresponding parameters are presented in Table 1. The fits are close, with R-squared goodness-of-fit

[3] See, for example, Kenneth J. Rothman, Epidemiology, Ch. 6, or other textbooks on epidemiology. Our linearized version of the Reed-Frost model uses the formula $n_{i+1} = n_i * R_{0_e} * s_i / n_{tot}$, where n_i is the number of infected people for the time period i, s_i is the number of susceptible people without immunity for the time period i, and n_{tot} is the total population. The generation time t_g is the duration of each time period in the model. Expected reported cases are equal to $p * n_i$ for each time step i in the model, where p is the case detection rate.

[4] One more parameter is required in fitting the curve, namely the difference between the actual first case and the first reported case. This parameter can only shift the model curve to earlier or later times, without changing the shape of the curve and without impacting any of the other parameters.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.26.20201814; this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

measures between 0.93 and 0.99. For South Africa, R_{0_e} and t_g are determined directly from the slopes and width of the observed data, and the regression produces a fit with R-squared of 0.99 with only one fitting parameter, namely the detection rate p which scales the Reed-Frost model results but does not alter the shape of the curve. For the other 11 countries, we use the value of t_g derived analytically for South Africa, and R_{0_e} and p are used as fitting parameters. These results demonstrate that the observed case patterns can indeed very accurately be described by an exponential outbreak halted by herd immunity.

The effective basic reproduction numbers, R_{0_e}, range between 1.5 (in Bolivia) and 2.0 (in Madagascar). Estimates for the basic reproduction number, R_0, the "natural" rate in the absence of social distancing, for SAR-CoV-2 (the virus that causes COVID-19) in Wuhan at the outset of the global epidemic, range from 1.4[5] up to 5.7[6]. R_0 might be expected to be higher in low-income countries due to factors such as dense living conditions, lack of access to clean water and sanitation facilities, and inability of most people to work from home. Thus, our findings suggest that, for the 12 LMICs studied, social distancing measures and practices likely reduced the effective basic reproduction number and slowed the spread of the disease – but with R_{0_e} above 1, they did not "crush the curve".

Table 1: Outbreak parameters for selected countries, which correspond to the best-fit curves shown in Figure 2.

Country	**Effective basic reproduction number (R_{0_e})**	**Detection rate (p)**	**% of population infected**	**R-squared for fitted curve** [perfect fit = 1]
Afghanistan	1.8	0.14%	72%	0.97
Bolivia	1.5	2.10%	51%	0.98
Central African Republic	1.9	0.13%	76%	0.93
Colombia	1.6	2.50%	52%	0.99
Egypt	1.6	0.16%	64%	0.99
Ethiopia	1.9	0.07%	67%	0.94
Kenya	1.9	0.08%	74%	0.96
Madagascar	2.0	0.07%	80%	0.96
Malawi	1.6	0.05%	63%	0.94
Namibia	1.8	2.09%	72%	0.97
Pakistan	1.8	0.18%	72%	0.94
South Africa	1.7	1.56%	71%	0.99

The numbers presented for R_{0_e} in Table 1 represent best estimates[7]. There are ranges of reasonably possible values because the effects of R_{0_e} and t_g on the observed case curves are hard to distinguish – especially when the observed data has more "noise" or when the observed data does not include many points after the peak, and when values of R_{0_e} are lower[8]. For example, for South Africa, R_{0_e} has a best-

[5] J Riou and CL Althaus, "Pattern of early human-to-human transmission of Wuhan 2019-nCoV," Euro Surveillance 2020;25(4):pii=2000058. https://doi.org/10.2807/1560-7917.ES.2020.25.4.2000058
[6] S Sanche et al., "High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2," Emerg Infect Dis. 2020;26(7):1470-1477. https://dx.doi.org/10.3201/eid2607.200282
[7] For the estimated values of R_{0_e}, we assume that the analytically derived t_g for South Africa, of 8 days, applies to all the other countries. Further details can be found in the supplementary materials.
[8] For a given shape of the curve of disease cases over time – characterised by the exponential rates of increase before the peak and of decrease after the peak and the width of the curve – there is a unique combination of R_{0_e} and t_g which leads to the observed curve. However, for any given curve, there are a range of combinations of R_{0_e} and t_g which can produce very similar curves. (This issue is why it is very hard to determine the basic reproduction number, R_0, for any new disease, even when the case doubling time is well-known.) In the case of South Africa, the observed data is sufficiently "clean" that we are able to calculate R_{0_e} and t_g. For the other countries, the observed data is more "noisy" or does not have much data after the peak, and there can be multiple combinations of R_{0_e} and t_g which produce reasonable fits to the observed data. For example, for Kenya, there are reasonable fits to observed data for the full range of viable possible values of t_g between 5 days and 15 days (given the timeframes after infection

medRxiv preprint doi: https://doi.org/10.1101/2020.09.26.20201814; this version posted September 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

estimate value of 1.74, and a range of reasonably possible values between 1.45 and 1.90. However, all of the realistically possible values of R_{0_e} and t_g produce similar shapes of the case curve, for each of the countries, with the same implication that total infections in the countries are above herd immunity levels[9].

Detection rates are estimated to be very low, ranging from 2.5% in Colombia to 0.05% in Malawi. These low detection rates explain how herd immunity could be reached despite low numbers of reported cases, relative to total population, in all the countries studied. These low detection rates are not surprising. Serological testing results in Kenya, Pakistan and South Africa suggest that the number of people with coronavirus antibodies substantially exceeds the reported cases – by factors of up to 500 in Kenya (based on data in mid-May)[10], of up to 370 in Pakistan (based on data from May to July)[11] and of 20-40 in South Africa (based on data in August)[12], which would correspond to detection rates of 0.02% for Kenya, 0.27% for Pakistan and 2.5-5.0% for South Africa. Furthermore, there is evidence that serological tests might underestimate infection rates and immunity by up to a factor of 2, because more people develop T-cell-mediated immunity than show antibodies[13].

In all these countries, the analysis indicates that over half of the population has been infected and has become immune – at least temporarily. Current total infection rates derived from the fitted curves range from 51% in Bolivia to 80% in Madagascar[14]. Note that the infection rates required for herd immunity normally quote the percentage of population infected at the peak of the curve, but significant numbers of people continue to be infected after this point, even as the numbers of new cases decline.

The infection fatality rate, or the percentage of deaths from COVID-19 among those infected with the SARS-CoV-2 virus, can be estimated for countries with reliable estimates of deaths. For Bolivia, Colombia and South Africa[15], the infection fatality rates calculated from reported deaths divided by the total number of infections derived from our analysis, are 0.13%, 0.09% and 0.04%, respectively. In these three countries, estimates have been made of excess deaths due to natural causes, and, if all of these excess deaths are due to COVID-19, the infection fatality rates for the three countries could be up to 0.50%, 0.12% and 0.11%, respectively[16]. All three countries are expected to have a lower overall infection fatality rate, compared to European countries, because their populations have a higher share of young people,

---

during which a person can transmit the disease), and the corresponding range of possible values of R_{0_e} is between 1.4 and 2.8; all associated with an R^2 of 0.973. See the supplementary materials for more details.

[9] As R_{0_e} determines the herd immunity threshold, it remains of interest to "untangle" R_{0_e} and t_g, which is possible when reported case data are fairly "clean", or which may be achieved by experimental confirmation of either parameter, as described in the supplementary materials.

[10] KEMRI/Wellcome Trust, Preliminary Report of SARS-CoV-2 antibody prevalence among blood donors in Kenya, 28 June 2020.

[11] S Zaidi et al. Seroprevalence of anti-SARS-CoV-2 antibodies in residents of Karachi—challenges in acquiring herd immunity for COVID 19, *Journal of Public Health*, fdaa170, https://doi.org/10.1093/pubmed/fdaa170.

[12] According to Dr. Zweli Mkhize, Minister for Health of South Africa, as reported by AP: https://apnews.com/cffcd4dfb1e3cbd810838fb9bde7a91d.

[13] T Sekine et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. Cell (2020), doi: https://doi.org/10.1016/j.cell.2020.08.017.

[14] All numbers apply to the portion of a country that has had some level of exposure. If specific physically concentrated groups in a country are fully shielded, the total proportion of the national population that has been infected to date, would be lower than the numbers listed here. For most countries, the numbers are as of 7 September; for some, they are as of 5 or 6 September.

[15] The other countries studied have not reported excess mortality, and are likely under-report deaths from COVID-19 significantly, and by differing proportions in each country. As a result, we limit our estimates of infection fatality rates to the cases of Bolivia, Colombia and South Africa.

[16] For all three countries, we use reported deaths as of 21 September. We estimate the infection fatality rates expected if all excess deaths from natural causes are due to COVID-19, from the ratios of excess deaths to reported deaths as of the dates for which estimates of excess deaths are available. For Bolivia, as of 31 August, reported deaths due to COVID-19 were 5,027 and underreported deaths attributed to COVID-19 were 14,508, according to the Servicio de Registro Cívico (SERECI) (as reported in https://muywaso.com/14-mil-muertes-que-no-se-pueden-ocultar-y-una-tasa-de-subregistro-del-75/). For Colombia, from 18 May to 2 August, reported deaths due to COVID-19 were 12,393 and total excess deaths were 15,728, as reported by Minsalud (https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/VSP/estimacion-exceso-mortalidad-colombia-covid19.pdf). For South Africa, reported deaths due to COVID-19 were 15,499 as of 15 September, while excess deaths due to natural causes (which may not all be due to COVID-19 but are assumed to be for this calculation) between 6 May and 15 September are estimated to be 44,481 by the South African Medical Research Council (SAMRC) (https://www.samrc.ac.za/reports/report-weekly-deaths-south-africa?bc=254).

medRxiv preprint doi: https://doi.org/10.1101/2020.09.26.20201814; this version posted September 28, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

who are significantly less likely to die from COVID-19 if they contract the virus. However, differences in population age profiles do not fully explain the estimated infection fatality rates. If reported infection fatality rates by age, based on data from several HICs[17], are valid for these countries, the expected overall infection fatality rates for Bolivia, Colombia and South Africa would be 0.57%, 0.63% and 0.33%, respectively. Possible explanations for why mortality risk for COVID-19 might be lower in these three countries, compared to the (mainly) European countries from which infection fatality rates by age are derived, could include differences in immune-system response (already observed, for example, between men and women in some HICs), partial immunity to COVID-19 due to prior exposure to other coronaviruses, differences in lethality and prevalence of different virus strains, and different infection rates for different age groups.

To test the robustness of our approach, we applied the same methodology to fit curves to the first peaks in the comparison countries shown in Figure 3 (represented as red lines in this figure). Researchers at Imperial College London and others argued convincingly in June that European countries have not reached herd immunity[18], and subsequent increases in cases have proven their point. Applying our model to fit curves just to the first peaks (shown with red circles in Figure 3)[19], and hence assuming that the peaks were due to herd immunity, we find that the best-fit curves match the observed data for the first peaks fairly well in all cases, but with values for the disease parameters that are implausible. For example, for France, R_{0_e} = 2.7 (near some of the best estimates for the basic reproduction number R_0 without any social distancing measures), and p = 0.22% (well below the detection rates of between 7% and 20% suggested by serological studies in European countries[20]). New Zealand is well known for "crushing" the curve – and the parameters associated with the best-fit "herd immunity curve" to its reported cases would be R_{0_e} = 2.6 and a completely unrealistic p = 0.0015%. Thus, our approach leads to a conclusion for the comparison countries that the evolution of reported cases was <u>not</u> due to herd immunity (but instead must have been due to control measures). With this check, we increase our confidence in the hypothesis that the outbreaks in the 12 LMICs studied <u>are</u> declining due to herd immunity, which generates well-fitting curves with plausible parameters.

Discussion

Prominent models of the epidemic from teams at Imperial College London (ICL)[21] and the University of Washington Institute for Health Metrics and Evaluation (IHME)[22] use SEIR simulations and determine key parameters – especially the effective reproduction number, which can vary over time – by fitting the models' results for deaths to the reported numbers of deaths from COVID-19 and utilising age-specific infection fatality rates from recent studies (in HICs). These models estimate that total infections to date

[17] AT Levin et al. "Assessing the Age Specificity of Infection Fatality Rates for COVID-19: Meta-Analysis & Public Policy Implications" - medRxiv 2020.07.23.20160895 doi: https://doi.org/10.1101/2020.07.23.20160895.
[18] LC Okell, R Verity, OJ Watson et al. Have deaths from COVID-19 in Europe plateaued due to herd immunity? The Lancet; 11-06-2020, doi: https://doi.org/10.1016/S0140-6736(20)31357-X.
[19] The fits are made to only the first 100 days of reported cases for each of France, Germany and Spain, the first 150 days of reported cases for the United Kingdom, and the first 75 days of reported cases for China and New Zealand.
[20] Calculated by comparing findings from seroprevalence studies in England (H Ward, CJ Atchison, M Whitaker, et al. Antibody prevalence for SARS-CoV-2 in England following first peak of the pandemic: REACT2 study in 100,000 adults. *medRxiv* 2020: 2020.08.12.20173690), Italy (Italy National Institute of Statistics. Primi risultati dell'indagine di sieroprevalenza sul SARS-CoV-2. 2020), Portugal (Portugal National Institute of Health. Relatório de Apresentação dos Resultados Preliminares do Primeiro Inquérito Serológico Nacional COVID-19. 2020) and Spain (R Pastor-Barriuso, B Perez-Gomez, MA Hernan et al. SARS-CoV-2 infection fatality risk in a nationwide seroepidemiological study. *medRxiv* 2020: 2020.08.06.20169722) with reported cases from around the times of the studies.
[21] The ICL model is described at https://mrc-ide.github.io/global-lmic-reports/parameters.html. See also PT Walker, C Whittaker, OJ Watson et al. The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. Science; 12-06-2020, doi: https://doi.org/10.1126/science.abc0035.
[22] The IHME model results are at https://covid19.healthdata.org/global. See also J Friedman, P Liu, E Gakidou, and IHME COVID-19 Model Comparison Team. Predictive performance of international COVID-19 mortality forecasting models. MedRxiv. 14 July 2020. doi:10.1101/2020.07.13.20151233.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.26.20201814; this version posted September 28, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

for the 12 LMICs we studied are much greater than reported, but much smaller than our analysis suggests. For example, the models estimate total infection rates for South Africa of 7.7% (ICL) and 9.5% (IHME)[23] – significantly lower than the rates of between 20% and 40% suggested by the recently announced serological study findings[24] – and detection rates of 13.9% and 11.2% respectively. It has previously been observed, by the IHME COVID-19 Model Comparison Team, that the predictive performance of seven COVID-19 models, including those of ICL and IHME, shows significantly higher errors for Sub-Saharan Africa, South Asia and Latin America and the Caribbean, compared to their performance for HICs[25]. Our research suggests why this might be the case. The ICL, IHME and other models are well-suited to HICs: reported deaths for such countries are likely to be reasonably close to actual deaths; the infection fatality rates used in the models are based mainly on studies conducted in HICs; and it is clear from the evolution of reported cases that these countries have not reached herd immunity[26] and that their effective basic reproduction numbers have varied significantly over time as disease control measures have been introduced and adjusted (necessitating the additional granularity of SEIR modelling). However, for some LMICs: reported deaths from COVID-19 are likely to understate actual deaths by large factors; age-specific infection fatality rates might differ substantially from those in Europe and North America; and a simple model, using the approximation that effective basic reproduction numbers and detection rates remain constant over time, may be sufficient to describe the evolution of reported cases well (at least for the 12 countries we studied).

Systematic studies of representative population samples should be conducted in each of the 12 LMICs discussed here, to determine the percentages of people who have been infected and are immune – and consequently test directly the primary conclusion that the overall infection rates are very high and exceed estimated herd immunity thresholds. Serological studies are ongoing in some counties, such as South Africa. However, these studies may still underestimate infection and immunity levels, because people may have immunity driven by T-cells without detectable antibodies[27]. However, it is unclear whether the presence of T-cells provides complete immunity and T-cell immunity is harder to determine than antibodies in blood.

Herd immunity means that the virus can no longer spread uncontrollably, but not that it is gone completely. Individuals may still contract the virus if they are not immune. Isolated communities with low infection rates may still experience localised outbreaks. Small general outbreaks might happen as control measures are relaxed, especially if this happens before cases have fully declined from the herd immunity peak.

It is not yet known how long immunity from SARS-CoV-2 lasts. Even if a population has herd immunity today, it is possible that this could be lost over time, which might lead to future outbreaks – the severity of which would depend on the share of people losing immunity, the amount of variation in timing of when people lose immunity, and whether susceptibility to reinfection is equal to the susceptibility to first infection. New strains of SARS-CoV-2 have emerged, and it is conceivable that future mutations could allow the virus to evade immune systems, and thus render previously immune populations susceptible again to the disease – but there is no evidence yet of any such immunity-evading strains.

---

[23] Derived from the total reported cases divided by the total mean estimates of new infections from the models, up to 8 September, as reported on https://ourworldindata.org/covid-models.
[24] According to Dr. Zweli Mkhize, Minister for Health of South Africa, as reported by AP: https://apnews.com/cffcd4dfb1e3cbd810838fb9bde7a91d.
[25] J Friedman, P Liu, E Gakidou, and IHME COVID-19 Model Comparison Team. Predictive performance of international COVID-19 mortality forecasting models. MedRxiv. 14 July 2020. doi:10.1101/2020.07.13.20151233.
[26] LC Okell, R Verity, OJ Watson et al. Have deaths from COVID-19 in Europe plateaued due to herd immunity? The Lancet; 11-06-2020, doi: https://doi.org/10.1016/S0140-6736(20)31357-X.
[27] T Sekine et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. Cell (2020), doi: https://doi.org/10.1016/j.cell.2020.08.017.

medRxiv preprint doi: https://doi.org/10.1101/2020.09.26.20201814; this version posted September 28, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Acknowledgements and declarations of interest

We wish to thank Muhannad Alramlawi and Yohann Sequeira for help with research on projected infection fatality rates and Othmane Fourtassi for support in reviewing and editing the paper. We also acknowledge many useful discussions with Partners and consultants at Dalberg Advisors, in particular Edwin Macharia.

C.J.A.N. and J.P.C. work at Dalberg Advisors, a management consultancy whose clients include multilateral agencies, foundations, international development agencies, governments, companies and NGOs. C.J.A.N. and J.P.C. conducted this work in a personal capacity, and the work was not funded by any client of Dalberg Advisors.