Supplementary information

Age-dependent effects in the transmission and control of COVID-19 epidemics

Authors: Nicholas G. Davies[1*], Petra Klepac[1^], Yang Liu[1^], Kiesha Prem[1], Mark Jit[1], CMMID COVID-19 working group, Rosalind M Eggo[1*]

The CMMID COVID-19 working group[1] is: Carl A B Pearson, Billy J Quilty, Adam J Kucharski, Hamish Gibbs, Samuel Clifford, Amy Gimma, Kevin van Zandvoort, James D Munday, Charlie Diamond, W John Edmunds, Rein MGJ Houben, Joel Hellewell, Timothy W Russell, Sam Abbott, Sebastian Funk, Nikos I Bosse, Fiona Sun, Stefan Flasche, Alicia Rosello & Christopher I Jarvis. Order of working group determined at random.

[1] Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, WC1E 7HT
^ these authors contributed equally
* correspondence to Rosalind M Eggo r.eggo@lshtm.ac.uk or Nicholas G Davies nicholas.davies@lshtm.ac.uk

1. Further details on fitting

For all Bayesian inference (i.e. shown in Figs. 1 and 2 of the main text), we use differential evolution Markov chain Monte Carlo[1], first running numerical optimization to place starting values for each chain near the posterior mode. We then run 2000-3000 samples of burn-in, and generate at least 10,000 samples post-burn-in. Recovered posterior distributions, with prior distributions overlaid, are shown in Figures S1-S9.

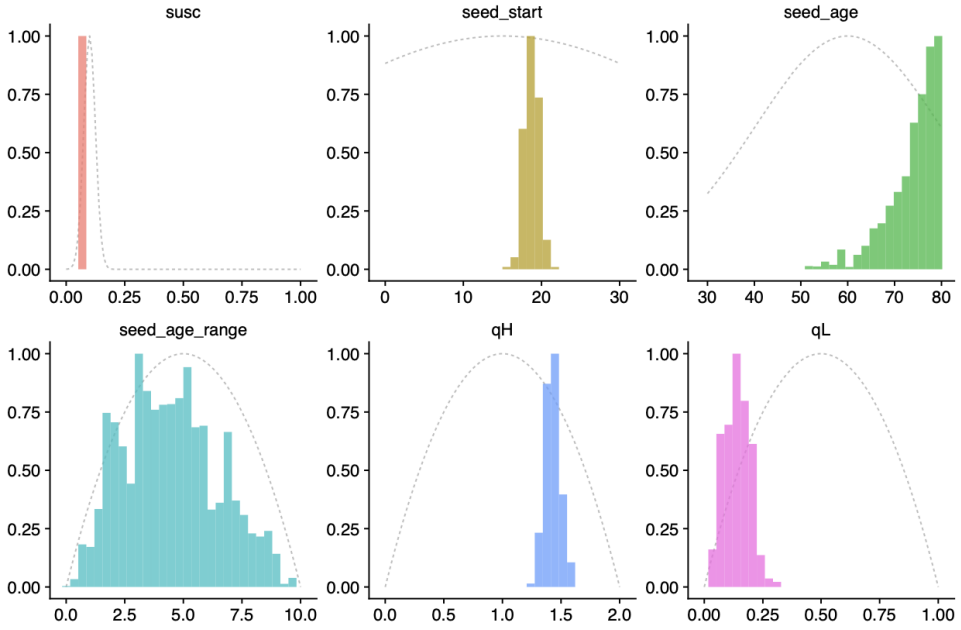

Fig. S1. Prior distributions (grey dotted lines) and posterior distributions (coloured histograms) for model parameters under hypothesis 1, fitting to the early epidemic in Wuhan; seed_start is measured in days after November 1st, 2019; this figure accompanies Fig. 1 of the main text.

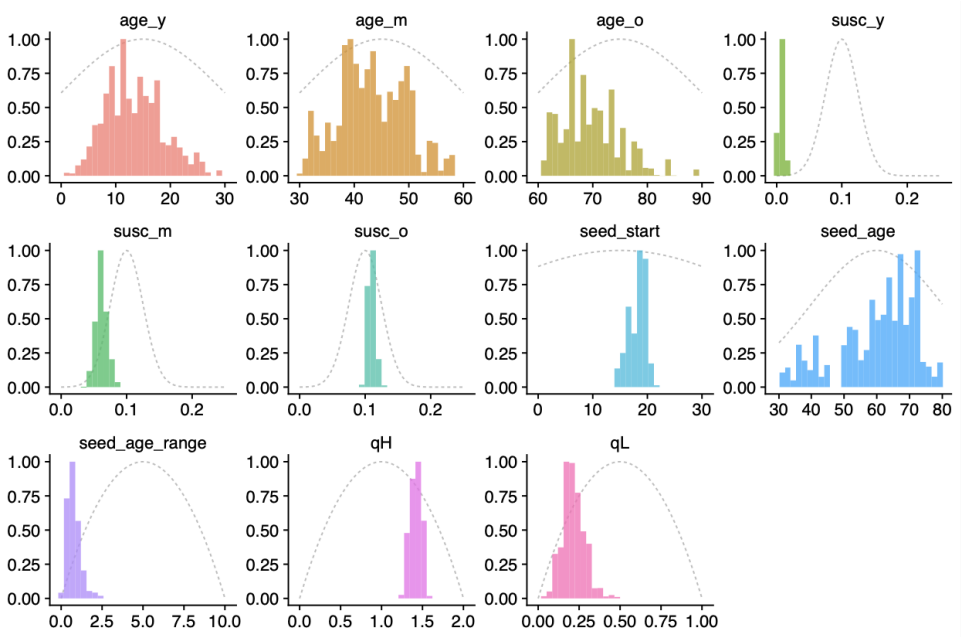

Fig. S2. Prior and posterior distributions for the epidemic in Wuhan, hypothesis 2. See caption for Fig. S1 for details.

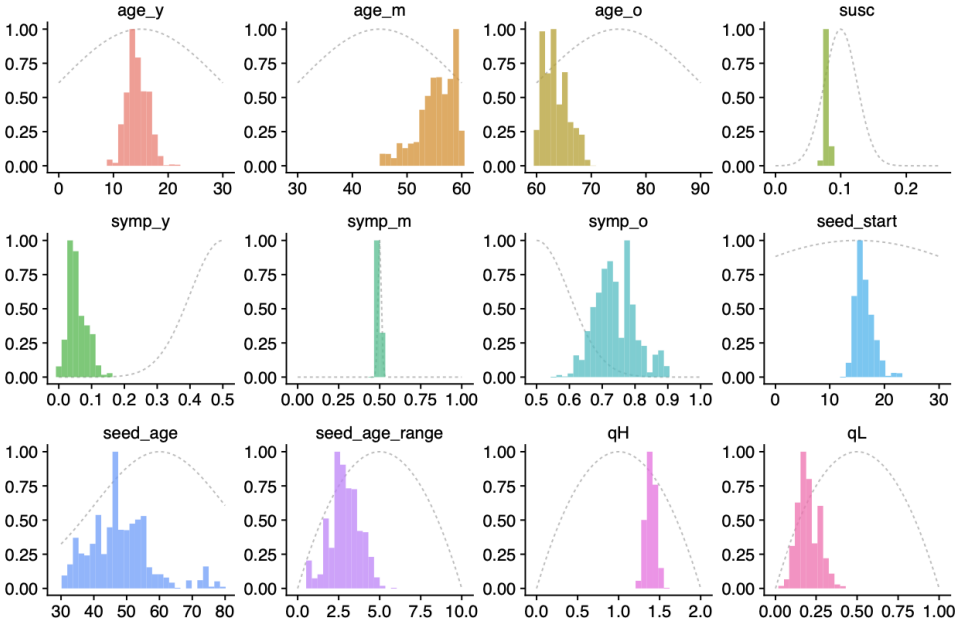

Fig. S3. Prior and posterior distributions for the epidemic in Wuhan, hypothesis 3. See caption for Fig. S1 for details.

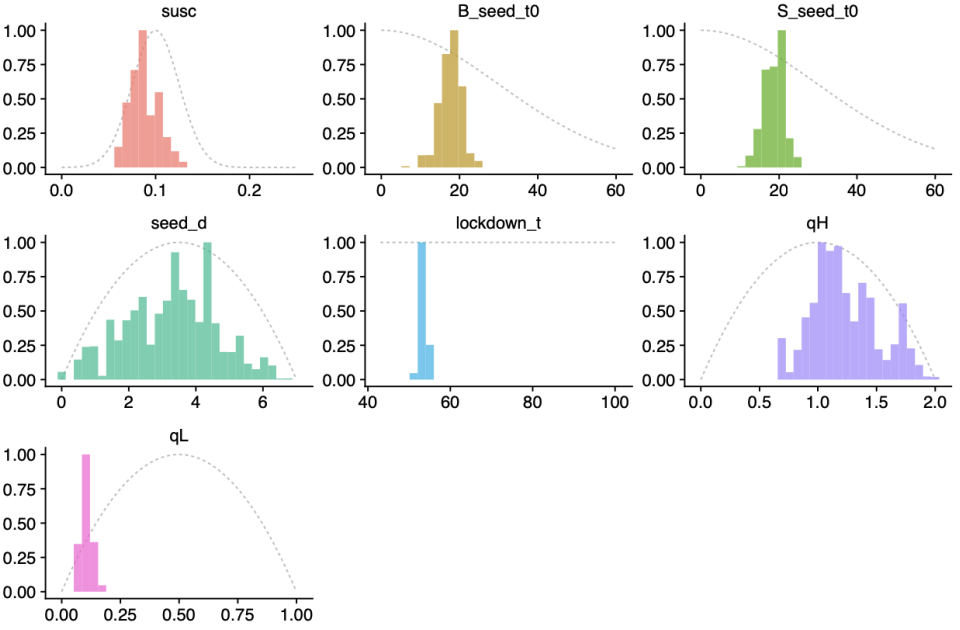

Fig. S4. Prior and posterior distributions for the epidemics in Beijing and Shanghai, using the "consensus" fit for age-specific clinical fraction and assuming subclinical infections are 50% as infectious as clinical infections. Times are in days after December 1st, 2019. Note, seed_d is the inferred duration of the seeding event. This figure accompanies Fig. 2c of the main text.

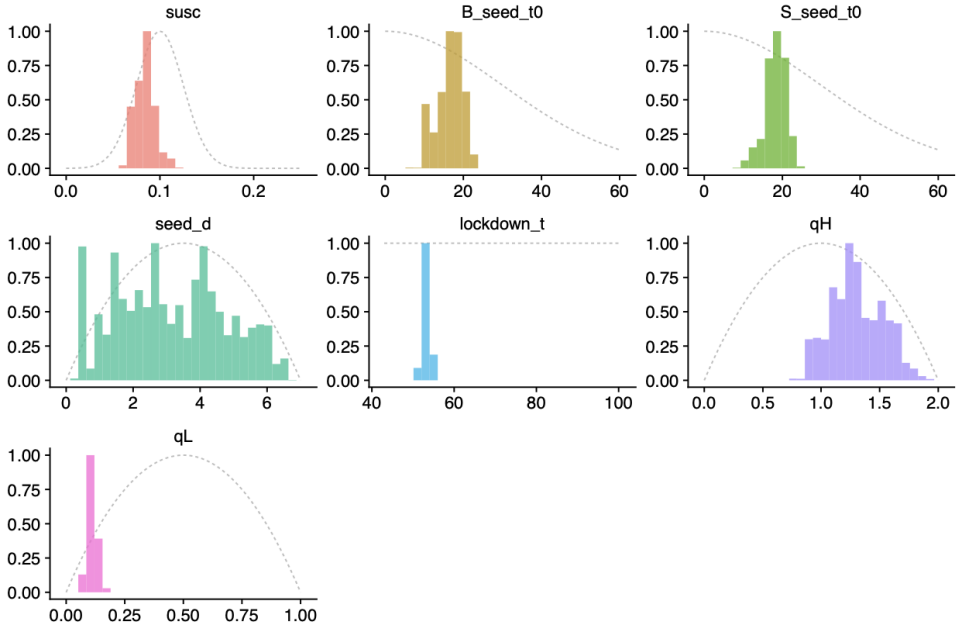

Fig. S5. Prior and posterior distributions for the epidemics in Beijing and Shanghai, using the fit for age-specific clinical fraction derived from China CCDC case numbers. See caption for Fig. S4 for details.

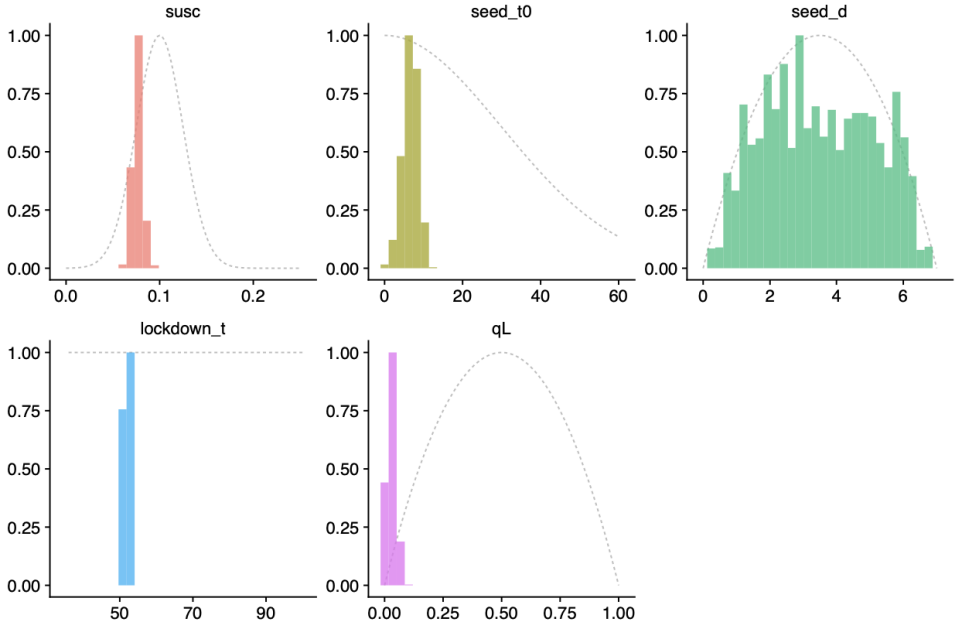

Fig. S6. Prior and posterior distributions for the epidemic in South Korea, using the "consensus" fit for age-specific clinical fraction and assuming subclinical infections are 50% as infectious as clinical infections. Times are in days after January 1st, 2019. This figure accompanies Fig. 2c of the main text.

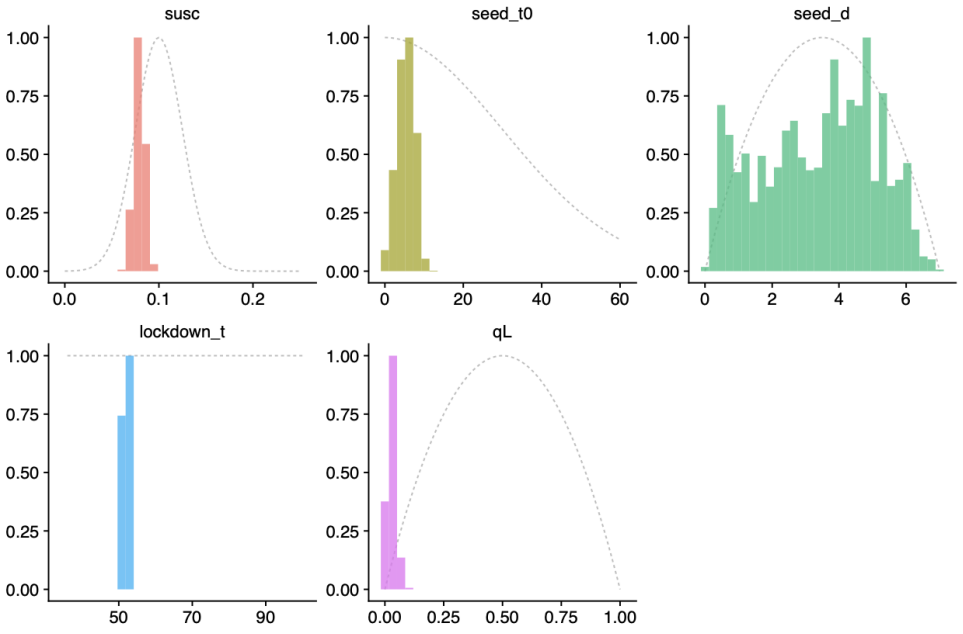

Fig. S7. Prior and posterior distributions for the epidemic in South Korea, using the fit for age-specific clinical fraction derived from case numbers for South Korea. See caption for Fig. S6 for details.

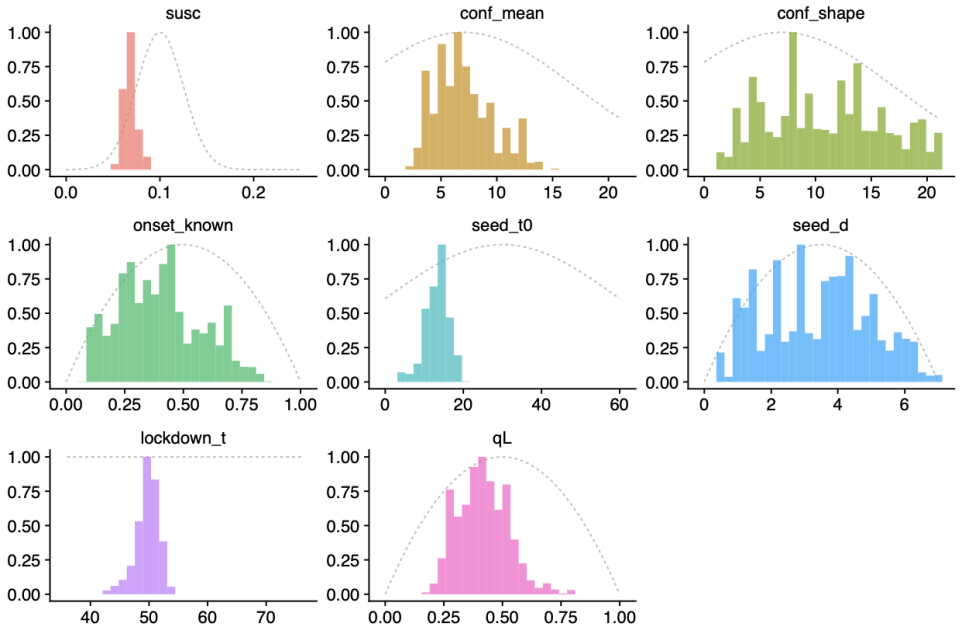

Fig. S8. Prior and posterior distributions for the epidemic in Lombardy, using the "consensus" fit for age-specific clinical fraction and assuming subclinical infections are 50% as infectious as clinical infections. Times are in days after January 1st, 2019. This figure accompanies Fig. 2c of the main text.

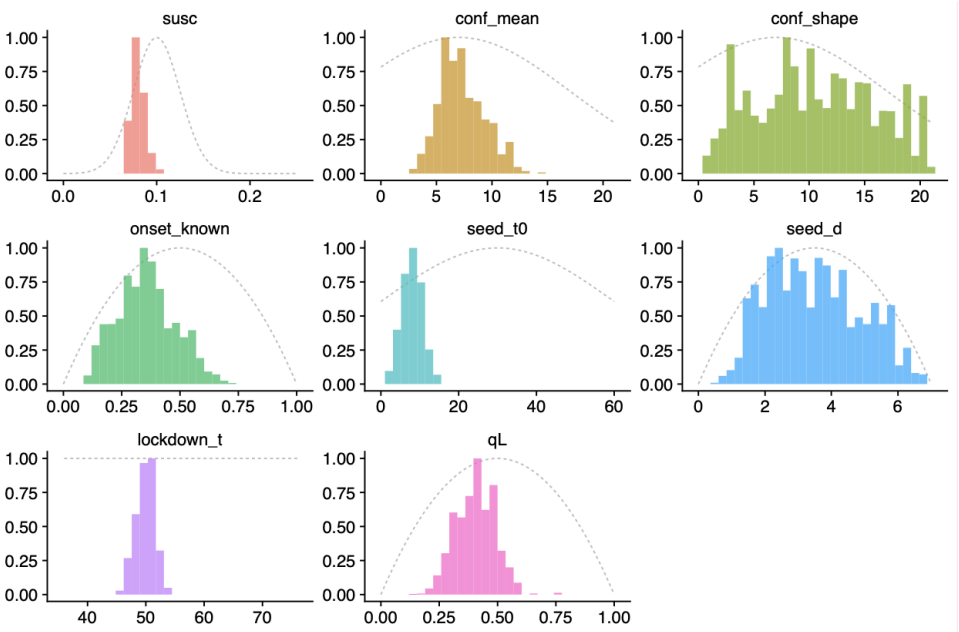

Fig. S9. Prior and posterior distributions for the epidemic in Lombardy, using the fit for age-specific clinical fraction derived from case numbers for Lombardy. See caption for Fig. S8 for details.

2. Contact matrices

Wherever possible, we use measured contact matrices; the measured contact matrices we use are those for Shanghai[2], Italy[3], the United Kingdom[3], and Zimbabwe[4]. We adapt each of these mixing matrices, using 5-year age bands, to specific regions of the countries they were measured in by reprocessing the original contact surveys that informed them with the population demographics of the regions we are analysing.

We use the contact survey conducted in Shanghai[2] to generate matrices for Wuhan (Fig. 1, main text) using demographic data for Wuhan prefecture and for 13 provinces of China using demographic data for Chinese provinces from ref.[3] . This survey allowed respondents to record both individual (one-on-one) and group contacts, the latter with approximate ages. While individual contacts were associated with a context (home, work, school, etc.) group contacts were not, and so we assumed that all group contacts which involved individuals aged 0-19 occurred at school. We also assumed that group contacts were of a lower intensity than individual contacts, weighting group contacts by 50% relative to one-on-one contacts. We use the demographics of Birmingham, Milan, and 12 regions of Italy along with contact surveys for the UK and Italy[3] to reconstruct contact matrices for the UK and Italy. Finally, we use the demographics of Bulawayo, Zimbabwe[5] along with the contact survey of residents of Manicaland province, Zimbabwe[4] to reconstruct the contact matrix for Bulawayo.

Where measured matrices are not available, we use synthetic matrices generated for 152 countries by Prem et al.[6]. We use these matrices for South Korea, Japan, Singapore, and Ontario (Fig. 2, main text), as well as for the capital cities of 146 countries (Fig. 4, main text). As a simplifying assumption, we use the age distribution of each country[7] in place of the specific demographics for each capital city, which were not readily available.

Each contact survey and synthetic matrix was characterised in terms of contacts made at home, work, school, and other locations. We split these into school and non-school components to account for school closures in each country. For Fig. 2, we assumed schools were closed during the epidemic in China, but open in Italy, Singapore, South Korea, Japan, and Canada, as the data we used to fit our models mainly come from times during which schools were open in these locations.

The contact matrices and demographics we used for Figs. 1-3 of the main text are shown in Supplementary Figure S10. Contact matrices used for Fig. 4 are available from Prem et al.[6].

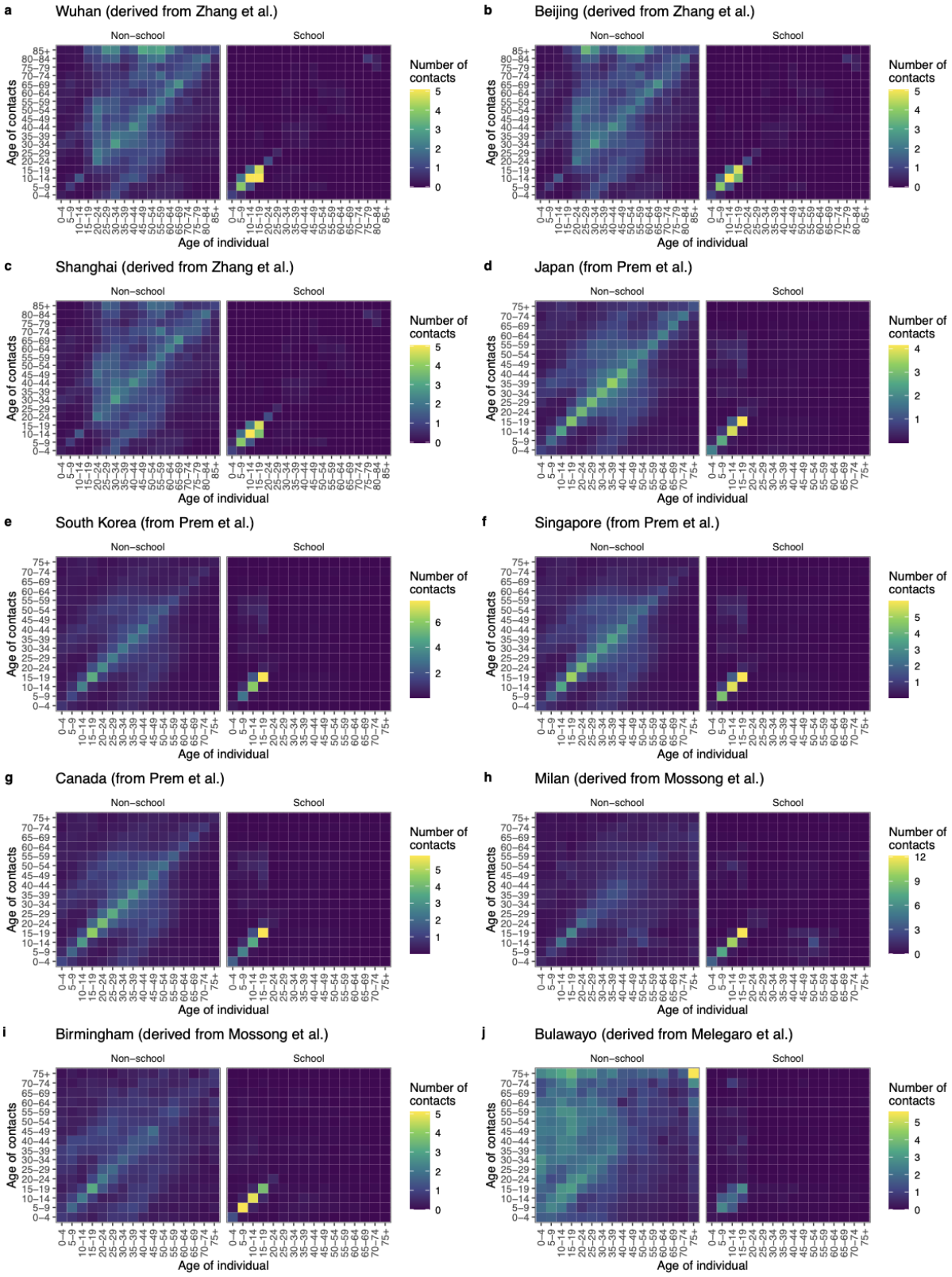

Fig. S10. Contact matrices used for Figs. 1-3 of the main text. We have not shown matrices for all 12 regions of Italy modelled, nor for all 13 provinces of China modelled, as these show similar patterns to the matrices for Milan and for Wuhan, Beijing and Shanghai, respectively.

3. Sensitivity to asymptomatic infectiousness

As the infectiousness of subclinical individuals is not identifiable from the data we have available, we adopt a baseline estimate of 50% relative to preclinical and clinical individuals. In Supplementary Figures S11-S13, we perform a sensitivity analysis by repeating our model runs with the alternative values of 25% and 75%.

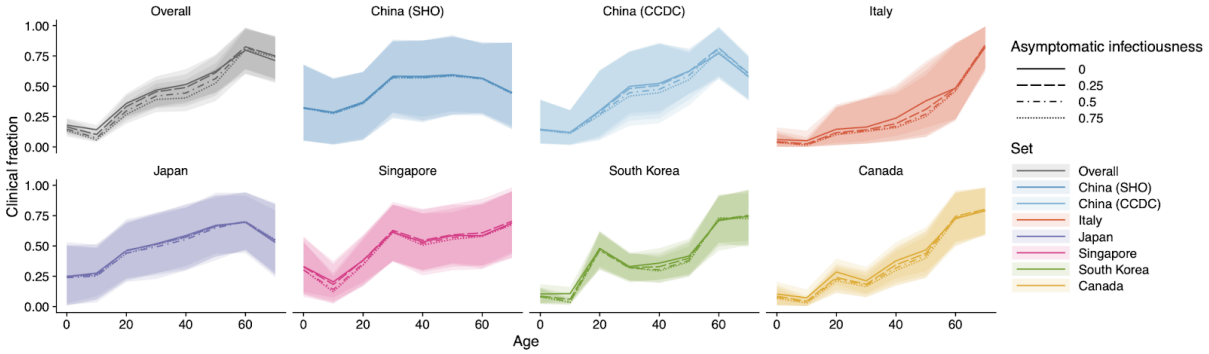

Fig. S11. Age-specific clinical fraction by country, assuming subclinical infections are 0%, 25%, 50%, or 75% as infectious as clinical infections.

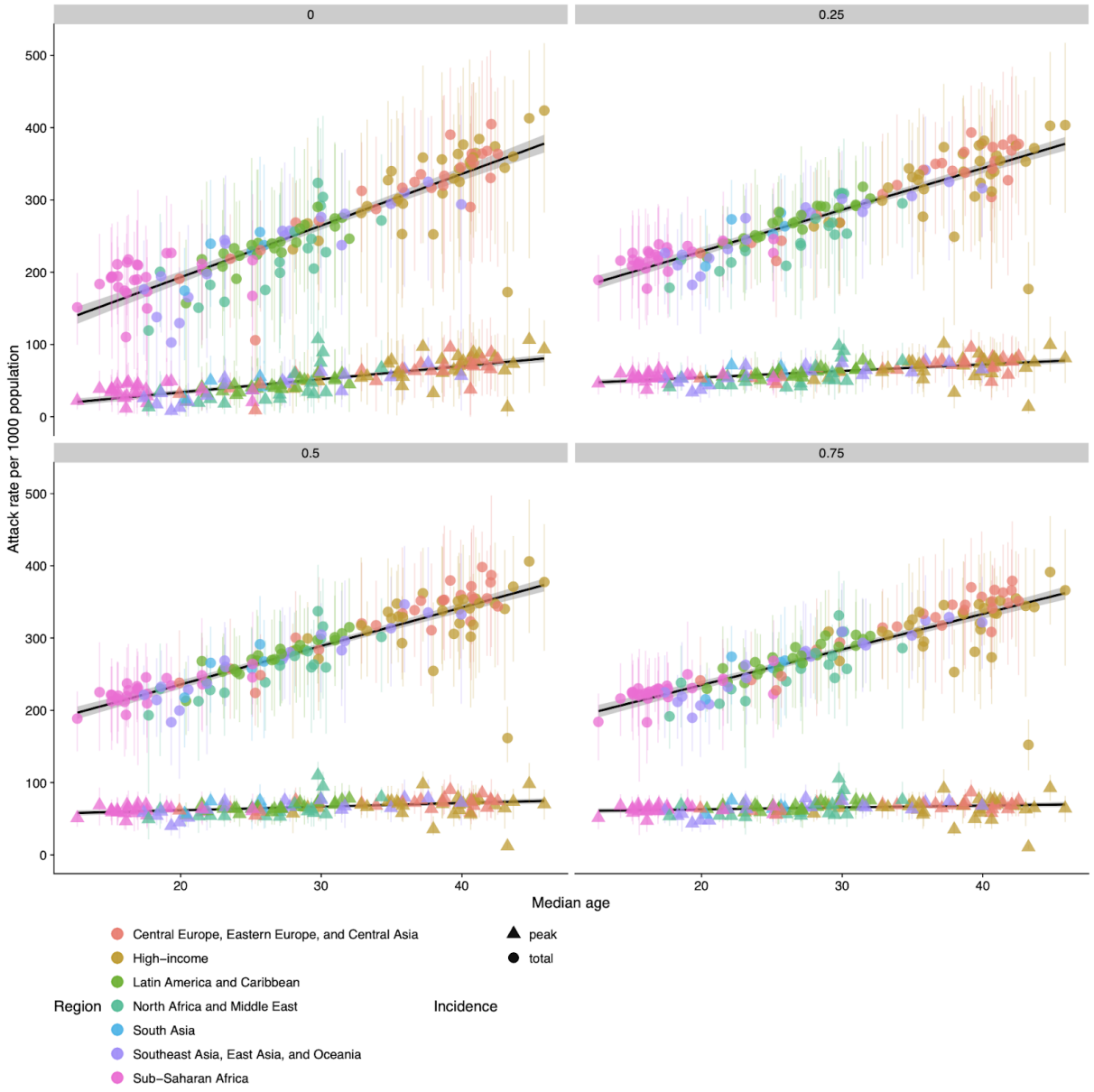

Fig. S12. Projected total and peak clinical case attack rate for 146 capital cities, under different assumptions for the infectiousness of subclinical cases.

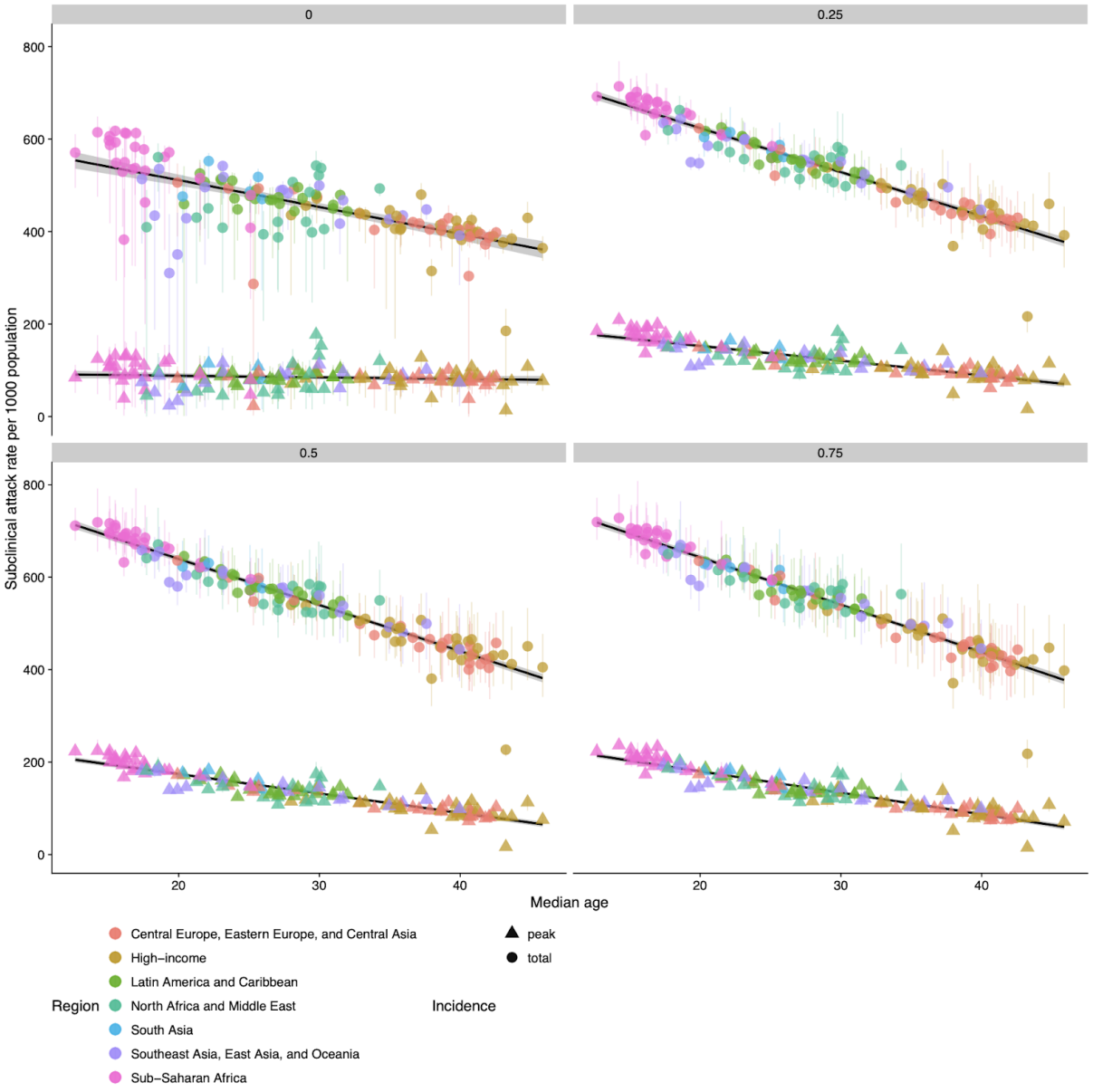

Fig. S13. Projected total and peak subclinical case attack rate for 146 capital cities, under different assumptions for the infectiousness of subclinical cases.

4. Sensitivity to age-specific clinical fraction in lower-income countries

In the main text, we present projections for 146 capital cities spanning both lower-income and higher-income settings. For this analysis, we assume that the age-specific clinical fraction is the same across all settings. However, a higher rate of comorbidities in lower-income countries could change the age-specific probability of developing clinical symptoms upon infection. To investigate this possibility, we construct a schematic alternative age-specific profile of clinical fraction by (1) increasing the age-specific probability of developing symptoms by 15% for individuals under the age of 20; (2) shifting the age-specific clinical fraction for individual over the age of 20 by 10 years older; and (3)

increasing the clinical proportion for individuals over 50 to its maximum amount, rather than retaining the slight decrease seen in individuals over 70 in higher-income countries (**Fig. S14**). Repeating our global projections analysis using this age-specific profile of severity across all low-income and lower-middle-income countries (**Fig. S15**) illustrates that comorbidities have the potential to substantially worsen the burden in lower-income countries, to the extent that they can readily exceed the burden in higher-income countries.

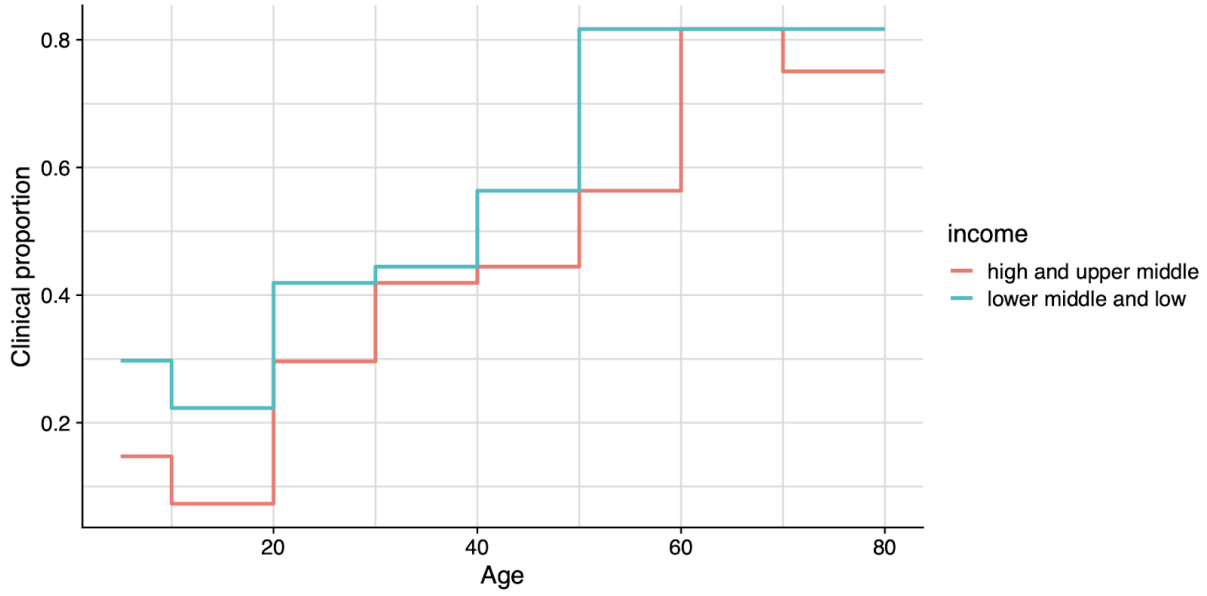

Fig. S14. Schematic age-specific clinical proportion for higher-income and lower-income countries.

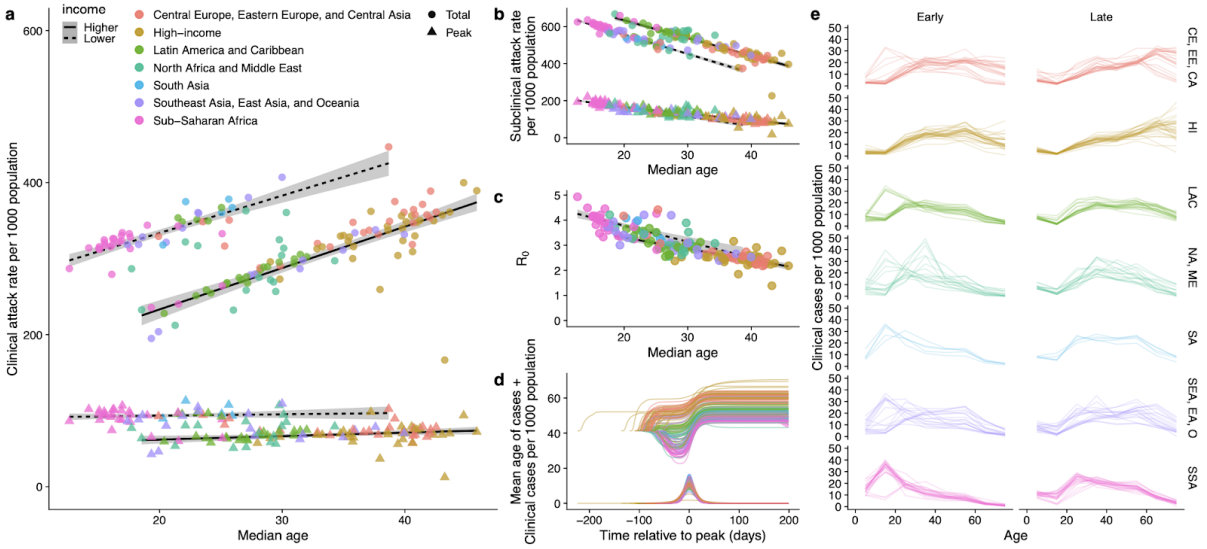

Fig. S15. Illustrative results of the projections for 146 capital cities assuming a higher age-varying clinical proportion in lower-income countries.

5. Fixing R_0 versus fixing susceptibility

For Fig. 3 of the main text, we illustrate differing effects of school closures for COVID-19 versus pandemic influenza by simulating an epidemic in three cities, Milan, Birmingham, and Bulawayo. To illustrate the impact that demographics alone have on the effectiveness of interventions, we fix the basic reproductive number R_0 at 2.4 in each of these settings. This means that higher rates of contact in Milan and Bulawayo compared to Birmingham, as captured by the contact surveys conducted in Italy and Zimbabwe, are not taken into account in this analysis. In **Fig. S16**, we show the results for a similar analysis in which susceptibility u_i is fixed across settings rather than R_0. Overall, the conclusions regarding the relative effectiveness of school closures for COVID-19 versus influenza are similar.

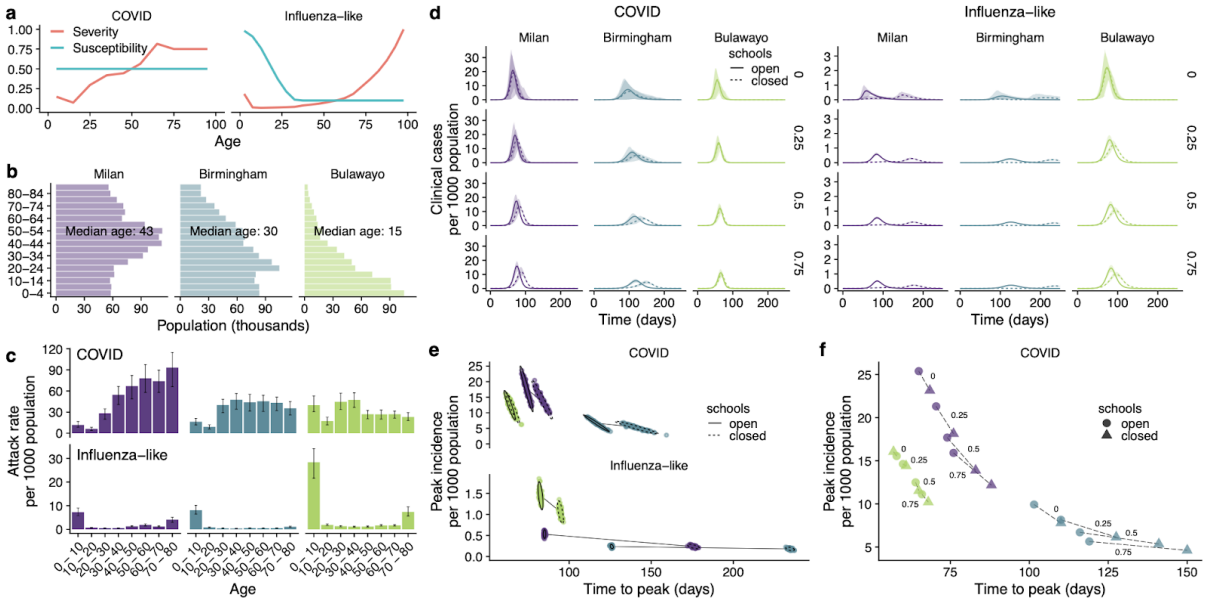

Fig. S16. Comparison of school closures in three exemplar cities when susceptibility u_i is fixed across settings.

References

1. Braak, C. J. F. T. A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. *Stat. Comput.* **16**, 239–249 (2006).
2. Zhang, J. *et al.* Patterns of human social contact and contact with animals in Shanghai, China. *Sci. Rep.* **9**, 1–11 (2019).
3. Mossong, J. *et al.* Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLOS Med.* **5**, e74 (2008).

4. Melegaro, A. *et al.* Social Contact Structures and Time Use Patterns in the Manicaland Province of Zimbabwe. *PLOS ONE* **12**, e0170459 (2017).
5. Bulawayo (City, Zimbabwe) - Population Statistics, Charts, Map and Location. http://citypopulation.info/php/zimbabwe-admin.php?adm1id=A.
6. Prem, K., Cook, A. R. & Jit, M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Comput. Biol.* **13**, e1005697 (2017).
7. World Population Prospects - Population Division - United Nations. https://population.un.org/wpp/.