nature communications

Article | Article in Press

https://doi.org/10.1038/s41467-026-78981-0

Fairness in infectious disease modeling

Received: 21 November 2025

Accepted: 8 October 2026

Published online: 02 November 2026

Check for updates

Yuhan Li[1], Nicoló Gozzi[2], Nicola Perra[1,3] & Michele Tizzoni[4]

The concept of fairness has been extensively examined within the domains of Machine Learning and Artificial Intelligence more broadly. It remains, however, largely underexplored in the field of Computational Epidemiology. Considering the substantial influence that epidemic models exert on public health policy, particularly in the context of outbreak preparedness and response, this shortcoming is of great relevance. Here, we propose a mathematical framework, grounded in core principles from Social Epidemiology, for evaluating the fairness of computational epidemic models. We begin by applying our framework to a range of epidemic modeling approaches and simulation scenarios, such as the initial spread of COVID-19 in London, New York, and Santiago de Chile, as well as the 2016 Zika virus outbreak in Colombia, demonstrating its consistent capacity to assess model fairness across diverse disease dynamics. Subsequently, we illustrate how our definition of fairness can be incorporated into the design of immunization strategies as a way to enhance health equity while simultaneously improving the overall effectiveness of such interventions. Overall, our results offer a new systematic methodology for quantifying fairness in computational epidemiology.

Infectious disease modeling is a key tool for informing the response to outbreaks and guiding policymakers in the adoption of interventions[1]. During the COVID-19 pandemic, for example, computational models have been widely used to assess the impact of non-pharmaceutical interventions (NPIs), evaluate future scenarios, and design mitigation strategies[2]. Due to a limited understanding of the interaction between the dynamics of the epidemic and the social determinants of health, in several cases, such intervention strategies have amplified disparities rather than reduced them[3–6]. Indeed, the COVID-19 pandemic has been a stark reminder that infectious diseases do not affect populations equally[5,7]. Outbreaks are often characterized by massive inequities, resulting in disproportionate health impacts among vulnerable groups and minorities[8,9].

Most of the epidemiological modeling frameworks available are unequipped to adequately capture the structural drivers of unequal health outcomes. Indeed, recent studies have highlighted how computational epidemiology lacks standard frameworks to address inequalities and health disparities systematically[10–12]. The urgent need for equitable models has led to a widespread call for epidemic models that can integrate aspects of equity into their mechanisms, leading to realistic descriptions of the epidemic burden among different socioeconomic and demographic groups[10,12–14].

The lack of equitable computational modeling frameworks to simulate infectious disease epidemics is accompanied by a lack of standard methods to define and evaluate a model's fairness. On the contrary, due to their increasing relevance in real-world decision making, the concept of algorithmic fairness has become central to machine learning and artificial intelligence[15,16]. Epidemic modeling has a significant impact on policy decisions, but a commonly agreed definition of the fairness of infectious disease models remains elusive[17].

Previous studies have addressed fairness in infectious disease modeling almost exclusively in two domains: pandemic forecasting[18–20] and vaccine allocation strategies[21–23]. In the first domain, the focus has mainly been on auditing subgroup-level predictive performance, i.e., whether forecast errors are systematically uneven between groups. In particular, previous studies mainly focused on addressing racial/ethnic disparities associated with the burden of COVID-19 in the United States. In the second domain, the emphasis has been on evaluating or optimizing allocation policies under fairness constraints, often balancing equity objectives against overall epidemiological effectiveness. Despite their importance, these contributions are generally

[1]School of Mathematical Sciences, Queen Mary University of London, London, UK. [2]ISI Foundation, Turin, Italy. [3]The Alan Turing Institute, London, UK. [4]Department of Sociology and Social Research, University of Trento, Trento, Italy. e-mail: michele.tizzoni@unitn.it

disconnected and do not provide a general, transferable definition of fairness for epidemic models. They are typically tailored to specific applications, making it difficult to compare fairness across diseases, model structures, and geographic contexts within one unified framework.

Here, we address this gap by proposing a general definition of fairness for models of infectious disease spread, grounded in the theory of Social Epidemiology[24,25]. We define the fairness of an epidemic model as its ability to capture the (un)equal distribution of disease burden according to one or more equity-related features. Given a demographic or socioeconomic trait that is a known determinant of health and the disease dynamics under study, an epidemic model will be considered fair if it captures the observed inequality in health outcomes according to such a variable. We translate our definition into practice by providing quantitative fairness measures for different epidemic scenarios, introducing a set of metrics that quantify the fairness of a model in various cases. Following the traditional literature of Social Epidemiology[24], we consider health inequalities in ordered social groups, unordered social groups, and spatially structured populations. In studying social inequalities, ordered social groups are those that exist along a clear hierarchical continuum—such as income levels—where members can be ranked from higher to lower status. Inequality in these groups is often analyzed in terms of relative positions and gradients of advantage or disadvantage. In contrast, unordered social groups—such as those based on gender, ethnicity, or religion—do not form a ranked continuum but are instead categorical distinctions. Inequality arises here from differential treatment, discrimination, or exclusion between groups, rather than from a measurable position along a social scale. Spatially structured populations can be considered a third case in which ordering may or may not be present. For instance, peri-urban or less urbanized districts may concentrate older and lower-income populations with reduced healthcare access. Hence, age or healthcare access could be used to rank spatial units along this gradient. At the same time, the distribution of diseases may not show a clear association with any sociodemographic dimension, in which case spatial units should be treated as unordered groups.

For each type of social inequality, we introduce a specific metric that quantifies a model's fairness: the relative concentration index (RCI) fairness, the Theil index fairness, and spatial fairness. The first quantifies the accuracy of the model in capturing the health gradients associated with naturally ordered social groups, such as income, education, or age. The second score assesses the fairness of the model in the case of unordered social groups, such as racial or ethnic groups. The third assesses a model's fairness in capturing the geographic disparities of disease burden, regardless of whether these are associated with ordered or unordered groups.

We first explore the application of our definitions of fairness through numerical simulations of the initial spread of COVID-19 in the United States, the United Kingdom, and Chile, as well as the Zika virus in Colombia. These case studies were chosen to maximize heterogeneity in disease dynamics, equity-related dimensions, and model definitions: age-stratified mortality in London, racial/ethnic disparities in New York City, socioeconomic spatial gradients in Santiago and Colombia. This design allows us to assess whether the fairness metrics are transferable across pathogens, countries, and model structures, such as compartmental, metapopulation, and agent-based models. The results show that our definitions of fairness can be consistently applied to a range of modeling approaches and disease scenarios. Second, we demonstrate how the proposed fairness scoring methods could inform policy responses by enhancing the equity of decisions during an outbreak. Indeed, by simulating a proof-of-concept vaccination scenario, we demonstrate that our fairness metrics not only represent a novel equity-aware subgroup model validation tool but can also provide guidance for a more equitable allocation of resources across groups.

Results

Fairness of infectious disease models and health inequalities

Mathematical and computational models of infectious diseases usually incorporate a representation of the population of interest. It is common to assume that some population traits are more relevant than others for disease transmission and health outcomes. The simplest models, like the well-known SIR model, assume individuals in a population to be indistinguishable apart from their health status (i.e., susceptible, infected, or recovered)[26]. As models grow in complexity, they can integrate more traits to capture the characteristics that are deemed important for explaining disease outcomes. A classic example is the introduction of age groups and mixing patterns as a way to capture the observed heterogeneities in contact rates and health outcomes by age[27].

We refer to socioeconomic and demographic characteristics as equity-related because including them in a model determines whether the model can capture the associated health inequalities, thereby enabling equity considerations to be subsequently assessed. In this study, we distinguish between model fairness and health equity. Model fairness refers to the extent to which an epidemic model accurately reproduces observed inequalities in health outcomes across equity-related groups. Health equity refers to the normative goal of reducing unjust and avoidable disparities through public health decisions. Under this distinction, fairness is an evaluative property of models, while equity is an objective of interventions informed by those models. From this stand point we propose a general definition of fairness for epidemic models as follows:

> We define the fairness of an epidemic model as the model's ability to capture the (un)equal distribution of disease-related health outcomes according to one or more equity-related features. Given any demographic or socioeconomic variable that describes the population under study, an epidemic model will be *fair* if it accurately captures the observed distribution of health outcomes according to such variables.

The above definition of fairness incorporates the principle of equal performance from the machine learning domain[28], requiring the maximization of a model's accuracy for all groups in the population. However, it is important to note that, unlike machine learning algorithms, infectious disease models do not generate outcomes or predictions at the individual level, but often only at the population level. This implies that many popular fairness constraints adopted in machine learning, such as Demographic Parity, Equalized Odds, and Equal Opportunity[29], cannot be easily translated to the context of epidemic models.

In the following, we apply our definition by providing quantitative measures of the fairness of a model for different diseases, epidemic scenarios, and modeling assumptions. We follow a classic distinction in Social Epidemiology[24], considering three different types of health inequality that correspond—in our framework—to three different measures of fairness. In particular, we consider the case of (i) ordered social groups, (ii) unordered social groups, and (iii) spatially structured populations. This distinction is not merely conceptual but has direct practical implications, as measures of fairness developed for ordered social groups typically cannot be applied to unordered groups. In all three cases, our approach to measuring the model fairness consists of three steps:

1. The identification of a measure of inequality to describe the health disparities observed in real data;
2. The characterization of the health disparities observed in a model output using such a measure of inequality;
3. The definition and application of a fairness metric that compares a model's results with the real data according to the chosen measure of inequality.Finally, we demonstrate how the proposed approach to quantifying fairness can inform policymaking by

https://doi.org/10.1038/s41467-026-78981-0

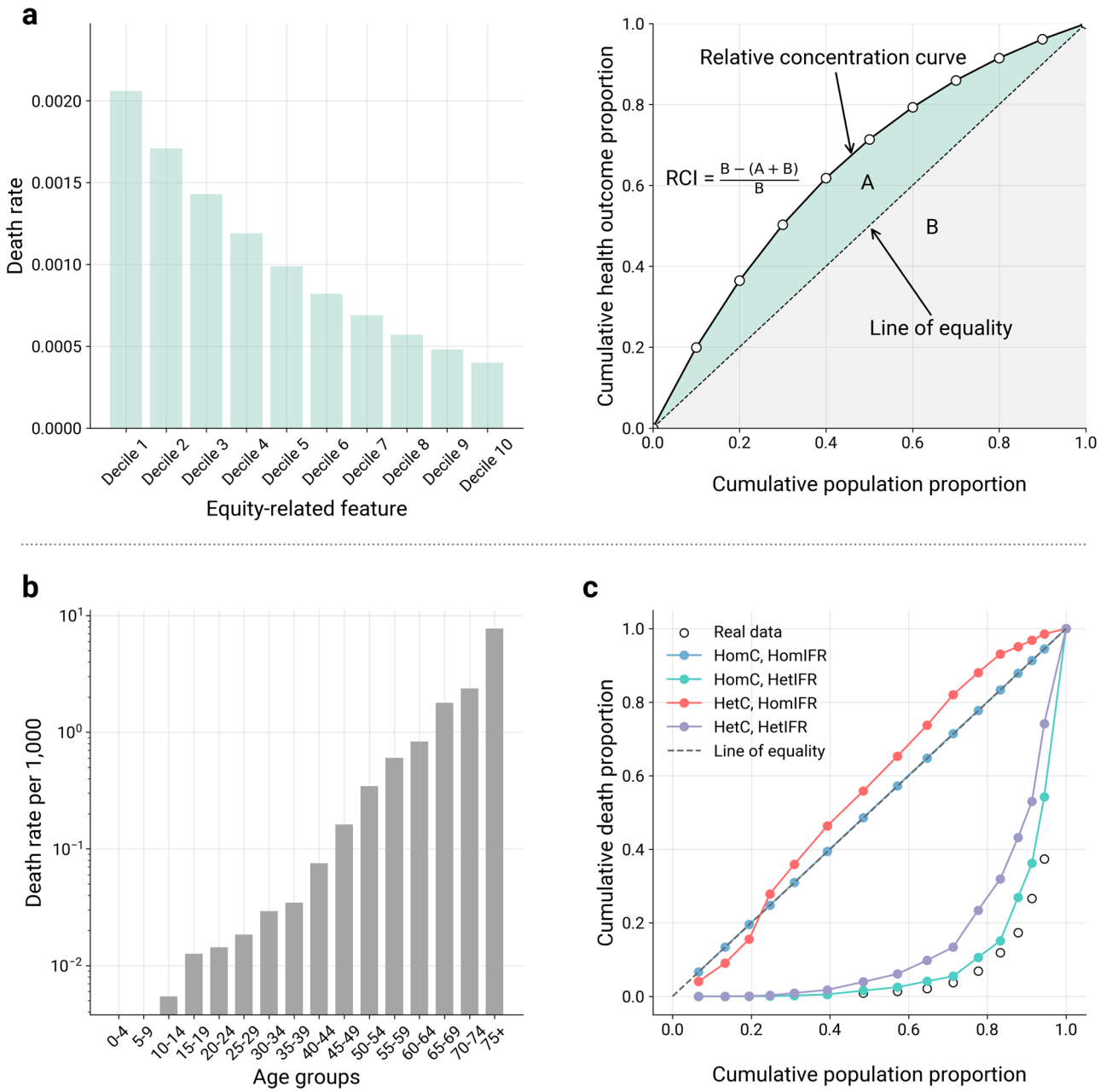


Fig. 1 | Models' fairness and ordered social groups. a The relationship between the relative concentration curve and the relative concentration index (right) obtained from a hypothetical distribution of deaths by population deciles ranked according to a generic equity-related feature (left). In this example, for illustrative purposes, we assume death rates to decrease as the ordered values of the equity-related feature increase. The formula in the right plot represents the geometric interpretation of the relative concentration index, and it is exactly equivalent to the definition of Eq. 1. **b** The reported COVID-19 death rates (per 1000) by age group in London, during the first pandemic wave between 2020/03/15 and 2020/07/05. **c** Comparison of the relative concentration curves (RCC) of death proportion from surveillance data (open black circles) and models' outcome, for four different epidemic models combining heterogeneous or homogeneous contact rates, and an age-dependent or flat IFR.

increasing the equity of health outcomes in public health interventions, such as vaccination campaigns.

Fairness in ordered social groups

We first consider the case of health inequalities associated with ordered social groups. Some demographic or socioeconomic traits imply an inherent ordering of the population subgroups, regardless of the health status of their members. For example, any trait measuring dimensions of the socioeconomic position of individuals implies a natural ordering, such as education, income, and wealth. Similarly, individuals' ages naturally define ordered groups within a population.

In the case of ordered social groups, a widely used measure of health inequality is the RCI[30,31]. The RCI is derived from a health concentration curve, where the population is ordered by social group status, and the cumulative fraction of the population is plotted against their share of total ill-health. Figure 1a illustrates the relationship between the RCI and the health concentration curve (right) for a fictitious disease, with death rates decreasing by deciles of the population ranked according to a generic equity-related feature (left). The relationship between the RCI and the health concentration curve is similar to the relationship between the Gini coefficient and the Lorenz curve, with the difference

that in a Lorenz curve, the population share is ranked by socio-economic variables on both axes[32].

Given a population that can be divided into N ordered social groups according to a given equity-related characteristic, the RCI is computed following the formula[24]:

$$\mathrm{RCI} = \frac{2}{\mu}\left[\sum_{i=1}^{N} p_i \mu_i R_i\right] - 1, \quad (1)$$

where p_i is the population share of group i, μ_i is the group average ill-health (i.e., death rate), R_i is the relative rank of the i-th group, and μ is the average ill-health of the whole population. The RCI ranges between −1 and 1, with extreme values corresponding to the case where the lowest (highest) ranked population group bears the full burden of the disease. Instead, an RCI = 0 indicates absolute equality between the groups. These extreme cases correspond, in Fig. 1a, to the area A taking values equal to B, $-B$, and 0, respectively.

Given the outcomes of an epidemic model and an observed distribution of the health burden related to an equity-related characteristic s, we propose to quantify the fairness of such a model as:

$$F(s) = 1 - \frac{|RCI_{m,s} - RCI_{d,s}|}{2} \quad (2)$$

where $RCI_{m,s}$ is the RCI of the models' outcomes, and $RCI_{d,s}$ is the RCI of real data when the population groups are ordered by s. $F(s)$ ranges between 0 and 1, with 1 indicating that the model perfectly reproduces the distribution of the health outcomes observed when ordering the population groups by s.

We show a practical application of our definition of fairness by computing the values of $F(s)$ for different epidemic models that describe the spread of SARS-CoV-2 in London, UK, during the first pandemic wave, between March and June 2020. In this case, we evaluate the fairness of each model in reproducing the observed inequality in the distribution of COVID-19 deaths by age. Indeed, as extensively documented, deaths from COVID-19 have been disproportionately concentrated in older age groups[33,34]. In London, the reported death rate per 1000 follows a highly skewed distribution, ranging from approximately 0 in infants aged 0–4 years to around 8 in individuals aged 75 and above (see Fig. 1b). This unequal distribution translates into the relative concentration curve shown in Fig. 1c (open circles), which corresponds to an $RCI_{d,age} = 0.85$. We explore the fairness of four age-structured SLIR-like compartmental models that combine different assumptions about the mixing patterns and infection fatality rate (IFR) adopted (see the "Methods" section for more details). In particular, we explore two types of mixing patterns feeding the models: homogeneous mixing across age groups (denoted as *HomC*) and synthetic age-stratified contact matrices[35,36] (denoted as *HetC*). The models also differentiate based on the IFR used. We consider two variations: a constant IFR equal to the population average and an age-varying profile obtained from ref. 34. The four models correspond to the possible combinations of contact patterns and IFR. All models are calibrated to the reported total weekly deaths in London (see the "Methods" and Supplementary Information).

Figure 1c shows the relative concentration curves of the median COVID-19 deaths simulated by each of the four epidemic models. As expected, the relative concentration curve produced by assuming both homogeneous mixing and a flat IFR falls on the line of equality, where the death rate is constant across groups (i.e., the absolute number of deaths is purely proportional to the group size). This model does not adequately capture the higher concentration of deaths in the elderly groups, as shown by its fairness score of $F(age) = 0.577$ [0.569, 0.584]. Similarly, the model considering a flat IFR and an age-stratified contact matrix has a fairness score of $F(age) = 0.528$ [0.521, 0.535]. The inclusion of an age-dependent IFR significantly improves model fairness. The model considering both age-stratified contacts and an age-dependent IFR has a fairness score of $F(age) = 0.923$ [0.919, 0.927] (Fig. 1b purple dots), while the one combining homogeneous mixing with a heterogeneous IFR has an even higher score of $F(age) = 0.976$ [0.973, 0.979] (Fig. 1b turquoise dots). Interestingly, the inclusion of age-stratified contacts slightly worsens model fairness, both when flat and heterogeneous IFR are considered. This can be explained by the fact that NPIs were applied as flat rescaling factors across all age groups, likely overestimating effective contacts among younger people whose interactions were comparatively more reduced by school, workplace, and leisure closures. Table S6 of the Supplementary Information summarizes these findings, showing the fairness scores based on Eq. 2, $F(\mathrm{age})$, for the four models.

Fairness in unordered social groups

Next, we examine the concept of model fairness applied to the case of health inequalities associated with unordered social groups. This is the case of sociodemographic variables that do not imply any natural ordering in population subgroups, such as race and ethnicity, gender, or religious affiliation. In such cases, health inequality is generally measured through information-theoretical metrics or entropy-based indices[24]. A popular choice among Economists and Social Epidemiologists is the Theil index[37]. The Theil index measures the distribution of health among N population subgroups divided according to a given feature s, as:

$$T(s) = \sum_{i=1}^{N} T_i \quad \text{with} \quad T_i = w_i \log\left(\frac{w_i}{n_i}\right), \quad (3)$$

where w_i is the share of ill-health in group i, and n_i is the share of the population in group i. Higher values of the Theil index correspond to a higher degree of inequality among the groups. We propose the following definition of the fairness score for unordered social groups, based on the Theil index:

$$F(s) = e^{-\sum_{i=1}^{N} |T_i^m - T_i^d|}, \quad (4)$$

where T_i^m is the component of the Theil index of group i computed on the model's outcome, while T_i^d is the component of the Theil index of group i computed in the data. It is important to note that, in our definition of $F(s)$, differences between the model and data are computed for each component of the corresponding Theil index (i.e., T_i^m and T_i^d) and not between the corresponding total indices. We use the exponential function to normalize $F(s)$ in the range (0, 1], with 1 indicating that a model perfectly reproduces the distribution of health among the N groups reported in the data.

As an illustrative example, we model the COVID-19 pandemic in New York City during the first wave using the approach proposed by Ma et al.[38]. Based on the code shared by the authors, we develop a compartmental SLIR model that describes the spread of SARS-CoV-2 by partitioning the New York City population into five racial and ethnic groups: non-Hispanic Whites, Hispanics or Latinos, Black/African Americans, Asians, and Others. Following the approach outlined in ref. 38, we compare a range of model structures by integrating social contact matrices stratified by racial and ethnic groups, hypothesizing different mixing behaviors and group susceptibilities. We investigated a *variable susceptibility model*, a *proportionate mixing model*, and an *assortative mixing model*. In the *variable susceptibility model*, individuals in each group have different susceptibilities to the virus, while contact rates remain the same among the groups. In the *proportionate mixing model* and *assortative mixing model*, susceptibility is the same in all groups, while contact rates vary. Proportionate mixing assumes that the contact intensity between each pair of groups is proportional to the product of the total contact rates (i.e., total number of contacts

https://doi.org/10.1038/s41467-026-78981-0

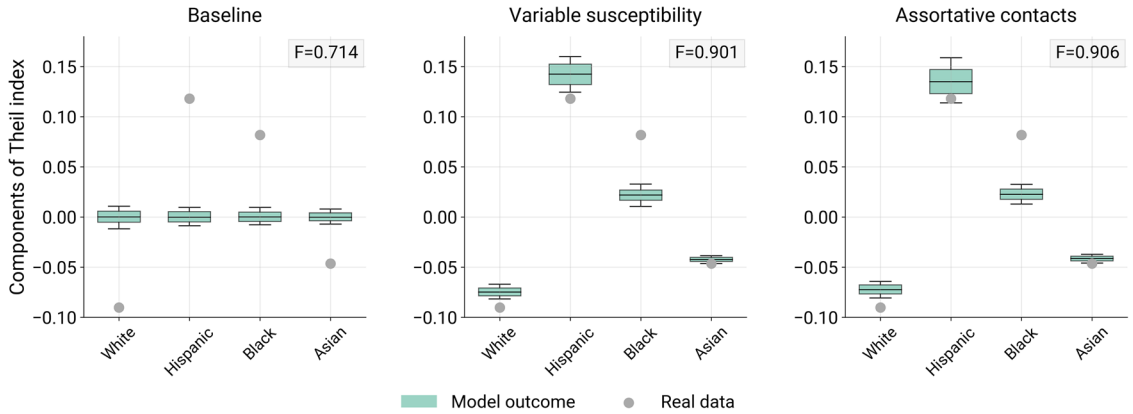


Fig. 2 | Models' fairness for racial and ethnic groups in New York City (unordered social groups). Components of the Theil index computed from Eq. 3, according to the distribution of COVID-19 deaths by race and ethnicity in New York during the first pandemic wave. In each panel, the Theil index components of real data are shown as gray circles. Boxplots represent the median, IQR, and 95% reference range computed from 1000 simulations of each model: baseline model (left), variable susceptibility model (center), assortative contacts model (right). The Theil index is computed from simulation outcomes at the end of the simulation period. The corresponding average fairness score of each model (i.e., Eq. 4) is shown in the top right corner.

per time period per individual) of those two groups. Assortative mixing extends proportionate mixing by partitioning a fraction f of contacts to be exclusively within the group while distributing the rest of the contacts according to proportionate mixing. For comparison, we define a *baseline model* with the same susceptibility and homogeneous mixing among different racial and ethnic groups. Moreover, we account for possible variations in the IFR. In contrast to the London age-stratified case, where the IFR is explicitly age-dependent, here we test two alternative assumptions for mortality risk across racial/ethnic groups: (i) a homogeneous IFR shared by all groups and (ii) a heterogeneous group-specific IFR (derived from age composition, see "Methods"). More details about the model implementation and calibration can be found in the "Methods" section and the Supplementary Information.

In total, we consider eight models, calibrating them to weekly COVID-19 deaths in New York City during the first wave, from March 15th to July 5th, 2020. We then assess their fairness concerning the racial/ethnic variable using the definition given in Eq. 4. Overall, the eight models exhibit a good fit for the total reported weekly deaths, with comparable performance shown in Table S4 of the Supplementary Information. However, their ability to capture the observed disparities in death rates by race or ethnicity varies significantly.

Figure 2 shows a comparison against the baseline of the two best models in terms of fairness (we refer the reader to Fig. S7 in the Supplementary Information for the complete comparison across all models). Each panel shows the four components of the Theil index, one for each racial or ethnic group, computed from the reported COVID-19 deaths and the model outcomes. As expected, the baseline model generates an equally distributed number of deaths among racial and ethnic groups, thus failing to capture the observed inequalities. Its fairness score is $F(race) = 0.714$ [0.705, 0.724]. By introducing variable susceptibility and maintaining a homogeneous IFR for all groups, the fairness of the model increases to $F(race) = 0.901$ [0.889, 0.912]. A model with a homogeneous IFR and assortative mixing achieves the highest fairness $F(race) = 0.906$ [0.893, 0.918]. Looking at the results for all models reported in Table S7 of the Supplementary Information, it emerges that models featuring heterogeneous IFR consistently display a lower fairness score than their counterparts with homogeneous IFR. For instance, a baseline model with equally distributed contact rates and heterogeneous IFR only achieves $F(race) = 0.621$ [0.613, 0.630] (Table S7). This suggests that the age-adjusted IFRs may not accurately reflect real-world disparities in COVID-19 death risk across racial and ethnic groups.

Fairness in spatially structured models

As a third case, we consider spatial epidemic models. When modeling the spatial spread of an epidemic, space is typically structured into a collection of units that naturally define subgroups of the population. Such subgroups often correspond to administrative regions, and they can either induce a population ordering or not. In this case, we show how both definitions of fairness proposed above (i.e., Eq. 2 and Eq. 4) can be applied to real-world epidemic scenarios involving spatial structure.

As a first example, we examine the spread of COVID-19 in the city of Santiago, Chile, during the first pandemic wave. We consider a previously published spatially structured epidemic model that integrates human mobility patterns derived from mobile phone data[39]. The model was shown to accurately capture the unequal infection patterns observed in Santiago in 2020. Indeed, changes in mobility patterns were strongly correlated with the Human Development Index (HDI) of Santiago's *comunas* (i.e., neighborhoods), leading to a higher burden of infection in the most deprived areas of the city (see Fig. 3a, b). To assess the role of human mobility in determining a model's fairness, we compare the outcomes of a model that integrates human mobility data with those of a model that disregards mobility (see the "Methods" and Supplementary Information for more details). We use the RCC as shown in Fig. 3c, where the population of Santiago's neighborhoods is ordered by HDI. The plot compares the cumulative proportion of COVID-19 cases reported in each community with the cumulative proportion of their population, ranked by HDI. While both models provide a good description of spatial inequalities, the RCCs show that the results of the SLIR model, which includes mobility, are closer to real data than those of the model that does not consider human movements. According to Eq. 2, the model with mobility is characterized by an average fairness score of $F(\text{HDI}) = 0.97$, while excluding mobility data decreases the score to 0.93. Alternatively, we can disregard the population ordering induced by the HDI and consider each comuna of Santiago as a population subgroup. Such a grouping is unordered, and we can compute the fairness score based on the Theil index (i.e., Eq. 4) of the two models. Figure 3d shows the components of the Theil index, one for each neighborhood, computed from the COVID-19 deaths and the average models' outcomes. Consistent with previous results, integrating mobile phone-derived mobility data increases the model's fairness, as evidenced by values of the Theil index components that are closer to the observed data. In this case, by including human mobility data, the average model's fairness increases from 0.63 to 0.79.

https://doi.org/10.1038/s41467-026-78981-0

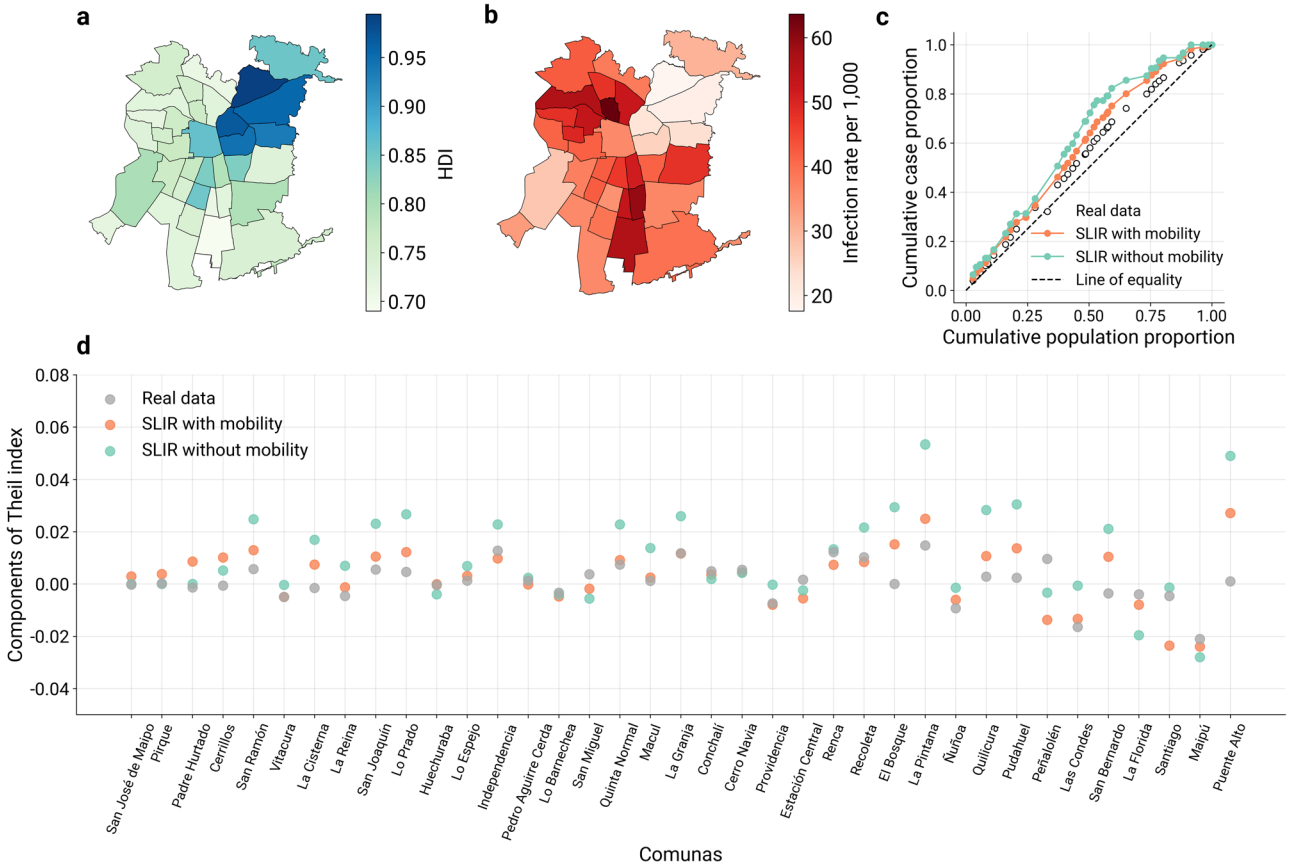


Fig. 3 | Fairness and the spatial spread of COVID-19 in Santiago, Chile (spatially structured models). **a** Choropleth map of the urban area of Santiago, where neighborhoods (i.e., comunas) are color-coded according to their Human Development Index (HDI). **b** COVID-19 infection rates reported in each comuna of Santiago during the first pandemic wave in 2020. **c** Comparison of the relative concentration curves of COVID-19 cumulative incidence reported by the surveillance (empty circles) and generated by two models: a spatial SLIR with mobility, and the same model without mobility flows. **d** Average value of T_i for each comuna computed from real data and models' outputs. In this case, comunas are considered as unordered units, and they are listed from left to right based on their administrative ID code.

In the Supplementary Information, as a second example, we examine the 2015–2016 Zika virus (ZIKV) outbreak in Colombia. Again, we investigate the role of human mobility in capturing the heterogeneous burden of infections in space and compare the fairness of two modeling approaches that integrate different mobility networks. To this end, we consider the framework proposed by Perrotta et al.[40], who modeled the spread of ZIKV in Colombia using a spatially structured model that incorporates human movements, either through mobile phone inferred mobility or through synthetic networks. Similarly to the results for Santiago, we observe that a model including mobile phone derived mobility data better captures the unequal distribution of disease across the departments of Colombia (see the Supplementary Information for more details).

We would like to stress that, although in the above examples we could consider the spatial partition of the population both as ordered and unordered, this is not always the case, and the fairness scores are not mutually exchangeable. Indeed, the RCI fairness (i.e., Eq. 2) can only be used if an equity-related variable allows for the ordering of subpopulations on a low-high gradient. In some cases, this may not be relevant for the disease under study or may be unfeasible due to the lack of data, thus leaving the Theil index score (i.e., Eq. 4) as the only option to evaluate a model's fairness.

Fairness across multiple dimensions

Finally, we extend our framework to models that include more than one equity-related feature at a time, where inequalities may arise jointly across dimensions. In this case, fairness should not be evaluated only marginally (e.g., by age or by race/ethnicity alone), but also across their intersection. Indeed, a model may be fair on a single dimension while still misrepresenting specific subpopulations.

To illustrate this point, we consider an agent-based model of COVID-19 in New York City stratified by age and neighborhood. We used Covasim[41], a popular open-source model developed in the midst of the COVID-19 pandemic. Covasim is a flexible framework for building a realistic age-structured synthetic population, including multiple contact layers across settings (see the Supplementary Information for more details). We contrasted different agent-based model structures incorporating two assumptions about the disease progression probabilities, which can be age-dependent or not, and the social contact structure, which can be random or hybrid. The random network connects agents through a single uniform layer. The hybrid network builds four layers: household, school, workplace, and community. Each layer features its own edge density and age mixing pattern, providing a more realistic representation of social interactions. We fitted the resulting four models to the weekly reported deaths in New York during the first COVID-19 pandemic (see the "Methods" and the Supporting Information for additional details). Each model accounts for both the population age structure and the spatial structure of the city, although inter-borough coupling is not explicitly represented since Covasim does not allow for spatial coupling. This allows us to compute the fairness scores separately for each dimension, using the RCI-based fairness for age and the Theil-based fairness for geography. To measure the total fairness of each model, we combined the two separate scores into a vector of two components:

https://doi.org/10.1038/s41467-026-78981-0

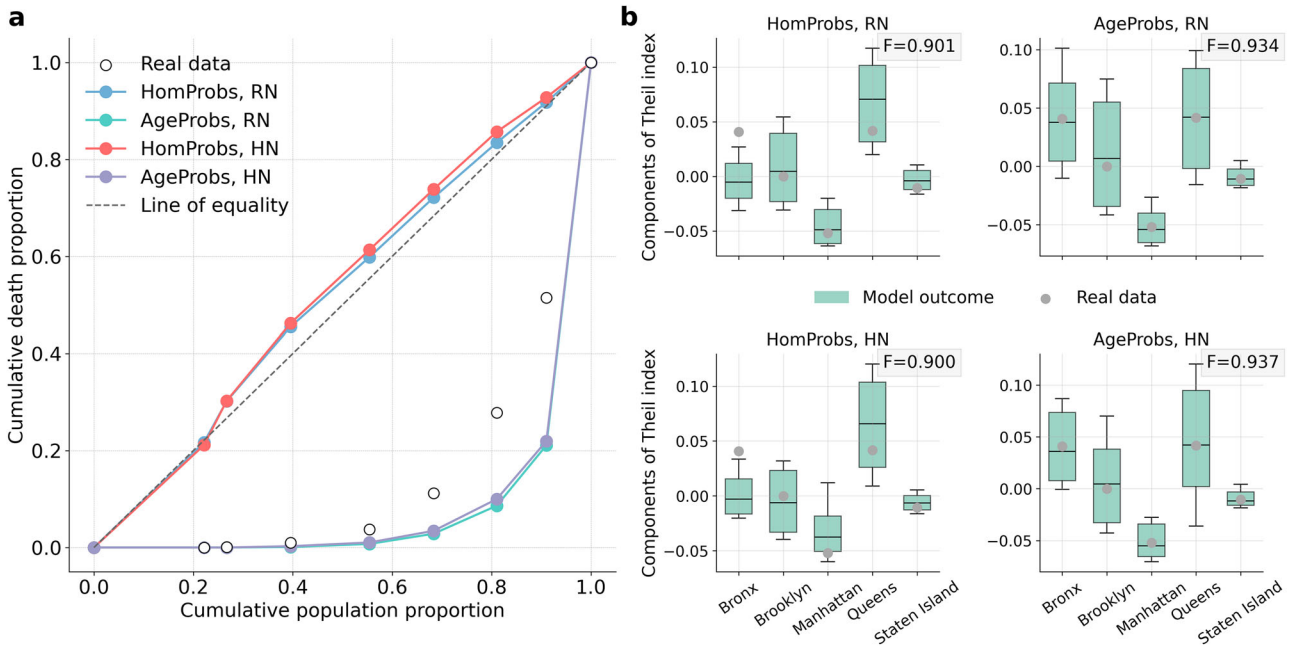


Fig. 4 | Fairness and agent-based models with multiple dimensions. **a** Comparison of the relative concentration curves of death proportion by age in New York City during the first COVID-19 wave. Open black circles correspond to the officially reported deaths. Colored curves correspond to four settings of Covasim[41], combining the assumption of age-dependent (AgeProbs) or constant (HomProbs) transitions between disease compartments and the assumption of a random contact network (RN) or hybrid network (HN). **b** Components of the Theil index from the four Covasim calibrated settings, measured according to the distribution of COVID-19 deaths in each NYC borough. In each panel, the Theil index components of real data are shown as gray circles. Boxplots represent the median, IQR, and 95% reference range computed from 1000 simulations of each model. The corresponding average fairness score of each model is shown in the top right corner.

$\mathbf{F} = (F(age), F(geo))$. The total fairness can be computed as the modulus of $\mathbf{F}$, i.e., $|\mathbf{F}|$, allowing us to compare different models across multiple dimensions.

Figure 4 shows the results of the agent-based models when evaluated by age (see Fig. 4a) and by borough (see Fig. 4b). The model with the highest fairness in both dimensions is the most expressive, which includes age-stratified disease progression and a hybrid network structure. Its fairness vector is $\mathbf{F}$ = (0.940, 0.937) (see Table S11 in the Supplementary Information for all models' results and uncertainties). As expected, the inclusion of age-stratified disease progression leads to the highest fairness values when measured through the RCI by age. Including a full hybrid network structure, however, does not provide a significant advantage when considering the spatial dimension, with the exception of the Bronx, the only borough where fairness improves with the hybrid model. This is exemplified by the Theil-index fairness scores of the age-stratified random network model ($F(geo)$ = 0.934 [0.868, 0.982]) compared to the age-stratified hybrid network model ($F(geo)$ = 0.937 [0.868, 0.981]). In conclusion, models that include age-proportional transition rates achieve very similar fairness scores on both dimensions, always higher than models assuming constant transition probabilities, and independent of the assumed contact network model.

Using model fairness to inform equitable vaccine allocation

So far, we have discussed the concept of fairness in terms of a model's ability to capture real-world disparities in health outcomes. However, fairness must relate not only to a model's realism but also to the potential for improving the equity of the decisions that policymakers will make based on the model's output. Decision making in the health domain faces significant trade-offs between algorithmic fairness constraints and the equity impact of decisions[29]. Here, we show how our definition of a model's fairness can guide the adoption of more equitable public health interventions.

To this end, we consider the COVID-19 epidemic in New York City and the model assuming assortative contacts and a homogeneous IFR, which shows the highest fairness, as described in Section "Fairness in Unordered Social Groups." We simulate a stylized intervention scenario in which public health officials must allocate a given number of protective resources to mitigate a second wave of infections. We compare the effectiveness of different allocation strategies in terms of the achieved relative reduction in deaths compared to an unmitigated scenario. For the sake of simplicity, we assume that preventive measures consist of a sterilizing vaccine; therefore, they effectively remove individuals from the pool of susceptible. Specifically, we consider a scenario in which policymakers must allocate V_T (which we assume covers 20% of the population) vaccine doses to the population, distributing them among the five racial and ethnic groups at the beginning of the second wave of the outbreak, triggered by the relaxation of NPIs (see Fig. 5) in July 2020. The model is calibrated and assessed in terms of fairness based on the reported COVID-19 deaths from the first wave. As for the second wave, we consider a purely hypothetical and simplified scenario obtained by simulating the relaxation of the NPIs and the emergence of a new variant with the calibrated model (see the Supplementary Information for more details).

Vaccine allocation can be done by assigning vaccines to each group proportionally to the relative size of the group population n_i. This is our baseline scenario that leads to an overall 68.2% [53.8–94.2%] relative reduction in the number of deaths (see Table 1). Alternatively, we consider allocating vaccines according to observed death rates in each group. A strategy that explicitly takes into account the risk of each group. In addition, we consider allocating the V_T doses using a formula that weights the population share of each group by the Theil index component introduced in Eq. 4. In this way, we directly introduce the fairness metric into the allocation, effectively linking our measure of the model's fairness to the decision-making process. Specifically, vaccines are distributed to each group according to their

https://doi.org/10.1038/s41467-026-78981-0

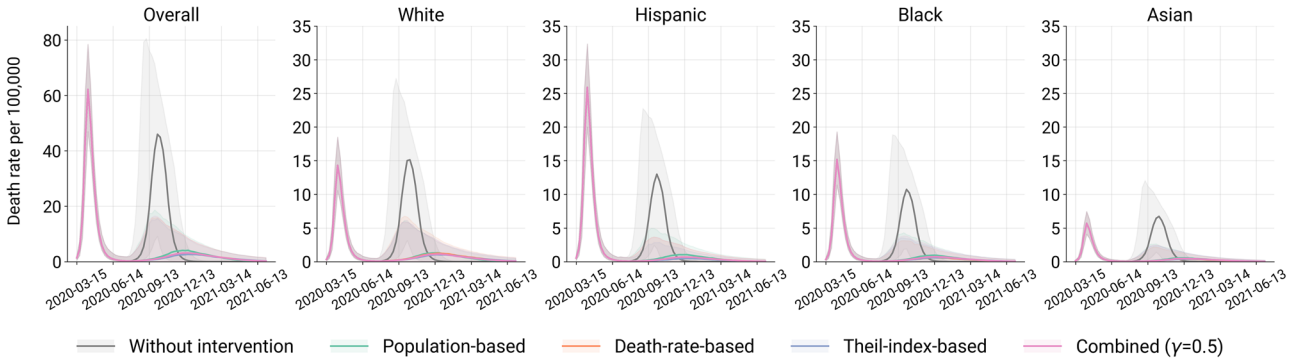


Fig. 5 | Comparison of vaccine allocation strategies on weekly deaths by race/ethnicity in NYC, given the number of vaccines V_T to cover 20% of the population in total. The plots show the complete deaths' trajectories during both waves. In the first wave, there is no vaccination; thus, the trajectories are overlapped. The vaccination is implemented after the first wave on 2020/07/06 in all scenarios. In the plot, we show the median of weekly deaths with a 95% confidence interval of 1000 samples.

Table 1 | Comparison of relative reduction in the number of deaths, by vaccine allocation strategy and racial/ethnic groups

Strategy	Relative reduction in deaths (%)				
	Overall	White	Hispanic	Black	Asian
Population-based	68.2 [53.8, 94.2]	68.9 [54.3, 94.4]	69.2 [54.7, 94.8]	67.7 [52.9, 94.3]	69.7 [55.0, 94.4]
Death-rate-based	72.1 [55.3, 97.1]	65.2 [45.7, 96.3]	77.1 [61.8, 97.9]	76.2 [61.5, 97.7]	75.3 [60.0, 97.6]
Theil-index-based	73.9 [56.6, 97.9]	69.2 [49.7, 97.6]	82.5 [68.7, 98.7]	73.2 [55.6, 97.9]	72.5 [53.9, 97.8]
Combined ($\gamma = 0.5$)	71.6 [55.3, 96.9]	69.4 [52.2, 96.6]	76.8 [62.2, 97.8]	70.8 [54.3, 97.0]	71.3 [54.8, 97.0]

Each entry reports the median reduction from 1000 samples and the 95% confidence interval.

corresponding values of $T_i = w_i \log\left(\frac{w_i}{n_i}\right)$, where w_i is the fraction of deaths reported by group i during the first wave. We then introduce T_i into the vaccine allocation, controlling the relative importance of the Theil index through a parameter γ ranging between 0 and 1 (see the "Methods" for more details). For $\gamma = 1$, we implement a strategy that is fully based on the Theil index components, while smaller values of γ take into account the group population size. Table 1 compares the relative reduction in deaths achieved by different vaccine allocation strategies based on the components of the Theil index for each racial/ethnic group. In general, allocation strategies that take disease burden into account achieve a larger overall reduction in deaths with respect to the population baseline. However, allocating vaccines according to group-specific death risk increases overall mortality reduction (72.1%) mainly in minority groups (Hispanic, Black, and Asian), while the majority group (Whites) experiences an increase in mortality relative to the population-based baseline. In contrast, by integrating the Theil index components into the allocation strategy, the Theil-index based strategy ($\gamma = 1$) and a combined strategy ($\gamma = 0.5$) achieve an overall median reduction of 73.9% and 71.6% in the number of deaths, respectively, with an increased median reduction in deaths across all groups relative to the baseline. In particular, for $\gamma = 1$, we achieve the most equitable outcome with the highest overall relative reduction of deaths and the best improvement in all racial/ethnic groups. The Theil-index based strategy consistently achieves the largest reduction in overall deaths when varying the value of V_T between 10% and 30% (see Tables S12 and S13 of the Supplementary Information). Finally, the results are robust to changes in the epidemic model and its assumptions about mixing (see Tables S10 and S11 in the Supplementary Information).

Discussion

Aiming to bring established principles and measures from Social Epidemiology[42] into Computational Epidemiology, we proposed a set of metrics to quantify and compare the fairness of models of infectious disease dynamics. In our framework, fairness is a diagnostic property of the model, quantifying how accurately it reproduces observed inequalities across population groups. Equity instead is understood as a normative objective that guides public health decisions. By contrasting different modeling approaches and disease settings, we showed that our fairness metrics can measure a model's ability to capture the observed inequalities in the distribution of illness between sociodemographic groups. Our results highlighted trade-offs between model complexity and fairness. We demonstrated that increasing model complexity through additional data layers or compartments, such as the presence of heterogeneous contact rates between groups, does not always enhance fairness. Similarly, models that achieved the best overall fit to disease data did not translate into the fairest outcomes when examining inequalities between sociodemographic groups. Therefore, our findings point to the idea that maximizing the fairness of an epidemic model cannot be simply achieved through total error minimization or by increasing models' complexity. Our fairness metrics offer an alternative to traditional accuracy measures by aligning more closely with the underlying concept of health inequality and its measures. In terms of practical guidance on how to include fairness metrics in epidemic modeling, our results suggest a three-stage workflow. First, for a set of candidate models with acceptable aggregate fit, compare fairness scores $F(s)$ for a pre-specified equity-related feature s (or a set of features). Second, prioritize models whose fairness distributions are consistently higher, while inspecting component-level disparities (RCI curves or Theil components) to identify which subgroups are misrepresented. Third, when informing public health strategies, use fairness-validated models to derive interventions from observed subgroup burden, and evaluate resulting equity and effectiveness outcomes under scenario uncertainty.

In Social Epidemiology, health inequalities are generally described as differences wherein socially disadvantaged groups, those who have consistently faced marginalization or discrimination, systematically

experience poorer health outcomes or higher health risks compared to more advantaged populations[25]. Determining which equity-relevant characteristics to include should rely on documented health or social heterogeneities across subpopulations, or align with the strategic interests of public health authorities and policymakers. In particular, the choice should prioritize those inequalities that can be addressed through intervention by policymakers. Indeed, as widely accepted in the field of machine learning and more broadly in artificial intelligence, algorithmic fairness goes beyond mere accuracy, placing greater emphasis on ensuring equitable outcomes[43]. To address this important point, we demonstrated that the definition of fairness in our framework provides a useful score to increase the equity of a hypothetical intervention, such as vaccine allocation during an outbreak. Indeed, our results show that when introduced as a guide for a vaccine allocation strategy, our measure of fairness reduces the overall disease burden while improving equity in disease outcomes between all groups. Prior work on vaccine allocation has extensively examined equity-aware intervention strategies[44,45], including fairness-efficiency trade-offs and prioritization of vulnerable populations, especially in response to the COVID-19 pandemic[22,46]. Recent reviews have further highlighted that many epidemic modeling studies primarily evaluate aggregate epidemiological outcomes, while subgroup-level disparities and equity considerations remain insufficiently assessed[23]. Our findings complement existing literature on equitable vaccination strategies by adding an explicit fairness criterion before policy optimization. In other words, the quality of subgroup representation in the model is treated as a prerequisite for trustworthy equity-oriented decisions. Furthermore, our results on vaccine allocation strategies are consistent with Chen et al.[22,46]. The equity/efficiency trade-off is weaker than often assumed, as it is possible to minimize total deaths while simultaneously reducing group-level disparities in mortality. Our experiments show that this pattern holds under minimal assumptions, and it can be achieved by incorporating our fairness-based weighting into the vaccination strategy. At the same time, our analysis remains stylized and does not resolve all ethical and implementation constraints of real campaigns (e.g., logistics, hesitancy, and governance). Future work should integrate these factors to evaluate fairness-informed allocation under operational conditions.

Our work is subject to limitations. First, we did not address the important issue of correctly identifying, from available data, the equity-related variables that must be included in an epidemic model and how a model can account for the multiple intervening mechanisms of disease exposure associated with one or more sociodemographic traits[47]. Indeed, the models we considered did not attempt to mechanistically link social causes with biological outcomes. This may explain why models showed low fairness scores in some settings, such as the ZIKV epidemic in Colombia. We expect that future modeling advances will improve the state-of-the-art through novel approaches that incorporate social drivers of inequity at their core, thus increasing models' fairness. Second, as already noted, we considered a simplified vaccine allocation scenario without aiming for realism. For instance, we did not address the ethical implications arising from a limited vaccine supply, the existence of adverse impacts from the vaccine[48], or the effects of partial/waning immunity. To effectively evaluate the relevance of our fairness metrics for decision making, a more extensive analysis is needed across different epidemic and intervention scenarios. This will be the subject of future work.

Finally, alternative inequality metrics could be considered to define the fairness scores. However, we chose those that satisfied three main requirements: being grounded in the theory of health inequalities and widely adopted in social epidemiology; being a summary of relative health inequality between social groups (and not absolute inequality); being parameter-free. Under these requirements, the set of available choices becomes narrow[24]. Nevertheless, we could have picked other mathematical definitions for $F(s)$. For instance, different distance functions or normalizations could translate inequality gaps into fairness scores while leaving RCI and Theil as the underlying measures of disparity. We selected the current definitions for interpretability and boundedness, and because component-level Theil comparisons better capture subgroup misrepresentation than comparisons based on the aggregate index alone.

Algorithmic fairness in the health domain, especially when related to clinical practice, represents a highly debated topic that has been extensively investigated[28,49]. In public and population health, algorithmic fairness has gained attention more recently, and its definition remains elusive as it intersects with complex social determinants, systemic inequities, and the limitations of available data[17]. For computational epidemic modeling, the debate is still in its infancy. Previous studies have examined the fairness of COVID-19 forecasting models in the USA[18–20], but they have not accounted for global measures of health disparities across subgroups, nor have they considered ways to evaluate the equity impact of policy decisions. Our study provides epidemic modelers with a principled and more general framework that can be easily adapted to different contexts and modeling approaches, as demonstrated by the examples we considered. We acknowledge that an important alternative to post hoc fairness evaluation is to incorporate equity-related metrics directly into model development. Indeed, for model selection, one could fit models to stratified surveillance data or include group-specific terms in the loss function, so that observed differences across age, race/ethnicity, geography, or socioeconomic groups are considered during estimation rather than assessed only after fitting. Likewise, for vaccine allocation, equity-aware optimization could target group-specific outcomes or explicit equity constraints could enter in the objective function. We view these approaches as complementary rather than competing. When stratified data are sufficiently informative, and the modeling framework supports group-specific estimation, incorporating fairness during fitting or optimization should generally be preferred. However, several practical considerations motivate a standardized post-estimation fairness metric. First, many outbreak analyses begin from models calibrated on aggregate outcomes, for which subgroup performance remains uncertain. Second, epidemic models differ substantially in structure (e.g., compartmental, metapopulation, and agent-based), making direct comparison across frameworks difficult; our metrics provide a common evaluation layer based on model outputs and observed subgroup burdens. Third, in many applications, stratified data are incomplete, noisy, or not used in calibration, so subgroup representation must be diagnosed after fitting. At the same time, we do not view our metrics as inherently limited to post-hoc evaluation. In principle, our fairness scores can be incorporated directly into model fitting or policy optimization as objective functions or constraints. We did not implement this integrated estimation-evaluation workflow here, but we consider it an important next step of our work.

Computational epidemiologists and mathematical modelers have recently begun to address the socioeconomic inequalities that characterize the spread of infectious diseases[10,13]. The COVID-19 pandemic has demonstrated that inequalities in epidemic outcomes are strongly associated with socioeconomic disparities, systemic racism, and (lack of) social justice[50–52]. Such inequalities were reflected by significant disparities in social contacts[53,54], the affordability of NPIs[45,55–57], access to healthcare[5], and job security[58]. To address these pressing issues, emerging literature has put forward novel approaches to integrate social mechanisms of exposure into traditional epidemic models, such as through generalized contact matrices[14,59] or the use of structural causal influence[60]. Our work adds to this current thread of research by providing modelers and public health officials with a framework that will help evaluate epidemic modeling outcomes through the lens of equity, thus reducing socioeconomic biases in outbreak management and response.

Methods

The study did not require ethical approval because it did not involve human subjects and it primarily relies on simulated data.

Epidemic models

In all cases, we considered computational epidemic models that are fully stochastic and simulated using chain binomial processes. In the following, we provide the details of each model structure, parameters, and underlying assumptions. Data and code to reproduce the simulations and analyses of this work are available on Zenodo[61]. All simulations and data analyses were performed using Python.

Simulating the spread of SARS-CoV-2 in London. We use a standard age-structured SLIRD (Susceptible, Latent, Infectious, Recovered, and Dead) compartmental model incorporating the impact of NPIs. Individuals are divided into compartments according to their health status. The transitions between compartments are as follows. Susceptible (S) individuals transition to the latent compartment after interacting with infectious individuals. Individuals in the latent (L) phase are infected but not infectious until they move to the infectious (I) phase after the latent period. Then, infectious individuals either die (D) with an infection fatality rate (IFR) or recover (R) from the disease at a rate of $1 - IFR$. To account for a delay in reporting deaths, we set a parameter of Δ days of delay, and the individuals in the D compartment are moved to D^o after Δ days. In calibration, the number of individuals in D^o is calibrated against the reported deaths.

The population is stratified into 16 age groups from 0 to 75+, with 5-year brackets except for 75+, featuring age-specific contact patterns and IFRs. We consider two types of mixing among age groups: (1) homogeneous and (2) heterogeneous mixing. In the models with homogeneous mixing, we assume that the number of contacts between any two age groups is proportional to the product of the populations in those two groups. Based on this assumption, the homogeneous matrix is constructed by redistributing the total contacts derived from ref. 35 to each pair of age groups (see Section 1.2 in the Supplementary Information for more details). In the models with heterogeneous mixing, contacts between age groups are described by contact matrices from Prem et al.[35]. To include the impact of NPIs, we use mobility data from the Google COVID-19 Community Mobility Report[62] and the Oxford Coronavirus Government Response Tracker (OxCGRT)[63] to estimate the reduction in overall contacts and adjust the contact matrices (see Section 1.1 in the Supplementary Information). We also consider two approaches to model deaths: (1) homogeneous and (2) heterogeneous IFRs. In the models with homogeneous IFR, infection fatality rates are the same across the age groups, and their values are computed as the population-weighted average of the age-specific values. In the models with heterogeneous IFR, we use the age-specific IFRs[34].

Simulating the spread of SARS-CoV-2 in New York City by race/ethnicity. Here, we adopt and extend the modeling framework proposed by Ma et al.[38]. While it shares many ingredients of the model we used in the case of London, such as the integration of the effects of NPIs, there are also some important differences. Indeed, instead of considering sub-groups by age, the population is divided into five racial and ethnic groups: non-Hispanic Whites, Hispanics or Latinos, Black/African Americans, Asians, and Others. Ma et al. proposed three models: the *variable susceptibility model*, the *proportionate mixing model*, and the *assortative mixing model*[38]. We adopt and extend these, building eight models by varying the assumptions regarding susceptibility, contact patterns, and IFR. In the *variable susceptibility model*, individuals from each racial/ethnic group have different susceptibilities to the virus, while the contact rates are identical across all groups. In the *proportionate mixing model*, individuals have the same susceptibility, while the contact rate between a pair of groups i and j is proportional to the product of the total contact rates of group i and j. In the *assortative mixing model*, the contact pattern is an extension of proportionate mixing, partitioning a fraction f of contacts to be exclusively within the group while distributing the rest of the contacts according to proportionate mixing. In addition to the above three models, we propose a baseline seeded with identical susceptibility across racial and ethnic groups, assuming a homogeneous mixing assumption. Then, we double the above four models by considering two alternative IFRs (homogeneous or heterogeneous), which results in a total of eight models (see Section 1.3 in the Supplementary Information for more details). We adopt the calibrated susceptibility estimates reported in Ma et al.[38] for the models with variable susceptibility and set susceptibility to 1 for the models assuming equal susceptibility. The contact matrices are also estimated from ref. 38. For each model, the corresponding formulations of the contact matrices and IFR are provided in the Supplementary Information.

Spatially structured metapopulation models. To simulate the spread of COVID-19 in Santiago de Chile and the spread of ZIKV in Colombia, we use two spatially structured metapopulation models that integrate human movement patterns and disease dynamics[64,65]. In both models, space is divided into patches corresponding to administrative subdivisions (i.e., neighborhoods in Santiago and departments in Colombia), which are connected by mobility fluxes. In each patch, individuals interact according to the specific compartmental model that describes the disease under study. Both models are stochastic, and transitions among compartments are simulated through chain binomial processes. In the case of Santiago, the results presented in this work are based on the data and simulations published in ref. 39. All the details regarding the model implementation, calibration, and analysis can be found therein. The model for Colombia, instead, is extensively described in ref. 40. The results presented in this work are based on the data and simulations published therein.

Agent-based model. We develop an agent-based model using Covasim (COVID-19 Agent-based Simulator), an open-source framework for simulating COVID-19 transmission dynamics. A detailed documentation of the model is provided in ref. 41. Here, we briefly summarize the configuration adopted for New York City. Individuals are assigned to the simulated population according to borough-specific age distributions derived from census data. Simulations run on 100,000 agents and are dynamically rescaled to match the true population size of each borough using Covasim's built-in dynamic rescaling technique. Household structure follows the default configuration provided by the model. Disease natural history follows Covasim's default SEIR-plus-severity progression, with default values for incubation, infectious period, and progression delays. Age heterogeneity is controlled through two mechanisms: (i) disease progression probabilities (i.e., the probabilities of developing symptoms, severe illness, critical illness, and death), and (ii) relative susceptibility. In the heterogeneous setting, both quantities are allowed to vary by age, following Covasim's default specifications. In the homogeneous setting, both progression probabilities and susceptibility are assumed constant across all age groups. We consider two contact structures. The random network connects agents through a single uniform layer. The hybrid network composes four layers which are household, school, workplace, community, each with its own edge density and age mixing pattern, providing a more realistic representation of social interactions. Combining these assumptions yields four model variants: (i) homogeneous progression probabilities with a random network (HomProb, RN), which serves as the baseline; (ii) homogeneous progression probabilities with a hybrid network (HomProb, HN); (iii) age-specific progression probabilities with a random network (HetProb, RN); and (iv) age-specific progression probabilities with a hybrid network (HetProb, HN).

Models calibration

We calibrate the models of COVID-19 spread in London and NYC by fitting the aggregated weekly deaths against the reported data. We use the Approximate Bayesian Computation Sequential Monte Carlo (ABC-SMC) algorithm[66] for calibration. The ABC-SMC algorithm is an extension of the ABC rejection method that utilizes generated intermediate distributions and decreasing rejection thresholds. The algorithm consists of T generations, and within each generation, M particles (sampled parameters) are maintained. In the first generation, parameters θ_i ($i = 1, 2, \cdots, M$) are sampled from the prior distribution $\pi(\theta)$, and θ_i are accepted if the distance $\rho(y^d, y^m)$ between the observed data y^d and the simulated results $y^m \sim f(\theta_i)$ is smaller than an initial threshold ϵ_1. In the following generations t ($t > 1$), particles are sampled from the previously accepted generation θ^*_{t-1} and perturbed by a kernel $K(\theta^*)$. They are accepted if the same distance $\rho(y^d, y^m)$ falls below ϵ_t, where ϵ_t is the threshold of generation t and $\epsilon_t < \epsilon_{t-1}$. We adopted a Python implementation of ABC-SMC from the library *pyabc*[67]. In our simulations, we set $T = 10$, $M = 1000$, and we take the last accepted generation as the posterior distribution of the parameters in our model. The settings for free parameters, along with their prior distributions and posterior distributions, are presented in the Supplementary Information.

In the case of the agent-based model, Covasim, calibration is performed against borough-level weekly mortality data using the built-in optimization framework Optuna within Covasim. We conduct 1000 optimization trials and retain the 100 posterior parameter sets that yield the lowest mean squared error (MSE) relative to observed data. The calibrated parameters are reported in the Supplementary Information.

Fairness and vaccine allocation strategies

We consider a counterfactual intervention scenario based on the model of COVID-19 spread in New York City. Specifically, we assume that public health officials allocate a given number of vaccines (V_T) to the population at the beginning of the second wave of the outbreak. At the end of the first wave, which is calibrated to the reported COVID-19 deaths, we simulate a second wave by assuming the emergence of a new variant with a 50% higher transmission rate. We also assume that the relaxation of NPIs leads to higher contact rates compared to the last day of the first wave (see Section 4.1 in the Supplementary Information for more details). For simplicity, we assume that only susceptible individuals are vaccinated and that the vaccines provide sterilizing immunity. In other words, we remove a number of V_T individuals from the pool of susceptible, and they cannot be infected. In this setting, we investigate three vaccine allocation strategies:

- *Population-based strategy*. Vaccines are distributed to each group proportionally to the population size of that group. The number of vaccines allocated in group i (denoted by V_i) is $V_i = V_T \cdot \frac{N_i}{\sum_i N_i}$
- *Death-rate-based strategy*. Vaccines are distributed to each group proportionally to the death rate of that group observed during the first wave. The number of vaccines allocated in group i (denoted by V_i) is $V_i = V_T \cdot \frac{DR_i}{\sum_i DR_i}$, where DR_i is the death rate of group i.
- *Theil index-based strategy*. We consider both the population size and the Theil index component of each group for vaccine distribution. The number of vaccines distributed to group i is

$$V_i = V_T \cdot \frac{\alpha_i}{\sum_i \alpha_i} \tag{5}$$

where we define α_i as

$$\alpha_i = \left(1 + \gamma \frac{T_i}{\sum_i |T_i|}\right) n_i \tag{6}$$

where T_i is the component of the Theil index of group i (Eq. 3), and γ is a parameter controlling the importance of Theil index components, ranging in $[0, 1]$. If $\gamma = 1$, the importance of the Theil index is maximized, and we define it as the full Theil-based strategy. If $\gamma = 0$, the strategy is reduced to a population-based strategy. Intermediate values of γ define allocation strategies that mix the population distribution and the Theil index components. In principle, γ can be treated as a tunable parameter within an optimization routine to identify allocation rules that balance overall effectiveness and equity objectives.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data used to generate all figures in the paper are available in the Zenodo database under accession code: https://zenodo.org/records/22753990.

Code availability

The code that reproduces the simulations of COVID-19 spread in London and New York City, including the vaccination campaigns, and all corresponding measures of fairness, is available in the Zenodo database under accession link: https://zenodo.org/records/22753990. Code for reproducing other simulation scenarios (COVID-19 in Santiago, ZIKV in Colombia) is publicly available through the corresponding cited works.

References

1. Lofgren, E. T. et al. Mathematical models: a key tool for outbreak response. *Proc. Natl. Acad. Sci. USA* **111**, 18095–18096 (2014).
2. Vespignani, A. et al. Modelling COVID-19. *Nat. Rev. Phys.* **2**, 279–281 (2020).
3. Zelner, J. et al. Racial disparities in coronavirus disease 2019 (COVID-19) mortality are driven by unequal infection risks. *Clin. Infect. Dis.* **72**, e88–e95 (2021).
4. Sweeney, S. et al. Exploring equity in health and poverty impacts of control measures for SARS-CoV-2 in six countries. *BMJ Glob. Health* **6**, e005521 (2021).
5. Bambra, C., Riordan, R., Ford, J. & Matthews, F. The COVID-19 pandemic and health inequalities. *J. Epidemiol. Community Health* **74**, 964–968 (2020).
6. Zaki, B. L., Nicoli, F., Wayenberg, E. & Verschuere, B. Contagious inequality: economic disparities and excess mortality during the COVID-19 pandemic. *Policy Soc.* **41**, 199–216 (2022).
7. Fox, S. J. et al. Disproportionate impacts of COVID-19 in a large US city. *PLoS Comput. Biol.* **19**, e1011149 (2023).
8. Diez Roux, A. V. Social epidemiology: past, present, and future. *Annu. Rev. Public Health* **43**, 79–98 (2022).
9. Farmer, P. Social inequalities and emerging infectious diseases. *Emerg. Infect. Dis.* **2**, 259–269 (1996).
10. Tizzoni, M. et al. Addressing the socioeconomic divide in computational modeling for infectious diseases. *Nat. Commun.* **13**, 1–7 (2022).
11. Buckee, C., Noor, A. & Sattenspiel, L. Thinking clearly about social aspects of infectious disease transmission. *Nature* **595**, 205–213 (2021).
12. Naidoo, M. et al. Incorporating social vulnerability in infectious disease mathematical modelling: a scoping review. *BMC Med.* **22**, 125 (2024).
13. Richard, D. M. & Lipsitch, M. What's next: using infectious disease mathematical modelling to address health disparities. *Int. J. Epidemiol.* **53**, dyad180 (2024).

14. Manna, A., Dall'Amico, L., Tizzoni, M., Karsai, M. & Perra, N. Generalized contact matrices allow integrating socioeconomic variables into epidemic models. *Sci. Adv.* **10**, eadk4606 (2024).
15. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K. & Galstyan, A. A survey on bias and fairness in machine learning. *ACM Comput. Surv. (CSUR)* **54**, 1–35 (2021).
16. Ferrara, E. Fairness and bias in artificial intelligence: a brief survey of sources, impacts, and mitigation strategies. *Sci* **6**, 3 (2023).
17. Mhasawade, V., Zhao, Y. & Chunara, R. Machine learning and algorithmic fairness in public and population health. *Nat. Mach. Intell.* **3**, 659–666 (2021).
18. Tsai, T. C. et al. Algorithmic fairness in pandemic forecasting: lessons from COVID-19. *NPJ Digit. Med.* **5**, 59 (2022).
19. Erfani, A. & Frias-Martinez, V. A fairness assessment of mobility-based COVID-19 case prediction models. *PLoS ONE* **18**, e0292090 (2023).
20. Abrar, S. M., Awasthi, N., Smolyak, D., Sigalo, N. & Martinez, V. F. Auditing the fairness of the US COVID-19 forecast hub's case prediction models. *PLoS ONE* **20**, e0319383 (2025).
21. Yi, M. & Marathe, A. Fairness versus efficiency of vaccine allocation strategies. *Value Health* **18**, 278–283 (2015).
22. Stafford, E., Dimitrov, D., Ceballos, R., Campelia, G. & Matrajt, L. Retrospective analysis of equity-based optimization for COVID-19 vaccine allocation. *PNAS Nexus* **2**, pgad283 (2023).
23. Rumpler, E. & Lipsitch, M. Equity considerations in COVID-19 vaccine allocation modelling: a methodological study. *Interface Focus* **15**, 20240037 (2025).
24. Oakes, J. M. & Kaufman, J. S. *Methods in Social Epidemiology* (John Wiley & Sons, 2017).
25. Braveman, P. Health disparities and health equity: concepts and measurement. *Annu. Rev. Public Health* **27**, 167–194 (2006).
26. Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2008).
27. Wallinga, J., Teunis, P. & Kretzschmar, M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. J. Epidemiol.* **164**, 936–944 (2006).
28. Rajkomar, A., Hardt, M., Howell, M. D., Corrado, G. & Chin, M. H. Ensuring fairness in machine learning to advance health equity. *Ann. Intern. Med.* **169**, 866–872 (2018).
29. Chohlas-Wood, A., Coots, M., Goel, S. & Nyarko, J. Designing equitable algorithms. *Nat. Comput. Sci.* **3**, 601–610 (2023).
30. Wagstaff, A., Paci, P. & Van Doorslaer, E. On the measurement of inequalities in health. *Soc. Sci. Med.* **33**, 545–557 (1991).
31. Konings, P. et al. Analysis of socioeconomic health inequalities using the concentration index. *Int. J. Public Health* **55**, 71–74 (2010).
32. Nguyen, Q. D., Chang, S. L., Jamerlan, C. M. & Prokopenko, M. Measuring unequal distribution of pandemic severity across census years, variants of concern and interventions. *Popul. Health Metr.* **21**, 17 (2023).
33. Yanez, N. D., Weiss, N. S., Romand, J.-A. & Treggiari, M. M. COVID-19 mortality risk for older men and women. *BMC Public Health* **20**, 1742 (2020).
34. Verity, R. et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect. Dis.* **20**, 669–677 (2020).
35. 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).
36. Mistry, D. et al. Inferring high-resolution human mixing patterns for disease modeling. *Nat. Commun.* **12**, 323 https://doi.org/10.1038/s41467-020-20544-y (2021).
37. Theil, H. & Finizza, A. J. A note on the measurement of racial integration of schools by means of informational concepts. *J. Math. Sociol.* **1**, 187–194 (1971).
38. Ma, K. C., Menkir, T. F., Kissler, S., Grad, Y. H. & Lipsitch, M. Modeling the impact of racial and ethnic disparities on COVID-19 epidemic dynamics. *Elife* **10**, e66601 (2021).
39. Gozzi, N. et al. Estimating the effect of social inequalities on the mitigation of COVID-19 across communities in Santiago de Chile. *Nat. Commun.* **12**, 1–9 (2021).
40. Perrotta, D. et al. Comparing sources of mobility for modelling the epidemic spread of Zika virus in Colombia. *PLoS Negl. Trop. Dis.* **16**, e0010565 (2022).
41. Kerr, C. C. et al. Covasim: an agent-based model of COVID-19 dynamics and interventions. *PLoS Comput. Biol.* **17**, e1009149 (2021).
42. Kakwani, N., Wagstaff, A. & Van Doorslaer, E. Socioeconomic inequalities in health: measurement, computation, and statistical inference. *J. Econ.* **77**, 87–103 (1997).
43. Corbett-Davies, S., Gaebler, J. D., Nilforoshan, H., Shroff, R. & Goel, S. The measure and mismeasure of fairness. *J. Mach. Learn. Res.* **24**, 1–117 (2023).
44. Dai, Z.-X., Lan, H.-J., Hai, N., Wang, J.-Y. & Wang, H.-H. Balancing fairness and efficiency in dynamic vaccine allocation during major infectious disease outbreaks. *Sci. Rep.* **15**, 1371 (2025).
45. Gozzi, N. et al. Estimating the impact of COVID-19 vaccine inequities: a modeling study. *Nat. Commun.* **14**, 3272 (2023).
46. Chen, L. et al. Strategic COVID-19 vaccine distribution can simultaneously elevate social utility and equity. *Nat. Hum. Behav.* **6**, 1503–1514 (2022).
47. Zelner, J. et al. There are no equal opportunity infectors: epidemiological modelers must rethink our approach to inequality in infection risk. *PLoS Comput. Biol.* **18**, e1009795 (2022).
48. Zarebski, A. E. et al. Including frameworks of public health ethics in computational modelling of infectious disease interventions. *Interface Focus* **15**, 20250004 (2025).
49. Xu, J. et al. Algorithmic fairness in computational medicine. *EBioMedicine* **84** (2022).
50. Klein, B. et al. COVID-19 amplified racial disparities in the US criminal legal system. *Nature* **617**, 344–350 (2023).
51. Nande, A. et al. The effect of eviction moratoria on the transmission of SARS-CoV-2. *Nat. Commun.* **12**, 2274 (2021).
52. Mena, G. E. et al. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science* **372**, eabg5298 (2021).
53. Manna, A., Koltai, J. & Karsai, M. Importance of social inequalities to contact patterns, vaccine uptake, and epidemic dynamics. *Nat. Commun.* **15**, 4137 (2024).
54. Tizzani, M. & Gauvin, L. Socioeconomic determinants of protective behaviors and contact patterns in the post-COVID-19 pandemic era: a cross-sectional study in Italy. *PLoS Comput. Biol.* **21**, e1013262 (2025).
55. Kollepara, P. et al. Socioeconomic correlates of urban mobility trends in two Australian cities during transitional periods of the COVID-19 pandemic. *R. Soc. Open Sci.* **12**, 241463 (2025).
56. Gauvin, L. et al. Socio-economic determinants of mobility responses during the first wave of COVID-19 in Italy: from provinces to neighbourhoods. *J. R. Soc. Interface* **18**, 20210092 (2021).
57. Perra, N. Non-pharmaceutical interventions during the COVID-19 pandemic: a review. *Phys. Rep.* **913**, 1–52 (2021).
58. Pangallo, M. et al. The unequal effects of the health–economy trade-off during the COVID-19 pandemic. *Nat. Hum. Behav.* **8**, 264–275 (2024).
59. Di Domenico, L., Reichmuth, M. L. & Althaus, C. L. Individual-based and neighbourhood-based socio-economic factors relevant for contact behaviour and epidemic control. *Commun. Med.* **6**, 26 (2026).
60. Surasinghe, S., Manivannan, S. N., Scarpino, S. V., Crawford, L. & Ogbunugafor, C. B. Structural causal influence captures the forces

of social inequality in models of infectious disease. *Biol. Lett.* **22**, 20250655 (2026).
61. Li, Y. & Tizzoni, M. Fairness in computational models of infectious disease epidemics. Software, version v1.0.0. Preprint at https://doi.org/10.5281/zenodo.22753990 (2026).
62. Google LLC "Google COVID-19 Community Mobility Reports" https://www.google.com/covid19/mobility/.
63. Oxford COVID-19 Government Response Tracker https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker#data.
64. Colizza, V. & Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *J. Theor. Biol.* **251**, 450–467 (2008).
65. Tizzoni, M. et al. On the use of human mobility proxies for modeling epidemics. *PLoS Comput. Biol.* **10**, e1003716 (2014).
66. Minter, A. & Retkute, R. Approximate Bayesian computation for infectious disease modelling. *Epidemics* **29**, 100368 (2019).
67. Schälte, Y., Klinger, E., Alamoudi, E. & Hasenauer, J. pyABC: efficient and robust easy-to-use approximate Bayesian computation. *J. Open Source Softw.* **7**, 4304 https://doi.org/10.21105/joss.04304 (2022).

Acknowledgements

We thank Daniela Perrotta for sharing simulation results from her previously published work, which were helpful in developing the results presented in this paper. We gratefully thank Alessandro Vespignani for providing useful comments on an early version of the manuscript. M.T. thanks the Mathematical Research Institute MATRIX in Australia, where part of this research was initially conceived.

Author contributions

M.T. conceived the project; N.G., N.P., and M.T. designed the study and supervised the research; Y.L. investigated and visualized the results; Y.L., N.G., N.P., and M.T. contributed to the methodology; M.T. contributed to the writing-original draft; Y.L., N.G., N.P., and M.T. contributed to the writing, review, and editing.

Funding

N.G. acknowledges support from the Lagrange Project of the Institute for Scientific Interchange Foundation funded by Fondazione Cassa di Risparmio di Torino. M.T. acknowledges funding from the MATRIX-Simons Travel Grant and from the Open Access Initiative of the Department of Sociology and Social Research of the University of Trento.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains Supplementary Material available at https://doi.org/10.1038/s41467-026-78981-0.

Correspondence and requests for materials should be addressed to Michele Tizzoni.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

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

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

© The Author(s) 2026

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.