Strategic COVID-19 vaccine distribution can simultaneously elevate social utility and equity

Lin Chen1,2,10, Fengli Xu3,4,10✉, Zhenyu Han2, Kun Tang5, Pan Hui1,6,7,8✉, James Evans3,9✉, and Yong Li2✉

Balancing social utility and equity in distributing limited vaccines is a critical policy concern for protecting against the prolonged COVID-19 pandemic and future health emergencies. What is the nature of the trade-off between maximizing collective welfare and minimizing disparities between more and less privileged communities? To evaluate vaccination strategies, we propose an epidemic model that explicitly accounts for both demographic and mobility differences among communities and their associations with heterogeneous COVID-19 risks, then calibrate it with large-scale data. Using this model, we find that social utility and equity can be simultaneously improved when vaccine access is prioritized for the most disadvantaged communities, which holds even when such communities manifest considerable vaccine reluctance. Nevertheless, equity among distinct demographic features may conflict; for example, low-income neighbourhoods might have fewer elder citizens. We design two behaviour-and-demography-aware indices, community risk and societal risk, which capture the risks communities face and those they impose on society from not being vaccinated, to inform the design of comprehensive vaccine distribution strategies. Our study provides a framework for uniting utility and equity-based considerations in vaccine distribution and sheds light on how to balance multiple ethical values in complex settings for epidemic control.

Sweeping the globe since its outbreak, the COVID-19 pandemic continues to impose harm and loss on human societies everywhere1. This has revealed and intensified disparities in the health conditions of different communities because of systemic inequities in COVID-19 exposure and access to health systems2–4. In the fight against the pandemic, vaccines are regarded as the most critical medical resources, but they still face prominent shortages in many countries and communities5–6. As a result, vaccine prioritization has become a critical policy task in every public health system7–9, with well-considered strategies to balance multiple ethical values. Facing the pandemic as a shock to the whole society, we argue that all members should have an equal right and opportunity to attain the best protection from the pandemic. In this light, our aim is to achieve a desirable balance between social utility and equity, where social utility is defined as the prevention of mortality in the entire population, and equity is defined as the mitigation of mortality disparities in disadvantaged demographic groups10–13. These two goals represent the most visible metrics considered by health authorities and organizations worldwide14–16, undergirded by the contrasting ethical values of utilitarianism and egalitarianism. As J. Bentham put it, “the well-being of a portion of individuals” can sometimes be sacrificed to achieve “the greatest happiness of the greatest number”17. Previous research has identified trade-offs between social utility and equity in the distribution of health-care resources ranging from disease screening to treatment18–22. In COVID-19 vaccine distribution, the most recent studies have focused on the trade-off between minimizing the years of life lost and minimizing the number of lives lost23–27, both of which reflect social-utility-oriented values and neglect disparities across the population. In light of this gap, we aim to reveal the relationship between social utility and equity in COVID-19 vaccine distribution, with critical implications for designing vaccine prioritization.

Examining the social utility and equity of vaccine distribution strategies requires an epidemiological model that can capture the uneven risks faced by different communities18— for example, older persons and those with greater mobility have higher COVID-19 risk26–30. However, standard epidemiological models (for example, susceptible–infectious–recovered (SIR)31,32 and susceptible–exposed–infectious–recovered (SEIR)33,34) are built on the assumption of homogeneous population mixing, which prohibits them from capturing heterogeneity in the spread of coronavirus. Some recent work35–37 has aimed to augment standard epidemiological models with empirical mobility data, but these neglect inherent vulnerability differences embedded in demographic profiles24. Here we propose an epidemiological model that simultaneously captures heterogeneity in the mobility patterns and demographic profiles of different communities. Calibrated with large-scale mobility and census data covering more than 75 million residents in the United States, our model automatically tunes the dynamics of coronavirus spread within each community on the basis of its demographic profile and traces dispersion among communities with time-varying empirical mobility flows. Our model can accurately predict the number of daily deaths and reconstruct its uneven distribution among communities, enabling the evaluation of equity among communities defined by different demographics—that is, older adult ratio, average household income, essential worker ratio and racial-ethnic minority ratio,
which have garnered considerable attention throughout the pandemic\cite{4,28–30,35,36,39}. Accordingly, we examine four vaccine distribution strategies that prioritize communities on the basis of vulnerabilities defined by these four demographic features. We find that social utility and equity can be simultaneously improved when prioritizing the most disadvantaged communities in each demographic dimension for vaccine access. Such a result holds even when low-income communities show considerable vaccine hesitancy, which contrasts with the conventional view of inevitable trade-offs\cite{13–22}. Nevertheless, elevating the equity across one demographic dimension can degrade it across others, suggesting that demographic features alone are insufficient to guide vaccine distribution. To overcome this, we propose two demography-and-behaviour-aware indices, community risk and societal risk, designed to measure the effect of prioritizing each community for vaccination to reduce (1) its own mortality risks and (2) the mortality risks it imposes on society as a whole. On the basis of these two indices, we design a framework for vaccine prioritization that simultaneously improves social utility and equity in all dimensions across scenarios of different vaccination rates and timing. By providing a general framework to tease out utilitarian and egalitarian values in COVID-19 vaccine distribution, our findings carry broad implications for the design of vaccination strategies.

Results

Behaviour- and demography-informed epidemic modelling. To capture the heterogeneity in health risk faced by different communities\cite{3,4,28–30,35,38,39}, we propose an epidemic model that principally integrates two important factors in the spread of coronavirus—that is, demographic profiles and mobility behaviours (Fig. 1a). Specifically, demographic profile is found to be substantially correlated with the fatality rate of SARS-CoV-2 infection\cite{28,40}, while mobility behaviour determines the likelihood of exposure to coronavirus\cite{30,35,37,39,41}. The proposed behaviour- and demography-informed epidemic model (BD model) therefore divides the studied population on the basis of the minimum geographical units defined by the United States Census Bureau, known as census block groups (CBGs)\cite{42}, and maintains a local SEIR model for each of them to characterize the dynamics of intra-CBG epidemic spread, where the infection-fatality rate (IFR) is adjusted on the basis of the demographic profile and age-specified risks estimated in previous medical research\cite{28} (Supplementary Table 1). To capture inter-CBG transmission resulting from urban mobility, the proposed BD model constructs a bipartite network linking CBGs and points of interest (POIs) with time-varying edges to track hourly movements extracted from the SafeGraph dataset\cite{43}, where the edge weights reflect temporal mobility intensity. New infections occur in POIs and CBGs with different probabilities determined by environmental characteristics, and infected populations travel to other communities proportional to their extracted movements (Methods, ‘Epidemic model, calibration and preliminary analysis’).

We evaluate the proposed BD model in nine large metro statistical areas (MSAs) in the United States covering over 75 million people and compare it with two baseline models: a standard SEIR model and a metapopulation model that only considers heterogeneous demographic profiles. A Homogeneous strategy, which provides vaccine access to each community with uniform probability\cite{49}. Our experiments show that strategies that prioritize the worst-off communities drastically improve the equity defined on the corresponding demographic

Consequences of alternative vaccine distribution strategies. Social utility and equity represent the two most important concerns considered by public health policy makers in the COVID-19 pandemic\cite{11–16}. These account for the collective welfare of the entire society and disparities among individual communities. As a critical policy concern during the COVID-19 pandemic, discussions of vaccine distribution strategies have centred on the trade-off between social utility and equity\cite{4}, but this has been inadequately evaluated or supported with empirical data. With the proposed BD model, we aim to reveal mechanisms behind social utility and equity in vaccine distribution with large-scale empirical data. In vaccine distribution, we quantify social utility as a reduction in the overall fatality rate, and we quantify equity as a reduction in the Gini coefficient of fatality rates among communities (Methods, ‘Quantification of social utility and equity’). On the basis of previous analyses that reveal heterogeneous health risks faced by populations with different demographic profiles (Fig. 1d), we focus on four dimensions of equity among communities: equity among age groups, income groups, occupational groups and racial/ethnic groups.

Prioritizing the least advantaged populations is acknowledged as a fundamental value in health-care resource allocation\cite{46–48}. We therefore construct four vaccine distribution strategies that prioritize the most disadvantaged communities defined on four dimensions of demographic profiles: older adult ratio (Prioritize by Age), average household income (Prioritize by Income), essential worker ratio (Prioritize by Occupation) and minority ratio (Prioritize by Race/Ethnicity). As a baseline for comparison, we also construct a Homogeneous strategy, which provides vaccine access to each community with uniform probability\cite{47}. Our experiments show that strategies that prioritize the worst-off communities drastically improve the equity defined on the corresponding demographic

ARTICLES
dimension (Fig. 2a). Specifically, equity defined on age, income, essential worker ratio or minority ratio can be improved by 26.2% (95% CI, 0.168–0.356; P < 0.001), 40.5% (95% CI, 0.269–0.541; P < 0.001), 43.1% (95% CI, 0.315–0.547; P < 0.001) or 31.8% (95% CI, 0.179–0.456; P < 0.001) compared with the Homogeneous baseline, respectively. Moreover, in most cases, all four strategies also achieve improvement in social utility compared with the baseline (Fig. 2a), suggesting that overall social utility and equity defined on a specific demographic dimension are likely to be simultaneously improved by prioritizing the most disadvantaged communities within that dimension. This sharply contradicts the conventional view of an inevitable trade-off between social utility and equity19–22. Detailed analysis reveals that prioritizing communities under greater risk consistently results in larger improvements for both social utility and equity, which further highlights the effectiveness achieved by prioritizing the worst-off communities (Supplementary Table 2).

We further consider and incorporate the potential impacts of vaccine hesitancy and administration capability into our experiments, both of which can substantially undermine the benefit of distributed vaccines. Vaccine hesitancy refers to the phenomenon wherein people mistrust and refuse to take vaccines despite availability,
especially in low-income communities\(^{25,31}\). Administration capability refers to the limited capability of local facilities to roll out vaccines, which can be constrained by storage capacity, human force and other institutional factors\(^{33-34}\). First, we construct an Estimated Hesitancy scenario by setting vaccine hesitancy rates for different communities according to a population sample-based national assessment\(^{35}\). Second, we construct a Hesitancy + Capability scenario by further superimposing administration capability onto communities, which is estimated from empirical vaccination data provided by the US Centers for Disease Control and Prevention (CDC)\(^{36}\). Specifically, administration capability is estimated as the difference between vaccine acceptance hesitancy derived from surveys and the empirical vaccination rate. It captures the percentage of residents willing to take vaccines but who have not done so, probably due to limited administration capacity (for example, no nearby vaccine administration facility). To investigate how different patterns of vaccine hesitancy affect vaccination strategy results, we further design three hypothetical scenarios, where we set vaccine hesitancy rates from the bottom to top income groups at 0.4, 0.3, 0.1 and 0 (Hypothetical-1); 0.8, 0.6, 0.4, 0.2 and 0 (Hypothetical-2); and 0.9, 0.7, 0.5, 0.3 and 0 (Hypothetical-3) (Methods, 'Vaccination scenarios'). These scenarios sequentially reflect larger differences across income groups and thus allow us to explore more extreme disparities in vaccine hesitancy than we currently observe. In general, improvements in social utility diminish as differences in vaccine acceptance rates grow larger (Fig. 2b). Among the tested prioritization strategies, Prioritize by Income is most sensitive to changes in vaccine hesitancy rates. Nevertheless, its improvement to social utility does not vanish until the vaccine hesitancy rate in the bottom income group rises to five times that in the top group, as in Hypothetical-2, explained by vastly disproportionate risks facing different income groups in the pandemic. This hypothetical hesitancy is far larger than what we observe from data. As expected, in all five scenarios, prioritizing the most disadvantaged communities consistently and significantly improves equity. By following the robust guideline to prioritize the most disadvantaged, social utility and equity can both be improved even if the most disadvantaged groups manifest the highest hesitancy. This demonstrates the out-sized protective impact that would accrue to society from far greater investments in vaccination outreach, education and incentive for our most disadvantaged communities.

Fig. 2 | Social utility and equity under different vaccine distribution strategies. a, Changes in social utility and equity, compared with the Homogeneous baseline. The red and blue points represent the strategies prioritizing the most and least disadvantaged communities, respectively. In each plot, the first to the fourth quadrants represent (1) simultaneously improving utility and equity, (2) improving equity but damaging utility, (3) simultaneously damaging utility and equity, and (4) improving utility but damaging equity. b, Change in social utility under different scenarios of vaccine hesitancy (nine MSAs). The bottom and top of each box indicate the 25th and 75th percentile values. The line inside the box represents the median value of the results. When vaccine hesitancy in low-income communities is stronger, the benefit to social utility brought by prioritizing disadvantaged communities diminishes and is eventually erased, making it inferior to the baseline.
simply prioritizing the worst-off communities (Table 1). To reveal the relationships between disadvantages along different demographic dimensions, we analyse correlations among the older adult ratio, average household income, essential worker ratio and minority ratio for each CBG (Methods, ‘Correlation analysis of demographic features’). We observe a positive correlation between older adult ratio and average household income ($r = 0.14$), indicating that populations with older demographics tend to have higher incomes. Essential worker ratio negatively correlates with older adult ratio ($r = -0.2$) but positively correlates with average household income ($r = 0.28$), indicating that populations with larger proportions of essential workers tend to be younger with higher household incomes. Minority ratio negatively correlates with the other three demographic features, but to different extents ($r = -0.31$ with older adult ratio, $r = -0.59$ with average household income and $r = -0.16$ with essential worker ratio; Supplementary Fig. 2). These correlations highlight the mismatch of disadvantaged populations across different demographic dimensions, which results in conflicts between equities that cannot be settled on the basis of demographic features alone. This suggests the need to explore more essential mechanisms underlying demographic features and health that can forecast vaccination outcomes and guide vaccine distribution.

**Indices for estimating vaccine prioritization outcomes.** To inform the design of vaccine distribution, it is critical to accurately estimate the outcomes that would result from prioritizing certain communities for vaccine distribution. Specifically, for optimal design we must be able to estimate changes in overall social utility and equity when vaccinating each community. Policy designers typically rely either on a single demographic feature or on indicators computed solely on the basis of demographic data, such as the social vulnerability index (SVI) designed by the US CDC. Nevertheless, Fig. 2 shows that such approaches will probably degrade equity along different dimensions, due primarily to complex associations between distinct demographic profiles, infection-fatality risks and per capita mobility (Supplementary Fig. 3). We therefore design two vaccine outcome indices, community risk and societal risk, that capture the underlying mechanisms. To evaluate changes in equity when vaccinating a community, we design a community risk index as the expected mortality rate, calculated as the product of the estimated contact frequency due to average community movement and the infection-fatality risk associated with community demographic profiles. To evaluate changes in risks to society when vaccinating a community, we design a societal risk index as the expected number of deaths caused by infection in that community and the secondarily infections they impose on persons from other communities (Methods, ‘Quantification of community risk and societal risk’). Societal risk thus captures the number of lives saved in the whole population by vaccinating certain communities, providing a proxy for social utility. Our proposed community risk and societal risk indices capture different characteristics of communities that could result in trade-offs between social utility and equity. For example, older people with lower mobility face greater mortality risk once infected, but they are less likely to spread the disease than young people with high mobility. This is manifested by their high community risk and low societal risk indices.

We perform regression analysis to examine the power of these indices for estimating outcomes associated with alternative vaccination distribution strategies. Specifically, we generate numerous vaccine distribution instances, each of which vaccinates a set of randomly selected communities covering 2% of the total population. We obtain the impact on social utility and equity

---

### Table 1 | Changes in four dimensions of equity quantified by the Gini index, compared with the Homogeneous baseline

| Policy | Dimension of equity | Atlanta | Chicago | Dallas | Houston | Los Angeles | Miami | Philadelphia | San Francisco | Washington, DC |
|--------|---------------------|---------|---------|--------|---------|------------|-------|--------------|--------------|----------------|
| Prioritize by age | Age | 34.9%** | 38.8%** | 18.1%** | 18.3%** | 34.9%** | 12.3%** | 41.7%** | 8.1%** | 28.5%** |
| | Income | 18.0% | 7.2%** | 3.1% | 3.1% | (9.0%)* | (6.9%)* | (0.9%)** | 0.4% | 2.8% |
| | Occupation | (10.0%)** | 28.6%** | 1.6% | (71%) | 4.4% | 2.8% | 1.6%** | (13.3%) | 31.4%** |
| | Race/ethnicity | 28.6%* | (5.2%)* | 14.9%** | 17.1%* | 10.6% | 3.0%** | (1.7%)* | (6.4%) | (39.6)%** |
| Prioritize by income | Age | 5.0% | 3.5%** | 0.9% | 0.9% | (7.7%)* | 2.4%* | (15.7%)* | (0.7%) | (3.8%)** |
| | Income | 55.6% | 38.5%** | 30.8%** | 51.8%** | 63.9%** | 39.9%** | 14.3%** | 15.8%** | 55.0%** |
| | Occupation | 0.0% | 46.3%** | (2.9%) | (0.1%) | 33.8%** | 17.5% | 22.6%** | 6.0% | 22.8%** |
| | Race/ethnicity | (26.7%)* | 29.3%* | (4.4%) | (26.7%)* | (58.4%)* | 7.5%* | 19.2%* | 11.7% | 35.8%* |
| Prioritize by occupation | Age | (5.1%)* | 6.2%** | (4.2%)* | (1.3%) | 0.5% | 1.2% | (12.9%)* | (8.0%)* | 6.9%** |
| | Income | 36.1% | 14.3%** | (6.5%) | (3.6%) | (1.0%) | 4.0% | 4.7%** | 7.4%** | 6.0%* |
| | Occupation | 46.7%** | 51.6%** | 35.4%** | 28.5%** | 65.4%** | 50.7% | 14.3%** | 51.8%* | 43.0%** |
| | Race/ethnicity | (15.6%)* | 12.1%* | (2.9%) | (15.9%) | 23.2%* | 4.5% | 4.9%** | 0.5% | (2.2%) |
| Prioritize by race/ethnicity | Age | 11.3%** | (4.5%)* | 5.3%** | 1.5% | 7.7%** | 4.4%* | (22.3%)* | (5.3%) | (5.8%)* |
| | Income | (29.2%) | 17.8%** | (0.8%) | 0.1% | (14.0%)* | (13.6%)* | 2.8%** | 5.3%* | 23.9%** |
| | Occupation | 0.7% | 14.1%** | (71%) | (10.2%) | (6.9%) | 7.4% | 0.4% | 1.0% | (12.4%)* |
| | Race/ethnicity | 45.7%** | 28.8%** | 18.3%** | 27.5%** | 66.8%** | 25.2%** | 8.6%** | 18.6%** | 46.6%** |

Values with both parentheses and asterisks indicate significant degradation in the corresponding dimension of equity. Double-sided paired t-test: *P < 0.05; **P < 0.01. For detailed statistics, please refer to Supplementary Table 3.
for each vaccine distribution through simulation, and we perform ordinary least-squares (OLS) regression to estimate changes with demographic features and the proposed indices (Methods, ‘Quantification of community risk and societal risk’). Results show that demographic features alone explain only 38.7% (95% CI, 0.365–0.409; P < 0.001) of the variance in fatality rate reduction on average, but the incorporation of our societal risk index raises that value to 67.5% (95% CI, 0.656–0.693; P < 0.001) (Fig. 3b). In addition, demographic features alone explain on average only 62.9% (95% CI, 0.614–0.643; P < 0.001), 46.0% (95% CI, 0.431–0.489; P < 0.001), 41.4% (95% CI, 0.391–0.437; P < 0.001) and 48.5% (95% CI, 0.465–0.505; P < 0.001) of the variances in equity defined on age, income, essential worker ratio and minority ratio, respectively, but incorporating the community risk index raises those values to 70.4% (95% CI, 0.688–0.720; P < 0.001), 57.7% (95% CI, 0.549–0.605; P < 0.001), 52.1% (95% CI, 0.493–0.548; P < 0.001) and 57.9% (95% CI, 0.558–0.599; P < 0.001), respectively (Fig. 3c; the detailed regression results are presented in Supplementary Tables 4–12). The indices of societal risk and community risk significantly improve the estimation of changes in social utility and equity under any community prioritization scheme. These two indices also shed light on the simultaneous improvement of social utility and equity with vaccination. Specifically, we discover a positive correlation between community risk and societal risk (Fig. 3d), which indicates a non-negligible overlap between communities experiencing large community risk and those imposing large societal risk. Therefore, if a vaccine distribution strategy succeeds in targeting such overlapping communities, it can simultaneously achieve improvement in both social utility and equity. 

Informing the design of vaccine distribution strategies. On the basis of the proposed indices, we design a flexible framework to generate well-rounded vaccine distribution strategies (that is, a Comprehensive strategy) that can improve both social utility and equity in all demographic dimensions. Our framework integrates community risk, societal risk and demographic profiles with learned weights to generate a comprehensive index of vaccine priority for each community, then distributes vaccines by community accordingly (Methods, ‘Design of vaccine distribution strategies’). Besides the Homogeneous baseline and the four strategies examined in Fig. 2, we construct two additional strategies for comparison. First, an SVI-Informed strategy is designed to prioritize vaccines to communities according to the SVI released by the US CDC19, which is recommended for use in vaccine prioritization25,26. Second, to further justify the necessity of community risk and societal risk, we construct a Comprehensive-Ablation strategy that utilizes demographic features without our indices (Methods, ‘Design of vaccine distribution strategies’).
Results show that the Comprehensive strategy yielded by our framework achieves greater improvements in social utility and all dimensions of equity (Fig. 4a–c). Visualizing the epidemic patterns clearly illustrates that this strategy slows down the increase of deaths, flattening the daily death curve (Supplementary Fig. 4). Compared with the No-Vaccination scenario, distributing vaccines with the Homogeneous strategy reduces death rates by only 6.3% (95% CI, 0.049–0.077; P < 0.001) on average, while our Comprehensive strategy reduces death rates by 20.9% (95% CI, 0.167–0.252; P < 0.001) on average. As for impacts on equity by age, income, occupation and race/ethnicity, the Homogeneous strategy results in little improvement, if not deterioration (−0.1% (95% CI, −0.018 to 0.015; P = 0.343), −2.4% (95% CI, −0.072 to 0.025; P = 0.004), −0.1% (95% CI, −0.035 to 0.032; P = 0.096), 1.8% (95% CI, −0.007 to 0.042; P = 0.104)), while our Comprehensive strategy improves these dimensions of equity by 22.5% (95% CI, 0.167–0.252; P < 0.001) on average. As for impacts on equity with different levels of quality present in the population, the Homogeneous strategy improves these dimensions of equity by 22.5% (95% CI, 0.167–0.252; P < 0.001) on average, while our Comprehensive strategy improves these dimensions of equity by 22.5% (95% CI, 0.167–0.252; P < 0.001) on average. As for impacts on equity with different levels of quality present in the population, the Homogeneous strategy improves these dimensions of equity by 22.5% (95% CI, 0.167–0.252; P < 0.001) on average, while our Comprehensive strategy improves these dimensions of equity by 22.5% (95% CI, 0.167–0.252; P < 0.001) on average.

In sum, our experiments demonstrate that our vaccine distribution strategy successfully yields comprehensive strategies that simultaneously elevate social utility and the four dimensions of equity. We note that there are multiple ways to conceptualize inequity in public health undergirded by different theories of distributive justice. To our knowledge, there are at least three distributive justice theories distinguished by their views on what subset of health inequalities should be considered inequity. Cause-oriented theory classifies health inequalities into those caused by nature versus those caused by society and regards only the latter as health inequity59. Action-oriented theory shifts focus to avoidability, proposing that health inequity includes health inequalities “amenable to positive human intervention”, regardless of cause4. This view emphasizes agency-people’s ability to shape the future—acknowledging that humans are able to reduce inequality in many dimensions, even if some result from nature or luck40,53,56. The absolute theory measures health inequity in ungrouped individuals to avoid imposing prejudice and regards virtually all health inequalities as equity59.1. While each of these theoretical approaches relies on a distinctive mixture of ideological foundations, methodological limitations and pragmatic insights, we build our equity framework primarily on the basis of action-oriented theory. We define health inequity as those health inequalities that could have been mitigated by vaccination, regardless of whether the cause is social or pre-existing health. This research design echoes the framework proposed by National Academies of Sciences, Engineering, and Medicine4, and it identifies age, income, occupation and race/ethnicity as four critical dimensions of equity in our study.

Different from the traditional view of unavoidable trade-offs between social utility and equity, our BD model reveals that prioritizing the most disadvantaged communities for COVID-19 vaccine access can simultaneously improve social utility and equity. This outcome is driven by underlying community heterogeneity in both mobility behaviour and demographic profile. We resolve the tension of equity across different demographic features by designing two indices, community risk and societal risk, to estimate vaccination outcomes. The effectiveness of both indices reveals the necessity to jointly consider both demographic and behavioural heterogeneity in epidemic modelling and policy design.

The vaccine distribution framework we propose provides clear guidance to policymakers. Currently, the prioritization of vaccines is usually based on a rigid stratification of age or occupation43,45,46 and set for the state or country as a whole. In contrast, our framework enables the design of flexible distribution strategies aware of joint effects from mobility behaviours and demographic profiles, which can be tailored to local conditions. Moreover, our framework possesses the following two benefits. First, our method provides meso-scale policy guidance by achieving a balance between effectiveness and ease of implementation. With awareness of heterogeneous risks faced by different CBGs, we can maximize the benefits to society with limited vaccination dosages. Meanwhile, because vaccine priorities are determined on the CBG level, people within the same CBG are not discriminated against, providing local administrations with greater flexibility in the actual vaccine roll-out. Second, our method is privacy-preserving. For both demographic features and mobility records, we use only aggregate data on the CBG level, without revealing any individual information, presenting a minimal invasion of personal privacy. Our framework is therefore not only theoretically informative but also instructive for real-world practice. More broadly, equitable access to immunization is viewed by many as a critical part of the right to health, a fundamental human right endowing every person with the ability to pursue and claim their highest attainable health status54,55. Although it has been listed among the six principles of the Global Vaccine
Fig. 4 | Performance of the Comprehensive distribution strategy under various vaccination rates and timings. a. Changes in social utility and four dimensions of equity under eight vaccine distribution strategies. Values are normalized by the result of the Comprehensive strategy. The Comprehensive strategy (red) surpasses or is comparable to all other strategies in the five metrics, indicating its well-rounded effectiveness. In contrast, SVI-Informed (grey) and Comprehensive-Ablation (violet) result in degradation in certain dimensions of equity. b. Changes in social utility in each MSA. The bottom and top of each box indicate the 25th and 75th percentile values. Whiskers indicate 1.5× the interquartile range below and above the 25th and 75th percentile values. c. Changes in equity by age, income, occupation and race/ethnicity in each MSA. d. Overall performance of strategies under different vaccination rates. Overall performance is the sum of relative improvements in social utility and the four dimensions of equity compared with the Homogeneous baseline. The star shows overall performance if a vaccine is distributed proportionally to its real-world distribution, with a vaccination rate of 56% (that is, close to the current rate in the United States). e. Overall performance of strategies under different vaccination timings.
Action Plan since 2011\textsuperscript{18}, there is still a long way to go before we reach this goal. Our research provides insights to settle concerns regarding ethical values in the global health crisis with experiments on large, real-world data. Our study has several limitations. First, because medical studies on the etiology and pathogenesis of coronavirus are still ongoing, we only consider the widely acknowledged heterogeneity in IFR associated with age. Second, we focus solely on vaccine distribution within a country, but as the COVID-19 pandemic is a global public health emergency, it is of equal necessity to quantitatively study how to coordinate social utility and equity at the international level\textsuperscript{14,16}. Third, our study focuses solely on the demand side without investigating supply-side issues\textsuperscript{46}. Future research should focus on the impact of vaccination on people’s daily movement and economic recovery; interactions between vaccine manufacturing, transportation and distribution; and vaccine utility and equity issues on a global scale. Nevertheless, our study provides powerful guidance for vaccine prioritization even under circumstances of extreme vaccine hesitancy, recommending far greater societal investments in vaccination outreach, education and incentives for disadvantaged and undervaccinated communities than have hitherto been explored. Vaccinating those worst off represents the best step towards societal protection.

**Methods**

**Epidemic model, calibration and preliminary analysis.** Our BD model principally combines demographic profiles and mobility behaviours to simulate epidemic spread in urban communities. First, to reflect the heterogeneity of demographic profiles across communities, we calculate CBG-specific IFRs according to their demographic structure. Here we specifically focus on age structure, as the quantitative influence of age on IFR is the most widely recognized. We divide the population of each CBG into 17 age groups (0–4, 5–9, 10–14, 15–19, 20–24, 25–30, 31–35, 36–40, 41–45, 46–50, 51–55, 56–60, 61–65, 66–69, 70–74, 75–79 and 80+), associated with different levels of infection-fatality risk. To evaluate vaccine distribution strategies, we set IFRs according to recent, well-established findings\textsuperscript{36}. Because the infection-fatality transition for different individuals is independent, CBG-specific IFRs can be calculated as the weighted average of individual IFRs, where weights are determined by the proportion of each age group in the CBG.

Second, to reflect behavioural heterogeneity, populations from different communities do not homogeneously mix but connect via bipartite mobility networks representing visits from CBGs to POIs each hour, estimated from SafeGraph data in previous research\textsuperscript{2}. Specifically, each hourly mobility network is represented by a two-dimensional matrix, whose columns correspond to CBGs and rows to POIs. In the simulation, $S \rightarrow E$ transitions take place in both POIs and CBGs, while $E \rightarrow I$ and $I \rightarrow R$ transitions occur only in CBGs. To account for the impacts of different physical environments on transmission characteristics, we use the following adjacency matrices, $\beta_{poi}$, which are further customized by multiplying $\beta_{poi}$ by the median dwelling time of visitors ($T_i$) and the reciprocal of the squared area ($S_i$):

$$\beta_{poi(i)} = \frac{\beta_{poi} \times T_i}{S_i}. \tag{1}$$

Our simulation period lasts 63 days, which provides us enough time to simulate the impact of vaccination when the pandemic is at the stage of community transmission. Meanwhile, heterogeneity observed in the simulation, which reflects the effect of both non-pharmaceutical interventions (such as shop closing) and citizen reactions (such as a reduction in loitering). As our main objective is to tease out the direct outcomes of vaccine distribution strategies, we have not made assumptions about how vaccination in turn impacts mobility.

To introduce flexibility in adapting to different MSAs, we use a vector of age-stratified IFRs to indicate relative risks, and we estimate the actual IFR by fitting a scaling factor specific to each MSA. Such a practice is justified by the consistency in how IFRs increase with age\textsuperscript{40}. To calibrate the parameters, we use assumptions about how vaccination in turn impacts mobility. To analyse how mortality risks associate with demographic features, we stratify the whole population into deciles according to their older adult ratio, average household income, essential worker ratio and minority ratio. The fatality rate in each decile is calculated as the total number of deaths divided by the total population in decile. According to the occupation data released by the US Bureau of Labor Statistics, it is estimated that in all states, the lowest essential worker ratio in the labour force is 39.3%\textsuperscript{1}. Combined with the World Bank’s estimate in 2020 that the labour force accounts for about half of the total population\textsuperscript{2}, a reasonable essential worker ratio lies beyond 20%. Therefore, when analysing communities stratified by essential worker ratios, we filter those CBGs with a percentage below 20% to mitigate the impact of outliers.

**Vaccination scenarios.** We assume that vaccines are administered at a single point in time, and we characterize various vaccination scenarios with different vaccine rates (how many vaccines are administered in total) and vaccination timings (when these vaccines are administered). In our first experimental scenario, to allocate a fixed number of vaccines to all communities, we filter those CBGs distributed to communities are effectively used without waste (Fully Accepted), although we alter this in later simulations to model varying levels of vaccine hesitancy.

 Corresponding to Fig. 2b, we consider five additional vaccination scenarios to reflect the waste of distributed vaccines due to residents’ vaccine hesitancy and communities’ limited access to vaccines. In the Estimated Hesitancy and Hypothetical Capabilities scenario, we set vaccine acceptance rates to a hypothetical distribution to a sample-based national assessment\textsuperscript{17}. Communities with an average annual household income in the range of [0, 30,000], [30,001, 60,000], [60,001, 99,999] or [99,999, $\infty$] are associated with a vaccine acceptance rate of 72%, 74%, 81% or 86%, respectively. In Hesitancy + Capability, we further combine our estimated vaccine acceptance rate with real vaccination data from the US CDC\textsuperscript{5}, to calculate the vaccine accessibility in each community. The CDC provides the percentage of persons of different ages and ethnic groups that have been fully vaccinated. We take the vaccination data from 15 October 2021 (as shown in Supplementary Tables 13 and 14), calculate the weighted average over CBGs’ age structure and ethnicity composition, and multiply the age-determined rates by the ethnicity-determined rates to obtain each CBG’s current vaccination rate. If a CBG’s vaccination rate is lower than its residents’ vaccine acceptance rate, this means that some residents are willing to take vaccines but remain hindered by vaccine accessibility in the CBG. We thus divide obtained vaccination rates by real-world vaccine acceptance rates\textsuperscript{30} to obtain each CBG’s vaccine accessibility. In three hypothetical scenarios, we assign different vaccine acceptance rates to five income groups of CBGs to reflect the observed positive correlation between income and vaccine acceptance. Specifically, vaccine acceptance rates from the bottom to top income groups are set to be 0.6, 0.7, 0.8, 0.9 and 1 (Hypothetical-1); 0.2, 0.4, 0.6, 0.8 and 1 (Hypothetical-2); and 0.1, 0.3, 0.5, 0.7 and 1 (Hypothetical-3), sequentially reflecting larger differences across income groups.

Corresponding to Fig. 4d and Supplementary Figs. 6–19, we construct two series of vaccination scenarios by varying vaccination rates and vaccination timings to examine the generalizability of our Comprehensive strategy. First, we set the vaccination rate to 5%, 15%, 20%, 40% and 56% of the total population (Fig. 4d) and to 3%, 8%, 13% and 18% of the total population (Supplementary Figs. 6–8 and 12–15). Second, we set the vaccination timing to the 26th, 36th and 41st days (Fig. 4e) and to the 24th, 29th, 34th and 39th days (Supplementary Figs. 9–11 and 16–19).

In all scenarios, we assume that vaccination is fully effective—that is, people that have been vaccinated will not become infected or die from the disease during the observation period. Because we model the transmission on the CBG level, we obtain each CBG’s transmission and fatality risks by multiplying their original values by the vaccination rate—that is, if the vaccination rate in a CBG is 50%, the transmission and induced fatality risks its residents experience will halve.

**Correlation analysis of demographic features.** With demographic data from SafeGraph and the American Community Survey covering more than 42,000 CBGs for nine large MSAs in the United States, we analyse the pairwise correlations among four demographic features: older adult ratio, average household income, essential worker ratio and minority ratio. To eliminate scale differences across demographic features as well as examine systematic differences across MSAs, we first transform the absolute values of demographic features into percentile ranks for each MSA to reflect their relative levels compared with other CBGs of the same MSA. Next, we aggregate data from different MSAs and calculate the Spearman correlation between each pair of demographic features. In visualizing association patterns between demographic features (Supplementary Fig. 2), the probability...
Quantification of social utility and equity. We define metrics of social utility and equity in accordance with the fundamental ethical principles of maximum benefit and mitigation of health inequities in the National Academies of Sciences, Engineering, and Medicine's proposal\textsuperscript{50}. The maximum benefit principle aims to reduce the total damage of the pandemic. As loss of life represents severe and irreversible damage\textsuperscript{51}, we quantify social utility as an overall reduction in the fatality rate. The mitigation of health inequities principle aims to address higher risks faced by certain disadvantaged demographic groups. Specifically, we consider health inequity as health inequalities that could have been mitigated with vaccines, regardless of whether the cause is pre-existing health conditions or social deprivation\textsuperscript{52-56}. The COVID-19 pandemic has imposed substantially higher mortality risk on older adults, low-income households, essential workers and racial/ethnic minorities\textsuperscript{6,57-59}. Unlike inevitable mortality due to incurable disease, COVID-19-induced mortality risk does not evade prevention and can be effectively reduced with vaccination. We therefore define equity on four demographic dimensions: age, income, occupation and race/ethnicity. First, equity by age is calculated among CBGs with different rates of older adults, where people over 70 years old are defined as ‘older adults’ in our case, because they face significantly higher risks of death upon infection. Second, equity by income is calculated among CBGs with different average household incomes. Third, equity by occupation is calculated among CBGs with different essential worker ratios. We calculate essential worker ratios by combining the essential occupation list released by Delaware and Minnesota\textsuperscript{60} with the CBG-level employment data from SafeGraph, following the practice of relevant studies\textsuperscript{6,61}. Finally, equity by race/ethnicity is calculated among CBGs with different rates of racial/ethnic minority population. Here we follow the widely adopted approach to estimate ‘racial/ethnic minority’ with the complementary percentage of non-Hispanic white residents\textsuperscript{60}. More specifically, we adopt the widespread Gini coefficient\textsuperscript{77-80}, which reflects the relative mean absolute difference between all pairs of objects. We take the negative of the Gini coefficient for fatality rates among demographic groups as the inequity in that dimension, calculated as follows: for each demographic dimension, we rank all CBGs according to the corresponding feature and then divide them into N groups covering populations of virtually identical sizes. We denote the fatality rate of the ith group as $f_i$ and the average fatality rate as $\bar{f}$. The CBG groups are first placed in ascending order of fatality rate. The Gini coefficient $G$ is then calculated as follows:

$$G = \frac{\sum_{i=1}^{N} (2i - N - 1)f_i}{N\sum_{i=1}^{N} f_i}.$$  

Quantification of community risk and societal risk. The design of the community risk and societal risk indices integrates demography with empirical mobility data. First, the infection risk for each CBG (denoted $\Phi$) is calculated as the sum of two parts: the within-CBG infection risk ($\Phi_{\text{home}}$) and the infection risk from POIs that the CBG’s residents visit ($\Phi_{\text{poi}}$). To obtain the average mobility level for each CBG, we perform calculations on the average CBG–POI visiting matrix over all hours. For each CBG, the average population staying at home is denoted as $N_{\text{stay}}$, while the average population visiting the i-th POI is denoted as $N_{\text{vis}}$. Because the effect of age on infection risk is not clearly understood\textsuperscript{81}, we do not make assumptions about its relation with residents’ demography. For residents from any CBG, we denote the within-CBG transmission rate as $\beta_{\text{home}}$ and POI-specific transmission rates as $\beta_{\text{poi}}$.

People staying in CBGs rather than visiting POIs are more likely to stay at home instead of interacting with all others present in the same CBG through uniform mixing. We therefore divide CBG populations by total households according to census data from the American Community Survey to estimate the average household size for each CBG (denoted as $N_{\text{household}}$), and we assume that each individual staying in the CBG makes contact with that average. To deal with outliers, the quotient is further clipped to be at most 10, in accordance with the distribution of household sizes in the United States\textsuperscript{62}. The calculations are shown below:

$$\Phi = \Phi_{\text{home}} + \Phi_{\text{poi}},$$

$$\Phi_{\text{home}} = N_{\text{stay}} \times \frac{N_{\text{household}}}{N_{\text{stay}}} \times \beta_{\text{home}},$$

$$\Phi_{\text{poi}} = \sum_{i} N_{\text{vis}(i)} \times \beta_{\text{poi}(i)}.$$  

The fatality rate specific to each CBG is denoted $\alpha$, and the average fatality rate for all CBGs belonging to the same MSA is denoted $\bar{\alpha}$. For each CBG, community risk is equal to the expected number of deaths among its residents, and societal risk is equal to the expected deaths caused by this CBG across the whole MSA, which consists of two parts: deaths among its own residents and deaths of residents from other CBGs. In a population with a minor fraction of healthy people but a majority of healthy, susceptible people, the probabilities of being infected and of infecting others are asymmetric, although they depend on the same mobility process. To account for this asymmetry, we further adjust the two terms with the proportion of infected population (denoted $\gamma$) and that of susceptible population (denoted $\delta$) obtained by running our simulation until the examined time—that is, the 31st day (Supplementary Table 15). The calculations are shown below:

$$\text{community risk} = \Phi \times \gamma \times \alpha$$

$$\text{societal risk} = \Phi \times \gamma \times \alpha + \Phi \times (\Phi \times \delta) \times \bar{\alpha}.$$  

To justify community risk and societal risk as indices of vaccination outcomes, we perform OLS regressions of the change in social utility and the four dimensions of equity in each MSA. In simulating various vaccination strategies, we randomly sample 2% of the total population to achieve a range of vaccination results and then compute health outcomes. To generate more diverse samples in the sampling phase, we first divide the CBGs into 3$^3$ = 729 groups, corresponding to their levels (high/median/low) in the four demographic features and the two indices. After merging groups with too few CBGs, we generate a fixed number of samples for each group. For each experiment, we perform 30 stochastic simulations and take the average as our final result.

To examine whether societal risk improves the prediction of the impact of vaccination on social utility, we regress change in social utility on (1) the average and standard deviation of each of the four demographic features (that is, eight independent variables) and (2) the average and standard deviation of each of the four demographic features plus societal risk (that is, ten independent variables). Next, to examine whether community risk improves the prediction of vaccination impact on equity, we regress changes in equity by age, equity by income, equity by occupation and equity by race/ethnicity, respectively, on (1) the average and standard deviation of the four demographic features and (2) the average and standard deviation of the four demographic features plus community risk. To estimate goodness of fit for different regression models, we compare the values of adjusted $R^2$, which reflects the proportion of variance in the dependent variable explained by the independent variables, accounting for the number of independent variables.

Design of vaccine distribution strategies. General method of vaccine distribution. The central step in vaccine distribution is to generate a priority index for each community (CBG) and sequentially distribute vaccines to CBGs according to those priorities. A CBG will not receive vaccines unless those with higher priority are fully vaccinated, so as to prevent discrimination within the same CBG and to focus on CBG-level distribution free from additional health parameters. Specifically, when constructing vaccine distribution priorities according to a single demographic feature (that is, Prioritize by Age, Prioritize by Income, Prioritize by Occupation or Prioritize by Race/Ethnicity), demographic groups are ranked in descending order of predicted fatality rate, and only CBGs belonging to the group with the largest average fatality rate are vaccinated. In simulating various strategies, we randomly sample 2% of the total population to achieve a range of vaccination results and then compute health outcomes. To generate more diverse samples in the sampling phase, we first divide the CBGs into 3$^3$ = 729 groups, corresponding to their levels (high/median/low) in the four demographic features and the two indices. After merging groups with too few CBGs, we generate a fixed number of samples for each group. For each experiment, we perform 30 stochastic simulations and take the average as our final result.

To examine whether societal risk improves the prediction of the impact of vaccination on social utility, we regress change in social utility on (1) the average and standard deviation of each of the four demographic features (that is, eight independent variables) and (2) the average and standard deviation of each of the four demographic features plus societal risk (that is, ten independent variables). Next, to examine whether community risk improves the prediction of vaccination impact on equity, we regress changes in equity by age, equity by income, equity by occupation and equity by race/ethnicity, respectively, on (1) the average and standard deviation of the four demographic features and (2) the average and standard deviation of the four demographic features plus community risk. To estimate goodness of fit for different regression models, we compare the values of adjusted $R^2$, which reflects the proportion of variance in the dependent variable explained by the independent variables, accounting for the number of independent variables.

Society-Informed vaccine distribution strategy. The SVI is released and maintained by the US CDC / Agency for Toxic Substances and Disease Registry, which combines multiple socio-economic features to assess community resilience in the face of hazardous events, including epidemics\textsuperscript{83}. We use this index to construct an SVI-Informed strategy, which distributes vaccines according to SVI ranking in each MSA. We use the current (2020) release of the SVI 2018 for census tracts, which represents the most up-to-date version of the data. Estimated at the level of census tracts, which are larger geographical units than CBGs, SVI does not distinguish CBGs from the same census tract. Accordingly, our SVI-Informed strategy associates CBGs of the same census tract with the same priority and assigns vaccines to them indistinguishably.

Real-World vaccine distribution strategy. The current strategy is to vaccinate people in their geographical areas, according to age and ethnicity-determined vaccination rates for each CBG. Then, to ensure a fair comparison, we proportionally scale up vaccination rates in all CBGs so that the total number of vaccines is equal to that in other distribution strategies.

Framework for Comprehensive vaccine distribution strategies. Given the complex relationships not only among demographic features but also between demographic features and underlying mechanisms that determine epidemic impact, we must carefully devise strategies for vaccine distribution to obtain simultaneous improvement of social utility and equity along different demographic dimensions.
We therefore propose a flexible framework to automate the design of such a Comprehensive strategy via joint consideration of community risk, societal risk and multiple demographic features. To construct our Comprehensive strategy, we first rank the CBGs according to each of the four demographic features (older adult ratio, average household income, essential worker ratio and minority ratio) and the two indices (community risk and societal risk). We then use TOPSIS, a widely used multi-criteria ranking method\(^{1}\), to obtain a comprehensive index of vaccine priority via a weighted combination of the above six features. The initial weights of the six features are set to be equal. To adapt to the specific demographic and mobility patterns in different MSAs, optimal weights are determined through a greedy process to combine multiple features. Starting from an equally weighted combination, we perform simulations to estimate outcomes in social utility and the four dimensions of equity, and we then adjust the weights according to the following heuristic guidelines: (1) when improvement in any dimension of equity is unsatisfactory, the weight of the corresponding demographic feature will be increased; and (2) when improvement in social utility is unsatisfactory, the framework will require trials to increase the weight of either community risk or societal risk. Pseudocode describing the decision process is provided in Algorithm 1 (Supplementary Note 1).

To justify the critical roles played by the proposed behaviour- and demography-aware indices, we also construct an ablation of our Comprehensive strategy (Comprehensive-Ablation). In this version, the vaccine priorities of the CBGs are calculated only on the basis of a weighted combination of the four demographic features, removing their community risk and societal risk. We examine the generalizability of our Comprehensive strategy under scenarios with different vaccination rates and timings (see Methods, ‘Vaccination scenarios’ for the details).

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

**Data availability**

COVID-19 daily death data are available from the New York Times (https://github.com/nytimes/covid-19-data). Mobile phone mobility data and demographic data for CBGs are freely available upon request for academic, non-profit use from SafeGraph (https://www.safegraph.com/academics), which is a data company that aggregates anonymized location data from numerous applications to provide insights about physical places, via the SafeGraph Community. Estimated mobility networks are retrieved from https://covid-mobility.stanford.edu. Social vulnerability indices for communities are available at the website of the US Agency for Toxic Substances and Disease Registry (https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_downloaded.html). Vaccination data are available at the website of the CDC (https://covid.cdc.gov/covid-data-tracker/#/vaccination-demographic). The processed data have been deposited in Zenodo (https://doi.org/10.5281/zenodo.6608362).

**Code availability**

The code used in this research has been deposited in Zenodo (https://doi.org/10.5281/zenodo.6609393).

Received: 26 November 2021; Accepted: 12 July 2022; Published online: 25 August 2022

**References**

1. WHO et al. Monitoring Equity in Health: A Policy-Oriented Approach in Low- and Middle-Income Countries. Tech. Rep. (World Health Organization, 1998).
2. Norheim, O. F. & Asada, Y. The ideal of equal health revisited: definitions and measures of inequity in health should be better integrated with theories of distributive justice. *Int. J. Equity Health* 8, 40 (2009).
3. Riger, J. P. Health and social justice. *Lancet* 364, 1075–1080 (2004).
4. McClung, N. et al. The Advisory Committee on Immunization Practices’ ethical principles for allocating initial supplies of COVID-19 vaccine–United States, 2020. Morb. Mortal. Wkly Rep. 69, 1782–1786 (2020).
5. National Academies of Sciences, Engineering, and Medicine Framework for Equitable Allocation of COVID-19 Vaccine (National Academies Press, 2020).
6. WHO SAGE Roadmap for Prioritizing the Use of COVID-19 Vaccines in the Context of Limited Supply: An Approach to Inform Planning and Subsequent Recommendations (based on pandemic epidemiology, setting and vaccine supply; Scenarios, 13 November 2020 Tech. Rep. (World Health Organization, 2020).
71. Rowland, M. A. et al. COVID-19 infection data encode a dynamic reproduction number in response to policy decisions with secondary wave implications. Lancet 373, 423–431 (2009).
72. Labor Force in the United States, Y emen, China, France, Nigeria and Russia with different rate of cases. Lancet 372, 147–149 (2021).
73. Delaware List of Essential and Non-essential Businesses (WPVI, 2020); https://dig.abclocal.go.jp/wpvi/pdf/ Delaware-list-of-essential-and-nonessential-businesses-March-22-2020.pdf
74. Sanchez, D. G., Parra, N. G., Oxen, C. & Rijkers, B. Which Jobs Are Most Vulnerable to COVID-19? What an Analysis of the European Union Reveals Res. and Policy Brief No. 148384 (World Bank, 2020).
75. Van Voorhees, B. W., Walters, A. E., Prochaska, M. & Quinn, M. T. Reducing health disparities in depressive disorders outcomes between non-Hispanic whites and ethnic minorities. Med. Care Res. Rev. 64, 1575–1945 (2007).
76. Fujishiro, K. & Koenig, E. Comparing self-reported and O*NET-based assessments of job control as predictors of self-rated health for non-Hispanic whites and racial/ethnic minorities. PLoS ONE 15, e0237026 (2020).
77. Illesley, R. & Le Grand, J. The Measurement of Inequality in Health (Palgrave Macmillan UK, 1987).
78. Leclerc, A., Lert, F. & Fabien, C. Differential mortality: some comparisons between England and Wales, Finland and France, based on inequality measures. Int. J. Epidemiol. 19, 1001–1010 (1990).
79. Berndt, D. J., Fisher, J. W., Rajendrababu, R. V. & Studnicki, J. Measuring healthcare inequities using the Gini index. In Proc. 36th Annual Hawaii International Conference on System Sciences (ed. Dennis, E.) https://doi. org/10.1109/HICSS.2003.1174353 (IEEE, 2003).
80. Dixon, P. M., Weiner, J., Mitchell-Olds, T. & Woodley, R. Bootstrapping the Gini coefficient of inequality. Ecology 68, 1548–1551 (1987).
81. Dudley, J. P. & Lee, N. T. Disparities in age-specific morbidity and mortality from SARS-CoV-2 in China and the Republic of Korea. Clin. Infect. Dis. 71, 863–865 (2020).
82. Distribution of Households in the United States from 1970 to 2020, by Household Size (US Census Bureau, 2020); https://www.statista.com/ statistics/242189/distribution-of-households-in-the-us-by-household-size/ 83. Cheng, C.-H. & Huang, J.-J. Multiple Attribute Decision Making: Methods and Applications (CRC, 2011).

Acknowledgements
This work was supported in part by the National Key Research and Development Program of China under grant no. 2020AAA010600 to Y. and the National Natural Science Foundation of China under grant no. U1936217 to Y. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions
F.X., P.H., Y.L. and J.E. jointly launched this research and provided the research outline. L.C., F.X. and Y.L. designed the research methods. L.C. performed the experiments and prepared the figures. F.X., P.H., Y.L. and J.E. jointly analysed the results and participated in writing the manuscript. F.X., Z.H., K.T., Y.L., P.H. and J.E. provided critical revisions. All authors jointly analysed the results and participated in writing the manuscript.

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/s41562-022-01429-0.
Correspondence and requests for materials should be addressed to Fengli Xu, Pan Hui, James Evans or Yong Li.

Peer review information Nature Human Behaviour thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2022.
Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- [x] Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
  - Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  - Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
  
  Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection  No software was used for data collection.

Data analysis  Data analysis was performed using Python (version 3.6.12). The data is available at https://doi.org/10.5281/zenodo.6608362. The source code is available at: https://doi.org/10.5281/zenodo.6609393.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

COVID-19 daily death data are available at the New York Times(https://github.com/nytimes/covid-19-data). Mobile phone mobility data and demographic data for census block groups are freely available upon request for academic, non-profit use from SafeGraph (https://www.safegraph.com/academics), which is a data company that aggregates anonymized location data from numerous applications in order to provide insights about physical places, via the SafeGraph Community. Estimated mobility networks are retrieved from http://covid-mobility.stanford.edu.

Social vulnerability indices for communities are available at the website of U.S. Agency for Toxic Substances and Disease Registry (https://www.atsdr.cdc.gov/)
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences  ☑ Behavioural & social sciences  ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

Based on mobility data and demographic data, we designed an epidemic model, and used it to analyze the outcomes of alternative vaccine distribution strategies on social utility and equity. We propose two indices to capture the impact from vaccinating a certain neighborhood. We propose a framework to design vaccine distribution strategies that can simultaneously improve social utility and equity.

Research sample

We study existing data on nine of the largest American metropolitan statistical areas that have consistent death counting method during the period. COVID-19 daily death data are collected from the New York Times (https://github.com/nytimes/covid-19-data). Mobile phone mobility data and demographic data for census block groups are collected from SafeGraph (https://www.safegraph.com/academics). Estimated mobility networks are collected from http://covid-mobility.stanford.edu. Social vulnerability indices for communities are collected from the website of U.S. Agency for Toxic Substances and Disease Registry (https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html). Vaccination data are collected from the website of U.S. CDC (https://covid.cdc.gov/covid-data-tracker/#/vaccination-demographic). We did not perform sampling as they were comprehensive. SafeGraph mobility data have been shown to be representative of the population (https://docs.safegraph.com/docs/faq#section-what-is-the-breakdown-of-safe-graphs-mobility-device-panel-is-there-a-skew-toward-certain-demographics). Other data were designed to cover the entire population. The median age and median household income in MSA Atlanta, Chicago, Dallas, Houston, Los Angeles, Miami, Philadelphia, San Francisco, and Washington is 37, 37, 32, 32, 37, 42, 37, 37, 37 and 58,820, 62,683.5, 60,050, 67,321, 52,083, 65,667, 93,571, 96,506, respectively. These statistics distribute around the national average of 38.2 and 64,324, suggesting the data collection is representative.

Sampling strategy

We did not perform sampling, but included all data.

Data collection

We did not collect data, but utilized existing datasets.

Timing

2020-03-01 ~ 2020-05-02

Data exclusions

No data were excluded from analysis.

Non-participation

No participants were involved in the study.

Randomization

This study does not include randomized controlled trials, and data were not organized into experimental groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a  Involved in the study

☒ Antibodies
☒ Eukaryotic cell lines
☒ Palaeontology and archaeology
☒ Animals and other organisms
☒ Human research participants
☒ Clinical data
☒ Dual use research of concern

Methods

n/a  Involved in the study

☒ ChIP-seq
☒ Flow cytometry
☒ MRI-based neuroimaging