



Opinion

What's next: using infectious disease mathematical modelling to address health disparities

Danielle M Richard ^{1,*} and Marc Lipsitch ^{1,2}

¹Center for Forecasting and Outbreak Analytics, Centers for Disease Control and Prevention, Atlanta, GA, USA and ²Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

*Corresponding author. Center for Forecasting and Outbreak Analytics, Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA 30333, USA. E-mail: daniellemzr@gmail.com

Keywords: Infectious disease epidemiology, mathematical modelling, social determinants of health, social epidemiology, health equity.

Before and during the COVID-19 pandemic, an individual's age and race/ethnicity have been highly predictive of their risk of infectious diseases and their health consequences. Disparities were evidenced in COVID-19 incidence rates and in hospitalization, severity and mortality metrics in the USA¹ and in other countries.^{2,3} Identifying these disparate outcomes associated with demographic variables is valuable mainly if it prompts investigation into what mechanisms generate the disparities and inform how they can be reduced.⁴ A prominent report from the UK succinctly outlined that social determinants such as occupation, household characteristics, surrounding population density, urbanicity and social deprivation were all associated with increased risk of COVID-19 infection.³ Others have noted that social determinants can play a role in all stages of an outbreak, providing pathways for unequal exposure, transmission, susceptibility and treatment that produce and escalate disparities in health outcomes.⁵

In addition, the use of mathematical and computational infectious disease transmission models expanded during the COVID-19 pandemic. Researchers and decision makers used infectious disease models to estimate key outbreak parameters, inform how policy changes could affect the ongoing outbreak and predict what would happen next in various settings. Although social determinants were considered on occasion, few infectious disease transmission models were used to answer questions on the structural role of social determinants in the proliferation of the COVID-19 outbreak.⁶ To understand the origin of inequities in future pandemics and inform policies to reduce them, it is important to incorporate these determinants not only in data-gathering efforts, but also in model structure and analysis, as is appropriate for a given research question.

We aim to outline how infectious disease transmission modelling has been used to identify the mechanisms that create health disparities, to describe how mechanisms can change over the course of an epidemic and to propose policy recommendations that are informed by epidemiological principles. Although this review is not a systematic analysis of all

existing literature, we aim to inform the reader on the current state of the literature and on what is missing. We describe examples of infectious disease models that address health equity questions that fall into three categories: deterministic compartmental models, agent-based models and extensions to deterministic models that incorporate mobility data. We conclude by providing recommendations for how we can work to address what is missing in the infectious disease modelling field, so that we can be better prepared moving forward, for all respiratory diseases and not just COVID-19.

Deterministic compartmental models

Simple, equation-based infectious disease models structure the population into categories (compartments), most fundamentally those of Susceptible–Exposed–Infectious–Recovered (SEIR). SEIR models that have addressed health equity questions have considered methods to mechanistically incorporate social determinants into the structure of their models by parameterizing contact, susceptibility and other determinants of transmission separately for different social groups. A study by Menkir and colleagues is one example. The research team used income quintile-specific values for disease-related parameters such as disease transmission rates, case-fatality rates and vaccine coverage in a susceptible–infectious–recovered (SIR) model to examine the impact of varying vaccination strategies. By incorporating parameters stratified by income level, the researchers were able to assess the extent of differential infectious disease outbreaks spread in distinct subgroups of the population. Of note, findings indicated that, for lower-income quintiles, maximum prevalence was reached earlier than for other income groups. Further, higher within-quintile transmission translated into higher infection peaks, and higher vaccination coverage yielded the greatest benefit to the lowest-income quintiles.⁷ The effort to construct such models also pointed to limitations in the availability of data to estimate stratum-specific transmission and

Received: 7 July 2023. Editorial Decision: 17 November 2023. Accepted: 14 December 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of the International Epidemiological Association.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

severity parameters, highlighting the close dependence of models on high-quality data stratified by social determinants.

Other studies have addressed other questions of health equity by structuring transmission models to incorporate specific determinants. One team incorporated age-structured models for two social groups with high and low transmission into their SEIR model and used social contact data to parameterize contact rates for different social groups. This study varied parameters for contact intensity between the two groups and for susceptibility to infection, and derived predictions on how interventions can impact inequalities. Perhaps counter-intuitively, introduction of a vaccine with equal coverage in interacting high- and low-risk groups led to a greater relative inequality in disease incidence between the groups than in the absence of vaccination; this occurs because a given level of coverage comes closer to eliminating transmission in the low-risk group with lower transmission. This theoretical study did not specify explicitly the reason for the high- and low-contact groups, but rather assumed their existence and derived the consequences. This method could be extended to include particular social determinants, and still provides an example of thinking about inequalities in model production.⁸ To quantify how social exposure to infection varies across race and ethnicity, Ma *et al.* created separate sets of compartments for a selection of race and ethnicity groups, and used a social contact matrix estimated from housing segregation data to account for between-racial-group mixing. Findings indicated that the herd immunity threshold was reached after the cumulative incidence disproportionately increased in certain minority groups and that, even if the same groups continued to have higher contact rates over time, disparities in incidence could shrink and even reverse over time as a result of higher immunity in such communities.⁹ Such approaches described thus far expand our analytic capacity and allow researchers to consider mechanisms by which disparities may exist, how they depend on the phase of the epidemic and how different interventions, such as vaccinations, may ameliorate or inflate existing disparities. Improving the quality and level of detail of data to parameterize these models would enhance their usefulness.

There are also examples of models that have been used to examine outbreaks of a disease in specific vulnerable populations, such as those experiencing homelessness, or in specific settings, such as correctional facilities or nursing homes. To better understand the role of correctional facilities in the proliferation of disease, one team utilized compartmental modelling of infections within and between correctional facilities and the surrounding community for people who are incarcerated, correctional staff and surrounding community. The team found that, when an infectious disease outbreak occurs in a correctional facility, the outbreak gets reflected and magnified back to the surrounding community. Additional analyses were used to inform critical factors in determining the magnitude of an outbreak, and how movement of staff and people who are incarcerated can play a role.¹⁰ In the UK, Lewer and colleagues used SEIR models to understand how homelessness policies, such as hotel accommodations or infection control in homeless settings, affected overall case counts, deaths and hospital admissions in England. Findings indicated that lifting preventive measures led to large increases in cases and deaths for people experiencing homelessness.¹¹ Modelling efforts such as these not only help

decision makers to better understand the marked disparities in infectious disease outbreaks, but also can better represent infectious disease outbreaks by incorporating setting-specific transmission parameters. This merits not only separate modelling efforts, but also an effort to better understand how outbreaks within a subset of the population that is more vulnerable can extend to the larger population.

Agent-based models

Several studies have employed agent-based models to demonstrate how population structure and transmission dynamics help explain observed inequities. To better understand the cause of differences in influenza across census tracts of varying poverty levels, Kumar and colleagues used an agent-based model to simulate influenza attack rates across census tracts and with varying neighbourhood contact rates and income levels. Census tracts with higher poverty had earlier and steeper increases in infection rates—a pattern observed in both real-world data and in model output. Further, the team estimated that including population structure and population mixing in simulation modelling accounted for 33% of the observed inequality in infection prevalence between census tracts of high and low poverty levels.¹² To assess the relationship between influenza burden and social deprivation, Hyder and Leung used a spatially explicit model of Montreal that incorporated household composition as well as other social determinants with spatial arrangements of neighbourhoods. The authors found that heterogeneity of neighbourhood composition and spatial arrangement of neighbourhoods contributed to the observed relationship between influenza burden and social deprivation.¹³

There are also examples of agent-based models that are intended to guide policy recommendations. Focusing on infection control in nursing homes, another team used an agent-based model of SARS-CoV-2 transmission to highlight how different non-pharmaceutical interventions would affect disease incidence. The team found that daily antigen testing, compared with other interventions such as cohorting or staffing interventions, reduced the cumulative incidence the most, whereas combining screening testing with resident cohorting and immunity-based staffing interventions reduced the cumulative incidence even more.¹⁴ In another example, Nande *et al.* used a network model incorporating within- and between-household mixing to assess the impact of evictions on household transmission of COVID-19, accounting for evictions that led to households moving in with other households. Findings indicated that evictions would increase COVID-19 cases across the population and that a simulated increase in the eviction rate of 1% per month resulted in an infection level that was ~4% higher than baseline. Further, evictions were predicted in the simulations to increase the disparity of infection prevalence between low- and high-socio-economic status (SES) neighbourhoods, where the risk of infection increased more for the low-SES neighbourhoods. This study exemplifies how an infectious disease model can be adapted to consider the impact of a policy change on vulnerable populations and inform decision makers with estimates of policy impacts.¹⁵ Given sufficient data to constrain their assumptions, agent-based models and simulation studies such as these can be powerful tools to consider how policies

and changing epidemic quantities can affect outbreaks in specific settings or subpopulations.

Using mobility data in deterministic models

In recent years, the field of infectious disease modelling has started incorporating mobility data into infectious disease models to inform parameter estimation and as a proxy for social characteristics. One team incorporated SafeGraph mobility data into an underlying SEIR model to estimate the effects of specific reopening strategies and to predict infection rates in areas with different demographic compositions. The authors found that lower-income areas did not have as much reduced mobility and that larger proportions of individuals from lower-income areas were infected at specific locations due to higher contact density; the role of higher contact density in promoting transmission was a model assumption and the fact that lower-income individuals had more exposure to high-density locations thereby offered a hypothesis to explain the higher incidence of infection among such individuals.¹⁶

Another way to incorporate individual contacts into a model is by using contact survey data. Many other countries have contact survey data that are derived from POLYMOD—a paper-diary survey with data for >7000 participants in eight European countries.^{17,18} Because the USA does not have readily available and nationally representative contact data, one research team utilized an egocentric exponential random graph model based on POLYMOD data in combination with an SEIR model to assess inequalities in influenza transmission in the USA. The researchers proposed five potential immediate drivers of influenza transmission inequities: social contact differences, low vaccine uptake, high susceptibility to infection, low healthcare utilization and low sickness absenteeism. Findings indicated that, when SES-specific parameters were used, all five factors were associated with an increase in burden for the group of people with low SES.¹⁹ Notably, there are contextual reasons for why each of these factors may be seen more often in people with low SES. Describing this is beyond the scope of this discussion.

Although mobility data and other data sources from novel technologies are proving to be powerful in some applications, they also come with inherent biases that must be recognized. More research is needed to identify how applicable mobility data is at larger spatial scales, where there is considerable within-population heterogeneity in mixing patterns and other social drivers,²⁰ and how changes in contact patterns may be proxied by mobility data.

Recommendations for embedding health equity principles into infectious disease modelling

Although there is ongoing research and interest in the use of infectious disease modelling to address health equity concerns,^{21,22} there is much to be done. First, more studies that incorporate social determinants into compartmental, network and agent-based models are needed. As discussed, this can be accomplished by tracking outcomes based on social determinants and by incorporating social determinants into model processes (transmission, outcomes of infection, etc.). Further, improved models for specific settings such as shelters, schools, prisons and jails, healthcare settings and nursing homes and other congregate settings are needed to better

prepare us for future emerging infectious disease threats. Researchers can more effectively assess health disparities by creating models for vulnerable populations that incorporate the effects of policies on transmission. Our review demonstrates that there are quality examples of using infectious disease modelling to better understand health disparities and that we must continue to prioritize such models as the burden of infectious disease outbreaks continues to fall on disadvantaged groups.

Second, a renewed focus on gathering representative incidence and seroprevalence data in emerging outbreaks is needed. Public health surveillance in the USA remains less likely, for several reasons, to include individuals of lower SES. For example, those of lower SES are less likely to receive healthcare and therefore be included in local surveillance efforts. Some people of colour and people with diverse gender identities are more likely to experience stigma or racism in healthcare and may be less likely to seek healthcare when needed.²³ There are ongoing activities across the USA public health landscape to improve upon the existing surveillance systems to reduce these biases. Even so, accounting for and, where possible, adjusting for existing surveillance biases in analytics is critical, as is properly communicating our analytic findings with the correct caveats to interpretation. To further address these biases, random sampling of the prevalence and incidence of infection can provide key data to answer questions about existing disparities. Without these data, it is difficult to identify where in the cascade of prevention and care inequities occur because higher rates of disease are indistinguishable from higher rates of infection in vulnerable populations. The benefit of random sampling can be seen in the Real-time Assessment of Community Transmission-1 (REACT-1) study, conducted in the UK.^{24,25}

In newly emerging outbreaks, such as in the early days of COVID-19, gathering detailed data on proximal, mechanistic social risk factors for infection or a history of infection (seropositivity) can directly inform efforts to equitably reduce risk. Further, re-collecting and reanalysing these data iteratively over the course of an outbreak are also critical. In particular, when racial/ethnic minorities and/or those of lower SES are harder hit, as is often the case, choice of interventions depends on knowing the contribution of occupational exposure, household crowding, rurality or inability to isolate in disparities in exposure and the contribution of access to care or other causes of disparities in outcome. Such data, which are gathered in routine serologic studies such as the US National Health and Nutrition Examination Survey and sporadically in other studies, are much more helpful than simple race and ethnicity data in informing solutions.²⁶ Gathering these proximate risk factors is consistent with the increasingly accepted concept that race is not an exposure variable and should not be treated as such in modelling. Rather, a race variable can serve as a proxy for social determinants, as well as for lived experiences of racism and stigmatization.^{4,27} Knowing what specific mechanistic risk factors are associated with the outcome is necessary if the research is to inform solutions, rather than simply document disparities.³ A better understanding of how social determinants can change over the course of an outbreak for susceptible, infected and recovered populations can also be illuminating. If the epidemic is primarily in one subgroup of the population at first, but then is predominantly in other subgroups as the outbreak continues, the roles of these subgroups in driving transmission can be

inferred and serve as a prompt for intervention.²⁸ Conversely, understanding how disparities will evolve over the course of an epidemic even without interventions can provide appropriate baselines for assessing how well interventions work to curb disparities.⁹

Third, we must identify epidemic quantities of interest for subgroups of populations. A key limitation across many of the aforementioned studies was that there was little to no availability of data for key epidemic quantities in specific settings of interest or for specific subsets of the population. Improving our case data infrastructure so that we have data for social determinants would allow us to calculate parameters stratified by social determinants.

Towards this end, micro-scale studies of transmission²⁹ and contact patterns¹⁸ as a function of household size and density are needed to provide an empirical basis for the detailed structure and parameterization of these models, e.g. to decide how transmission differs in large and small households, or as a function of density in public places. Given the uptick in interest in ventilation as an intervention to reduce transmission, better understanding of how transmission rates vary across settings could have the additional benefit of informing investments in ventilation to make them more effective and more equitable.

Lastly, we need country-specific contact surveys that are stratified, as appropriate for the specific population, by race and ethnicity and by social determinants. In the USA, although researchers have been able to adapt the POLYMOD contact study to their needs, investing in a comprehensive USA-based contact survey with a priority placed on obtaining data for participants' social determinants could greatly expand the possibilities for mathematical modelling there.⁷ As noted above, stratifying these data by household sizes and/or by setting would also be beneficial.

Although the fields of infectious disease modelling and social epidemiology have expanded greatly in the past several decades, there is much to be done to formally embed epidemiologic principles into mathematical modelling methods. As the field of infectious disease mathematical modelling continues to expand and adapt to new and emerging public health threats, it is critical that we prioritize the development of methods to mechanistically account for social determinants and produce models that inform policies that can address the root causes of health disparities.

Ethics approval

Ethics approval is not required for this study because human patients were not involved.

Data availability

No new data were generated or analysed in support of this research.

Author contributions

D.M.R. and M.L. conceptualized the research topic. D.M.R. conducted literature review and drafted the manuscript. M.L. supervised and reviewed the manuscript.

Funding

This work was supported by the Andrew and Corey Morris-Singer Foundation to M.L.

Acknowledgements

We thank Dr. Eric Mooring for helpful comments and feedback.

Conflict of interest

M.L. reports paid service on an advisory board to Janssen Pharmaceutical prior to CDC affiliation. D.M.R. has no conflicts of interest to declare.

References

1. Mackey K, Ayers CK, Kondo KK *et al*. Racial and ethnic disparities in COVID-19-related infections, hospitalizations, and deaths: a systematic review. *Ann Intern Med* 2021;174:362–73.
2. Mena GE, Martinez PP, Mahmud AS, Marquet PA, Buckee CO, Santillana M. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science* 2021; 372:eabg5298.
3. Race Disparity Unit, Cabinet Office, United Kingdom. *Final Report on Progress to Address COVID-19 Health Inequalities*. 2021. <https://www.gov.uk/government/publications/final-report-on-progress-to-address-covid-19-health-inequalities/final-report-on-progress-to-address-covid-19-health-inequalities> (3, May 2023 date last accessed).
4. Jones CP. Invited commentary: “race,” racism, and the practice of epidemiology. *Am J Epidemiol* 2001;154:299–304.
5. Bamba C. Pandemic inequalities: emerging infectious diseases and health equity. *Int J Equity Health* 2022;21:6.
6. Zelner J, Masters NB, Naraharisetti R, Mojola SA, Chowkwanyun M, Malosh R. There are no equal opportunity infectors: epidemiological modelers must rethink our approach to inequality in infection risk. Papin JA, ed. *PLoS Comput Biol* 2022; 18:e1009795.
7. Menkir TF, Jbaily A, Verguet S. Incorporating equity in infectious disease modeling: case study of a distributional impact framework for measles transmission. *Vaccine* 2021;39:2894–900.
8. Munday JD, van Hoek AJ, Edmunds WJ, Atkins KE. Quantifying the impact of social groups and vaccination on inequalities in infectious diseases using a mathematical model. *BMC Med* 2018; 16:162.
9. Ma KC, Menkir TF, Kissler S, Grad YH, Lipsitch M. Modeling the impact of racial and ethnic disparities on COVID-19 epidemic dynamics. *Elife* 2021;10:e66601.
10. Weyant C, Lee S, Andrews JR, Alarid-Escudero F, Goldhaber-Fiebert JD. Dynamics of respiratory infectious diseases in incarcerated and free-living populations: a simulation modeling study. *Med Decis Making* 2023;43:42–52.
11. Lewer D, Braithwaite I, Bullock M *et al*. COVID-19 among people experiencing homelessness in England: a modelling study. *Lancet Respir Med* 2020;8:1181–91.
12. Kumar S, Piper K, Galloway DD, Hadler JL, Grefenstette JJ. Is population structure sufficient to generate area-level inequalities in influenza rates? An examination using agent-based models. *BMC Public Health* 2015;15:947.
13. Hyder A, Leung B. Social deprivation and burden of influenza: testing hypotheses and gaining insights from a simulation model for the spread of influenza. *Epidemics* 2015;11:71–79.
14. Holmdahl I, Kahn R, Hay JA, Buckee CO, Mina MJ. Estimation of transmission of COVID-19 in simulated nursing homes with

- frequent testing and immunity-based staffing. *JAMA Netw Open* 2021;4:e2110071.
15. Nande A, Sheen J, Walters EL *et al.* The effect of eviction moratoria on the transmission of SARS-CoV-2. *Nat Commun* 2021; 12:2274.
 16. Chang S, Pierson E, Koh PW *et al.* Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* 2021; 589:82–87.
 17. Mossong J, Hens N, Jit M *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. Riley S, ed. *PLoS Med* 2008;5:e74.
 18. Mousa A, Winskill P, Watson OJ *et al.* Social contact patterns and implications for infectious disease transmission—a systematic review and meta-analysis of contact surveys. *Elife* 2021;10:e70294.
 19. Zipfel CM, Colizza V, Bansal S. Health inequities in influenza transmission and surveillance. Perkins A, ed. *PLoS Comput Biol* 2021;17:e1008642.
 20. Buckee C, Noor A, Sattenspiel L. Thinking clearly about social aspects of infectious disease transmission. *Nature* 2021; 595:205–13.
 21. Abuelezam NN, Michel I, Marshall BD, Galea S. Accounting for historical injustices in mathematical models of infectious disease transmission: an analytic overview. *Epidemics* 2023;43:100679.
 22. Williams TG, Brown DG, Guikema SD *et al.* Integrating equity considerations into agent-based modeling: a conceptual framework and practical guidance. *JASSS* 2022;25:1. doi: [10.18564/jasss.4816](https://doi.org/10.18564/jasss.4816).
 23. Hamed S, Bradby H, Ahlberg BM, Thapar-Björkert S. Racism in healthcare: a scoping review. *BMC Public Health* 2022;22:988.
 24. Elliott P, Whitaker M, Tang D *et al.* Design and Implementation of a National SARS-CoV-2 Monitoring Program in England: REACT-1 Study. *Am J Public Health* 2023;113:545–54.
 25. Jernigan DB, George D, Lipsitch M. Learning From COVID-19 to Improve Surveillance for Emerging Threats. *Am J Public Health* 2023;113:e1–e3.
 26. McQuillan GM, Kruszon-Moran D, Kottiri BJ, Curtin LR, Lucas JW, Kington RS. Racial and ethnic differences in the seroprevalence of 6 infectious diseases in the United States: data from NHANES III, 1988–1994. *Am J Public Health* 2004;94:1952–58.
 27. Krieger N. Theories for social epidemiology in the 21st century: an ecosocial perspective. *Int J Epidemiol* 2001; 30:668–77.
 28. Goldstein E, Pitzer VE, O'Hagan JJ, Lipsitch M. Temporally varying relative risks for infectious diseases: implications for infectious disease control. *Epidemiology* 2017;28:136–44.
 29. Cauchemez S, Donnelly CA, Reed C *et al.* Household Transmission of 2009 Pandemic Influenza A (H1N1) virus in the United States. *N Engl J Med* 2009;361:2619–27.