Perspective

Identifying outbreak risk factors through case-controls comparisons

Check for updates

Nina H. Fefferman [1,2,3,4] ✉, Michael J. Blum[1], Lydia Bourouiba[5], Nathaniel L. Gibson[1], Qiang He [6,7], Debra L. Miller [4], Monica Papeș[1], Dana K. Pasquale [8,9], Connor Verheyen[5] & Sadie J. Ryan [10,11,12] ✉

Outbreaks are typically investigated using approaches that aim to identify place- and context-dependent causative factors. As the focus is on understanding the basis of a specific outbreak, the resulting narratives are rarely suitable for forecasting risk or developing generalizable predictive and preventative measures. This Perspective article proposes applying a case-control framework as an outbreak epidemiological study design to promote evidence-based decision-making for prevention and response to outbreaks. The approach involves identifying counterfactuals, with case-control comparisons drawn to test hypotheses about conditions that manifest outbreaks. First, a framework is described for iterative multidisciplinary interrogation to elucidate and identify minimally sufficient sets of factors that lead to disease outbreaks. Next, example case-control comparison frameworks are discussed, centered on pathogen(s), influential contributor(s), or landscape(s), illustrated with examples focused on pathogen transmission.

Investigations of infectious disease outbreaks often focus on identifying place- and context-dependent factors responsible for the emergence and spread of that specific outbreak. This makes it hard to develop generalizable predictive and preventive measures. Understanding disease emergence and spread is a complex challenge that extends beyond the current modes of investigation. While it is widely understood that factors rarely act in isolation, conventional approaches can fall short when disease dynamics arise from synergies among multiple factors, warranting study from multiple disciplinary perspectives. Contributing factors can be firmly rooted in virology, immunology, and pharmacology[1], or may describe ecological and environmental influences[2], while others might relate to social and engineering domains involving human actions, behaviors, and interventions[3,4]. Phenomenological investigations of particular outbreaks also can fail to capture the contingent nature of multifactorial dynamics, where the importance of one factor can be contingent on a "perfect storm" of other concurrent factors.

Progress has been made in developing new approaches to investigate outbreaks. For example, capturing complex dynamics in landscape-based epidemiological modeling of outbreaks[5–7]. Whereas initial models considered a particular outbreak scenario of interest and hypothesized the contributions of observable sets of features[8–12], more recent efforts examined multiple outbreaks[13–15] and have utilized techniques such as fore- and hindcasting to test predictive models relying on either statistical patterns[5,6,16] or inferred causal dynamics, as described previously[7]. While these approaches have provided valuable insights, it is unclear whether they will provide a robust basis for studying and forecasting dynamics of outbreaks that rely on many different contributing, potentially synergistic, factors.

This Perspective article proposes adopting a multidisciplinary case-control hypothesis testing framework, at the scale of an outbreak rather than of an individual, to substantively advance understanding of disease emergence and spread. This provides a key component of a case-control analytical toolkit[17]. Existing case-control epidemiological analysis methodology can serve as a platform for iterative rounds of hypothesis generation and testing. By analyzing combinations of conditionally dependent factors from across disciplines, the approach allows identification of minimally sufficient sets of factors (Box 1) present in early stages of outbreaks that turn emergent infections into outbreaks. In this Perspective, we propose a case-control approach as a way of increasing the capability of epidemiological studies to

[1]Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, USA. [2]National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, TN, USA. [3]Department of Mathematics, University of Tennessee, Knoxville, TN, USA. [4]One Health Initiative, School of Natural Resources and College of Veterinary Medicine, University of Tennessee, Knoxville, TN, USA. [5]The Fluid Dynamics of Disease Transmission Laboratory, Fluids and Health Network, Massachusetts Institute of Technology, Cambridge, MA, USA. [6]Department of Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, USA. [7]Institute for a Secure and Sustainable Environment, University of Tennessee, Knoxville, TN, USA. [8]Duke Department of Population Health Sciences, Duke University School of Medicine, Durham, NC, USA. [9]Duke Network Analysis Center, Duke University, Durham, NC, USA. [10]Department of Geography, Quantitative Disease Ecology and Conservation (QDEC) Lab, University of Florida, Gainesville, FL, USA. [11]Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA. [12]College of Life Sciences, University of KwaZulu-Natal, Durban, South Africa. ✉e-mail: nfefferm@utk.edu; sjryan@ufl.edu

Box 1 | Minimally sufficient sets of potential causal factors

Consider any case in which the objective is to study what leads to a positive versus negative outcome (i.e., an outbreak versus contained infection). Suppose also that a set, S, can be described, of elements 1 through N of possible causal factors that we hypothesize might together determine the outcome. A sufficient set, M_j, is any subset S that, taken together, leads to a positive outcome. Mathematically, if $S = \{x_1, ..., x_N\}$ where each element $x_i = \{0\ if\ absent/false\ 1\ if\ present/true\}$, then a sufficient set is a subset $M \subseteq S$ such that $if\ \forall y_i \in M_j,\ y_i = 1$ then the studied outcome is guaranteed to be positive. Note, that this means that the satisfaction of each element of the specific set is required for an outbreak, not the total number of elements in the sufficient set. Further, M is minimally sufficient if for any $y_i \in M_j$, $y_i = 0$, then the studied outcome

is guaranteed to be negative (i.e., a subset is minimally sufficient if the exclusion of any one of the elements changes the outcome from positive to negative). Note that one minimally sufficient set cannot be a subset of another, but there is otherwise no restriction on overlap between minimally sufficient sets, nor on the number of elements contained in the minimally sufficient set. For example, if $N = 10$, then we could have $M_1 = \{x_1, x_2, x_3\}$, $M_2 = \{x_1, x_2, x_4\}$, $M_3 = \{x_1, x_2, x_5, x_6, x_7\}$, $M_4 = \{x_8\}$, and $M_5 = \{x_9, x_{10}\}$ all as minimally sufficient sets. If those are all of the minimally sufficient sets contained within S, then a maximal set of minimally sufficient subsets exists, and, therefore, a full description of all the ways in which we could have a positive outcome.

isolate and predict causal scenarios that increase the likelihood of outbreaks occurring, presenting three types of controls, with examples as illustrations.

Modern medical research serves as a valuable precedent, illustrating the merits of adopting case-control comparisons to investigate multifactorial dynamics. Medical research shifted from investigating pathology in the human body as a set of independent tests of hypotheses based on expert opinion and careful observation towards employing a framework for evidence-based inquiry that forces an inclusion of otherwise unanticipated drivers and that provides a pathway for synthesis within a hierarchy of evidence[18]. Within that hierarchy are tools designed specifically to move scientific understanding from enigmatic, heterogeneous, multifactorial data to a well-supported, evidence-based understanding of the drivers of pathology. A basic tool of evidence-based medicine is the concept of the Case-Control Study[17]: a retrospective analysis in which individuals who exhibit an observed outcome (cases) are "paired" in study design with individuals who do not exhibit the observed outcome (controls), with analysis then performed to identify which exposures (i.e., potential driving factors) they share and which are distinct only among the cases. This allows for the calculation of measures such as odds ratios[19] to estimate the likely impact of combinations of particular features on health outcomes.

Analogous to clinical medicine concerning pathologies of the body, epidemiology considers outbreaks as pathologies of a population. Unlike physiology, however, a multitude of potential types of controls may be matched to outbreak cases for exploration and analysis. To isolate and understand epidemiological risk and the factors that drive (re)emergence of pathogens of concern, three main categories of case-control formulation are proposed for outbreaks that can serve as a general foundation for evidence-based epidemiological analysis of outbreak risk: factors; pathogens; and landscapes. In each of these categories, the observed outcome that exists for cases and that does not for controls is a circulating outbreak, but what constitutes a well-matched control is partitioned into a narrower equivalence class, concentrating on the nature of the comparators considered.

Influential contributors as case controls

Evidence-based epidemiological analysis of outbreak risk focuses on factors that may limit or foster a disease outbreak (Box 2). Abiotic factors such as temperature, humidity, and precipitation can define the distribution and abundance of a pathogen. Contextual or geographic factors capturing socioeconomic conditions, aspects of the built environment, or human demography and behavior (e.g., lockdowns) may also influence pathogen transmission risk.

Temperature. Dengue fever virus is a flavivirus transmitted by *Aedes aegypti* and *Aedes albopictus* mosquitoes. Originally of animal origin, the virus is now characteristically found in the human environment, with seasonal to perennial circulation among container-breeding mosquitoes and humans[20]. When symptomatic in humans, infection causes high

fever, joint aches, and often there is a characteristic macular rash[21]. Typically thought of as a tropical disease, dengue has been (re)emerging in new extra-tropical locations as climate change has been shifting temperature boundaries for transmission[22]. Accordingly, temperature could limit transmission to locations where the environment is sufficient to support the vector, *Aedes spp.* mosquitoes. A case-control pairing could be established on the basis of the presence or absence of suitable temperature for the *Aedes* vector(s). By extension, a related factor in the prevention of outbreaks, such as knowledge of dengue risk, could be assessed by exploring it in a city well within the temperature range of suitability, versus a city whose average temperatures never or rarely exceed the minimum threshold for transmission.

Sociopolitical and economic (in)stability. Diphtheria, caused by the bacterium Corynebacterium diphtheriae, can form an obstructive biofilm that hinders breathing and swallowing. It can also produce a blood-borne toxin that can cause fatal heart and nerve damage. Disruption of childhood vaccination protocols during the fall of the Soviet Union led to widespread and unanticipated outbreaks of diphtheria among adults due to waning immunity[23]. The duration of childhood immunity also degraded more rapidly and for a greater percentage of vaccination recipients than anticipated[24]. This effectively served as a natural case-control experiment, contrasting the same population against itself in a scenario of demographically dependent herd immunity. When outbreaks were sufficiently prevented among children, the entire population was largely protected against widespread transmission, but as the demography of vaccine protection shifted, so did outbreak dynamics. Socioeconomic instability was also an obvious contributing factor, limiting access to transmission-blocking care. Disruptions of an intervention aimed at population-level protection (e.g., vaccination, water treatment, food aid) as a function of sociopolitical and economic instability are not unusual. Treating the "before and after" in a time-series of health outcomes as case-control comparisons can provide evidentiary support for long-term investment in population-level health and nutrition programs.

Pathogens as case controls

Evidence-based epidemiological analysis of outbreak risk might focus explicitly on a pathogen of interest, with case-control comparisons examining biological conditions that may limit or foster a disease outbreak. Here, the focus might be on the nature of distributions, relative abundance, or ecological interactions that shape pathogen transmission or exposure risk.

Trypanosoma cruzi. Chagas disease is caused by *Trypanosoma cruzi*, a protozoan parasite transmitted to humans and other mammals via triatomine "kissing" bugs. Chagas is a global health threat, with the current majority of cases primarily concentrated in Central and South America. It is estimated that 8 million individuals are currently infected,

Box 2 | Case control approach in pandemic response—seasonality, relative humidity and covid-19

The severity of infectious respiratory diseases, such as Influenza or COVID-19, is often suggested to be seasonal, though the evidence supporting a relationship between weather and respiratory infectious disease outbreaks remains controversial[39,40]. Any causal relationship between weather and respiratory infectious diseases is expected to be complex and multifaceted, involving rich biophysical and fluid physics processes as well as host and pathogen physiology modulated by environmental factors[41]. Research on this topic has thus far primarily focused on geographically localized datasets and outdoor conditions. Case-control analysis could be a more suitable and even necessary approach to shed light on the nature of respiratory disease dynamics. For example, the pre-vaccine COVID-19 pandemic stage covered a wide range of climates and is thus suitable for contrasting regional conditions. Verheyen and Bourouiba conducted a large-scale retrospective analysis blending both case-control and cohort methodologies across a comprehensive global dataset (Box fig. 1) comprising COVID-19 statistics, outdoor meteorological variables, and experimentally validated indoor relative humidity (RH) estimates[38]. The study revealed a robust relationship between regional indoor RH and COVID-19 outbreak severity. The study also revealed that COVID-19 outcomes during the early stages of the pandemic (i.e., pre-vaccination) were consistently less severe at

intermediate indoor RH between 40% and 60% and more severe at extreme indoor RH outside of this range. The association remained robust even after controlling for outdoor weather conditions, government response, and the statistical methodologies and data processing approaches used. Case-control and cohort methods were used to further refine the potential relationship between COVID-19 outbreaks and indoor climate, with odds ratios supporting further association between intermediate indoor RH and better global outcomes, such as fewer deaths, a smaller increase in deaths, or a smaller percentage change in deaths. The consistent findings[38] may be a function of the high-level nature of the analysis, where data aggregation masks lower-level variations while enabling the extraction of the general most robust patterns in the data[42]. Caution is essential, however, when conducting global case-control analyses. Inter-country variability in outbreak magnitudes and reporting conditions can lead to noisy, error-prone datasets, indicating a need for normalization to ensure validity of comparisons. Thus, the very notion of robustness ought to include extensive sensitivity analyses (e.g.,[38]), covering differing treatments and assumptions for both the underlying datasets and the methodology. Emerging associations should persist even when many of the assumptions or methods of analysis are perturbed.
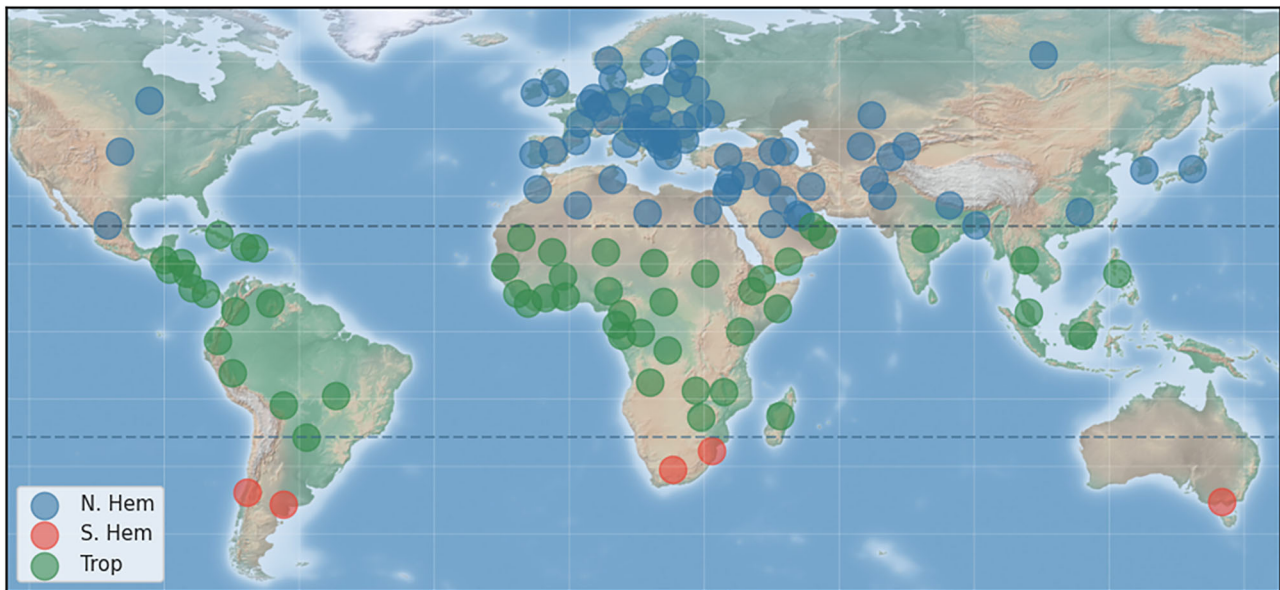
Box Fig. 1 | The global dataset described in[38] established to understand the role of indoor climate in COVID-19 deaths. The aggregated global dataset was stratified in color-coded regions (map) to investigate the effects of low indoor relative humidity (RH) in contrast to moderate and high RH.

and upwards of 100 million people live in areas of high infection risk[25]. While Chagas has historically been considered a disease of rural communities, there is growing evidence that urban populations are at risk of *T. cruzi* infection[26,27]. Evidence of *T. cruzi* in cities implies the presence of triatomine vectors that sustain infection in peridomestic reservoir hosts (e.g., rodents, raccoons, dogs), but triatomines have been sparingly observed in urban landscapes. A case-control approach could test the hypothesis of transient infection of *T. cruzi* in cities, with consideration given to conditions that can sustain infection in hosts where triatomines are rare or absent. This could reveal novel mechanisms contributing to

infection (e.g., vertical transmission) and offer new insights into the eco-epidemiology of one of the world's most burdensome diseases.

Hantaviruses. Hantaviruses that cause human disease persist in some species of rodents, maintained via horizontal transmission among reservoir species[28,29]. The virus is excreted in urine and feces, with human infection occurring most commonly via inhalation of aerosolized matter containing the virus inside dwellings[29] and, as treatment comprises only supportive care, can lead to death[29]. Each of the hantaviruses known to infect humans corresponds to a single rodent host genus[30], but viral

assortment occurs, potentially generating new strains[30]. Factors associated with human infection include environmental conditions (e.g., rainfall, temperature), reservoir population density affecting viral prevalence, and variation in human-reservoir interactions across different landscapes[28]. A case-control approach for hantavirus could therefore test for the presence (i.e., external to and inside of human dwellings) and population size of a particular rodent reservoir. This approach could reveal human-reservoir interaction tipping points as a function of the number and nature (e.g., duration, frequency) of exposures leading to human infection.

Landscapes as case controls
Finally, evidence-based epidemiological analysis of outbreak risk might consider landscape-level phenomena that shape disease emergence and spread. As illustrated below, consideration might be given to landscape functionality and features such as, degree and type of habitat connectivity.

Public water systems. *Cryptosporidium* is a parasite that causes the diarrheal disease cryptosporidiosis. Humans may be infected as a result of exposure to fecal matter shed by animals carrying certain species of *Cryptosporidium*[31]. Waterborne *Cryptosporidium* is very resistant to chlorine, the most common disinfectant in the United States[32]. Consequently, outbreaks of cryptosporidiosis can result from exposure to *Cryptosporidium* in drinking water, potentially impacting large populations served by public water systems in the United States. The landscape can be a critical determinant of the transport of fecal matter from infected animals, such as livestock in concentrated animal feeding operations, to the sources of public water systems. Landscape features may include proximity of the source water to infected animals, weather and land cover conditions that are conducive to stormwater runoff (i.e., connectivity and transport), and inadequate source water management practices (i.e., containment or removal). Accordingly, a case-control analysis could compare populations served by public water systems with distinctive landscape features to identify the causes of cryptosporidiosis outbreaks.

Social-ecological mosaics. Cities can be described as social-ecological landscapes, wherein there is an admixture of natural and built spaces adjoining or overlapping with one another. Within social-ecological landscapes, habitat mosaics can reflect differences in land use, habitation, disturbance (e.g., natural disasters), and economic development, sometimes reflecting disparities driven by discriminatory public policy[33–35]. Spatial patterns in outbreak preparedness, perhaps driven by regional beliefs, can shape exposure risks[36]. Phenomena such as property abandonment and persistent vacancy can create social-ecological mosaics that influence risk of exposure to peridomestic animals that host pathogens of concern[35]. It has been demonstrated, for example, that the diversity, distribution, and prevalence of rodents and associated pathogens can reflect patterns of abandonment in neighborhoods within a city[33,34]. Notably, these patterns can directly tie exposure risk to spatial socioeconomic landscapes, creating risk hotspots in already historically underserved areas. Case-control analysis of rodent demography in different neighborhoods could reveal how abandonment shapes exposure risk among neighborhoods within cities[33,34] and between cities (i.e., exhibiting distinct patterns of abandonment) to consider how social-ecological forces manifest risk across distinct geographies.

Conclusion
While many factors driving outbreak risk are well-studied, more work is warranted to better understand what allows an infectious disease to emerge and then tip from endemic to epidemic to pandemic. Greater knowledge and predictive capacity can be achieved by employing a case-control framework for inquiry that forces consideration of multidisciplinary drivers of disease outbreaks, providing a pathway towards synthesis across a hierarchy of evidence (Box 2). Case-control formulations allow for more rigorous and explicit consideration of complex interactions among sets of factors (Box 1) that may not yet be fully identified or understood; generate hypotheses for

further investigation; and bridge gaps among disciplines in ways that are otherwise difficult to synthesize. This illustrates the versatility and merits of a case-control approach to promote evidence-based decision-making. Further, the merits of isolating minimally sufficient sets of factors allow us to move away from a more traditional focus on identifying and reducing one large single-source risk, since that may inadvertently create equal, if not greater, risks from a diffuse suite of other, now slightly altered contributory factors[37]. While many of the factors driving outbreak risk have been well studied, and many case-control studies have focused on individual risk factors at the scale of the individuals who are infected (or not) during an outbreak, further work is warranted to better understand the many ways in which an infectious disease can tip from endemic to epidemic to pandemic. Employing a more robust analytical framework and investigative toolkit can only aid in ongoing efforts to progress beyond phenomenological interrogation and establish a consistent basis for developing generalizable preventable measures that respond to dynamic, multifactorial risks.

References
1. Alizon, S. & van Baalen, M. Acute or chronic? Within-host models with immune dynamics, infection outcome, and parasite evolution. *Am. Nat.* **172**, E244–E256 (2008).
2. Carlson, C. J. et al. Climate change increases cross-species viral transmission risk. *Nature* **607**, 555–562 (2022).
3. Arthur, R. F., Gurley, E. S., Salje, H., Bloomfield, L. S. P. & Jones, J. H. Contact structure, mobility, environmental impact and behaviour: the importance of social forces to infectious disease dynamics and disease ecology. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160454 (2017).
4. Delabouglise, A. et al. Cultural practices shaping zoonotic diseases surveillance: the case of highly pathogenic avian influenza and Thailand native chicken farmers. *Transbound. Emerg. Dis.* **64**, 1294–1305 (2017).
5. Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proc. Natl. Acad. Sci. USA* **101**, 15124–15129 (2004).
6. Gharbi, M. et al. Time series analysis of dengue incidence in Guadeloupe, French West Indies: forecasting models using climate variables as predictors. *BMC Infect. Dis.* **11**, 166 (2011).
7. Plowright, R. K., Sokolow, S. H., Gorman, M. E., Daszak, P. & Foley, J. E. Causal inference in disease ecology: investigating ecological drivers of disease emergence. *Front. Ecol. Environ.* **6**, 420–429 (2008).
8. Keeling, M. J. & Gilligan, C. A. Bubonic plague: a metapopulation model of a zoonosis. *Proc. Biol. Sci.* **267**, 2219–2230 (2000).
9. Tildesley, M. J. & Ryan, S. J. Disease prevention versus data privacy: using landcover maps to inform spatial epidemic models. *PLoS Comput. Biol.* **8**, e1002723 (2012).
10. Sørensen, J. H., Mackay, D. K., Jensen, C. O. & Donaldson, A. I. An integrated model to predict the atmospheric spread of foot-and-mouth disease virus. *Epidemiol. Infect.* **124**, 577–590 (2000).
11. Nolan, P. M., Hill, G. E. & Stoehr, A. M. Sex, size, and plumage redness predict house finch survival in an epidemic. *Proc. Biol. Sci.* **265**, 961–965 (1998).
12. Eisenberg, J. N. et al. An analysis of the Milwaukee cryptosporidiosis outbreak based on a dynamic model of the infection process. *Epidemiology* **9**, 255–263 (1998).
13. Poletti, P., Caprile, B., Ajelli, M., Pugliese, A. & Merler, S. Spontaneous behavioural changes in response to epidemics. *J. Theor. Biol.* **260**, 31–40 (2009).
14. Preisler, H. K., Hicke, J. A., Ager, A. A. & Hayes, J. L. Climate and weather influences on spatial temporal patterns of mountain pine beetle populations in Washington and Oregon. *Ecology* **93**, 2421–2434 (2012).

15. Hill, S. C. et al. Comparative micro-epidemiology of pathogenic avian influenza virus outbreaks in a wild bird population. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **374**, 20180259 (2019).

16. Viboud, C., Simonsen, L. & Chowell, G. A generalized-growth model to characterize the early ascending phase of infectious disease outbreaks. *Epidemics* **15**, 27–37 (2016).

17. Lewallen, S. & Courtright, P. Epidemiology in practice: case-control studies. *Community Eye Health* **11**, 57–58 (1998).

18. Burns, P. B., Rohrich, R. J. & Chung, K. C. The levels of evidence and their role in evidence-based medicine. *Plast. Reconstr. Surg.* **128**, 305–310 (2011).

19. Case-Control and Matched Studies. In *Biostatistical Methods: The Assessment of Relative Risks*, 2nd Edition (ed. Lachin, J. M.) (Wiley, 2011).

20. Stewart Ibarra, A. M. et al. A social-ecological analysis of community perceptions of dengue fever and Aedes aegypti in Machala, Ecuador. *BMC Public Health* **14**, 1135 (2014).

21. Guzman, M. G., Gubler, D. J., Izquierdo, A., Martinez, E. & Halstead, S. B. Dengue infection. *Nat. Rev. Dis. Primers* **2**, 16055 (2016).

22. Ryan, S. J., Carlson, C. J., Mordecai, E. A. & Johnson, L. R. Global expansion and redistribution of Aedes-borne virus transmission risk with climate change. *PLoS Negl. Trop. Dis.* **13**, e0007213 (2019).

23. Vitek, C. R. & Wharton, M. Diphtheria in the former Soviet Union: reemergence of a pandemic disease. *Emerg. Infect. Dis.* **4**, 539–550 (1998).

24. Hardy, I. R., Dittmann, S. & Sutter, R. W. Current situation and control strategies for resurgence of diphtheria in newly independent states of the former Soviet Union. *Lancet* **347**, 1739–1744 (1996).

25. Lidani, K. C. F. et al. Chagas disease: from discovery to a worldwide health problem. *Front. Public Health* **7**, 166 (2019).

26. Bern, C. Chagas' disease. *N. Engl. J. Med.* **373**, 456–466 (2015).

27. Ghersi, B. M. et al. In the heart of the city: Trypanosoma cruzi infection prevalence in rodents across New Orleans. *Parasit. Vectors* **13**, 577 (2020).

28. Jonsson, C. B., Figueiredo, L. T. M. & Vapalahti, O. A global perspective on hantavirus ecology, epidemiology, and disease. *Clin. Microbiol. Rev.* **23**, 412–441 (2010).

29. Mitchell, C. PAHO/WHO. https://www3.paho.org/hq/index.php?option=com_content&view=article&id=14911:hantavirus&Itemid=0&lang=en#gsc.tab=0 (2019).

30. Muyangwa, M., Martynova, E. V., Khaiboullina, S. F., Morzunov, S. P. & Rizvanov, A. A. Hantaviral proteins: structure, functions, and role in hantavirus infection. *Front. Microbiol.* **6**, 1326 (2015).

31. Bouzid, M., Hunter, P. R., Chalmers, R. M. & Tyler, K. M. Cryptosporidium pathogenicity and virulence. *Clin. Microbiol. Rev.* **26**, 115–134 (2013).

32. Rose, J. B., Huffman, D. E. & Gennaccaro, A. Risk and control of waterborne cryptosporidiosis. *FEMS Microbiol. Rev.* **26**, 113–123 (2002).

33. Ghersi, B. M. et al. Flooding and abandonment have shaped rat demography across post-Katrina New Orleans. *Landsc. Urban Plan.* **215**, 104218 (2021).

34. Peterson, A. C. et al. Rodent virus diversity and differentiation across post-Katrina New Orleans. *Sustain. Sci. Pract. Policy* **13**, 8034 (2021).

35. Gulachenski, A., Ghersi, B. M., Lesen, A. E. & Blum, M. J. Abandonment, ecological assembly and public health risks in counter-urbanizing cities. *Sustain. Sci. Pract. Policy* **8**, 491 (2016).

36. Suarez, G. P. et al. A generic arboviral model framework for exploring trade-offs between vector control and environmental concerns. *J. Theor. Biol.* **490**, 110161 (2020).

37. Buckell, J. et al. COVID-19 vaccination, risk-compensatory behaviours, and contacts in the UK. *Sci. Rep.* **13**, 8441 (2023).

38. Verheyen, C. A. & Bourouiba, L. Associations between indoor relative humidity and global COVID-19 outcomes. *J. R. Soc. Interface* **19**, 20210865 (2022).

39. Fisman, D. N. Seasonality of infectious diseases. *Annu. Rev. Public Health* **28**, 127–143 (2007).

40. Carlson, C. J., Gomez, A. C. R., Bansal, S. & Ryan, S. J. Misconceptions about weather and seasonality must not misguide COVID-19 response. *Nat. Commun.* **11**, 4312 (2020).

41. Moriyama, M., Hugentobler, W. J. & Iwasaki, A. Seasonality of respiratory viral infections. *Annu. Rev. Virol.* **7**, 83–101 (2020).

42. GAO. *Covid-19: data quality and considerations for modeling and analysis.* https://www.gao.gov/products/gao-20-635sp (2020).

Acknowledgements

Author contribution

N.H.F., S.J.R., M.J.B. convened the author group; N.H.F., S.J.R., M.J.B., L.B., N.L.G., Q.H., D.L.M., M.P., D.K.P. and C.V. all conceived of and wrote the manuscript; N.H.F., S.J.R., M.J.B., L.B., N.L.G., Q.H., D.L.M., M.P., D.K.P. and C.V. all reviewed and edited drafts, and contributed to the final version.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Nina H. Fefferman or Sadie J. Ryan.

Peer review information *Communications Medicine* thanks Prateek Kunwar and the other anonymous reviewer(s) for their contribution to the peer review of this work.

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

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