Understanding human behaviour for pandemic preparedness with epigames

Andrés Colubri[1,2,*], Dmitri Williams[3], Thomas Valente[4], Chris T. Bauch[5], John M. Drake[6,7], Melinda C. Mills[8], John Drury[9], Christophe Fraser[7], Luca Ferretti[7], Jasmina Panovska-Griffiths[7,10,*]

[1] Department of Genomics and Computational Biology, UMass Chan Medical School, Worcester MA 01605, United States of America
[2] Broad Institute of Harvard and MIT, Cambridge MA 02142, United States of America
[3] USC Annenberg School for Communication and Journalism, University of Southern California, Los Angeles CA 90007, United States of America
[4] Department of Population and Public Health Sciences, Keck School of Medicine of University of Southern California, Los Angeles CA 90032, United States of America
[5] Department of Applied Mathematics, University of Waterloo, Waterloo ON N2L 3G1, Canada
[6] Center for the Ecology of Infectious Diseases and Odum School of Ecology, University of Georgia, Athens GA 30602, United States of America
[7] Pandemic Sciences Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7LF, United Kingdom
[8] Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, United Kingdom
[9] School of Psychology, University of Sussex, Brighton BN1 9QH, United Kingdom
[10] The Queen's College, University of Oxford, Oxford OX1 4AW, United Kingdom
*=corresponding authors

Infectious diseases that spread from person to person by direct transmission, respiratory pathogens such as influenza and coronaviruses among them, impose a large global health burden and remain the most likely causative agents for future devastating pandemics[1]. For many such diseases, transmission occurs when individuals are in close proximity for a sufficient time and through highly structured social contact networks[2,3]. Data on the properties of these networks, including their temporal and spatial structure, how pathogens spread on them, and how interventions may alter this spread are scarce[4] or inconsistent[5], and seldom incorporate behavioural features. This produces a knowledge gap between policy-relevant models of pathogen transmission and the data they require: contact networks in high spatial and temporal resolution and their variability and malleability under different conditions.

We propose a solution that leverages mobile technology to measure contact networks across social settings, environmental conditions, and various contexts, while explicitly integrating behavioural data. Digital smartphone-based platforms that enable collection of human behaviour and environmental and contextual factors during experimental epidemic games ("epigames") can generate data on real-life dynamic social contact networks, arguably the closest proxy for observing pathogen transmission in human populations. Epigames are controlled situations in which participants join a simulated epidemic via a gamified smartphone app. Over the course of the game, participants interact with each other, enabling the Bluetooth signal between phones to measure mutual proximity, contact duration, and social connection. As the game progresses, participants may become infected by a hypothetical pathogen, moving through susceptible, infectious, and then recovered states. Furthermore, they can make decisions that affect their own and others' quantifiable outcomes. For example, whether they choose to isolate after being told they are infected, or if they use a vaccine or antivirals, in response to the simulated spread and the rewards/penalties built-into the app (such as earning points for isolating, losing points for getting infected or infecting others, and winning prizes at the end of the game). Participants can also respond to survey questions administered through the app about their knowledge and attitudes regarding public health interventions, disease risk, perceived norms, shared identity, and other factors. This would make epigames uniquely capable of gathering not only real-life contact networks, but also behavioural and attitudinal data from the participants, and thus give insight beyond the game context into how people *may* respond during real outbreaks.

This approach originates in classic behavioural experimentation in economics and game theory (e.g., prisoner's dilemma)[6] but has recently been gaining recognition as a valuable source of data in both environmental science[7] and public health[8]. Moreover, during the COVID-19 pandemic, it was applied to study attitudinal and behavioural choices by means of online psychological experiments seeking to replicate disease spread[9,10]. The epigames approach is different from these predecessors in that epigames are a form of field experiment[11] that take place in naturalistic settings, including college campuses, scientific conferences, community organisations, and congregate workplaces. Furthermore, epigames have a high degree of mechanistic realism (i.e., the simulated disease spreads over proximal interactions just as direct contact pathogens do). This provides real-world action fidelity, the correspondence between actions in a real and simulated environments[12], addressing many concerns about the external validity of the data from these experiments[13], an issue affecting all behavioural experimentation methods[14].

Epigames' data can be analysed with established game theory and network science techniques, as well as with novel machine learning methods like mechanistically informed deep learning[15], to characterise the different networks that form across settings and contexts and identify the key features that influence their construction. Epigames also provide an adaptable mechanism for testing a wide range of interventions to reduce disease spread. By implementing epigames with different incentive structures to "nudge" participants to take decisions that modify their susceptibility to infection or transmission rates during the game, researchers would be able to test hypothesis on individual perceptions[9] and network factors[16] that influence behaviour. As mentioned earlier, the game could involve the decision to wear a mask, take a diagnostic test, or receive a vaccine, with varying costs and benefits associated to each one. From a network intervention perspective, epigames could be used to evaluate the effect of individual measures (e.g., messages targeting to the most connected individuals), group-based strategies (e.g., behavioural nudges addressing all member of densely knit cliques), and induction approaches (e.g., introducing opinion leader "seeds") to stimulate peer-to-peer diffusion of protective behaviours. Finally, the high-resolution data from these intervention experiments can beget environmentally- and behaviourally calibrated realistic agent-based models of epidemics that will enable an unprecedented assessment of the role of behaviour in pathogen transmission and response to interventions.

The concept of app-based epidemic games builds upon prior efforts, while explicitly designed to overcome their limitations to study contact networks, human behaviour, and the interplay between the two. Traditional methods like manual contact diaries have been successful in measuring the social mixing patterns of large numbers of individuals[2] across age cohorts and geographical locations[17], but they are limited by their reliance on self-reporting, which is subject to recall bias and low temporal resolution. Sensor-based technologies offered a way to address these limitations. Research projects using Radio Frequency Identification (RFID) tags such as SocioPatterns[18] and Bluetooth-enabled mobile phones like the FluPhone Project[19] and the Copenhagen Networks Study[20], made it possible to capture high-resolution, objective proximity data. However, all these projects had a narrower focus on the physical contact network itself; they could measure potential transmission pathways but were not designed to simultaneously capture health-related preferences, attitudes, and decisions of the individuals forming those networks. Furthermore, many of the studies required specialised hardware (e.g., RFID tags, loaner phones) or were confined to specific settings like schools[21] or cruise ships[22], restricting their scalability and generalizability.

More recent projects leveraging participants' own smartphones, such as SafeBlues[23] and Operation Outbreak (OO)[24], can be considered the closest predecessors to app-based gamified epidemiology experiments. Yet, their primary objectives were different: SafeBlues focused on using virtual spread to forecast epidemics in real-time, and OO was first designed for experiential education in outbreak science and public health. While subsequent work[25,26] demonstrated the potential of the OO platform for epidemiology research, neither was originally conceived as a flexible, reusable infrastructure for systematically testing hypotheses about human behaviour and networks during epidemics.

The epigames approach addresses these gaps by generating four interrelated data streams: (1) high-resolution, real-life contact networks, (2) quantifiable behavioural data from in-game decisions (e.g., choices to self-isolate or vaccinate), (3) attitudinal data from integrated questionnaire surveys, and (4) environmental data that may influence person-to-person interactions (e.g., weather). While this approach is not without its own limitations, such as the Hawthorne effect (study-induced changes in participant behaviour) and network sampling inaccuracies due to sensor heterogeneity in consumer devices, it represents a significant advance by creating a unified experimental framework to study not only how a pathogen could spread, but also why it spreads in the patterns it does, driven by human choices. More broadly, open and reproducible science standards make it an essential requirement that the infrastructure for the epigames be built to avoid reimplementation of the basic methods to reproduce and extend past results. Constructed correctly, such infrastructure can provide a powerful tool to test hypotheses and validate new models that incorporate human behaviour and environmental factors in a more principled and data-driven manner, bridging epidemiological modelling with the biological, social, and ecological sciences.

To illustrate the potential of an experimental platform addressing these gaps, **Figure 1** describes a hypothetical epigame, the "Quarantine Game" (QG), that could be designed with such a platform to empirically investigate quarantine-seeking behaviour. This QG would facilitate the generation of high-resolution data on evolving disease risk, social influence, and protective health behaviours. In terms of external validity, this experimental design also addresses the issue of indicator mapping: how much in-game behaviours reflect, or map over to, real-world counterparts, by proposing pre- and post-game surveys that allow researchers to quantify behaviour parallelism and the attitude-behaviour gap[13] by comparing survey responses to in-game behavioural choices. Large-scale OO research pilots conducted over the past five years were either observational[25] or exploratory in nature rather than employing rigorous experimental design[26]; however, they successfully collected both network and behavioural data, supporting the feasibility of this approach.

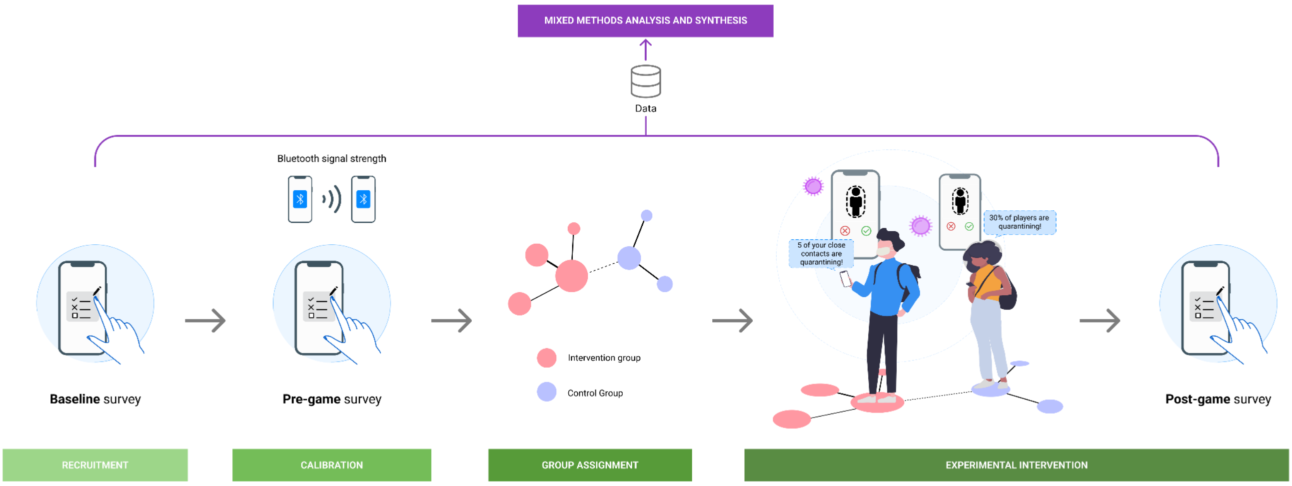

Figure 1: *Proposed study protocol for the Quarantine Game (QG), a scenario-based experimental epidemic game (epigame) in which participants face decisions about whether to self-isolate after exposure to a simulated digital pathogen, with varying levels of peer behaviour information and in-game incentives. In the QG, physical proximity and contact network structure determine transmission of the pathogen among participants; however, they do not need to physically isolate to quarantine, as it would be unrealistic to ask them to do so. Rather, they isolate in the world of the game by choosing the quarantine behaviour in the app. Thus, the costs and benefits of the gamified quarantine need to map to some extent to those of real-life quarantine. One such possible mapping (but not the only one) could be introduced via a point system where an immediate penalty due to quarantine would reflect lost opportunities for work, socialization, etc., and a steeper penalty for getting sick later in the game, representing the health burden of disease. Final point scores could result in prizes (e.g., gift cards) via a lottery to introduce a tangible incentive. The general study protocol would comprise the following stages: (1) recruitment, during which participants install the epigames app and complete baseline attitudinal surveys, (2) calibration, when participants complete parallel surveys reframed in the epigame context, and the app constructs the empirical distribution to Bluetooth signal strengths (RSSI) between participants to estimate the parameters of the RSSI-to-distance mapping by leveraging the known peaks of socially-determined pairwise physical separation, (3) group assignment, via graph-cluster randomization methods to minimize cross-talk between groups, (4) experimental manipulation, where each group is exposed to different experimental conditions (e.g., presence or absence of information about peer behaviour; aggregate vs. individual-level peer compliance data, local vs. global risk of infection or number of cases), and (5) data analysis comprising a mixed methods and synthesis approach to evaluate behavioural parallelism and network data fidelity prior to developing downstream network and behavioural models. (Figure credits: Yinan Dong and Mansi Khandpekar from Colubri Lab)*

Concluding remarks

The COVID-19 pandemic underscored a significant asymmetry in our scientific arsenal against infectious diseases. On one hand, we have powerful experimental and computational tools to sequence new viral genomes in a matter of weeks, study the immune response of human cells in minute detail, and determine the structure of host receptor molecules with unprecedented precision, all while we aim to develop vaccines in just 100 days from the moment a novel pathogen emerges. On the other hand, we lack robust, scalable methods to study the human behaviours and social structures that mediate transmission. This is a major gap that needs to be addressed to enable the construction of more realistic and sophisticated models, e.g., agent-based models that include individual behaviour and social interactions.

Gamified app-based studies carried out with an experimental platform for infectious disease research—one that enables interdisciplinary teams to design, deploy, and refine studies integrating behavioural, network, environmental, and biological data at scale—can fill this gap. The platform will not only allow generation of dynamic real-life social contact networks that are crucial for modelling pathogen spread but also enable linking with behavioural and attitudinal data, as well as more rigorous testing of public health interventions. Writing in 2013, Edmunds, Eames, and Keogh-Brown identified the challenge of constructing epidemiological models that meaningfully incorporates human behaviour, given the trade-offs between observational studies with limited generalizability and mechanistic models without a clear theoretical basis due to the multiplicity of behavioural change theories. They concluded that *"understanding which of these [theories] are most applicable and in what circumstances will be a significant undertaking for which high-quality empirical information will be essential if these models are to be used to guide future decision-making"*[27]. More than 12 years later and in the aftermath of a worldwide pandemic, the challenge remains. Developing a platform that addresses this persistent challenge and allows collation of dynamic real-life contact networks will represent a significant advance for the scientific community in bridging the gap between experiments and models in infectious disease research.

Epigames offer a promising path forward. They can provide a unique combination of control, realism, and scalability. By combining behavioural decision-making, real-time network dynamics, and structured incentives within simulated outbreak scenarios, epigames can investigate critical questions: What drives protective action? How do social norms and peer influence shape health decisions? What interventions change behaviour and network structure? To answer these questions and be better prepared for the next pandemic, we need investment in interdisciplinary research that brings together epidemiologists, behavioural scientists, game designers, computer scientists, among other experts, as well as in the development and maintenance of the necessary infrastructure. Just as genome sequencing became a global priority during COVID-19 and the avian influenza outbreaks, behavioural-epidemiological experimentation via digital technology must now be elevated to a core capability in pandemic preparedness.

References

1. Mallapaty, S. The pathogens that could spark the next pandemic. *Nature* **632**, 488 (2024).
2. Mossong, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74 (2008).
3. Block, P. *et al.* Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nat Hum Behav* **4**, 588–596 (2020).
4. Grassly, N. C. & Fraser, C. Seasonal infectious disease epidemiology. *Proc Biol Sci* **273**, 2541–2550 (2006).
5. Kummer, A. G. *et al.* Evaluating Seasonal Variations in Human Contact Patterns and Their Impact on the Transmission of Respiratory Infectious Diseases. *Influenza Other Respir Viruses* **18**, e13301 (2024).
6. Becker, G. S. *The Economic Approach to Human Behavior*. (University of Chicago Press, 1976).
7. Wang, Z. *et al.* Communicating sentiment and outlook reverses inaction against collective risks. *Proc Natl Acad Sci U S A* **117**, 17650–17655 (2020).
8. Pouliquen, A. *et al.* An experimental game to assess hunter's participation in zoonotic diseases surveillance. *BMC Public Health* **24**, 342 (2024).
9. Woike, J. K., Hafenbrädl, S., Kanngiesser, P. & Hertwig, R. The transmission game: Testing behavioral interventions in a pandemic-like simulation. *Sci Adv* **8**, eabk0428 (2022).
10. Nunner, H., Buskens, V., Corten, R., Kaandorp, C. & Kretzschmar, M. Disease avoidance

threatens social cohesion in a large-scale social networking experiment. *Sci Rep* **13**, 22586 (2023).
11. van Dijk, E. & De Dreu, C. K. W. Experimental Games and Social Decision Making. *Annu Rev Psychol* **72**, 415–438 (2021).
12. Stoffregen, T. A., Bardy, B. G., Smart, L. J. & Pagulayan, R. On the nature and evaluation of fidelity in virtual environments. in *Virtual and adaptive environments* 111–128 (CRC Press, 2003).
13. Williams, D. The mapping principle, and a research framework for virtual worlds. *Commun. Theory* **20**, 451–470 (2010).
14. Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. & Podsakoff, N. P. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol* **88**, 879–903 (2003).
15. Kuwahara, B. & Bauch, C. T. Predicting Covid-19 pandemic waves with biologically and behaviorally informed universal differential equations. *Heliyon* **10**, e25363 (2024).
16. Valente, T. W. Network interventions. *Science* **337**, 49–53 (2012).
17. Prem, K., Cook, A. R. & Jit, M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Comput Biol* **13**, e1005697 (2017).
18. Cattuto, C. *et al.* Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS One* **5**, e11596 (2010).
19. Yoneki, E. FluPhone study. in *Proceedings of the 6th ACM workshop on Challenged networks* 65–66 (ACM, New York, NY, USA, 2011).
20. Sapiezynski, P., Stopczynski, A., Lassen, D. D. & Lehmann, S. Interaction data from the Copenhagen Networks Study. *Sci Data* **6**, 315 (2019).
21. Stehlé, J. *et al.* High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS One* **6**, e23176 (2011).
22. Pung, R. *et al.* Using high-resolution contact networks to evaluate SARS-CoV-2 transmission and control in large-scale multi-day events. *Nat Commun* **13**, 1956 (2022).
23. Dandekar, R. *et al.* Safe Blues: The case for virtual safe virus spread in the long-term fight against epidemics. *Patterns (N Y)* **2**, 100220 (2021).
24. Colubri, A. *et al.* Preventing Outbreaks through Interactive, Experiential Real-Life Simulations. *Cell* **182**, 1366–1371 (2020).
25. Specht, I. *et al.* Analyzing the impact of a real-life outbreak simulator on pandemic mitigation: An epidemiological modeling study. *Patterns (N Y)* **3**, 100572 (2022).
26. Musa, S. S. *et al.* The effect of behavioral factors and intervention strategies on pathogen transmission: Insights from a two-week epidemic game at Wenzhou-Kean University in China. *medRxiv* (2024) doi:10.1101/2024.12.14.24318955.
27. Edmunds, W. J., Eames, K. & Keogh-Brown, M. Capturing human behaviour: is it possible to bridge the gap between data and models? in *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases* (eds. Manfredi, P. & D'Onofrio, A.) 311–321 (Springer New York, New York, NY, 2012).