A call for citizen science in pandemic preparedness and response: beyond data collection

Yi-Roe Tan,1 Anurag Agrawal,2 Malebona Precious Matsoso,3 Rebecca Katz,4 Sara L M Davis,4,5 Andrea Sylvia Winkler,6,7 Annalena Huber,6 Ashish Joshi,8 Ayman El-Mohandes,8 Bruce Mellado,9,10 Caroline Antonia Mubaira,11 Felipe C Canlas,12 Gershim Asiki,13 Harjyot Khosa,14 Jeffrey Victor Lazarus,15 Marc Choisy,16,17 Mariana Recamonde-Mendoza,18,19 Olivia Keiser,20 Patrick Okwen,21 Rene English,22 Serge Stinckwich,23 Sylvia Kiwuwa-Muyingo,13 Tariro Kutadza,11 Tavpritesh Sethi,24 Thuso Mathaha,9 Vinh Kim Nguyen,5 Amandeep Gill,1 Pelling Yap1

ABSTRACT

The COVID-19 pandemic has underlined the need to partner with the community in pandemic preparedness and response in order to enable trust-building among stakeholders, which is key in pandemic management. Citizen science, defined here as a practice of public participation and collaboration in all aspects of scientific research to increase knowledge and build trust with governments and researchers, is a crucial approach to promoting community engagement. By harnessing the potential of digitally enabled citizen science, one could translate data into accessible, comprehensible and actionable outputs at the population level. The application of citizen science in health has grown over the years, but most of these approaches remain at the level of participatory data collection. This narrative review examines citizen science approaches in participatory data generation, modelling and visualisation, and calls for truly participatory and co-creation approaches across all domains of pandemic preparedness and response. Further research is needed to identify approaches that optimally generate short-term and long-term value for communities participating in population health. Feasible, sustainable and contextualised citizen science approaches that meaningfully engage affected communities for the long-term will need to be inclusive of all populations and their cultures, comprehensive of all domains, digitally enabled and viewed as a key component to allow trust-building among the stakeholders. The impact of COVID-19 on people’s lives has created an opportune time to advance people’s agency in science, particularly in pandemic preparedness and response.

INTRODUCTION

The world struggled to respond to the COVID-19 pandemic in a timely and effective manner due to a lack of accurate, continuous real-time global primary data availability on outbreaks, and the absence of globally consistent evidence-based responses. WHO tasked the Independent Panel for Pandemic Preparedness and Response to conduct a review of the global health response to the pandemic, and in their final report, they called for a fundamental transformation of the international system to prevent a future pandemic.1 However, this must be done with care: the pandemic led many states to use the public health crisis to justify sweeping and intrusive surveillance powers, weakening
public trust. A neutral and trusted global pandemic preparedness and response collective that is culturally and contextually informed, science-based, digitally enabled, respectful of human rights and works across the continuum of pandemic phases: namely preparedness, response, recovery and surveillance, needs to be created by engaging all stakeholders, including communities most affected by the infectious diseases in question.

A crucial research priority is trust-building among the stakeholders that are interdependent in the response to a pandemic and creating direct and efficient vehicles of communication with communities directly affected by the pandemic. An analysis has shown that higher levels of trust in the government and interpersonal trust had significant associations with lower COVID-19 infection rates and higher acceptance rates of the vaccine. It is evident that there is a lack of mutual trust and a common language among the public, researchers and policymakers particularly in the areas of development and implementation of public health interventions, as observed during the current and previous outbreaks. This has undermined public compliance with and effectiveness of mitigation strategies. Similarly, in the response to HIV, and other epidemics, community engagement has proven key to an effective response; in the response to HIV, tuberculosis and malaria, this recognition of the right to participation has extended to roles for community representatives at high levels of decision-making. One proven pathway to build and maintain public trust is through community engagement in the three steps of collective knowledge-making, namely problem definition and data generation, simulation and modelling, clear communication and decision-making informed by data-generated insights.

In this analysis, we focus on the involvement of communities in an iterative process from participatory data generation and modelling to visualisation and communication as well as public validation and hypothesis generation, maximised through digital enablers, including artificial intelligence. Here, we define the ideal state of citizen science as a practice of public participation, collaboration and co-creation in all aspects of scientific research to increase knowledge, build trust, generate accurate and timely data, and develop strategies in pandemic preparedness and response (figure 1). Box 1 summarises our methodological approach.

**KEY ASPECTS OF CITIZEN SCIENCE IN PANDEMIC PREPAREDNESS AND RESPONSE**

The literature revealed non-uniform and arguable definitions of community engagement. In this review, we included all mention of community engagement, from passive surveillance of the public to active collaboration with communities. Adapted from Arnstein’s original ladder of citizen participation, the five levels of citizen engagement, namely inform, consult, involve, co-create and empower, describe the iterative process and extent to which governments can engage the community in decision-making (figure 2). In the typology by Shirk et al, citizen science projects between researchers and communities can be classified into five main categories: (1) ‘contractual’, where communities request researchers to perform scientific investigations; (2) ‘contributory’, which are designed by researchers and for which communities primarily contribute data; (3) ‘collaborative’, which are designed by researchers and

---

**Figure 1** Spectrum of citizen science activities which the public can participate in.
for which communities contribute data and help to refine project design, analyse data and disseminate findings; (4) ‘co-created’, which are designed by researchers and communities working together and for which public participants are actively engaged in the research process and (5) ‘collegial’, where non-credentialed individuals perform research independently with varying degrees of expected recognition by researchers. Using COVID-19 as the main use case, we highlighted the successes of and barriers to current participatory approaches used in infectious diseases outbreak and the priority research areas to realise the potential of citizen science to mount an effective response for future pandemics.

Citizen science approaches in participatory data generation

Community engagement in the field of participatory health data gathering has played an important role in public health, particularly in the field of infectious diseases. With the increased penetration of mobile phones and technology, digital-based participatory data generation systems are fast evolving. For example, ‘Mo-Buzz’ was launched in Sri Lanka to combat dengue by allowing the public to report dengue symptoms and potential dengue mosquito-breeding sites. Using global positioning system (GPS) technology, real-time dengue hotspot maps and targeted alerts can be sent to users and communities at risk. Similarly, “Flu Near You”, a real-time influenza symptom tracking system was developed in the USA for early detection and response to influenza by leveraging on voluntary symptom reporting from the public. When COVID-19 struck, this influenza surveillance system was rapidly adapted to “Outbreaks Near Me” for COVID-19 surveillance. Expanding beyond humans to animal diseases, the Participatory One Health Disease Detection project enables early detection of zoonotic diseases. This mobile application implemented in Thailand enabled volunteers to report suspected outbreaks, which contributed to a faster and more coordinated response among government officials, veterinarians and public health experts. By enabling the generation of large volumes of data over short timescales and extensive geographical areas, digital-based participatory systems can better support the early detection of outbreaks through reporting of early symptoms and health-seeking behaviours. With the increase in volumes of data so grows the complexity of information that needs to be extracted from the data. Where classical analytical techniques are unable to cope with the volume and complexity, machine learning and big data tools can be leveraged to better harness data resulting from multiple data sources.

During the initial phases of the COVID-19 pandemic when case numbers were low, many countries adopted the shoe-leather epidemiology surveillance method. Collection of epidemiological data relevant to the investigation was done by community health volunteers physically visiting places to investigate a representative sample of the affected people. This was feasible when numbers were low. However, due to the rapid speed at which the pandemic unfolded, these traditional methods of outbreak investigation were insufficient to ensure timely data collection, and to accurately evaluate and contain situations on the ground. Sole reliance on this type of manpower-intensive surveillance method proved to be unsustainable, particularly in resource-limited settings and high-density populations. Advanced technology and existing digital infrastructure enabled countries to tap into new data sources for surveillance. A spectrum of approaches to collect data from the community, from involuntary mobility monitoring and social media surveillance to participatory syndromic data generation, has been adopted by countries. For example, South Korea uses electronic health records, credit card transactions, GPS data and closed-circuit television to address limitations in memory recall during contact tracing. Japan, Germany, Singapore, New Zealand and Norway, among others, have launched smartphone applications that use Bluetooth signals, GPS tracking or recording of location-specific QR codes to identify individuals who have come into close proximity with a patient with COVID-19.
Numerous participatory surveillance applications have sprung up, including the ‘Zoe COVID-19 symptom study’, which collected voluntarily self-reported information from consenting participants on COVID-19 in Sweden, the UK and the USA. This allowed monitoring of the spread of COVID-19 when access to testing was limited, and also contributed to research questions such as the identification of anosmia as a key symptom of COVID-19 and delirium as a frequent symptom in older people.27

The above highlights a major gap in the current landscape for participatory data generation. Most of the community engagement efforts are limited to the ‘inform’ and ‘consult’ stages where communication is unidirectional, giving communities little or no meaningful input into decision-making about what kinds of data are gathered and how. This raises problems of power and voice: if communities are only low-paid data enumerators or passive producers of data over which they have no control, the result risks widening power imbalances and exploiting vulnerable groups through data mining for private profit. Future participatory approaches should engage communities in collaborative problem definition and corresponding data generation to enable them to play a more active and larger role that goes beyond data provision to empowerment. For example, the Digital Health and Rights Project engages transnational networks of people living with HIV, human rights lawyers and AIDS advocates in designing, implementing and analysing qualitative data in a multicountry study on how young adults use digital platforms to access health information and services, using the results to develop global health policy recommendations.28

**Citizen science approaches in participatory modelling**

Although participatory approaches in data generation have gained popularity over the years, there has not been significant progress in participatory modelling in health, especially in pandemic preparedness and response. Participatory modelling and simulation approaches have mostly been used in the field of natural resources and environmental research.29 30 However, given growing evidence of biases encoded in algorithms and models, and the challenges in quantifying structural factors that can shape health outcomes, it is crucial to engage communities in the review and audit of both the assumptions and data used in modelling.8

The equitable participation of stakeholders, including communities, scientists and policymakers, in the modelling process aims to achieve two outcomes. First, their expertise can help ensure the assumptions behind models are more attuned to the social, cultural and political context. From a technical perspective, participatory modelling provides an appropriate interaction setting for participants to co-produce models which make sense to them and generate useful discussions in framing, designing of settings, assumptions, elements and outputs.31 32 Second, their involvement fosters a collective learning environment resulting in a shared understanding of health issues as well as identification of data sources that can fit better into the models, or data gaps that may need to be addressed.33 34 Through collective reasoning and tools for scenario-based simulations, communities can evaluate the effects of implementing different policies and share policy recommendations, thus facilitating better decisions and improved trust, acceptance and ownership of the modelling outcomes and subsequent interventions.31 35

Participatory modelling has shown some successes in COVID-19, particularly in two domains. First, it was demonstrated that by engaging policymakers and public health professionals throughout all stages of the modelling process, under the guidance of modellers and researchers, modelling outputs can be translated to more tailored policy decision-making.36 37 Second, citizen scientists were deployed to help in simulations through their own personal digital devices. For example, in the online game ‘EVE Online’, a mini-game ‘Project Discovery’ doubles as a citizen science platform studying the human immune system’s response to COVID-19.38 Participants take on data analysis through gameplay that helps researchers isolate specific patterns as predictors of disease severity. In the ‘COVID-19 Moonshot’ project, >200 000 volunteers around the world are helping to find potential treatments for COVID-19 by running simulations on their home computers.39 However, there exists a gap in engaging all groups of stakeholders in the same participatory modelling process for consensus and trust-building, and the risk of only including people who have access to technology.

Although there are growing interests in participatory modelling for pandemic preparedness and response, more can be done to leverage this approach to address the disconnect between science, policy and implementation by uniting communities, modellers, researchers and policymakers towards a common goal and under a common platform. By bringing more transparency to the process of how insights are generated from data, participatory modelling could be the tool of engagement and the bridge to build trust in a collective decision-making process.

**Citizen science approaches in data visualisation and communication**

To drive actionable changes informed by data-generated insights, data visualisations need to be tailored to the accessibility and understanding of different stakeholders, from the public to policymakers. Through the co-design process and a human-centred approach, collective knowledge-making can then be possible to build consensus and drive decision-making through empowerment. One example of citizen science approaches involving the community is the design of the visualisations for the ‘Participatory Tracking Project’ in Tamil Nadu. To present the survey findings in a way that is understandable by even the least literate women in the districts, an iterative feedback mechanism was adopted to fine-tune the
visualisation prototypes. By incorporating cultural and context-specific designs, the research team and villagers produced a final visualisation which the villagers could identify with, empowering them to make informed decisions through data. In the realm of pandemic preparedness and response, participatory visualisation is often an afterthought in which new visualisations are added or revised only after end users start using the systems. During the COVID-19 pandemic, most infographics and illustrations are designed by individual stakeholders with varying success. To truly facilitate the translation of data into applied knowledge and informed decision-making, more collaborative citizen science approaches need to be adopted among groups of stakeholders and at all levels of data and digital literacy to identify the best ways of representing data. Data should be presented in a way to allow community stakeholders to give feedback to its relevance and adequacy, thus validating the analysis and enabling an iterative cycle of reflection for action, as evidently adopted in participatory action research. In this era where social media dominates, public communication of science and policy decisions should be robust, yet understandable by the public, to address issues of misinformation and disinformation. Engaging the community in myth-busting campaigns requires more research on their format and delivery. Citizen science is one possible method to allow for the management of misinformation in the public and social media by facilitating an exchange of balanced information and the engagement of communities as necessary active participants in a complex health information environment. This creates an interactive and bidirectional dissemination of health information, raises awareness of public health issues and improves health literacy of communities. For example, sub-Saharan countries initiated public engagement communication to help the public understand about COVID-19 spread and precaution measures. This accelerated the implementation of acceptable community interventions such as the setting up of hand washing stations at locations convenient to the people. Through effective public engagement, there can be more targeted health messaging and communication, including valuable feedback loops.

**POTENTIAL TO MAXIMISE CITIZEN SCIENCE IMPACT**

Based on the reviewed literature, we noted that although citizen science approaches have been used for data generation during COVID-19, there are only a handful of examples for participatory modelling, and data visualisation and communication. By focusing on the gaps along the data-models-visualisations workflow, more clarity and transparency to how data inputs translate to actionable outputs will be made possible, thus enabling trust building among the various stakeholders (figure 3). Furthermore, COVID-19 has accelerated the development and application of digital technologies and innovation in health. Building on previous work in digital-based participatory systems, digitally enabled citizen science in pandemic preparedness and response has the potential to scale in both breadth and depth to overcome some of the challenges faced by traditional participatory approaches. However, one of the most systemic barriers is digital inequity, particularly, but not reserved to, low-to-middle-income countries. Digital inequities include barriers to access to digital tools, networks and communication channels, as well as conflicting life priorities, particularly among women, girls and vulnerable and marginalised populations. There exists a risk of digitally enabled citizen science approaches in widening existing health disparities by excluding those who do not have access to digital technologies. To truly tap the potential of digitally enabled citizen science in pandemic preparedness and response, countries and agencies should address these gaps through support for digital literacy and empowerment among the populations who need to engage in citizen science, as well as promoting affordable access to technology, internet and mobile data.

Capacity development efforts to improve the quality of data generated from citizen scientists is another area of importance highlighted by researchers. Researchers have voiced concerns about the quality of data generated from communities. However, with sufficient oversight, training and guidelines in place, a collaborative model of science with a balance between citizen-initiated and researcher-initiated projects can be possible with bidirectional trust established. Capacity development efforts are

---

**Figure 3** Research gaps in the continuum of participatory data generation-modelling-visualisation.
growing in this area, such as the citizen science project launched by the University of Illinois Chicago School of Public Health. Members of community-based organisations can participate in the school’s new citizen scientist certificate programme to build research skills, before being deployed to the community to explore public health issues such as that of vaccine hesitancy in the USA.59

CONCLUSION
The impact of COVID-19 on the lives of people has increased the public’s query into the process of policymaking and their attention on informed decision-making. It is an opportune time to advance people’s agency in science, particularly in the field of pandemic preparedness and response. To do that, we will need to first raise awareness among the community on their important role in the above-mentioned three steps of collective knowledge-making, as well as elucidate the acceptability and feasibility of participatory citizen science approaches. Active engagement of communities should be viewed as a pivotal, rather than an ancillary component in all research efforts to empower and build trust among the public, researchers and policymakers. Ultimately, this will hopefully convert to more timely and efficient evidence-based responses to future pandemic risks, reducing their impact on individuals, communities and countries.

Author affiliations
1International Digital Health & AI Research Collaborative (I-DAIR), Geneva, Switzerland
2Trivedi School of Biosciences, Ashoka University, Sonepath, Haryana, India
3Pharmacy & Pharmacology, University of Witwatersrand, Member of IPPR, Johannesburg-Braamfontein, South Africa
4Center for Global Health Science and Security, Georgetown University, Washington, DC, USA
5Global Health Centre, Graduate Institute Geneva, Geneva, Switzerland
6Centre for Global Health, Department of Neurology, Technical University of Munich, Munich, Germany
7Centre for Global Health, Institute of Health and Society, University of Oslo, Oslo, Norway
8Graduate School of Public Health and Health Policy, City University of New York, New York, New York, USA
9School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand, Johannesburg, South Africa
10Subatomic Physics, iThemba Laboratory for Accelerator Based Sciences, Somerset West, South Africa
11Zimbabwe National Network of People Living with HIV (ZNNP+), Harare, Zimbabwe
12Wireless Access for Health, Tarlac, Philippines
13African Population and Health Research Center, Nairobi, Kenya
14International Planned Parenthood Federation, New Delhi, India
15Hospital Clinícal, University of Barcelona, Instituto de Salud Global de Barcelona, Barcelona, Spain
16Centre for Tropical Medicine and Global Health, University of Oxford Nuffield Department of Medicine, Oxford, UK
17Oxford University Clinical Research Unit, Ho Chi Minh City, Ho Chi Minh, Viet Nam
18Institute of Informatics, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil
19Bioinformatics Core, HCPA, Porto Alegre, Brazil
20Institute of Global Health, Universite de Geneve, Geneva, GE, Switzerland
21Effective Basic Services (eBASE), Bamenda, Cameroon

REFERENCES
1The independent panel for pandemic preparedness & response. Available: https://theindependentpanel.org/
2Intrusive, omnipresent surveillance growing during COVID-19 pandemic, un expert warns. Available: https://www.who.int/en/news-room/detail/26-09-2022-Intrusive,-omnipresent-surveillance-growing-during-COVID-19-pandemic,-un-expert-warns
3COVID-19: an exploratory analysis of infection and fatality rates, and contextual factors associated with preparedness in 177 countries, from Jan 1, 2020, to Sept 30, 2021. Lancet 2022;399:1489-1512.
4Glasziou P, Moore H, Kennedy D, et al. A snapshot of the practicality and barriers to COVID-19 interventions: public health and healthcare workers’ perceptions in high and low- and middle-income countries. PLoS One 2021;16:e0260041.
5Bangerter A, Investigating and rebuilding public trust in preparation for the next pandemic. 2021;14: 19-1-3.
6van der Weerd W, Timmermans DR, Beaulieu DJ, et al. Monitoring the level of government trust, risk perception and intention of the general public to adopt protective measures during the influenza A (H1N1) pandemic in the Netherlands. BMC Public Health 2011;11:575.
7Bargain O, Aminjonov U. Trust and compliance to public health policies in times of COVID-19. J Public Econ 2020;192:104316.
8Sara MD. The uncounted: politics of data in global health. New York Cambridge University Press, 2020.
9Henderson J, Ward PR, Tonkin E, et al. Developing and maintaining public trust during and Post-COVID-19: can we apply a model developed for responding to food scares? Front Public Health 2020:8:359.
10Adams S, Rhodes T, Lancaster K. New directions for participatory modelling in health: redistributing expertise in relation to localised matters of concern. Glob Public Health 2021;1:1–15.
