UCL

AI for People & Planet

Policy Commentary: Emerging Epidemics

AI for People and Planet – Emerging Epidemics

The **Artificial Intelligence (AI) for Emerging Epidemics** policy roundtable discussion formed part of the UCL roundtable series 'AI for People and Planet.' Adequately managing outbreaks and diseases, as we have seen with the ongoing COVID-19 pandemic, is often hindered by a deluge of misinformation and disinformation. The aim of this roundtable was to convene a group of experts in AI, public health, epidemiology and infectious diseases to clarify thinking about the relationship between AI and emerging epidemics.

Executive Summary

For the purposes of this document, we adopted the definition of AI as *a reasonable combination of data and mathematically smart algorithms.* The key points highlighted during the discussion include:

- The descriptive, predictive and prescriptive role of AI prior to the current pandemic has rapidly evolved since the outset of the outbreak, leading to innovations, capabilities and enthusiasm for data science that have not previously been seen.
- The current and past focus has been on human-to-human transmission, but there needs to be an equal emphasis on integrating research on the epizootic cycles of pathogens and epidemic transmission cycles.
- A major challenge to the widespread adoption of AI is lack of standardisation, integration and availability of health data.
- Automation of the way in which datasets are brought together would assist in making informed predictions about future events.
- There are multiple streams of data that feed into decision-making. Determining which datasets are appropriate to inform decisions, and also understanding the aim of decisions, is crucial if AI is to be used effectively.
- Alignment between organisations in different policy areas and sectors and interdisciplinary approaches are key for ensuring that data are used ethically and democratically.

These key points are discussed in greater detail below.

Discussion 1: How did we see the role of AI for epidemic preparedness and response before the COVID-19 pandemic?

Descriptive, predictive and prescriptive roles

AI had three main roles before the COVID-19 pandemic, which are outlined in Table 1.

Table 1: AI's roles pre-COVID-19

Role	Description of role prior to COVID-19	Example of role during COVID-19
Descriptive - what is happening	AI's descriptive capabilities gave access to a certain amount of data to monitor people's health and inform us of other diseases. However, other kinds of rich data were missing that would have provided more details on specific contexts.	During the COVID-19 pandemic, mobility data reports from Google, Apple and Facebook etc. have helped map the spread of the virus and understand patterns in transmission.

Predictive - what is about to happen	The majority of existing research has been based on classic AI models developed before COVID-19 within an environment where time was available to understand models.	During the COVID-19 pandemic, the criticality of the results has been much more important.
Prescriptive - what do we need to do to prepare for what is going to happen	Previously, AI was not able to assist in mapping out the potential effects of policy decisions.	New AI tools and algorithms are now being used to help determine which kind of policy interventions are needed to achieve the best outcome possible.

A 'One Health' approach to data

The pre-COVID-19 conceptualisation of models and AI tools for predicting the jump from animals to humans was limited, as there were little data available. Whilst previously the focus has been on human-to-human transmission, there now needs to be an equal emphasis on understanding the epizootic cycles of pathogens and epidemic transmission cycles. Adopting a One Health perspective means looking at the interaction of human health, animal health and the environment. However, there is currently no direct or regular exchange of information and data between these areas, beyond high-burden diseases, such as malaria or dengue fever.

While datasets are not currently adequately integrated, there is scope to automate the way we bring these together to start making informed predictions about future events. Greater public awareness and platforms for data scientists working in ecology to communicate their findings are also needed.

Limitations due to lack of data access

Methods for modelling infectious diseases using online user activity data, such as web searches, had been in development before the current pandemic. However, there have been criticisms of past efforts due to perceived biases, as well as unreliability in models caused by barriers in access to datasets. For example, models created during the Swine Flu pandemic overestimated flu prevalence and led to the decommissioning of Google Flu Trends.

This issue has been partially solved by encouraging Google to offer access to aggregate search activity data as well as research groups developing more advanced machine learning models that can mitigate the issues of their predecessors. However, this body of work has focused on diseases for which there is a well-established evidence base. Limited focus has been placed on forecasting, which meant that existing modelling methods were unprepared to manage the onset of the COVID-19 pandemic. More rigorous modelling approaches for estimating influenza prevalence based on web searches and their incorporation into national health surveillance systems served as the foundation for building novel models for COVID-19.

Discussion 2: How has AI been useful in mitigating direct and indirect effects of COVID-19, and where did it fall short (e.g. processes, frameworks, ethics, data safety, funding, regulation, access)?

Challenge – Validating models

Since the outset of the COVID-19 pandemic, there has been a rapid progress in AI's capacity to manage, monitor and forecast outbreaks. However, the issue of missing data, especially for people from Minority Ethnic Groups in the Global North or people in the Global South, is still posing a major challenge to the validity of models. Lack of access to appropriate data is particularly problematic for resource-limited countries where routine, electronic health record systems for human health do not exist at the national level.

Initially, models were also being trained on highly skewed datasets, such as those with disproportionately high numbers of those without the disease. As clinical understanding of the virus has improved, models have been adapted locally, which has been key in the context of very little and often low-quality data. Additionally, having greater domain expertise in public health and epidemiology is also crucial for the relevance of models and to contextualise results.

Challenge – Value-based data

Value-based data (i.e. social and behavioural practices and personal beliefs) are not easily collected in datasets. Being able to systematically feed this kind of data into models still poses a challenge.

UCL COVID-19 Social Study

The UCL COVID-19 Social Study is the UK's largest study into the psychological and social impact of the pandemic. With over 70,000 participants, the study provides weekly data to decision-makers both nationally and internationally. Negative attitudes towards vaccines are a major public health concern and the Social Study was able to identify predictors of COVID-19 vaccine refusal.

Opportunity – Leveraging enthusiasm for data science

There has been a newfound interest in data science amongst both the research and policymaking communities. The rapid genetic sequencing of the virus and subsequent sharing of this bioinformatics data was a remarkable feat that has continued into the second year of the pandemic. Genome sequencing has enabled researchers to explore different diagnostic, treatment and vaccine interventions and played a large role in the management of the outbreak. As the amount of data continues to grow, policymakers and decision makers have called for dashboards to help make sense of the data in real time.

i-sense COVID RED

To address gaps in available data, UCL researchers have led the rapid adaptation of existing i-sense technologies, which aim to identify outbreaks of infectious disease, to create the i-sense COVID Response Evaluation Dashboard (COVID RED). COVID RED collates and presents data from the Office for National Statistics (ONS), Public Health England and NHS and is currently the only dashboard that explores the entire COVID-19 response system as a whole.

Opportunity – Evidence synthesis

At the time of writing (March 2021), over 80,000 papers had been published on COVID-19. Using AI and natural language processing (the analysis of language in text and speech) to synthesise this massive body of evidence could assist decision-makers by identifying effective interventions and key patterns and trends across studies.

Discussion 3: What is needed to enable AI to help build equitable, resilient health systems that are prepared to respond to future epidemics?

Comprehensive data streams and frameworks

There are multiple streams of data that feed into decision-making, including health management information systems, routine disease-specific records, city information, latency operation data and satellite imagery. Determining what datasets are appropriate to inform decisions, as well as understanding the aim of these decisions, is crucial if AI is to be used effectively. Additionally, a greater understanding of which health and ecological data should be monitored is needed, especially data related to zoonotic origin diseases.

A precursor to the widespread adoption of AI is greater standardisation in health data. Robust policy frameworks are needed to create environments that include regulatory measures, incentive programs and research streams to address public health priorities.

Accounting for national readiness

An AI maturity model provides a framework for assessing an organisation or country's current AI readiness and capabilities. Assessing at which stage (exploring, experimenting, formalising, optimising and transforming) in the maturity model a country resides is crucial for determining what AI and digital health technologies can be implemented.

Democratising AI

Alignment between organisations in different policy areas and sectors and interdisciplinary approaches are key for ensuring that data are used ethically, with considerations for how the data are likely to be used and accessed in the future. Collaboration between academia and industry to increase the recruitment of people from diverse backgrounds into the field of AI would assist in democratising the field. Crucially, issues of equity and social justice must be considered at every point to ensure that AI models do not exacerbate existing inequalities.

Conclusion

The ongoing COVID-19 pandemic has led to rapid advancements in the use of AI modelling and technologies for controlling and managing its impacts. In order for these advancements to continue to progress effectively, system-wide changes are needed to better integrate and standardise health datasets. While the past focus has been on human-to-human transmission, there now needs to be an equal emphasis on examining epizootic cycles of pathogens and epidemic transmission cycles. Beyond managing disease outbreaks, using AI-based algorithms for analysing health data will require policy and governance frameworks to ensure that AI is used ethically and democratically.

AI has the capability to play a significant role in predicting and managing emerging epidemics; what is now needed is alignment across sectors to realise its full potential. However, there remains significant variance in how different populations and communities either benefit from AI or are harmed by it. An interdisciplinary agenda of research and action is needed which acknowledges the importance of focusing on people, processes, and politics, going beyond the technical discussions of AI towards how it can be used to achieve impact in the real world.

Participants

UCL Participants

Dr Lele Rangaka (Institute for Global Health, *Chair*)
Dr Yipeng Hu (Faculty of Engineering Science)
Dr Vasileios Lampos (Faculty of Engineering Science)
Dr Audrey Prost (Institute for Global Health)
Professor Stephen Roberts (Institute for Global Health)

External Participants

Mr Amane Dannouni (Boston Consulting Group)
Mr Nazir Halliru (Columbia University)
Dr Nick Jackson (Coalition for Epidemic Preparedness Innovations)
Dr Ramesh Krishnamurthy (World Health Organization)
Professor Trudie Lang (University of Oxford)
Dr Maimuna Majumder (Harvard Medical School)
Dr Mosa Moshabela (University of KwaZulu-Natal)
Dr Leah Mwai (UK FCDO East Africa Research and Innovation Hub)
Professor Barbara Prainsack (King's College London)
Dr David Redding (Zoological Society of London)
Dr Tara Vishwanath (The World Bank)

This document was prepared under the Chatham House rule by Ms Audrey Tan and Dr Lele Rangaka. Please get in touch if you would like to contribute to these discussions.

The AI for People & Planet: Emerging Epidemics Roundtable is supported by:

UCL PUBLIC POLICY

UCL INNOVATION & ENTERPRISE

UCL INSTITUTE FOR GLOBAL HEALTH

GRAND CHALLENGES