# Perspective

# Artificial intelligence for modelling infectious disease epidemics

Moritz U. G. Kraemer[1,2,36,37 ✉], Joseph L.-H. Tsui[1,2,36], Serina Y. Chang[3,4,36], Spyros Lytras[5], Mark P. Khurana[6], Samantha Vanderslott[1,7], Sumali Bajaj[2], Neil Scheidwasser[6], Jacob Liam Curran-Sebastian[6], Elizaveta Semenova[8], Mengyan Zhang[9], H. Juliette T. Unwin[10], Oliver J. Watson[11], Cathal Mills[1,12], Abhishek Dasgupta[1,13], Luca Ferretti[1], Samuel V. Scarpino[14,15], Etien Koua[16], Oliver Morgan[17], Houriiyah Tegally[18], Ulrich Paquet[19], Loukas Moutsianas[20], Christophe Fraser[1], Neil M. Ferguson[11], Eric J. Topol[21], David A. Duchêne[6], Tanja Stadler[22,23], Patricia Kingori[24], Michael J. Parker[1,24], Francesca Dominici[25], Nigel Shadbolt[9,26], Marc A. Suchard[27], Oliver Ratmann[28,29], Seth Flaxman[9], Edward C. Holmes[30], Manuel Gomez-Rodriguez[31], Bernhard Schölkopf[32], Christl A. Donnelly[1,12], Oliver G. Pybus[1,2,33,37], Simon Cauchemez[34,37] & Samir Bhatt[6,11,35,36,37 ✉]

Infectious disease threats to individual and public health are numerous, varied and frequently unexpected. Artificial intelligence (AI) and related technologies, which are already supporting human decision making in economics, medicine and social science, have the potential to transform the scope and power of infectious disease epidemiology. Here we consider the application to infectious disease modelling of AI systems that combine machine learning, computational statistics, information retrieval and data science. We first outline how recent advances in AI can accelerate breakthroughs in answering key epidemiological questions and we discuss specific AI methods that can be applied to routinely collected infectious disease surveillance data. Second, we elaborate on the social context of AI for infectious disease epidemiology, including issues such as explainability, safety, accountability and ethics. Finally, we summarize some limitations of AI applications in this field and provide recommendations for how infectious disease epidemiology can harness most effectively current and future developments in AI.

AI[1] is transforming many aspects of contemporary science[2] and has the potential to similarly change the landscape of infectious disease epidemiology. AI can be defined as intelligent behaviour exhibited by machines and computers and has been an active area of research since the 1950s[3]. Over the past decade, the focus of AI methods has shifted substantially from logic-based approaches[4] to those associated with deep learning[5]. In this Perspective, we define AI and related data science approaches broadly and therefore include methods from machine learning (ML)[6], probability theory[7], numerical optimization[8] and new directions in scalable computation[9,10].

Infectious disease epidemiology is the study of why infectious diseases emerge, how they transmit within and among populations, and of the strategies that can be used to prevent, control and mitigate their spread[11]. Mathematical, computational and statistical modelling is an essential component of this interdisciplinary field, and quantitative models are used to inform public health policies and responses at local and global scales[11]. Although much attention has been paid to the application of AI to problems in human health, such as patient diagnosis[12], individual-level disease risk prediction[13] and decision support for doctors[14], there have been fewer demonstrations of the

[1]Pandemic Sciences Institute, University of Oxford, Oxford, UK. [2]Department of Biology, University of Oxford, Oxford, UK. [3]Department of Electrical Engineering and Computer Science, University of California Berkeley, Berkeley, CA, USA. [4]UCSF UC Berkeley Joint Program in Computational Precision Health, Berkeley, CA, USA. [5]Division of Systems Virology, Department of Microbiology and Immunology, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [6]Section of Epidemiology, Department of Public Health, University of Copenhagen, Copenhagen, Denmark. [7]Oxford Vaccine Group, University of Oxford and NIHR Oxford Biomedical Research Centre, Oxford, UK. [8]Department of Epidemiology and Biostatistics, Imperial College London, London, UK. [9]Department of Computer Science, University of Oxford, Oxford, UK. [10]School of Mathematics, University of Bristol, Bristol, UK. [11]MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, UK. [12]Department of Statistics, University of Oxford, Oxford, UK. [13]Doctoral Training Centre, University of Oxford, Oxford, UK. [14]Institute for Experiential AI, Northeastern University, Boston, MA, USA. [15]Santa Fe Institute, Santa Fe, NM, USA. [16]World Health Organization Regional Office for Africa, Brazzaville, Congo. [17]WHO Hub for Pandemic and Epidemic Intelligence, Health Emergencies Programme, World Health Organization, Berlin, Germany. [18]Centre for Epidemic Response and Innovation (CERI), School for Data Science and Computational Thinking, Stellenbosch University, Stellenbosch, South Africa. [19]African Institute for Mathematical Sciences (AIMS) South Africa, Muizenberg, Cape Town, South Africa. [20]Genomics England, London, UK. [21]Scripps Research, La Jolla, CA, USA. [22]Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland. [23]Swiss Institute of Bioinformatics, Lausanne, Switzerland. [24]The Ethox Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK. [25]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [26]The Open Data Institute, London, UK. [27]Department of Biostatistics, University of California, Los Angeles, Los Angeles, USA. [28]Department of Mathematics, Imperial College London, London, UK. [29]Imperial-X, Imperial College, London, UK. [30]School of Medical Sciences, The University of Sydney, Sydney, New South Wales, Australia. [31]Max Planck Institute for Software Systems, Kaiserslautern, Germany. [32]Max Planck Institute for Intelligent Systems and ELLIS Institute Tübingen, Tübingen, Germany. [33]Department of Pathobiology and Population Sciences, The Royal Veterinary College, London, UK. [34]Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Université Paris Cité, U1332 INSERM, UMR2000 CNRS, Paris, France. [35]Pioneer Centre for Artificial Intelligence University of Copenhagen, Copenhagen, Denmark. [36]These authors contributed equally: Moritz U. G. Kraemer, Joseph L.-H. Tsui, Serina Y. Chang, Samir Bhatt. [37]These authors jointly supervised this work: Moritz U. G. Kraemer, Oliver G. Pybus, Simon Cauchemez, Samir Bhatt. ✉e-mail: moritz.kraemer@biology.ox.ac.uk; samir.bhatt@sund.ku.dk

## Box 1

# Glossary

**Active learning** is a ML framework designed to minimize data labelling efforts; the model iteratively asks a human user to label relevant samples and the model learns from the replies.

**Bayesian optimization** is used to numerically optimize a function that is difficult to optimize or evaluate using conventional approaches.

**Causal inference** refers to a diverse suite of statistical methods that investigate and exploit the cause–effect relationships that underlie observed associations.

**Explainable AI** is a set of methods that enable human users to comprehend the output created by 'black box' ML algorithms (that is, algorithms and models of which the internal workings are not easily accessible or interpretable).

**Deep generative models** are ML models that use deep neural networks to generate/simulate new data that are similar to the training data.

**Fine-tuning** refers to the training of a previously fitted model (that is, a pre-trained model) on a new task. It is a type of transfer learning and is typically used in cases in which data are scarce.

**Foundation models** refers to very large deep learning models that are pre-trained on vast, general-purpose, datasets—for example, LLMs.

**Graph neural networks** are a family of artificial neural networks for processing graph-structured data (for example, social networks or protein structures).

**Mechanistic epidemic models** are models in which a plausible mechanism for the infection process is represented in an analytical or simulation framework (for example, using differential or difference equations).

**Multimodal learning** refers to approaches that attempt to leverage multiple types of data, such as images and language.

**Recurrent neural networks** are a family of artificial neural networks designed for processing temporal, sequential data. They have been used extensively for problems in speech, natural language and genetics.

**Reinforcement learning** is a ML paradigm in which an agent learns how to take actions conditioned on its state in an environment, so as to maximize a reward signal. With human feedback, these systems can be aligned to match human preferences.

**Retrieval-augmented generation** is a technique to assist LLMs for text generation.

**Self-supervised learning** is a ML method in which a model learns from data without labels for guidance (compare with active learning).

**Surrogate model (or emulation)** is a model used to approximate another, typically more complex or computationally expensive, model.

**Transfer learning** is a process whereby the knowledge acquired by a model from an initial task is leveraged for a later task.

**Transformers** are neural network architectures that process sequences (such as text) in parallel and compute the relationships of each part of the sequence to others, enabling an understanding of the context of that sequence.

**Zero-shot learning** refers to the ability of an AI/ML model to generalize to a new task or problem that is not represented in the dataset used to train the model.

---

use of AI models for infectious disease epidemiology—mirroring the lower uptake of AI methods in population health more generally, including for non-communicable diseases[15]. Yet, effective and sustainable improvements to public health must address determinants of health at both the individual and population levels, as well as the interactions between them[16]. New AI approaches that span public health and individual outcomes have the potential to transform infectious disease responses and enable more targeted, equitable and robust interventions.

One possible reason for the relatively lower penetrance of AI in infectious disease epidemiology is the challenge of acquiring large-scale, standardized and representative data[17] for the training and evaluation of AI/ML models with many parameters. However, new AI approaches perform increasingly well with limited data. For example, approaches based on fine-tuning or transfer learning[18] (Box 1) of large pretrained networks can achieve strong performance without the need for months of initial training or terabytes of data, and recent advances in self-supervised learning[19] can facilitate further zero-shot learning on a range of epidemiological questions (Box 2).

We consider two main topics in this perspective: first, we outline how current and anticipated advances in AI can change how we answer key questions in infectious disease modelling. For example, new AI methods can improve the collection and integration of heterogeneous data sources, and AI/ML could be included in policy and decision-making frameworks to improve population health. Second, we discuss some social and ethical contexts of the application of AI to infectious disease epidemiology; we consider the explainability, safety and accountability of AI in public health, and provide recommendations for how infectious disease epidemiology can harness developments in AI effectively.

## AI to answer epidemiological questions

One of the key challenges during the early stages of an epidemic is to understand the severity and epidemic potential of the infectious agent. This involves estimating fatality ratios, serial and generation intervals, transmissibility and epidemic growth rates, as well as inferring infection networks and transmission heterogeneity in different settings[20] (Box 2). In traditional epidemiology, these questions are often answered using observational data (such as case–control studies, cohort studies or household surveys)[21]. However, the small scale of such studies and the idiosyncrasies of data collection mean that the true transmission process is observed imperfectly and may not be representative of the whole epidemic. The actual chain of infection events and locations where infections occur is often uncertain (for example, individuals may visit multiple locations and contact many people, some of whom might be infectious but asymptomatic), hindering efforts to accurately measure quantities such as the incubation period or transmission intensity from observational data alone. Other issues include underreporting, censoring, truncation, non-random data omission and uneven data reporting. Here, Bayesian data augmentation[22] has proven valuable in improving parameter inference in the context of missing data, and AI approaches can help in the scalability and inference of such models[23] (Fig. 1a). A promising direction involves approximate Bayesian inference such as that based on variational inference with normalizing flows[24]. This reframes the problem from one of sampling to optimization, which can be addressed efficiently using gradient-based deep learning techniques. Use of these tools makes fast and accurate inference more achievable, especially when inferences have to be completed quickly.

Given the difficulties in generating a complete description of the transmission tree that underlies an epidemic, model-based analyses

# Important epidemiological concepts

**Fatality ratio:** the proportion of those affected by a disease who do not survive, calculated for different subpopulations (for example, for cases/reported infections or hospitalizations) and often stratified by factors such as age, time and location. The infection fatality ratio (that is, the proportion of those infected who subsequently die from that infection) is an important population-wide measure of disease severity. Given the time lags inherent in both the infection process and health reporting systems, it is not appropriate in an early, growing epidemic to estimate the infection fatality ratio as the number of deaths to date divided by the number of infections to date and parametric modelling is required[160].

**Forecasting:** the prediction of a disease's future trajectory using mathematical and/or statistical tools. Generally, trends in summary statistics such as incidence or prevalence are predicted forward in time on the basis of current and past data. Many approaches blend a mix of time-series modelling with mechanistic models of disease spread.

**Generation interval:** the duration between the infection of an individual and the infection of one of its secondary infections; this length of time varies among infector–infectee pairs and is therefore represented by a probability distribution. This value depends on factors such as the duration and timing of infectiousness and may be affected by the epidemic growth rate or local prevalence of infection[161]. Parameters such as the generation interval and the reproduction number strongly determine the growth rate of an epidemic and are used to determine the planning, timing and evaluation of public health interventions. Precise infection times are almost never directly observed; thus, the generation interval distribution is usually estimated by a combination of modelling and partial observation.

**Infection networks:** the spread of a contagion through a population, represented as a weighted, directed network between infectors and infectees[162]. Although true, complete infection networks are rarely observed directly, they can be estimated through a mixture of epidemiological and contact pattern data. Key epidemiological parameters can then be estimated using generation intervals and other data, such as pathogen genome sequences or demographic information.

**Nowcast:** the reconstruction of the current state of the epidemic as accurately as possible, when contemporaneous data may be affected by reporting delays. Even in the absence of delays, integrating information from heterogeneous sources requires data synthesis. Nowcasting uses statistical models to debias and synthesize multiple, noisy data sources to present a current picture of an epidemic.

**Reproduction number:** the average number of secondary cases arising from a single case in a population, also known as the $R$ value. If estimated for a situation in which all individuals are susceptible to infection, this is termed the basic reproduction number, $R_0$. The effective, or time-varying reproduction number, $R_t$, relaxes this assumption and considers the average number of cases generated in the current state of the population. The reproduction number is typically estimated from data using methods including regression, mechanistic models or networks[163].

**Transmissibility:** the propensity of a pathogen to be transmitted between individuals, which is a combination of pathogen-related factors relating to the pathogen (for example, the size of the inoculum and the pathogen's ability to evade host immune responses) and host-related factors (for example, the number and intensity of contacts a host makes). Capturing change in transmissibility over time is important for planning interventions.

**Vaccine effectiveness:** a measure of how well vaccination protects against infection, transmission and/or symptomatic and severe disease, typically measured as one minus the relative risk of the outcome in vaccinated versus unvaccinated persons. When measured in a randomized trial, the term 'vaccine efficacy' is often used.

---

are performed on routinely collected aggregated data, such as counts of cases, hospitalizations or deaths. Mechanistic and semi-mechanistic disease transmission models are commonly used to test hypotheses and estimate key epidemiological parameters (for example, transmissibility and virulence) with associated uncertainty[25]. Although these models provide substantial mechanistic insights into transmission and can be used to construct counterfactual scenarios, they often come with considerable computational costs due to the complexity of the numerical methods and inference in a high-dimensional parameter spaces[26]. Recent advances in deep generative modelling (Box 1) have shown that inference can be accelerated by approximating the original model with a generative model surrogate[27,28] or by using variational inference[29] (Box 1 and Table 1). Improvements in the speed of model computation offers additional opportunities to increase the model complexity and realism, including potential linkages between individual transmission heterogeneity and population-level outcomes[30]. As AI-accelerated methods have the potential to reduce model run times from weeks to hours, they could substantially speed up the iterative process of informing policy decisions and model refinement (Table 1).

## Prediction and forecasting

During an infectious disease epidemic, policy decisions regarding the deployment and timing of interventions rely on reliable estimates of the number of current cases and forecasts of future cases[31]. For a given disease, the number of cases, hospitalizations and deaths reported by public health surveillance systems over time are analysed to inform key policy decisions. Although epidemiological datasets may include a rich variety of demographic (such as ethnicity and age), clinical, geospatial and pathogen genomic information, at their simplest, they consist solely of quantitative trends and can therefore be analysed using a range of approaches, including statistical time-series techniques[32], modern ML methods and advanced deep learning models[5] including newer foundation models (Table 1).

However, surveillance data cannot tell us directly how many individuals are infectious at any given day, or what the expected future trajectory of cases might be. Epidemic surveillance data are almost always noisy and affected by biases in reporting (for example, time-variable delays), testing (for example, targeting of limited capacity or willingness to test in certain groups) and sampling (for example, towards groups that can more readily access healthcare)[33,34]. For example, data may appear to suggest a declining epidemic, but this trend could be an artefact of time-dependent changes in reporting or declining testing availability[35]. These data issues could lead to ineffective and delayed responses.

Current approaches to nowcasting[35–37] attempt to address these systematic biases and enable robust estimation of the contemporary epidemic situation[38]. However, choosing an optimal nowcasting model, which can both generalize to new data and accurately capture
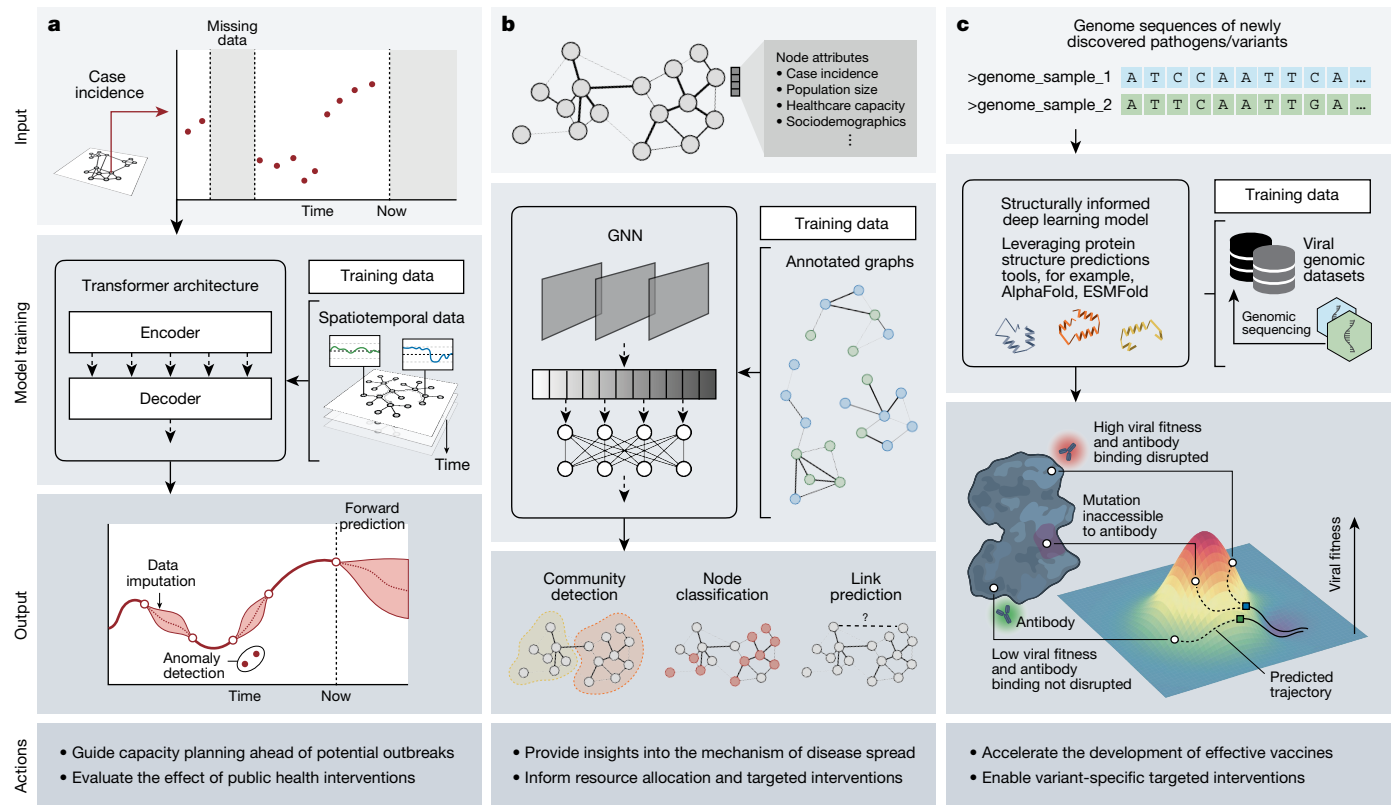
**Fig. 1 | AI approaches to tackling key epidemiological questions. a**, Time-series analysis using foundation models. Pre-training large-scale foundation models with transformer architecture using observations (empirical and/or simulated) from time-series epidemic data enables zero-shot applications (Box 1) for forward prediction, data imputation and anomaly detection. The output from these models can be used to evaluate the effect of public interventions and guide capacity planning for future outbreaks. **b**, Modelling infectious disease spread using GNNs. Pathogen dissemination can be represented by annotated graphs, in which the nodes correspond to locations or individuals and the edges represent potential transmission pathways (such as human or vector interactions); each node is associated with a set of features (for example, case incidence and population size) that are either indicators or drivers of spread. GNNs have the ability to learn complex patterns from such

data, enabling node classification (predicting disease prevalence), community detection (identifying infection clusters) and link prediction (revealing cryptic transmission pathways). These insights can provide a detailed understanding of the underlying mechanism of disease spread and inform resource allocation and targeted interventions. **c**, Predicting immune-escape mutations using biologically informed deep learning models. By taking advantage of recent advances in protein structure prediction (such as AlphaFold and ESMFold), these models could enable the early detection of pathogens or variants that are likely to develop mutations conferring resistance to existing vaccines or therapeutics through antibody-binding disruption. Such predictions can inform the design of next-generation vaccines and guide the prioritization of containment strategies targeted at emerging lineages.

complex patterns, is challenging, especially from noisy and incomplete data. The simplest solution is to choose a model from a large, complex class of models and then try to mitigate any potential overfitting or over-parameterization (Fig. 1a). Recent work in which nowcasting and forecasting are improved with ensemble methods shows promising results and provides successful routes to create model diversity[39,40]. During the COVID-19 pandemic, multiple research groups undertook and evaluated epidemic forecasts, accelerating progress towards more standardized and rigorous models for informing public-health decision making.

Foundation models from large deep neural networks[41] represent new and powerful ways to explore and understand time-series surveillance data and go beyond current approaches. However, these foundation models need to be pretrained on hundreds of billions of datapoints from a diverse range of time series (across a range of domains); consequently, the patterns that they learn are effective at zero-shot[42] (Box 1) performance on a range of new surveillance data. The models restrict the fitted functions to plausible shapes, helping to accurately quantify uncertainty[42]. Historically, AI models were unable to include rigorous measures of uncertainty, but recent work in deep learning shows promise in using activation functions[43], ensembles[44] and conformal prediction[45] for the quantification of uncertainty.

## Modelling mechanisms of disease spread

Infectious disease modelling has relied traditionally on mechanistic models that represent the disease transmission process as a simplified but definable process. Examples include mathematical susceptible–infected–recovered (SIR) equations and individual/agent-based models[46], in which a simulation of transmission is constructed and each agent is assigned specific attributes and events. When fitted correctly to rich data, individual-based models generate insights into the transmission process and can be used to evaluate interventions. However, parameter inference is much harder to perform using agent-based models than for standard statistical techniques owing to their lack of analytical components[29]. Here, AI approaches are providing new ways to facilitate efficient inference[29] through a range of innovative approximations that use deep neural networks. For example, by taking advantage of variational Bayesian inference, new neural network architectures can accommodate large and dynamic agent-based models[47]. Without AI approaches, inference on very large agent-based models is challenging as the parameter spaces are too large, and generating samples or realizations from such models can be time consuming. An alternative solution to computational intractability is provided by gradient-based methods such as variational inference, as discussed above.

**Table 1 | New AI approaches to answering key questions in infectious disease epidemiology, their potential effects and the level of maturity in the scientific and public health community**

| Problem domain | Method | Potential effects | Level of maturity |
|---|---|---|---|
| Inference of epidemiological parameters | Generative Bayesian and surrogate models[36] | Incremental improvements in the speed and accuracy inferring epidemiological estimates, especially when data are missing. Faster generation and iteration of model results for policy. | Bayesian methods can be implemented easily in available software packages. Surrogate models need to be developed for specific applications[153]. |
| Epidemic forecasting/ nowcasting | Time-series foundation models[154] and ensemble techniques[39] | Potential for improved speed and accuracy in estimating future trajectory of cases. Better generalization of trends for medium term forecasts | Ensemble forecasting is standard practice in influenza and COVID-19 modelling. Several foundation time-series models exist and are easy to access but have yet to be implemented. |
| Scenario modelling | Compartmental, mechanistic models[155] and RL agent-based models[156] | AI's impact may be limited because mechanistic models better capture the impact of interventions. AI models can be used for RL[156] for de novo discovery of individual behavioural adjustments and policy options, and to speed up the generation of realistic future scenarios. | RL neural network architectures exist, but bespoke versions need to be created for specific applications in infectious disease modelling. |
| Understanding mechanisms of disease spread | GNNs[94] and graph foundation models[60] | The ability to incorporate large amounts of structured data with the potential for improved inference of disease transmission routes. | Early research outputs |
| Infectious disease surveillance | Active learning[157] and Bayesian optimization | Improved cost-effectiveness through the focus on collection of the most valuable data. | Early research with limited implementation to date |
| Risk prediction | Multimodal AI[158] | Improved inference of risk factors of severe disease from multiple data types. | Early research with potential applications in population health |
| Analysing pathogen genomes | Protein language models[68,71] | New applications are possible for improved design of vaccines and therapeutics, and for anticipating and evaluating pathogen variants with new phenotypes (for example, immune evasion). | Early research outputs |
| Public health decision making | Markov decision processes and RL[159] | Improved ability to learn optimal decisions from human–machine interactions. | Conceptual, theoretical |

Dynamical 'compartment' models are one of the most commonly used analytical tools in mathematical epidemiology, the best known of which is the SIR model, which partitions the population into a small number of groups (compartments). SIR dynamical models are comparatively easy to fit to data but often lack the detail necessary to bridge the gap between individual behaviour and population-level dynamics. Alternative models based on branching processes[48], partial differential equations or self-excitatory Hawkes processes[49] are growing in popularity and there are strong similarities between these approaches[48,50]. As previously noted, foundation time-series models can be used to constrain function space by restricting the set of functions a priori and embedding these constraints within a chosen mechanistic model (termed semi-mechanistic models). The overall objective is to base modelling on plausible epidemiological realities. However, model misspecification can still introduce errors. Using these approaches, AI models have the potential to not only achieve high predictive accuracy within current mechanistic frameworks, but also to learn more complex hidden mechanisms that can improve existing models.

Graph neural networks (GNNs) that operate on discrete-structured data[51] (Fig. 1b) present a particularly promising type of AI for detailed understanding and forecasting of infectious disease dynamics[52]. Graphs (or networks) emerge naturally in many areas of infectious disease epidemiology, including contact networks of the spread of a disease through individuals and populations[53], phylogenetic trees and networks to track pathogen evolution, and social and information networks to understand health behaviours (for example, masking and vaccine uptake), the spread of information and misinformation and social influence[54]. Through their ability to represent rich relationships, GNNs have recently been shown to accurately predict COVID-19 cases per region[55], forecast influenza-like illness[56] and nowcast vaccine uptake from online information networks[57]. GNNs also show promise in expanding dynamical compartmental models to capture complex spatiotemporal dynamics among discrete geographical regions[58]. GNNs can expand agent-based models by enabling differentiability of all model components, facilitating joint inference over complex models[59]. Finally, future graph-based foundation models may be able to learn transferrable representations of graphs that generalize to any unseen graph[60], therefore potentially enabling the transfer of knowledge from data-rich to data-poor settings. Many important discrete problems on graphs are NP-hard (non-deterministic polynomial time−finding a solution is impossible given current knowledge) and need to be solved heuristically. AI approaches through GNNs offer powerful continuous optimization tools such as gradient descent, which can improve the solving of discrete problems on graphs.

## Immunological and genomic interactions

Disease preparedness and response requires the timely detection of emerging zoonotic pathogens or new variants of known pathogens. Even though genomic surveillance is essential for monitoring the circulating genetic diversity, determining the epidemiological and disease phenotype of the pathogens identified usually requires time- and resource-consuming experimental work. Advances in AI models change this paradigm by using genetic sequences as their input, capturing information about pathogen proteins[61,62]. These can pave the way for high-throughput prediction of protein structures and pathogen phenotypes from genomic data, enabling rapid preliminary assessment of new phenotypes and ultimately reducing the amount of experimental work needed to obtain good predictive performance[63] (Fig. 1c). However, predicting from small labelled datasets to large unlabelled datasets can be error prone, and recent cross-prediction approaches show great promise in improving downstream inferences[64]. AI models applied to genomic data can also be used to classify virus lineages[65], infer when, where and how a pathogen emerged[66], predict the pathogen traits such as transmissibility[67], escapability and spread[68], and predict host specificity to identify probable cross-species spillovers[69,70]. AI models can also enhance the accuracy of phylogenetic inference, enabling a more precise characterization of the infection process.

# Perspective

In addition to predicting complex and multifaceted phenotypes from genetic sequences, AI approaches can help to infer the evolutionary trajectory of pathogens that are already circulating in a host population and can potentially anticipate escape variants[71]. For example, human respiratory viruses that cause great global health burden, such as SARS-CoV-2 and influenza viruses, accumulate mutations continuously owing to replication errors that, through natural selection, enable them to evade host antibody immunity mounted against previous virus variants[72] (Fig. 1c). AI-driven variant forecasting can inform the development of targeted vaccines, monoclonal antibodies and drugs targeting future pathogen strains[73,74] could be paired with recent advances in accurately predicting protein structures as monomers and in complex with other molecules[75]. However, the variant that will dominate global or local spread will not only depend on predictions of pathogen evolution, but also on underlying dynamical processes influenced by human behaviour, cross-immunity, climatic conditions and previous exposure to other, related pathogens[76]. Failing to account for the complexity of these interacting processes could lead to false predictions from AI models alone. For example, recent work showed that epidemiological features predict SARS-CoV-2 variant spread better than those based on evolutionary and immunological drivers[77].

Sequence-based AI methods also enable the exploration of as-yet unseen pathogen diversity. The recent metagenomic and metatranscriptomic discovery of RNA viruses has exponentially expanded our knowledge of the diversity of viruses found in animals, some of which could potentially spillover into humans and cause disease, or interact with other co-infecting pathogens[78]. The default tool for studying this diversity is molecular phylogenetics, which comprises a diverse field[79] of methods spanning the biological, statistical and computational sciences. Variational Bayesian inference empowered through probabilistic programming could deliver a step change in the scale of phylogenetic tree estimation possible from molecular sequences[80]. Furthermore, ML/AI models may be used to improve predictions of spillover potential between reservoirs and human hosts using pathogen genomic data[78].

## AI-aided integration of key data

### Data science and representative sampling

Despite a step-change since the COVID-19 pandemic in the ability to perform large-scale, interdisciplinary data collection during disease outbreaks, data scarcity and geographical biases remain major obstacles in global epidemiological research[81] (Fig. 2). These have been shown to exacerbate racial biases[82] and lead to non-optimal public health policy decisions. In particular, the uneven and inequitable distribution among locations of diagnostic testing resources (including genotyping)[83] results in biased sampling, limiting the representativeness and use of the data collected[84]. It is not only the volume of data that necessarily provides the most insights, but also the quality of them (for example, sampling design and representativeness); this big data paradox[85] has important implications for our ability to infer key epidemiological parameters, to make accurate forecasts and to respond to emerging threats[86].

As the global community continues to scale up infrastructure for infectious disease data collection and integration of data types that help to understand the drivers of transmission (for example, socio-economic, census and satellite-derived climate/environmental data; Fig. 2), AI might be used in tackling problems such as improving the cost-effectiveness of surveillance and mitigating uneven representation in surveillance data. For example, active learning[87] and Bayesian optimization[88] (Box 1 and Table 1) can be used to inform the design of adaptive disease surveillance strategies (such as selecting the next set of locations, aircrafts or individuals to test based on current observations) that are tailored to answer specific epidemiological questions[89]. They are particularly useful when the costs of data collection and processing are high, as is the case of genomic surveillance, for example. However,
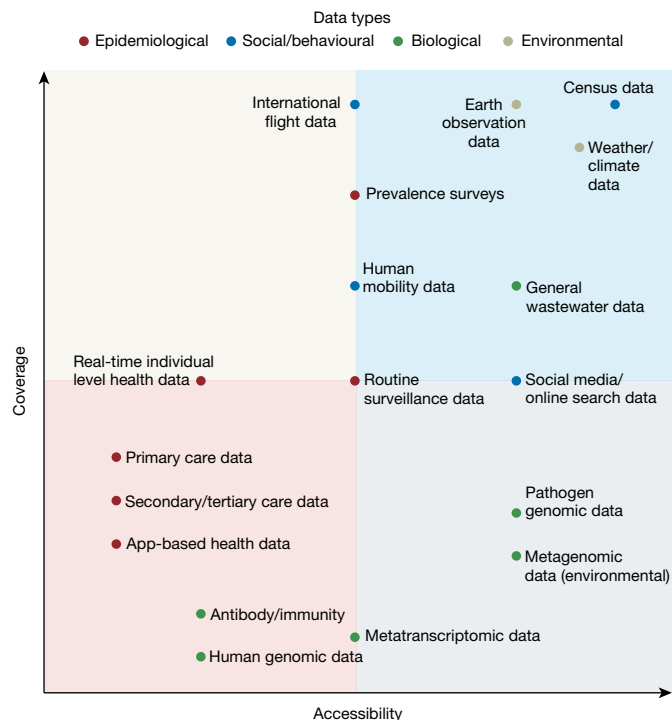


**Fig. 2 | Classification of data types to investigate infectious diseases.** Data used to inform epidemiological modelling are placed in terms of their accessibility to the research community and their population-level coverage. Data are classified based on their data types, including epidemiological (red), environmental (brown), social–behavioural (blue) and biological (green).

implementing an adaptive surveillance system for outbreak detection, tracking and monitoring is challenged by the logistics and time scales for data collection, processing and translation into actionable insights[89]. Tightly linking computational and practical approaches to disease surveillance will be necessary if AI models are to offer tangible benefits. With openly accessible datasets on the drivers of transmission, it might become possible using AI to predict disease distributions without access to highly resolved epidemiological data.

### Integrating new data sources

A plethora of new data sources and digital tools is creating new approaches to the modelling of infectious disease epidemics. During the COVID-19 pandemic in the UK, the National Health Service COVID-19 app was downloaded to >21 million mobile phone devices, providing users with timely notifications of infectious disease exposures and instructions in case their contact was diagnosed positive[90]. Such mobile applications are a major new source of individual-level behavioural data for understanding and informing public health interventions. Research has shown how digital contact tracing apps could transform our ability to reduce transmission at a fraction of the cost of manual contact tracing[91], especially when pathogens transmit rapidly and only a minority of infected individuals show early symptoms. Furthermore, contact-tracing apps could provide new, real-time data about the variation in transmissibility as a function of proximity and duration of contact, and how this variation changes through time and among pathogen variants[92].

More generally, it is now feasible to aggregate individual-level mobility information from mobile phones, and this new source of data can substantially enhance the precision and use of mathematical models of infectious disease[93,94]. For example, such data are used to characterize population-level trends in human movement during emergencies[95], and to understand how frequently individual venues are visited by

individuals. During the COVID-19 pandemic, a variety of sources of smartphone-derived mobility data were used to monitor population behaviours and how they shifted in response to public health interventions or perceived risk. Data governance and privacy-preserving aggregation of location signals can ensure that individuals are not identifiable, but challenges remain in ensuring that sensitive data are not leaked or used unintentionally to reinforce existing biases. Although mathematical descriptions of how pathogens spread through networks of susceptible, infected and recovered individuals are well established, further work is needed to integrate these epidemiological principles with new data sources and move beyond describing patterns in big datasets, especially when the contact networks that underlie transmission are transient and highly structured[86]. GNNs (see the section above) provide an approach to the linkage of high-precision epidemiological models with new data sources.

Integrating individual- and population-level indicators of disease activity using combinations of wastewater surveillance, population surveys, syndromic and digital surveillance, body-worn physiological sensor data (such as heart rate, step counts and sleep) and contact network information could enhance our ability to detect outbreaks, monitor disease spread and follow the long term sequelae of outbreaks[96]. For example, sensory-based surveillance data can detect early signs of long COVID from sleep and physical activity patterns[97]. Reinforcement learning (RL) systems (Box 1) have been used to support testing decisions at airports when limited testing was available[98]. There is the potential to apply multimodal AI to images, text, speech and other data to interrogate and predict outcomes at the individual and population level[99].

In addition to consensus genome sequences of sampled pathogens, the genetic diversity captured in metagenomic or deep-sequencing raw data can provide insights into disease spread and pathogen spillover potential[100]. Modern tools based on AI/ML algorithms harness the frequency with which mutations are introduced into pathogen genomes to predict the timing of infection, identify nucleotide signatures that characterize epidemiologically linked individuals or resolve directionality in transmission events[101]. During the COVID-19 pandemic, pathogen genome datasets grew in size substantially, from a few thousand genomes in previous epidemics to tens of millions of genomes. As a result, tools for the rapid analysis of large genomic datasets were built and are now available for future pandemics[102,103].

As the global climate changes, the distribution and incidence of climate-sensitive infectious diseases such as dengue and cholera are expanding[104] and the analysis of these dynamics could be enhanced by AI. Collection and incorporation of past, current and likely future climate data into epidemiological models can improve our understanding of key epidemic dynamics and the reliability of predicted short- and long-term epidemic trajectories. Other opportunities arise when AI is used to improve short- and medium-term weather forecasts; recent research has demonstrated the potential for better climate modelling using physics-based models that incorporate aspects of AI[105]. This will improve the evaluation of the effect of climate on infectious diseases, which results from changes in host and vector species distributions, or from changes in human behaviour and contact patterns.

## AI-driven disease modelling for policy
### Public health decision making
During international health crises, such as the emergence of a novel pathogenic infectious disease, countries face complex decision-making challenges across sectors (for example, health, economics, education and social well-being) to contain and mitigate disease spread. International travel and urbanization mean that today's epidemics occur in tightly connected populations and can lead to pandemics within weeks of their emergence[106]. Decision makers must assess the cost and benefits of policies that are implemented to protect public health by maximizing their efficacy while minimizing negative and unintended consequences, all under great uncertainty. AI approaches that focus on developing generalized systems can potentially provide decision support[107]. Consequently, the mechanics of how public health policy decision making is done and how data and model results are interpreted may fundamentally change with the adoption of advanced AI.

It is common for there to be an interaction between infectious disease modellers and policymakers[31], with questions and feedback from policymakers being communicated to the modelling community, sometimes coordinated through scientific advisory committees. In these interactions, time has a critical role: decision-makers are often presented with large amounts of data and results that may be inadequately summarized at the appropriate level and that are not always directly relevant to the decision at hand. Meanwhile, modellers face the challenge of developing increasingly complex models to address a wide range of questions, for which characterizing a clear objective or task can be sometimes difficult (Fig. 3). Owing to the exponential nature of epidemic growth and consequent risks, decisions often need to be taken urgently, based on the best possible evidence available at the time. AI is likely to be a critical tool for improving the speed of modelling for informing policy making. New ML and AI approaches can reduce greatly the runtime of epidemiological models, enabling a more thorough exploration of scenarios and their statistical uncertainties. Furthermore, large language models (LLMs) could provide bespoke summaries of complex quantitative models that are tailored to a decision maker's preferences; not everyone likes graphs[108].

Markov decision processes and RL[109] are promising theoretical frameworks in which new data types, prediction models and expert feedback can be integrated to enable rational and more timely decision-making during epidemics (Fig. 3). This has become possible due to advances in the speed at which models can be run and updated (see the 'Modelling mechanisms of disease spread' section). Even though RL frameworks are attractive for tackling multicountry disease outbreaks, they require clear definitions of objectives and rewards, including enumeration of the costs and benefits of different policy actions (RL is particularly useful in robotics and strategic multiplayer games). Incorporating public health objectives, which often change and are less likely to be quantifiable into flexible models that can incorporate human feedback, might provide attractive opportunities for future science–society–policy interactions. Close collaboration between decision makers, civil society and the scientific community is therefore necessary to advance sociotechnical decision-making frameworks capable of commanding well-founded public trust and confidence for epidemiology, in advance of a future global pandemic[110].

It remains an open question whether today's RL models (or other algorithms for decision-making) will be able to reconstruct and predict outbreak trajectories with sufficient certainty to be practicable, and the answer will be contingent on the quality and timeliness of available data. Unintended and often unanticipated consequences from interventions, such as the evolution of drug resistance[111] or behavioural changes in response to perceived high infection risks, have yet to be incorporated systematically into epidemic models ahead of their deployment in decision frameworks[112]. Challenges remain in evaluating the costs of interventions, which are deployed in combinations and frequently change through time or among locations. Incorporating this heterogeneity requires better methods in AI and causal inference, including emulated trials, as well as the use of hybrid models that integrate observational and causal information with simulations, to form causal 'digital twins' that permit counterfactual questions to be answered[113].

### Ethics in the development of AI tools
There are strong ethical reasons to support scientific efforts to explore how AI might improve infectious disease prevention and control efforts. Such improvements could save lives and reduce the burden of infectious disease. However, the successful and appropriate deployment of AI tools will depend crucially on the integrated identification, careful
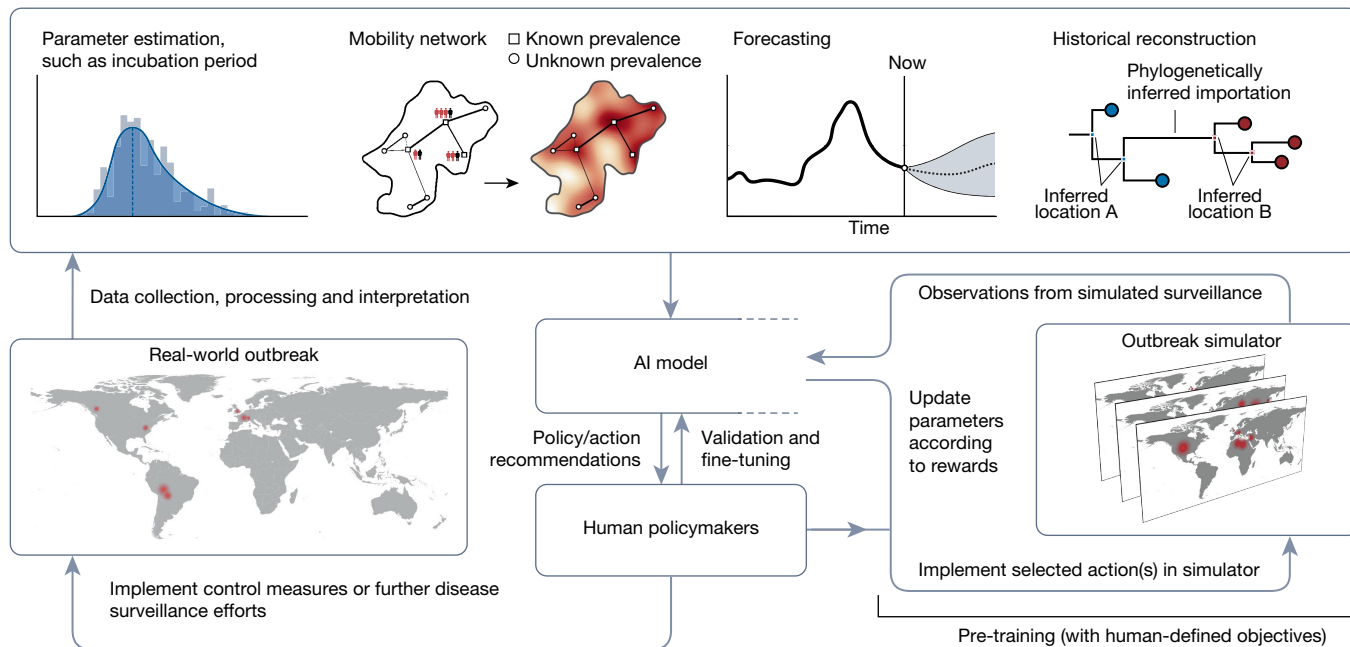
**Fig. 3 | Iterative approach to public health decision making.** Conjectured use of AI to optimize the design and implementation of effective control measures during a hypothetical multicountry disease outbreak. Data collected from disease surveillance are processed and analysed by an AI agent, potentially integrating other AI models for parameter estimation, nowcasting and forecasting of epidemic trajectories, as well as reconstructing historical transmission events based on current observations (top row). Before deployment of the AI model, it is trained using an outbreak simulator that simulates the spread of infectious diseases and the effects of different control measures; the model learns from these simulations by updating its model parameters according to the outcomes of the simulations and corresponding rewards (bottom right). Insights from the empirical data are used by the trained model to inform and recommend the most effective control measures, with validation and feedback from human policymakers and stakeholders to fine-tune alignment between model objectives and societally beneficial criteria. This is followed by the execution of public health actions (for example, whether to implement recommended or alternative control measures, or deploy further disease surveillance efforts) made ultimately by human policymakers (bottom left). The map shapefile was obtained from the Africa Data Hub under a Creative Commons licence CC BY 4.0.

analysis and resolution of key ethical challenges. Some of these are common to the application of AI in all fields and contexts, such as concerns regarding transparency, accountability, fairness, privacy and avoidance of deepening of existing inequities[114].

Others will be more specific to the use of AI tools for infectious disease in the context of pandemic preparedness and prevention. These include the importance of equitable and fair practices for the collection, storage and sharing of data, with a focus on protecting the interests of disadvantaged communities that have been historically affected disproportionately by infectious diseases, such as those in low- and middle-income countries. Some of these ethical concerns are not specific to AI, but related to the conduct of the infectious disease surveillance on which it depends, and its likely effect on lower-income and marginalized communities[115]. An important question in this regard will be what is, or should be, owed to communities that experience public health interventions with a potential benefit to public health globally, but that may cause harmful effects locally (for example, privacy infringements and cordons sanitaires). One key ethical consideration, whether AI tools are being shared equitably for use by public health authorities, is likely to raise broader questions of capacity building. Making AI meaningfully accessible to all who need it will require the development and sharing of expertise through collaboration and the access to computational resources. An important and interesting ethical question in the context of the opportunities outlined in this paper might be whether the use of AI in public health policy could potentially be used to better minimize the effects of interventions on disadvantaged communities.

A second broad class of ethical questions concerns the ways in which AI tools are to be deployed in the design and implementation of public health policy. An important lesson from COVID-19 was that all policy decisions implicitly or explicitly embody value judgements that have a moral component, for example, the questions of who should get vaccines first; under what conditions and to what extent personal liberties can be justifiably impinged on to reduce transmission; and the level of privacy that should be maintained in the use of digital contact tracing[116]. It is vital that such judgements are subject to deliberation, clearly justified and are accountable. The inclusion of AI in such processes amplifies the need for explainability, fairness and accountability in data-informed public health policy. AI must support rather than undermine local decision-making and autonomy in public health and no technical solution will be able to fully quantify the trade-offs between the potential harms and benefits of interventions. It is vital as a component of this, and a core requirement for well-founded trust and confidence, that the development of public policy and the uses of AI in it are informed by inclusive, meaningful public engagement, as well as inclusion of public health and clinical perspectives. There will probably be an interdependence between public health deployment and on-going research and development of AI tools for new sources of data, such as environmental/wastewater surveillance. This raises important questions about the role of research as a component of effective and responsible public health and the ethics of conducting research in global health emergencies.

## Public health communication

Misinformation online (and disinformation campaigns that deceive and mislead audiences) can threaten the implementation and success of public health strategies against epidemics[117]. The WHO highlighted the importance of fighting misinformation during the COVID-19 pandemic and has outlined plans for reporting online misinformation[118,119]. AI could be used to identify and address misinformation and to empower effective public health communication[117]. Owing to the variety, volume and speed of modern online communication, spanning text,

images, audio and video, it is impractical to use traditional statistical approaches to investigate the dynamics of digital data; instead, AI tools will support the analysis and synthesis of accurate and reliable information about health from the digital domain.

Generative AI models can provide real-time insights into information relevant to public health[120] and how public sentiment changes during an epidemic, potentially measuring the likelihood of adoption or adherence to public health measures and complementing more traditional survey information. Including the public's feedback into AI models through LLM-based generative agents in individual-based epidemic models has been proposed and might enable those models to learn the public's perception of, and behavioural responses to, epidemic events[121]. Researchers have used natural language processing to analyse data about exposure to misinformation and vaccine-sceptical messages shared on social media and to provide estimates of the effect of these messages on vaccine hesitancy[122]. However, extracting the complex reasons for vaccine hesitancy from quantitative data remains challenging[57].

There are several risks in using AI to inform or support public health communication. Careful attention is needed to avoid or reduce bias, which is typically categorized into data-driven, algorithmic and human biases[123]. AI algorithms are by definition dependent on large training datasets, but biomedical datasets have historically excluded certain populations, including women and minority ethnic groups[124]. Biases can therefore lead to ethical mistakes or misunderstandings[125]. There is a risk in damaging public trust in public health information, with AI chatbots, for example, being shown to be less popular with the public for delivering health advice[126]. Another risk relates to the propensity for LLMs to produce false information, often referred to as AI hallucinations[127], but principled statistical approaches show promise in the systematic identification of these[128].

### Open data and explainability

Open biological data linked to disease outcomes have powered many recent advances in the medical sciences and have improved disease at the individual level[129,130]. For some non-communicable diseases, the creation and maintenance of open databases using countries' centralized digital health data have been supported by governments, long-term research funding and industry cooperation. Rigorous safety and ethical reviews should accompany and govern the access to and use of such centrally maintained databases. Moreover, it is important to adopt safeguarding processes and controls for the safe release of AI models trained on sensitive data.

In infectious disease modelling, the data landscape is more fragmented and harmonization of infectious disease surveillance data across countries remains a major challenge[131]. The WHO highlighted the promotion and growth of digital health and innovation in their recent digital health agenda[132] and with the adoption of digital health, new routes for the use of AI are in reach. Recent studies have shown the potential for software tools to integrate modern optical character recognition, natural language processing and LLMs to enhance the accuracy of and speed of data extraction from semi-structured/unstructured sources, such as PDF files, press releases and situation reports[133,134]. These approaches promise to improve standard epidemiological practices, such as the meta-analysis of data from the medical literature[135] and multicountry analyses that leverage structured databases. Furthermore, basic epidemiological research tasks such as systematic reviews can be time consuming, and LLMs provide a simple interface to support the collection and summarization of research articles[136].

Some large foundation models have been made open-source by technology companies and there is potential for them to be rigorously audited[137] and used as pretrained models for further fine-tuning in infectious disease epidemiology. For example, fine-tuning pipelines such as retrieval-augmented generation could make AI models more reliable and support the extraction of valuable insights from privately stored data[138].

Further discussion is needed to understand how international, sensitive and multimodal data might be paired with novel AI models in safe and distributed hardware architecture, and how AI computing can be made more environmentally sustainable. Surveillance systems need to be linked centrally or through decentralized computational infrastructures in a secure manner, and with privacy and governance structures that can accommodate complex data sharing arrangements. An example of this is federated learning, in which partial model training occurs across multiple local servers without sharing sensitive data externally, with the results iteratively merged on a central server[139,140]. However, coordinating across multiple databases can become burdensome. Moreover, there is a need to follow existing principles to ensure that data and digital assets are comparable and can easily be deployed across multiple settings (for example, the FAIR principles: findability, accessibility, interoperability and reusability[141]). Importantly, data science approaches to health data require careful ethical oversight, as recently reviewed in the context of research in Africa[142]. The successful application of AI algorithms to individual health data requires trust and a shared understanding that AI use benefits the individuals contributing their data, as well as others. The challenges of data sharing and access are not limited to health data alone but extend also to data relevant in predicting health outcomes, including data on individual behaviours (such as human mobility data) that are often owned by companies and in some cases of commercial value[143].

Lastly, algorithms should be explainable, meaning that the predictions made by the model can be interpreted by users and stakeholders alike. In contrast to mechanistic models, which are explainable by construction, foundation models are not built with interpretable inductive foundations and remain 'black boxes' to users[144]. Quantitative measures for explainable AI have been developed, and efforts have been made in the infectious disease literature[145], but no comprehensive guidelines for the field have yet been published. Promising areas for future exploration are causal inference methods in statistics and ML[146].

### Limitations of AI in ID modelling

We note at least three fundamental limitations of AI that will limit its applications to epidemic modelling. First, current models struggle with explainability, which constrains their ability to provide mechanistic insights into the transmission process and their power to generalize beyond previously observed data and scenarios. Second, AI models that are designed to support general tasks, often with a text or voice chat interface, currently do not include specific models for answering or communicating about key epidemiological questions and concepts. Integrating single task models into more general foundation models might represent an approach to developing a future general 'AI-infectious disease' assistant. Third, capable AI models are currently trained by large technology companies at huge costs. Waiting for their deployment to then fine-tune them to answer specific tasks is not sustainable, nor equitable.

## Recommendations for AI in ID epidemiology

The risks of AI have been discussed at length[107] but their potential to transform science is unquestioned. When viewed through a broad lens, AI can be seen as a continuation of the rich legacy of mathematics, statistics and computer science in the field of infectious disease epidemiology. Advances in AI have the potential to both complement existing approaches and to inspire entirely new epidemiological methods. Examples of potential breakthroughs in AI epidemiology include pretraining large and foundation models, zero-shot and few-shot learning approaches that require minimal data, innovative new data architectures, and computational environments to run and disseminate these models efficiently. In part, these breakthroughs will be enabled by

# Perspective

the embedding of fundamental biological knowledge and constraints within AI models[71], which echoes developments in the application of AI to other fields[61,147].

However, the benefits of AI for public health will be critically dependent on the availability, accessibility and representativeness of data and a firm ethical framework. Even though more data have become available during health emergencies, especially for COVID-19, routine surveillance data for infectious diseases often remains siloed and inaccessible to the broader community, prohibiting the use of these data for improved disease modelling. However, the substantial volume of existing data means that advances aided by AI will probably be generated from existing openly available datasets. More broadly, the quality standard for statistical models used in health domains should be higher than for other fields, especially when understanding of cause-and-effect is relevant to treatment and policy decisions and when problems such as poor model calibration could cause considerable patient harm[148]. The development of benchmarks to evaluate new methods is therefore critical for building trust and for highlighting where AI might provide the greatest effect in the field. In part, this is why statistical models in medicine are guided not only by predictive accuracy, but also value calibration, explainability, causality and a rigorous theoretical basis. We join the AI community in recommending that the fundamental theory underpinning current AI models is studied further[149]. We advocate for robust data transparency and ethical sharing, and a detailed study of a range of biases. Finally, we support justifiable barriers of entry to including models into clinical practice, and research into robustness, training and evaluation of popular AI models.

Mechanistic models in infectious disease epidemiology continue to be developed independently of advances in AI. In this Perspective, we have highlighted key areas in which AI could provide either incremental or potentially transformative improvements in modelling epidemic dynamics. However, evaluation of the true value added by AI approaches is difficult, and should include often-overlooked considerations such as the environmental costs associated with training complex AI models. There have been recent examples in which simpler AI methods do not consistently outperform their mechanistic counterparts[37]. Even when AI models do show substantial empirical improvements, translating these models into effective government policy requires considering many factors beyond mere data fit[31]. Although the potential of AI models is evident, their ultimate value will depend on demonstrating a clear opportunity cost advantage over existing approaches—which has yet to be proven definitively.

Open access to and sharing of data, analyses and results in a collaborative and transparent environment can greatly benefit the success of public health campaigns for infectious disease control[150]. By contrast, the scale and cost of training foundation AI models is still increasing[151] and is prohibitive for most. Although we hope that new ideas to mitigate this barrier will emerge, the current suite of AI models should be more transparent in their code, architecture and data. Likewise, the development of transparent and open-access repositories of pathogen-specific data (for example, genomic sequences), environmental data (such as temperature and land cover) and behavioural data (such as human mobility patterns) will remain central for training models and reducing data access inequities. Finally, demonstrating the effectiveness of AI in improving policy decisions that benefit population health remains a notable challenge. For AI to be successful in that regard, the coming years must see growth in close collaborations between research, policy and society[152].

1. McCarthy, J., Minsky, M. L., Rochester, N. & Shannon, C. E. A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Mag.* **27**, 12 (2006).
2. Wang, H. et al. Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47–60 (2023).
3. Hodges, A. B. Jack Copeland (ed.), The Essential Turing: The Ideas that Gave Birth to the Computer Age. Oxford: Clarendon Press, 2004. Pp. viii+613. ISBN 0-19-825079-7. £50.00 (hardback). ISBN 0-19-825080-0. £14.99 (paperback). *Br. J. Hist. Sci.* **39**, 470–471 (2006).
4. Russell, S. J., Norvig, P. & Davis, E. *Artificial Intelligence: A Modern Approach* (Prentice Hall, 2010).
5. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
6. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, 2001).
7. Gelman, A., Carlin, J. B. & Stern, H. S. *Bayesian Data Analysis. Texts in Statistical Science* 696 (Chapman & Hall/CRC Press, 2003).
8. Nocedal, J. & Wright, S. *Numerical Optimization* (Springer New York, 2006).
9. Paszke, A. et al. PyTorch: an imperative style, high-performance deep learning library. *Proc. 33 International Conf. Neural Information Processing Systems* (eds. Wallach, H. M. et al.) 8026–8037 (Curran Associates, 2019).
10. Carpenter, B. et al. Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1–32 (2017).
11. Anderson, R. M. & May, R. M. Population biology of infectious diseases: part I. *Nature* https://doi.org/10.1038/280361a0 (1979).
12. Myszczynska, M. A. et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nat. Rev. Neurol.* **16**, 440–456 (2020).
13. Topol, E. J. Medical forecasting. *Science* **384**, eadp7977 (2024).
14. Moor, M. et al. Foundation models for generalist medical artificial intelligence. *Nature* **616**, 259–265 (2023).
15. Taylor-Robinson, D. & Kee, F. Precision public health-the Emperor's new clothes. *Int. J. Epidemiol.* **48**, 1–6 (2019).
16. Anderson, P. W. More is different. *Science* **177**, 393–396 (1972).
17. Krieger, N., Waterman, P. D., Chen, J. T., Testa, C. & Hanage, W. P. Missing again: US racial and ethnic data for COVID-19 vaccination. *Lancet* **397**, 1259–1260 (2021).
18. Wiens, J. & Shenoy, E. S. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clin. Infect. Dis.* **66**, 149–153 (2018).
19. Acosta, J. N., Falcone, G. J., Rajpurkar, P. & Topol, E. J. Multimodal biomedical AI. *Nat. Med.* **28**, 1773–1784 (2022).
20. Kraemer, M. U. G. et al. Monitoring key epidemiological parameters of SARS-CoV-2 transmission. *Nat. Med.* **27**, 1854–1855 (2021).
21. Monto, A. S., Koopman, J. S. & Longini, I. M. Jr Tecumseh study of illness. XIII. Influenza infection and disease, 1976–1981. *Am. J. Epidemiol.* **121**, 811–822 (1985).
22. Cauchemez, S., Carrat, F., Viboud, C., Valleron, A. J. & Boëlle, P. Y. A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. *Stat. Med.* **23**, 3469–3487 (2004).
23. Tran, T., Pham, T. T., Carneiro, G., Palmer, L. & Reid, I. A Bayesian data augmentation approach for learning deep models. *In Proc. 31st International Conference on Neural Information Processing Systems* (eds von Luxburg, U. et al.) 2794–2803 (ACM, 2017).
24. Rezende, D. & Mohamed, S. Variational inference with normalizing flows. In *Proceedings of the 32nd International Conference on Machine Learning* **37**, 1530–1538 (PMLR, 2015).
25. Knock, E. S. et al. Key epidemiological drivers and impact of interventions in the 2020 SARS-CoV-2 epidemic in England. *Sci. Transl. Med.* **13**, eabg4262 (2021).
26. Bouman, J. A. et al. Bayesian workflow for time-varying transmission in stratified compartmental infectious disease transmission models. *PLoS Comput. Biol.* **20**, e1011575 (2024).
27. Semenova, E., Mishra, S., Bhatt, S., Flaxman, S. & Unwin, H. J. T. in *Epistemic Uncertainty in Artificial Intelligence* (eds Cuzzolin, F. & Sultana, M.) 13–27 (Springer, 2024).
28. Mishra, S. et al. πVAE: a stochastic process prior for Bayesian deep learning with MCMC. *Stat. Comput.* **32**, 96 (2022).
29. Cranmer, K., Brehmer, J. & Louppe, G. The frontier of simulation-based inference. *Proc. Natl Acad. Sci. USA* **117**, 30055–30062 (2020).
   **This study highlights recent advancements in simulation-based inference, focusing on new machine learning techniques that improve inference quality in complex simulations across various scientific domains.**
30. Savcisens, G. et al. Using sequences of life-events to predict human lives. *Nat. Comput. Sci.* **4**, 43–56 (2024).
   **This study presents a predictive framework using transformer-based models to analyse life-event sequences, demonstrating enhanced accuracy in forecasting individual outcomes such as mortality and personality traits.**
31. Medley, G. F. A consensus of evidence: the role of SPI-M-O in the UK COVID-19 response. *Adv. Biol. Regul.* **86**, 100918 (2022).
32. Brockwell, P. J. & Davis, R. A. in *Time Series: Theory and Methods* (eds Brockwell, P. J. & Davis, R. A.) 1–41 (Springer, 1987).
33. Sherratt, K. et al. Exploring surveillance data biases when estimating the reproduction number: with insights into subpopulation transmission of COVID-19 in England. *Philos. Trans. R. Soc. Lond. B* **376**, 20200283 (2021).
34. Mena, G. E. et al. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science* **372**, eabg5298 (2021).
35. Hawryluk, I. et al. Gaussian process nowcasting: application to COVID-19 mortality reporting. In *Proc. 37th Conference on Uncertainty in Artificial Intelligence, UAI 2021* Vol. 161, 1258–1268 (PMLR, 2021).
   **This paper introduces a Gaussian process model for nowcasting COVID-19 mortality, correcting for reporting delays and providing robust uncertainty estimates to improve real-time epidemiological assessments.**
36. Lison, A., Abbott, S., Huisman, J. & Stadler, T. Generative Bayesian modeling to nowcast the effective reproduction number from line list data with missing symptom onset dates. *PLoS Comput. Biol.* **20**, e1012021 (2024).
   **This paper presents a unified generative Bayesian model to estimate the effective reproduction number from incomplete line list data, addressing limitations in traditional methods by considering missing symptom onset dates and right-truncated case counts.**
37. McGough, S. F., Johansson, M. A., Lipsitch, M. & Menzies, N. A. Nowcasting by Bayesian smoothing: a flexible, generalizable model for real-time epidemic tracking. *PLoS Comput. Biol.* **16**, e1007735 (2020).
   **This paper introduces a Bayesian smoothing approach for nowcasting that accurately estimates real-time epidemic case counts by incorporating temporal relationships and adapting to reporting delays across diseases.**

38. Brizzi, A., O'Driscoll, M. & Dorigatti, I. Refining reproduction number estimates to account for unobserved generations of infection in emerging epidemics. *Clin. Infect. Dis.* **75**, e114–e121 (2022).

39. Cramer, E. Y. et al. Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the United States. *Proc. Natl Acad. Sci. USA* **119**, e2113561119 (2022).
    **This study evaluates probabilistic forecasting methods for COVID-19 mortality, showing that ensemble approaches, which integrate predictions from multiple models, consistently yield more accurate forecasts.**

40. Wood, D. et al. A unified theory of diversity in ensemble learning. *J. Mach. Learn. Res.* **24**, 1–49 (2023).
    **This paper presents a framework for understanding ensemble diversity in supervised learning, shifting the focus from maximizing diversity to managing a bias–variance–diversity trade-off in ensemble methods.**

41. Das, A. et al. A decoder-only foundation model for time-series forecasting. In *Proceedings of the 41st International Conference on Machine Learning* 235, 10148–10167 (PMLR, 2024).

42. Pourpanah, F. et al. A review of generalized zero-shot learning methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 4051–4070 (2023).

43. Gal, Y. & Ghahramani, Z. Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In *Proc. 33rd International Conference on Machine Learning* (eds Balcan, M. F. & Weinberger, K. Q.) Vol. 48, 1050–1059 (PMLR, 2016).

44. Ganaie, M. A., Hu, M., Malik, A. K., Tanveer, M. & Suganthan, P. N. Ensemble deep learning: a review. *Eng. Appl. Artif. Intell.* **115**, 105151 (2022).

45. Angelopoulos, A. N. & Bates, S. *Conformal Prediction: A Gentle Introduction* (Now, 2023).

46. Hunter, E., Namee, B. M. & Kelleher, J. D. A taxonomy for agent-based models in human infectious disease epidemiology. *J. Artif. Soc. Soc. Simul.* **20**, 2 (2017).

47. Quera-Bofarull, A., Chopra, A., Calinescu, A., Wooldridge, M. & Dyer, J. Bayesian calibration of differentiable agent-based models. Preprint at arxiv.org/abs/2305.15340 (2023).
    **This paper introduces a method that combines variational inference with differentiable agent-based models for better Bayesian parameter calibration, addressing issues with complex likelihood functions and model inaccuracies.**

48. Pakkanen, M. S. et al. Unifying incidence and prevalence under a time-varying general branching process. *J. Math. Biol.* **87**, 35 (2023).

49. Unwin, H. J. T. et al. Using Hawkes processes to model imported and local malaria cases in near-elimination settings. *PLoS Comput. Biol.* **17**, e1008830 (2021).

50. Rizoiu, M.-A., Mishra, S., Kong, Q., Carman, M. & Xie, L. SIR-Hawkes: linking epidemic models and Hawkes processes to model diffusions in finite populations. In *Proc. 2018 World Wide Web Conference* 419–428 (International World Wide Web Conferences Steering Committee, 2018).

51. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. In *International Conf. on Learning Representations* (ICLR, 2017).

52. Liu, Z. et al. A review of graph neural networks inepidemic modeling. In *Proc. 30th ACM SIGKDD Conf. on Knowledge Discovery and Data Mining* 6577–6587 (2024).

53. Salathé, M. et al. A high-resolution human contact network for infectious disease transmission. *Proc. Natl Acad. Sci. USA* **107**, 22020–22025 (2010).

54. Al-Garadi, M. A., Khan, M. S., Varathan, K. D., Mujtaba, G. & Al-Kabsi, A. M. Using online social networks to track a pandemic: a systematic review. *J. Biomed. Inform.* **62**, 1–11 (2016).

55. Panagopoulos, G., Nikolentzos, G. & Vazirgiannis, M. Transfer graph neural networks for pandemic forecasting. *AAAI* **35**, 4838–4845 (2021).
    **This paper presents a method to use graph neural networks and transfer learning to predict infectious disease spread, modelling population movement and disease diffusion patterns, which improves forecasting accuracy across various regions.**

56. Deng, S., Wang, S., Rangwala, H., Wang, L. & Ning, Y. Cola-GNN: cross-location attention based graph neural networks for long-term ILI prediction. In *Proc. 29th ACM International Conference on Information & Knowledge Management* 245–254 (ACM, 2020).

57. Chang, S. et al. Measuring vaccination coverage and concerns of vaccine holdouts from web search logs. *Nat. Commun.* **15**, 6496 (2024).

58. Wang, L. et al. CausalGNN: causal-based graph neural networks for spatio-temporal epidemic forecasting. *AAAI* **36**, 12191–12199 (2022).

59. Aylett-Bullock, J. et al. June: open-source individual-based epidemiology simulation. *R. Soc. Open Sci.* **8**, 210506 (2021).

60. Liu, J. et al. Towards graph foundation models: a survey and beyond. Preprint at arxiv.org/abs/2310.11829 (2023).

61. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
    **This study introduces AlphaFold, a neural network approach that achieves an extremely high level of accuracy in predicting protein structure.**

62. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
    **This paper introduces AlphaFold 3, a model that uses a diffusion-based architecture to predict complex biomolecular interactions, enhancing the understanding of protein interactions and aiding in the design of targeted therapies.**

63. Hie, B., Zhong, E. D., Berger, B. & Bryson, B. Learning the language of viral evolution and escape. *Science* **371**, 284–288 (2021).

64. Zrnic, T. & Candès, E. J. Cross-prediction-powered inference. *Proc. Natl Acad. Sci. USA* **121**, e2322083121 (2024).

65. de Bernardi Schneider, A. et al. SARS-CoV-2 lineage assignments using phylogenetic placement/UShER are superior to pangoLEARN machine-learning method. *Virus Evol.* **10**, vead085 (2024).

66. Malik, A. J., Poole, A. M. & Allison, J. R. Structural phylogenetics with confidence. *Mol. Biol. Evol.* **37**, 2711–2726 (2020).

67. Voznica, J. et al. Deep learning from phylogenies to uncover the epidemiological dynamics of outbreaks. *Nat. Commun.* **13**, 3896 (2022).

68. Ito, J. et al. A Protein language model for exploring viral fitness landscapes. Preprint at *bioRxiv* https://doi.org/10.1101/2024.03.15.584819 (2024).

69. Mollentze, N. & Streicker, D. G. Predicting zoonotic potential of viruses: where are we? *Curr. Opin. Virol.* **61**, 101346 (2023).

70. Babayan, S. A., Orton, R. J. & Streicker, D. G. Predicting reservoir hosts and arthropod vectors from evolutionary signatures in RNA virus genomes. *Science* **362**, 577–580 (2018).

71. Thadani, N. N. et al. Learning from prepandemic data to forecast viral escape. *Nature* **622**, 818–825 (2023).
    **This study introduces EVEscape, a framework that uses deep learning from historical viral sequences and structural information to predict viral mutations that evade immune responses, providing early insights for vaccine design and pandemic preparedness.**

72. Koel, B. F. et al. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science* **342**, 976–979 (2013).

73. Shan, S. et al. Deep learning guided optimization of human antibody against SARS-CoV-2 variants with broad neutralization. *Proc. Natl Acad. Sci. USA* **119**, e2122954119 (2022).

74. Hie, B. L. et al. Efficient evolution of human antibodies from general protein language models. *Nat. Biotechnol.* **42**, 275–283 (2024).

75. Lin, Z. et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123–1130 (2023).

76. Mahmud, A. S., Martinez, P. P. & Baker, R. E. The impact of current and future climates on spatiotemporal dynamics of influenza in a tropical setting. *PNAS Nexus* **2**, gad307 (2023).

77. Maher, M. C. et al. Predicting the mutational drivers of future SARS-CoV-2 variants of concern. *Sci. Transl. Med.* **14**, eabk3445 (2022).

78. Mollentze, N., Babayan, S. A. & Streicker, D. G. Identifying and prioritizing potential human-infecting viruses from their genome sequences. *PLoS Biol.* **19**, e3001390 (2021).

79. Grenfell, B. T. et al. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* **303**, 327–332 (2004).

80. Zhang, C. & Matsen, F. A. IV. A variational approach to Bayesian phylogenetic inference. *J. Mach. Learn. Res.* **25**, 1–56 (2024).
    **This paper presents a variational framework for Bayesian phylogenetic inference that enhances exploration efficiency and reduces the number of required iterations compared to traditional Markov chain Monte Carlo methods.**

81. Bajaj, S. et al. COVID-19 testing and reporting behaviours in England across different sociodemographic groups: a population-based study using testing data and data from community prevalence surveillance surveys. *Lancet Digit. Health* **6**, e778–e790 (2024).

82. Obermeyer, Z., Powers, B., Vogeli, C. & Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**, 447–453 (2019).

83. 100,000 Genomes Project Pilot Investigators. 100,000 Genomes pilot on rare-disease diagnosis in health care—preliminary report. *N. Engl. J. Med.* **385**, 1868–1880 (2021).

84. Brito, A. F. et al. Global disparities in SARS-CoV-2 genomic surveillance. *Nat. Commun.* **13**, 7003 (2022).

85. Yang, Y. et al. Exploring the big data paradox for various estimands using vaccination data from the global COVID-19 Trends and Impact Survey (CTIS). *Sci. Adv.* **10**, eadj0266 (2024).

86. Dan, S. et al. Estimating fine age structure and time trends in human contact patterns from coarse contact data: the Bayesian rate consistency model. *PLoS Comput. Biol.* **19**, e1011191 (2023).

87. Settles, B. *Active Learning Literature Survey* (University of Wisconsin-Madison, 2009).

88. Garnett, R. *Bayesian Optimization Book* (Cambridge Univ. Press, 2023).

89. Tsui, J. L. et al. Toward optimal disease surveillance with graph-based active learning. *Proc. Natl Acad. Sci. USA* **121**, e2412424121 (2024).
    **This research proposes an active learning policy that optimally allocates limited testing resources across a network to improve disease surveillance and predictive accuracy while minimizing the number of required tests.**

90. Wymant, C. et al. The epidemiological impact of the NHS COVID-19 app. *Nature* **594**, 408–412 (2021).

91. Ferretti, L. et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936 (2020).

92. Ferretti, L. et al. Digital measurement of SARS-CoV-2 transmission risk from 7 million contacts. *Nature* **626**, 145–150 (2024).

93. Kraemer, M. U. G. et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).

94. Chang, S. et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2021).

95. Kraemer, M. U. G. et al. Mapping global variation in human mobility. *Nat. Hum. Behav.* **4**, 800–810 (2020).

96. Page, B. & Topol, E. J. Digitising the outbreak. *Lancet* **402**, 2186 (2023).

97. Radin, J. M. et al. Sensor-based surveillance for digitising real-time COVID-19 tracking in the USA (DETECT): a multivariable, population-based, modelling study. *Lancet Digit. Health* **4**, e777–e786 (2022).

98. Bastani, H. et al. Efficient and targeted COVID-19 border testing via reinforcement learning. *Nature* **599**, 108–113 (2021).

99. Quer, G. et al. Wearable sensor data and self-reported symptoms for COVID-19 detection. *Nat. Med.* **27**, 73–77 (2021).

100. Schmidt, B. & Hildebrandt, A. Deep learning in next-generation sequencing. *Drug Discov. Today* **26**, 173–180 (2021).

101. Skums, P. et al. QUENTIN: reconstruction of disease transmissions from viral quasispecies genomic data. *Bioinformatics* **34**, 163–170 (2018).

102. Turakhia, Y. et al. Ultrafast sample placement on existing tRees (UShER) enables real-time phylogenetics for the SARS-CoV-2 pandemic. *Nat. Genet.* **53**, 809–816 (2021).

103. Turakhia, Y. et al. Pandemic-scale phylogenomics reveals the SARS-CoV-2 recombination landscape. *Nature* **609**, 994–997 (2022).

104. Baker, R. E. et al. Infectious disease in an era of global change. *Nat. Rev. Microbiol.* **20**, 193–205 (2022).

105. Lam, R. et al. Learning skilful medium-range global weather forecasting. *Science* **382**, 1416–1421 (2023).

106. Chinazzi, M. et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).

# Perspective

107. Bengio, Y. et al. Managing extreme AI risks amid rapid progress. *Science* **384**, 842–845 (2024).

108. Transcript of module 2 public hearing on 23 November 2023—searchable transcripts of the UK COVID-19 Inquiry hearings documentation. *UK COVID-19 Inquiry* ukcovid19inquiry. dracos.co.uk/module-2/2023-11-23/ (UK COVID-19 Inquiry, 2023).

109. Sutton, R. S. & Barto, A. G. *Reinforcement Learning, Second Edition: An Introduction* (MIT Press, 2018).

110. Benz, N. C & Rodriguez, M. G. Human-aligned calibration for AI-assisted decision making. In *Proc. 37th International Conference on Neural Information Processing Systems* (eds Oh, A. et al.) 14609–14636 (ACM, 2023).

111. Gupta, R. K. et al. HIV-1 drug resistance before initiation or re-initiation of first-line antiretroviral therapy in low-income and middle-income countries: a systematic review and meta-regression analysis. *Lancet Infect. Dis.* **18**, 346–355 (2018).

112. Shea, K. et al. Harnessing multiple models for outbreak management. *Science* **368**, 577–579 (2020).
    **This article discusses the importance of integrating expert elicitation methods with decision-theoretic frameworks to enhance the effectiveness of multiple modelling approaches in managing outbreaks, assisting policymakers in navigating uncertainties**.

113. Kekić, A. et al. Evaluating vaccine allocation strategies using simulation-assisted causal modeling. *Patterns* **4**, 100739 (2023).

114. WHO. *Ethics and Governance of Artificial Intelligence for Health* (WHO, 2021).

115. Parker, M. *Ethical Hotspots in Infectious Disease Surveillance for Global Health Security Social Justice and Pandemic Preparedness* (Oxford Univ. Press, 2023).

116. Parker, M. J., Fraser, C., Abeler-Dörner, L. & Bonsall, D. Ethics of instantaneous contact tracing using mobile phone apps in the control of the COVID-19 pandemic. *J. Med. Ethics* **46**, 427–431 (2020).

117. Gradoń, K. T., Hołyst, J. A., Moy, W. R., Sienkiewicz, J. & Suchecki, K. Countering misinformation: a multidisciplinary approach. *Big Data Soc.* **8**, 20539517211013848 (2021).

118. WHO. How to report misinformation online. *WHO* www.who.int/campaigns/connecting-the-world-to-combat-coronavirus/how-to-report-misinformation-online (WHO, 2024).

119. WHO. Fides—a network of healthcare influencers. *WHO* www.who.int/teams/digital-health-and-innovation/digital-channels/fides (WHO, 2024).

120. Chen, C. & Stadler, T. GenSpectrum Chat: data exploration in public health using large language models. Preprint at arxiv.org/abs/2305.13821 (2023).

121. Williams, R., Hosseinichimeh, N., Majumdar, A. & Ghaffarzadegan, N. Epidemic modeling with generative agents. Preprint at arxiv.org/abs/2307.04986 (2023).
    **This study introduces an agent-based epidemic modelling framework using generative AI, enabling agents to make autonomous decisions based on contextual information and effectively incorporating human behaviour into epidemic simulations**.

122. Allen, J., Watts, D. J. & Rand, D. G. Quantifying the impact of misinformation and vaccine-skeptical content on Facebook. *Science* **384**, eadk3451 (2024).

123. Benjamin, R. *Race After Technology: Abolitionist Tools for the New Jim Code* (John Wiley & Sons, 2019).

124. Norori, N., Hu, Q., Aellen, F. M., Faraci, F. D. & Tzovara, A. Addressing bias in big data and AI for health care: a call for open science. *Patterns* **2**, 100347 (2021).

125. Morley, J. et al. The ethics of AI in health care: a mapping review. *Soc. Sci. Med.* **260**, 113172 (2020).

126. *Exploring Public Attitudes Towards the Use of Digital Health Technologies and Data* (The Health Foundation, 2023); www.health.org.uk/publications/long-reads/exploring-public-attitudes-towards-the-use-of-digital-health-technologies.

127. Zhang, Y. et al. Siren's song in the AI ocean: a survey on hallucination in large language models. Preprint at arxiv.org/abs/2309.01219 (2023).

128. Farquhar, S., Kossen, J., Kuhn, L. & Gal, Y. Detecting hallucinations in large language models using semantic entropy. *Nature* **630**, 625–630 (2024).

129. AlphaFold Protein Structure Database. *EBI* alphafold.ebi.ac.uk/ (EBI, 2024).

130. 100,000 Genomes Project. *Genomics England* www.genomicsengland.co.uk/initiatives/100000-genomes-project (Genomics England, 2022).

131. Xu, B. et al. Epidemiological data from the COVID-19 outbreak, real-time case information. *Sci. Data* **7**, 106 (2020).

132. WHO. Digital health and innovation. *WHO* www.who.int/teams/digital-health-and-innovation (WHO, 2024); .

133. Omar, M., Brin, D., Glicksberg, B. & Klang, E. Utilizing natural language processing and large language models in the diagnosis and prediction of infectious diseases: a systematic review. *Am. J. Infect. Control* https://doi.org/10.1016/j.ajic.2024.03.016 (2024).

134. Schmidt, L. et al. Data extraction methods for systematic review (semi)automation: update of a living systematic review. *F1000Res.* **10**, 401 (2021).

135. Wu, S. et al. Deep learning in clinical natural language processing: a methodical review. *J. Am. Med. Inform. Assoc.* **27**, 457–470 (2020).

136. Hasan, B. et al. Integrating large language models in systematic reviews: a framework and case study using ROBINS-I for risk of bias assessment. *BMJ Evid. Based Med.* https://doi.org/10.1136/bmjebm-2023-112597 (2024).

137. Kasy, M. & Abebe, R. Fairness, equality, and power in algorithmic decision-making. In *Proc. 2021 ACM Conference on Fairness, Accountability, and Transparency* 576–586 (ACM, 2021).

138. Lewis, P. et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Proceedings of the 34th International Conference on Neural Information Processing Systems* (eds Larochelle, H. et al.) 793 (Curran Associates Inc., 2020).

139. Dayan, I. et al. Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat. Med.* **27**, 1735–1743 (2021).

140. Hassan, C., Bon, J. J., Semenova, E., Mira, A. & Mengersen, K. Federated learning for non-factorizable models using deep generative prior approximations. Preprint at arxiv.org/abs/2405.16055 (2024).

141. Wilkinson, M. D. et al. The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).

142. Adebamowo, C. et al. Ethical oversight of data science health research in Africa. *NEJM AI* **1**, AIpc2400033 (2024).

143. Barreras, F. & Watts, D. J. The exciting potential and daunting challenge of using GPS human-mobility data for epidemic modeling. *Nat. Comput. Sci.* **4**, 398–411 (2024).

144. Bommasani, R. et al. The Foundation Model Transparency Index after 6 months. *Stanford CFM* crfm.stanford.edu/2024/05/21/fmti-may-2024.html (Stanford CRFM, 2024).

145. Gibb, R. et al. Interactions between climate change, urban infrastructure and mobility are driving dengue emergence in Vietnam. *Nat. Commun.* **14**, 8179 (2023).

146. Peters, J., Janzing, D. & Schölkopf, B. *Elements of Causal Inference* (MIT Press, 2021).

147. Wong, C. How AI is improving climate forecasts. *Nature* **628**, 710–712 (2024).

148. Van Calster, B. et al. Calibration: the Achilles heel of predictive analytics. *BMC Med.* **17**, 230 (2019).

149. Zhang, C., Bengio, S., Hardt, M., Recht, B. & Vinyals, O. Understanding deep learning (still) requires rethinking generalization. *Commun. ACM* **64**, 107–115 (2021).

150. Silva, D. S. & Smith, M. J. Data sharing during pandemics: reciprocity, solidarity, and limits to obligations. *J. Bioeth. Inq.* **20**, 667–672 (2023).

151. OpenAI et al. GPT-4 technical report. Preprint at arxiv.org/abs/2303.08774 (2023).

152. Burki, T. Crossing the frontier: the first global AI safety summit. *Lancet Digit. Health* **6**, e91–e92 (2024).

153. Charles, G. et al. Seq2Seq surrogates of epidemic models to facilitate Bayesian inference. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence* 14170–14177 (AAAI Press, 2023).

154. Garza, A. & Mergenthaler-Canseco, M. Nixtla: TimeGPT-1: production ready pre-trained time series foundation model for forecasting and anomaly detection. *Github* https://github.com/Nixtla/nixtla (Github, 2023).

155. Borchering, R. K. et al. Public health impact of the U.S. scenario modeling hub. *Epidemics* **44**, 100705 (2023).

156. Mehrjou, A. et al. Pyfectious: an individual-level simulator to discover optimal containment policies for epidemic diseases. *PLoS Comput. Biol.* **19**, e1010799 (2023).

157. Tsui, J. L.-H. et al. Toward optimal disease surveillance with graph-based active learning. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.2412424121 (2024).

158. Tu, T. et al. Towards generalist biomedical AI. *NEJM AI* https://doi.org/10.1056/AIoa2300138 (2024).

159. Ouyang, L. et al. Training language models to follow instructions with human feedback. *Proceedings of the 36th International Conference on Neural Information Processing Systems*, 2011–2025 (Curran Associates Inc., 2024).

160. Ghani, A. C. et al. Methods for estimating the case fatality ratio for a novel, emerging infectious disease. *Am. J. Epidemiol.* **162**, 479–486 (2005).

161. Kenah, E., Lipsitch, M. & Robins, J. M. Generation interval contraction and epidemic data analysis. *Math. Biosci.* **213**, 71–79 (2008).

162. Newman, M. E. J. Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128 (2002).

163. Gostic, K. M. et al. Practical considerations for measuring the effective reproductive number, $R_t$. *PLoS Comput. Biol.* **16**, e1008409 (2020).

**Competing interests** S. Bhatt is a paid member of the Academic Council of the Schmidt Science Fellows programme outside the scope of this work. This affiliation is unrelated to the submitted work, and the programme does not stand to benefit from this publication. M.A.S. receives grants from the US National Institutes of Health within the scope of this work, and grants and contracts from the US Food and Drug Administration, the US Department of Veterans Affairs, and Johnson and Johnson, all outside the scope of this work. C.F. is a member of two committees that advise the UK Department of Health on emerging epidemics, namely NERVTAG and SPI-M. The other authors declare no competing interests.

**Additional information**
**Correspondence and requests for materials** should be addressed to Moritz U. G. Kraemer or Samir Bhatt.
**Peer review information** *Nature* thanks Peter Klimek, Amalio Telenti 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.