nature machine intelligence

Review article

https://doi.org/10.1038/s42256-024-00895-7

# Machine learning for data-centric epidemic forecasting

Received: 23 March 2023

Accepted: 6 August 2024

Published online: 27 September 2024

Check for updates

Alexander Rodríguez[1,2,3], Harshavardhan Kamarthi[1,3], Pulak Agarwal[1], Javen Ho[1], Mira Patel[1], Suchet Sapre[1] & B. Aditya Prakash[1]

The COVID-19 pandemic emphasized the importance of epidemic forecasting for decision makers in multiple domains, ranging from public health to the economy. Forecasting epidemic progression is a non-trivial task due to multiple confounding factors, such as human behaviour, pathogen dynamics and environmental conditions. However, the surge in research interest and initiatives from public health and funding agencies has fuelled the availability of new data sources that capture previously unobservable aspects of disease spread, paving the way for a spate of 'data-centred' computational solutions that show promise for enhancing our forecasting capabilities. Here we discuss various methodological and practical advances and introduce a conceptual framework to navigate through them. First we list relevant datasets, such as symptomatic online surveys, retail and commerce, mobility and genomics data. Next we consider methods, focusing on recent data-driven statistical and deep learning-based methods, as well as hybrid models that combine domain knowledge of mechanistic models with the flexibility of statistical approaches. We also discuss experiences and challenges that arise in the real-world deployment of these forecasting systems, including decision-making informed by forecasts. Finally, we highlight some challenges and open problems found across the forecasting pipeline to enable robust future pandemic preparedness.

The devastating impact of the COVID-19 pandemic on human lives, economic development and society as a whole has illustrated our vulnerability to major epidemics. While the science of epidemic forecasting is, in many respects, still in its initial stages, the current pandemic and those before it (such as H1N1 and Ebola) have shown its crucial importance. Preventing and responding to such pandemics requires actionable epidemic forecasts to (for example) design effective healthcare policies and make optimal supply chain decisions. Generating such forecasts presents a range of interdisciplinary challenges, including understanding the biological processes driving pathogen evolution, assessing responses to immunization and drug resistance, and modelling population-level interactions among heterogeneous groups, among others[1]. In response to these challenges, there has been growing interest in data-centred solutions for epidemic forecasting[2], building on several initiatives in the past few years from both government public health agencies and funding agencies. For instance, in 2013 the US Centers for Disease Control and Prevention (CDC) introduced the FluSight challenge[3], which has not only aided in improving flu forecasting capabilities and public health decision-making, but also helped to build a community of researchers around this topic. Similar initiatives have followed for Ebola[4], dengue[5] and also COVID-19[6] led by institutions around the globe such as the European CDC[7], Intelligence Advanced

[1]College of Computing, Georgia Institute of Technology, Atlanta, GA, USA. [2]Computer Science and Engineering, University of Michigan, Ann Arbor, MI, USA. [3]These authors contributed equally: Alexander Rodríguez, Harshavardhan Kamarthi. e-mail: alrodri@umich.edu; hkamarthi3@gatech.edu; badityap@cc.gatech.edu

 https://doi.org/10.1038/s42256-024-00895-7

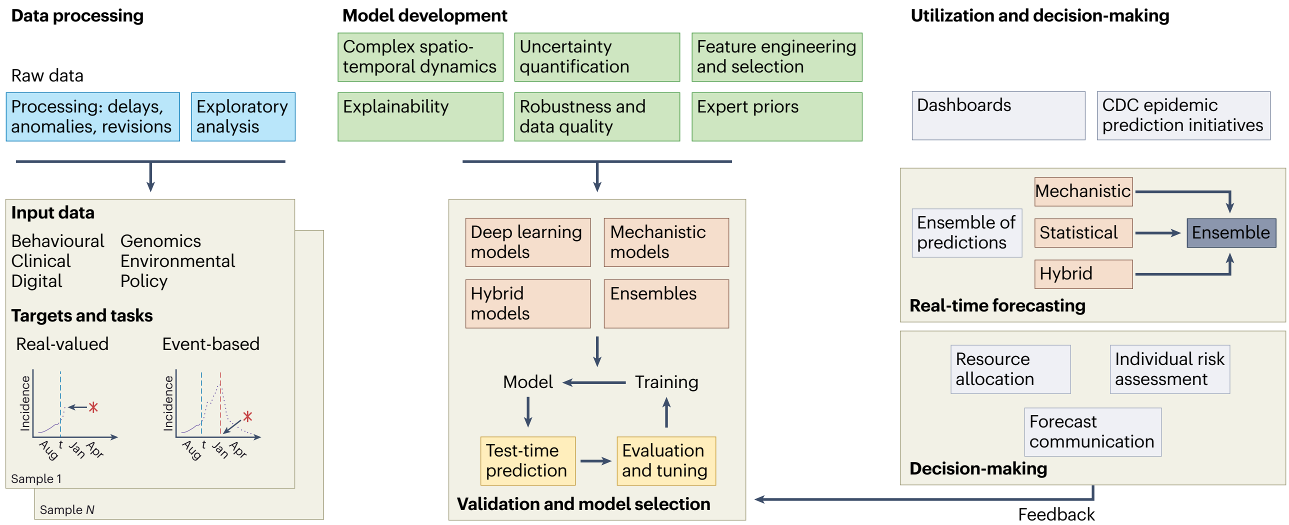


**Fig. 1 | Overview of the data-centric epidemic forecasting pipeline.** We begin by collecting data from heterogeneous sources, preparing these data and choosing from a set of epidemic targets and tasks. During model development, we consider various aspects of epidemic spread (such as spatiotemporal dynamics) and forecast utilization (for example, uncertainty quantification). Finally, real-time forecasts have multiple uses including dashboards, ensemble composition and other public health initiatives that serve as a platform for decision-making for resource allocation, risk assessment, general communication to the public and other domains. Red asterisks denote the forecasting tasks: the value of future incidence as an example of a real-valued task, and peak time as an example of an event-based task. *t* represents the forecasting week.

Research Projects Activity (IARPA) and Pan American Health Organization (PAHO) in Latin America. These forecasting initiatives have provided an unprecedented opportunity for researchers to observe both the successes and gaps in the current science of forecasting. Similarly, US agencies such as the National Science Foundation, the National Institutes of Health and US Army Research Laboratory have held a series of recent symposia and funding calls related to pandemic forecasting, which have given a much-needed impetus to this topic. This interest has also culminated in the establishment of the first Center for Forecasting and Outbreak Analytics by the US CDC in 2021.

This Review delves into data-driven computational methods driven by advances in machine learning and their capabilities to leverage new sources of data, from biological to behavioural. We found that the availability of data from reliable sources (several of them publicly accessible) is increasing, a trend that has only been accelerated by the COVID-19 pandemic. This includes richer epidemiological datasets and digital data streams such as mobility data[8], online surveys[9] and wastewater samples[10]. We also found a number of technical innovations in machine learning and deep learning motivated by this domain, which have not only opened new horizons in the science of epidemic forecasting but also contributed to a broader understanding of time-series analysis and scientific artificial intelligence (AI). This Review is an effort to encompass recent methodological and practical advances in machine learning at an opportune time to help and enable the broader computational and data/machine learning/AI communities to engage in this area. We refer to this recent body of work as data-centric epidemic forecasting. In the following sections, we will discuss the multiple components of data-centric epidemic forecasting, including datasets and problem formulation. We will then turn to modelling approaches: traditional machine learning, deep learning and hybrid methods. Finally, we discuss future challenges and opportunities in this emerging field.

## Data-centric epidemic forecasting

Here we briefly describe the components of the data-centric forecasting pipeline, which we conceptualize in Fig. 1 as comprising three stages: data processing, model training and validation, and utilization and decision-making (see the Supplementary Information for in-depth discussion of each of the components).

### Datasets

A diverse set of datasets has been used to better inform epidemic forecasting, offering benefits ranging from early-stage indicators to capturing complementary facets that help to explain disease spread dynamics. Figure 2 shows sources of such data; most sources on the left side are from clinical disease surveillance, whereas those on the right are from digital disease surveillance.

Clinical information has traditionally been the primary source of surveillance data for epidemiology, including aggregate markers from epidemiological line lists and testing records (for example, tracking people infected, recovered and deceased). Another important data source is syndromic surveillance, which is based on clinical features discernible before a diagnosis is confirmed or activities prompted by the onset of symptoms, serving as an alert of changes in disease activity[11]. Digital surveillance takes advantage of the widespread use of mobile devices and innovations in digital communication to provide real-time data without extensive human effort. Such technological developments can be used to track the large-scale mobility and disease exposure of populations[8]. Recent works have leveraged features from various unconventional data sources as well. For instance, images from remote-sensing satellites have been used to detect outbreaks and track prevalence by monitoring vacancies in car parks near hospitals and other sensitive locations[12]. Pharmaceutical sales records, other retail records that capture useful consumer signals, and restaurant reservation data have also previously proven valuable for epidemic forecasting[13]. By incorporating these diverse and complementary data sources, researchers have sought to gain a more comprehensive understanding of disease spread dynamics.

### Targets, tasks and evaluation

Forecasts can be divided into projections[11,14] (for a specific future scenario) and predictions[15] (for the most likely scenario). Common predictive targets are the number of symptomatic patients, hospitalizations[16] and mortality[17] across varying spatial (for example, country, region/state/province and county/city) and temporal (for example, weekly or daily) scales based on the requirements of decision makers and domain experts[5,18]. Tasks can be broadly classified as real-valued prediction tasks, such as the prediction of future incidences, nowcasting and

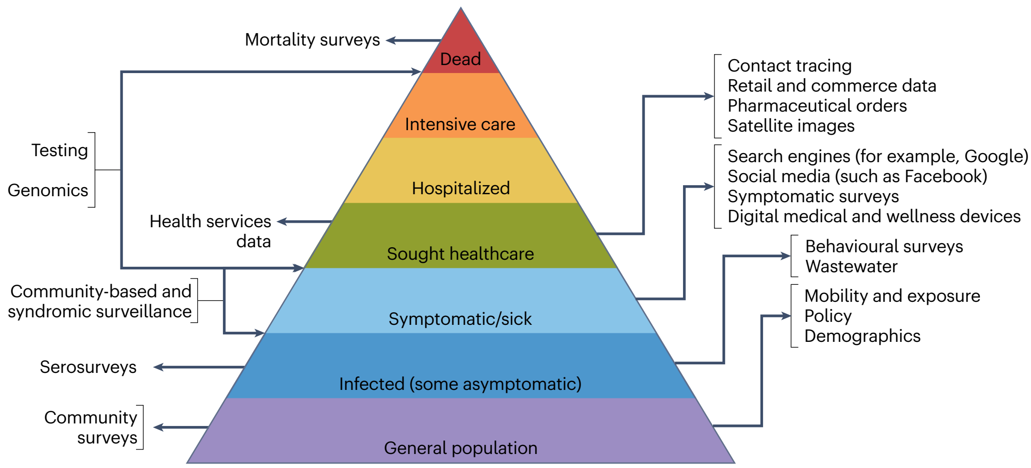


**Fig. 2 | Conceptualization of disease surveillance data sources.** The pyramid depicts the potential stages people may progress though during an illness from bottom to top, with the area of each level proportional to the number of people in that stage. We connect each of these levels with our proposed classification of datasets (a detailed classification can be found in the Supplementary Information) used in the literature to inform forecasting models and depict representative examples. Sources on the left side of the pyramid are from direct clinical surveillance; those on the right are digital sources of proxy indicators.

peak incidence/height, and event-based prediction tasks that focus on predicting the time/stages of an outbreak or forecasting season, such as peak time and onset[15]. One common example of a point forecast metric for the evaluation of real-valued tasks is the mean absolute error[17,19]. Evaluations of probabilistic forecasts often rely on log scores for targets with a predefined range (for example, percentages)[5,15,20,21]. When working with unbounded predictions (such as the numbers of cases, deaths and hospitalizations), probabilistic metrics such as interval score or the weighted interval score[17] are preferred, because they are 'proper scoring rules'[22]. It is pertinent to note that the choice of evaluation metric should take into account the varying requirements of public health agencies and decision makers[16] and that CDC evaluations often put more emphasis on probabilistic evaluations[5,15].

### Overview of methods
We mainly focus on predictions in this Review as this is where most data-driven work has been undertaken. We broadly classify methods into three categories: mechanistic; statistical, machine learning and AI; and hybrid. Our focus is on the latter two, where most recent data-centric work has been done. Nevertheless, we elaborate on mechanistic models in the Supplementary Information and refer the reader to other excellent surveys[2,23] for in-depth discussions.

We now discuss some key insights for different kinds of models in Table 1. First, mechanistic models explicitly encode the causal mechanisms of epidemic spread. These models have a rich history dating back to the eighteenth century and offer valuable insight and explainability. However, incorporating non-traditional data sources to complement mechanistic models is often challenging[24].

On the other hand, statistical, machine learning and AI models focus on learning empirical relationships between past and future phenomena. Early attempts involved traditional regression models that could incorporate various data sources, such as searches[25,26] and social media[27]. However, these methods often fall short due to their limited flexibility in capturing useful features and patterns from complex dynamics of the real world. This limitation is where deep learning models shine, which also allow us to apply recent advances to model spatiotemporal dynamics and address data sparsity—although at the cost of reduced explainability. Density estimation methods, grounded in Bayesian modelling, enable the incorporation of expert-based priors; however, this also makes them less amenable to incorporating new datasets, as specifying priors for each of these datasets becomes an additional challenge.

Lastly, the more recent hybrid models combine mechanistic models with the data-driven flexibility of AI. Mechanistic models may incorporate statistical components to address data sparsity and integrate new datasets. Conversely, statistical components can be used to refine the output of mechanistic models, enhancing modelling flexibility to learn expressive features while preserving explainability. Another hybrid paradigm involves machine learning models being informed by priors from mechanistic models through generating training data or serving as learning constraints. The last hybrid paradigm is the wisdom of crowds, where the objective is to harness multiple predictors, encompassing all previous perspectives and their unique aspects.

In the following sections, we elaborate on the machine learning and hybrid modelling approaches in greater detail. We also discuss model selection in the Supplementary Information.

## Statistical, machine learning and AI models
These models learn complex patterns from a wide variety of input signals that are assumed to influence the epidemic and then leverage them to forecast. Unlike mechanistic models, they typically require minimal modelling assumptions and offer a more flexible approach to extracting patterns from data.

### Traditional machine learning models
A diverse range of machine learning methods have been used and tailored to facilitate more data-driven approaches to epidemic forecasting. Some of these methods specialize in making predictions based on time-series data through standard regression techniques, whereas others incorporate the analysis of unstructured data (such as textual information). The development of methods for the effective quantification of forecast uncertainty is also a crucial aspect.

**Classical statistical methods.** These techniques rely on traditional regression models and take different sets of indicators as inputs, such as the volumes of search queries containing expert-curated keywords[25,26]. Common modelling approaches here are based on autoregressive models[28–30], nearest-neighbour regression[31], multi-task Gaussian processes and elastic nets[32], the minimum description length principle[33] and dynamic Poisson autoregression[34].

**Text-based methods.** Such methods leverage data sources in the form of text, which (as can be seen in Fig. 2) are important sources of digital surveillance. Stylography and part-of-speech detection[35], as well as

 https://doi.org/10.1038/s42256-024-00895-7

**Table 1 | Most prominent capabilities and gaps (weaknesses) of modelling paradigms**

| | Modelling paradigms | Reference(s) | Main capabilities | Main gaps |
|---|---|---|---|---|
| Mechanistic | Compartmental, meta-population, agent-based | 65,87,106–113 | X; UQ; ST | NTD; DAT; NFE |
| Statistical, machine learning and AI | Traditional machine learning models | | | |
| | Linear models | 25,26,114 | NTD; X | ST; NFE |
| | Autoregressive models | 19,28,30,115 | NTD | ST; NFE |
| | Nonlinear models | 31,34 | NTD | ST; NFE |
| | Hierarchical models | 29,32,33 | ST | NFE |
| | Vision and language models | | | |
| | Vision models | 12 | NFE; NTD | X; ST |
| | Language-based models | 35,36 | NFE; NTD | X; ST |
| | Probabilistic topic models | 37–39,116 | NTD; X; UQ | ST |
| | Deep learning models | | | |
| | Off-the-shelf sequential models | 45–47,86 | NTD; NFE | X; ST; DAT |
| | Temporal similarity modelling | 48–50 | ST; NFE; NTD | X; DAT |
| | Spatiotemporal modelling | 51–53 | NTD; NFE | X |
| | Multimodal learning | 54,55 | DAT; NFE; NTD | X |
| | Transfer learning | 56,57 | DAT; NFE; NTD | X |
| | Probability density estimation | | | |
| | Kernel density estimation | 41,42 | EP; UQ | NTD |
| | Parametric Bayesian inference | 40 | EP; UQ | NTD |
| | Non-parametric methods | 43,44 | EP; UQ | NTD |
| | Neural uncertainty quantification | 58,59 | UQ; NFE; NTD | X; EP |
| Hybrid | Mechanistic models with statistical components | | | |
| | Data assimilation | 60–62 | X | NTD; NFE |
| | Statistical estimation of mechanistic parameters | 64,66–68,117 | DAT; NTD; X | NFE |
| | Discrepancy modelling | 69–72,118 | DAT; NFE; X | NTD |
| | Priors from mechanistic models inform statistical models | | | |
| | Learning from synthetic and simulation data | 73 | DAT; EP | NTD |
| | Learning with mechanistic constraints | 74,75 | EP; DAT; NTD | ST |
| | Wisdom of crowds | | | |
| | Experts and prediction markets | 76–80,119,120 | MP; EP | NTD; ST |
| | Ensembles | 5,17,18,20,30,81–84,101,121 | MP; EP; NTD | X |

DAT, addresses data sparsity and quality issues; EP, incorporates expert priors and feedback; MP, leverages multiple prediction models; NFE, no manual feature engineering; NTD, incorporates non-traditional data sources; ST, incorporates complex spatial dynamics; UQ, uncertainty quantification; X, explainability.

using word embeddings[36], are popular techniques used to extract intermediate features that are useful for forecasting[36]. Others rely on topic models, which can automatically find topics and parts of the text that are often associated with symptoms and treatments[37]. More recent works also mapped tweets to stages of disease progression (for example, susceptible-infected-recovered), which led to improvements in population-level prediction[38,39].

**Probability density estimation.** These methods estimate the probability distribution of the forecasts, thereby determining their uncertainty, which is important for decision-making. Parametric inference methods, such as empirical Bayes, estimate the parameters of a distribution. They model epidemic curves as probabilistic functions of some specific characteristics selected by experts, including shape, peak height, peak week and pacing[40]. Non-parametric methods, which do not assume a fixed form or distribution for the underlying data, can also incorporate these epidemic curve characteristics as part of their kernel or similarity function. Prominent examples include kernel density estimation[41] and delta density[42], which were successfully used in past CDC flu forecasting initiatives[20]. Other non-parametric methods include Gaussian process[43] and other kernel-based methods to model temporal and spatial correlations[44].

## Deep learning models

Artificial neural networks facilitate the handling of high-dimensional data and offer a flexible, generalizable approach to extract useful representations from various data modalities. Here we present key ideas in deep learning for epidemic forecasting, including modelling spatial–temporal disease dynamics, handling data heterogeneity and multimodality, and taking advantage of transfer learning. We also discuss uncertainty quantification with these models, focusing on integrating multiple data modalities.

**Off-the-shelf sequence models.** These works propose to leverage general-purpose sequential models, such as recurrent neural networks and transformers, to model the temporal evolution of epidemics and incorporate time series of non-traditional data sources such as social media[45], search trends[46] and environmental factors[47].

**Modelling temporal dynamics via similarity.** Several methods have enhanced general-purpose architectures to tackle the challenges posed by limited data by capitalizing on similarities across epidemic dynamics. These works developed clustering-based deep learning architectures to explicitly model temporal similarity, comparing time-series segments across time (for example, past flu seasons with similar patterns to the current season)[48,49] and space (for example, regions with similar epidemic curves)[50].

**Incorporating structured spatial relationships.** Deep learning methods for spatiotemporal modelling aim to capture the underlying spatial propagation of disease. Therefore, in addition to considering the spatial structure of geographical proximity, these models also leverage cross-regional mobility flows. Some early works proposed neural architectures composed of recurrent neural networks and convolutional layers to encode the spatial distribution of locations[51]. More recent methods are based on graph neural networks and recurrent neural networks that capture latent co-evolving dynamics at multiple locations[52,53].

**Leveraging heterogeneous and multimodal data.** One key advantage of deep learning models is their ability to facilitate the learning of knowledge representations from various data modalities. Recent work has explored leveraging multiple static features (for example, demographics, healthcare vulnerability) and dynamic features (for example, indicators of government response)[54], while other studies investigated spatial features beyond proximity and mobility flows, such as travel restrictions[55].

**Transferring knowledge representations.** Transfer learning is another effective solution for challenges associated with limited data, and is particularly valuable in responding to emerging pandemics where data is inherently scarce. For example, a heterogeneous domain transfer learning framework was proposed to adapt a historical influenza forecasting model to the new scenario where COVID-19 and influenza coexist[56]. Another study formulated a meta-learning setting to leverage data from regions initially impacted by the COVID-19 pandemic, aiming to assist with predictions in regions on the verge of being impacted[57].

**Deep learning models for uncertainty quantification.** Uncertainty in epidemic forecasts has been addressed by developing non-parametric deep generative models that leverage similarities in historical data[58]. Specifically, they learn forecast distributions as a probabilistic combination of representations of past time-series segments (for example, flu seasons). Follow-up work extended the idea to multiple sources and modalities, where these are combined dynamically based on their relevance in the accuracy and uncertainty of forecasts[59].

## Hybrid models

Hybrid models bridge the gaps between the domain knowledge embedded in mechanistic models or guidance/predictions from experts, the predictive and pattern-mining capabilities of machine learning and predictions coming from experts, as well as the broader public. These types of model can be assembled in multiple ways, and we broadly classify them into the following types: (1) mechanistic models with statistical components, (2) statistical models informed by mechanistic models and (3) wisdom of crowd models.

### Mechanistic models with statistical components

These methods make inferences using mechanistic models, which are aided by statistical components that address challenges such as facilitating model recalibration, incorporating data features from unconventional sources and accounting for modelling limitations. Here we present multiple ways to accomplish this.

**Data assimilation.** Techniques such as filtering are used to regularly reinitialize and recalibrate mechanistic models using the most recent observations from diverse sources; importantly, these observations must measure a state that is explicitly represented in the mechanistic model. One of the earliest studies on incorporating digital data sources involved integrating Google Flu Trends into forecasts of a susceptible–infectious–recovered–susceptible compartmental flu model[60]. Specifically, they assimilated Google Flu Trends (a proxy measure for 'number of infectious cases') via ensemble adjustment Kalman filters. Follow-up works extended this method to other predictive targets (such as hospitalizations[61]), pathogens[62] and agent-based models[63].

**Estimating parameters of a mechanistic model from features.** These methods can integrate a broader range of data sources than data assimilation techniques. They are designed to alleviate the challenging optimization landscape that arises during the calibration of mechanistic parameters via machine learning modules that ingest the data and inform parameter calibration. Some studies leverage geo-localized tweets to inform the estimation of the initial conditions of mechanistic models[64]. Others differentiate through a numerical solver (for example, Runge Kutta) to calibrate mechanistic parameters, taking into account observations from adjacent geographical regions[65]. One recent line of research proposed an alternative approach through the development of differentiable simulators based on mechanistic models. These differentiable simulators enable end-to-end, gradient-based calibration of the mechanistic model parameters, circumventing the need for traditional numerical solvers. A mass action model (represented as an ordinary differential equation) can be seamlessly converted into a differentiable simulator. The simulator parameters can then be calibrated based on static (for example, demographic, policy) and time-varying (for example, mobility) features ingested by a machine learning module[66,67]. Recent work has extended this idea to agent-based models—which are inherently non-differentiable and orders of magnitude more complex—by combining tensorized operations, invariances in disease transmission and relaxations of discontinuous functions[68].

**Discrepancy modelling.** This line of work acknowledges the limitations of mechanistic models in modelling disease dynamics and proposes to incorporate a machine learning model to refine/correct predictions. Recent efforts introduced a hierarchical Bayesian model that incorporates prior knowledge and learns to correct mechanistic predictions based on trends, seasonal variations and state-specific deviations[69]. The same authors obtained first place in the 2018/2019 CDC FluSight challenge by further developing their model through the imposition of consistency and facilitation of information sharing across scales[70]. Similarly, other methods leverage spatiotemporal dependencies to correct mechanistic predictions of meta-population models[71]. Other studies[72] have proposed that real-time forecasts of any model could be refined via machine learning by leveraging important patterns that these models may have missed during development. Recent work introduced a deep learning architecture based on recurrent graph neural networks to refine predictions based on the temporal dynamics of data revisions[72]; the authors showed that this led to improvements in all of the top five models in the COVID-19 Forecast Hub, composed of both mechanistic and statistical modules.

### Priors from mechanistic models inform statistical models

In these methods, the statistical model incorporates components of mechanistic models as a source of prior knowledge that facilitates learning. One direct way to do this is via the use of high-resolution synthetic data from mechanistic models[73]. Specifically, a deep learning architecture that learns from simulation time-series data generated by a stochastic variant of the susceptible–exposed–infectious–recovered model was proposed. Another approach involves using mechanistic priors as integral components in the design of learning frameworks, and thus informing machine learning models about mechanistic epidemic dynamics. Recent work built on physics-informed neural networks to jointly learn partially observed epidemic dynamics and connect them with heterogeneous datasets[74]. Meanwhile, other studies have incorporated constraints from mechanistic equations into their latent representations, which are learned through tensor factorization[75].

### Wisdom of crowd models

Wisdom of crowd methods harness the collective insights from previously introduced modelling approaches, human predictions and domain heuristics—expertise often excluded from models—to leverage common sense and specialized knowledge. Methods based on expert consensus[76,77] and prediction markets[78] have been developed for epidemic forecasting[79], as well as predicting important variables such as vaccine efficacy, safety, timing and delivery[80].

An adjacent approach involves using ensembles of predictions from multiple models, which are widely used to obtain better predictive performance than their individual components. Indeed, ensemble methods have consistently outperformed most, if not all, individual methods in multiple CDC forecasting competitions—for example, influenza[20], Ebola[4,5] and COVID-19[17]. Aggregating ensemble models based on locations[18], as well as uncertainty[81], model complexity and diversity[82] seems to yield good results. However, a key limitation of using ensembles in real-time forecasting is the availability of forecasts for all models in all instances to gauge their relative importance. This motivated the use of adaptive ensemble approaches[83], using representative clustering[84] or a simple equally weighted average of forecasts[30]. The latter has yielded good results in practice and has been used by the official US CDC ensemble for the COVID-19 pandemic.

## Epidemic forecasting in action

We now describe some important initiatives for real-time epidemic forecasting, such as the CDC FluSight challenge and the COVID-19 Forecast Hub, which have helped to encourage research in this area. Despite their successes, challenges persist—from methodological problems to data quality issues. We also discuss recent methods that strive to integrate these forecasting tools with decision-making to help policymakers.

### Real-time outbreak response

Real-time forecasting initiatives have energized research in epidemic forecasting, fostering knowledge exchange among researchers and translating results into public health tools. One prominent example is the CDC FluSight forecasting challenge for seasonal influenza[15]. In response to COVID-19, the US CDC and partners organized the COVID-19 Forecast Hub[6], which received international attention, and their predictions were publicized on multiple web portals (including the official CDC website). Building on the COVID-19 Forecast Hub, the Scenario Hub focuses on projections for 6 months ahead, conditional on specific scenarios[14]. Among the most important insights from these initiatives is that no modelling approach by any single team was effective in all instances[15,17,20]. The teams that submit real-time forecasts need to deal with issues with operationalizing different methodologies, handling changing disease patterns, varying data availability and data quality issues. Some common problems are data access[20], data biases and inconsistencies across multiple geographies[85], data sparsity and noise[86]. These initiatives received international attention, sparking similar schemes in Europe[7].

### Performance on the ground

As mentioned, the 2013 US CDC FluSight challenge[15] represented a pioneering effort in real-time forecasting. This initiative and others were usually followed by a systematic evaluation of forecasting performance. Comprehensive evaluations showcased the usefulness and maturity of methods for predicting seasonal influenza[20]. These successes spanned several similar initiatives for other diseases such as Ebola and dengue[4,5]. However, the COVID-19 pandemic exposed multiple shortcomings in epidemic forecasting and performance was uneven. There was often a conflation between projections and predictions. These are different types of forecast, as noted in the 'Target, tasks and evaluation' section, and should be analysed separately. Projections and long-term predictions (6 months or longer) were frequently highlighted in the media in early stages of the pandemic; many turning out to be severely incorrect[87], thereby impacting public trust. Short-term predictions (4 weeks ahead) emerged as a more viable alternative. In April 2020, the US CDC consortium COVID-19 Forecast Hub led these efforts and had considerably more success than initial projections, especially with the ensemble of all models. Nevertheless, there were several gaps: sometimes performance during trend changes and uncertainty calibration fell short of expectations[17]. Indeed, the CDC occasionally removed 2- to 4-week-ahead case forecasts from their website due to concerns about well-calibrated uncertainty[88].

### Bridging forecasting and decision-making

Forecasting models have been leveraged to perform strategic and tactical decision-making by policymakers and public health workers. A strategic intervention involves decision-making at a large scale; for example, whether to cull or vaccinate animals against diseases[89] or how to determine optimal policies for lockdowns and travel restrictions[1]. A tactical intervention focuses on achieving a small-scale goal whose action space is large, such as the allocation of ventilators and optimizing for the supply chain constraints of medical supplies[90]. One complementary approach to strategic decision-making is active management[91], which proposes general principles for designing and adapting strategies. It has proved to be effective in applications such as controlling disease outbreaks[92] and developing vaccination strategies[91].

## Open challenges and opportunities

As discussed in this survey, there have been noteworthy advancements in data-centric forecasting. However, the COVID-19 pandemic highlighted that, despite these advances, the scientific community was not fully prepared for a rapidly evolving crisis, and readiness may still be lacking. Achieving an effective pandemic response may require combined efforts across different domains, including data, modelling and evaluation. We now outline our forward-looking perspective on these important challenges.

### Creating principled methods for data quality issues

Developing methods to address data-related issues could be a fruitful direction to improve our forecasting capabilities. We could formulate new statistical problems to improve data quality[72] and explicitly account for biases in data that may affect the fairness of downstream decisions[93]. Other problems arise in datasets containing sensitive information (for example, electronic health records) for which we could exploit advances in federated learning and differential privacy[94]. We emphasize here the importance of efforts to build publicly accessible data archives, such as that introduced by Reinhart et al.[95]. Building such infrastructure can accelerate research on data quality issues.

### Going beyond short-term forecasting

Determining and expanding our forecasting limits remain open challenges[96,97]. To exploit long-term patterns from data, some works have developed methods to bridge mechanistic and AI approaches[74], building on advances in scientific AI[98]. Incorporating scenario-based projections, causal inference and causal representation learning to connect multimodal data with interventions[99] is another promising research direction. On the spatial scale, we can develop more principled methods that use spatial dynamics and hierarchical relationships. Other orthogonal aspects that provide useful higher-order patterns include multi-scale behavioural models[2], the evolution (phylodynamics) of pathogens[100] and other biological indicators. Towards this goal, developing methods to leverage multilevel dynamics and relations at multiple scales could be another fruitful direction.

### Improving the combination of models and wisdom of crowds predictions

Combining ensembles of models to maximize their strengths remains difficult. The proficiency of a single model varies across temporal and spatial scales[17]. New models may be introduced, and some teams may change their methodologies or not publish predictions for a while[101]. In most studies, multiple weighting schemes are considered, but there is no consensus on which perform best. In this regard, it is important to explore new schemes using techniques such as optimal ensemble weighting and mixtures of experts[102]. Overcoming inefficiencies and misinformation[103], as well as developing better interfaces that leverage advances in human–computer interactions to elicit detailed and reliable human predictions for prediction markets, are pressing open challenges.

### Making forecasts explainable and actionable

For high-stakes decisions in public health, calibrated forecasts are crucial, but they can be challenging because there are multiple sources of uncertainty. Recent work has shown the usefulness of taking data-driven representations of uncertainty[58] from multiple sources into account. The explainability of predictions to domain experts is an important aspect to bridge forecasts and decision-making. Linear regression and mechanistic approaches can easily provide explanations, but other models (for example, artificial neural nets) require more sophistication to provide explanations[48,58,86]. In addition, the forecasting set-up and evaluation need to be constantly scrutinized to make forecasts suitable for public communication and decision-making. One possible direction could be defining new forecast targets[104]. For error metrics, we need a better understanding of which are most important for each decision type, and how to compare them across temporal and spatial scales. It is also necessary to standardize evaluation guidelines for presenting methodological advances[105].

### Evaluating ad hoc adjustments of real-world model deployment

The real-world deployment of forecasting systems requires a non-trivial amount of human involvement in (for example) handling anomalous data and predictions that look incorrect to the expert eye[85,86]. Other examples can be found when calibrating mechanistic models—one might need to adjust domain inputs, such as setting boundaries for parameter optimization[106]. Such interventions are crucial to the success of models in real-world situations, such as the epidemic prediction challenges discussed in the 'Real-time outbreak response' section. However, these critical interventions are not always formally accounted for, and contribute to the technical debt (the future costs arising from inadequate solutions to current issues). It is therefore important to find ways to disentangle the main factors that affect a model or team's performance, such as the methodology, ad hoc expert adjustments and technical debt. Otherwise, we may be uncertain about the true measurement of prediction challenges, which could compromise the generalizability and robustness of future pandemic responses.

To conclude, this Review provides an overview of data-centric methods for epidemic forecasting. We have highlighted recent data sources (the 'Data-centric epidemic forecasting' section) and discussed key approaches that utilize them across statistical, deep learning and hybrid models (the 'Statistical, machine learning and AI models' and 'Hybrid models' sections). We have also discussed the current capabilities and preparedness of the scientific and public health communities for responding to emerging epidemics like COVID-19 (the 'Epidemic forecasting in action' section). Despite the considerable progress made, numerous challenges across multiple domains remain that demand further attention, some of which are listed above. Addressing these challenges has the potential to substantially improve forecasting systems, ensure their extensibility to future data sources and ultimately enhance our preparedness for future crises.

## References


1. Holmdahl, I. & Buckee, C. Wrong but useful—what COVID-19 epidemiologic models can and cannot tell us. *N. Engl. J. Med.* **383**, 303–305 (2020).
2. Marathe, M. & Vullikanti, A. K. S. Computational epidemiology. *Commun. ACM* **56**, 88–96 (2013).
3. Biggerstaff, M., Cauchemez, S., Reed, C., Gambhir, M. & Finelli, L. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. *BMC Infect. Dis.* **14**, 480 (2014).
4. Viboud, C. et al. The RAPIDD ebola forecasting challenge: synthesis and lessons learnt. *Epidemics* **22**, 13–21 (2018).
5. Johansson, M. A., Apfeldorf, K. M., Dobson, S. & Devita, J. et al. An open challenge to advance probabilistic forecasting for dengue epidemics. *Proc. Natl Acad. Sci. USA* **116**, 24268–24274 (2019).
6. Cramer, E. Y. et al. The United States COVID-19 Forecast Hub dataset. *Sci. Data* **9**, 462 (2022).
7. Sherratt, K. et al. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations. *eLife* **12**, e81916 (2023).
8. Aktay, A. et al. *Google COVID-19 Community Mobility Reports* (Google, accessed 15 May 2024); https://www.google.com/covid19/mobility
9. Astley, C. M. et al. Global monitoring of the impact of the COVID-19 pandemic through online surveys sampled from the Facebook user base. *Proc. Natl Acad. Sci. USA* **118**, e2111455118 (2021).
10. Peccia, J. et al. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat. Biotechnol.* **38**, 1164–1167 (2020).
11. Biggerstaff, M. et al. Coordinating the real-time use of global influenza activity data for better public health planning. *Influenza Other Respir. Virus.* **14**, 105–110 (2020).
12. Butler, P., Ramakrishnan, N., Nsoesie, E. O. & Brownstein, J. S. Satellite imagery analysis: what can hospital parking lots tell us about a disease outbreak? *IEEE Ann. Hist. Comput.* **47**, 94–97 (2014).
13. Miliou, I. et al. Predicting seasonal influenza using supermarket retail records. *PLoS Comput. Biol.* **17**, e1009087 (2021).
14. Borchering, R. K. et al. Public health impact of the US Scenario Modeling Hub. *Epidemics* **44**, 100705 (2023).
15. Biggerstaff, M. et al. Results from the centers for disease control and prevention's predict the 2013-2014 Influenza Season Challenge. *BMC Infect. Dis.* **16**, 357 (2016).
16. Chakraborty, P. et al. What to know before forecasting the flu. *PLoS Comput. Biol.* **14**, e1005964 (2018).
17. 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).
18. Kandula, S. et al. Evaluation of mechanistic and statistical methods in forecasting influenza-like illness. *J. R. Soc. Interf.* **15**, 20180174 (2018).
19. Kandula, S. & Shaman, J. Near-term forecasts of influenza-like illness: an evaluation of autoregressive time series approaches. *Epidemics* **27**, 41–51 (2019).
20. Reich, N. G. et al. A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States. *Proc. Natl Acad. Sci. USA* **116**, 3146–3154 (2019).
21. Bracher, J. On the multibin logarithmic score used in the FluSight competitions. *Proc. Natl Acad. Sci. USA* **116**, 20809–20810 (2019).
22. Gneiting, T. & Raftery, A. E. Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.* **102**, 359–378 (2007).

23. Hethcote, H. W. The mathematics of infectious diseases. *SIAM Rev.* **42**, 599–653 (2000).
24. Viboud, C. & Vespignani, A. The future of influenza forecasts. *Proc. Natl Acad. Sci. USA* **116**, 2802–2804 (2019).
25. Polgreen, P. M., Chen, Y., Pennock, D. M., Nelson, F. D. & Weinstein, R. A. Using internet searches for influenza surveillance. *Clin. Infect. Dis.* **47**, 1443–1448 (2008).
26. Ginsberg, J. et al. Detecting influenza epidemics using search engine query data. *Nature* **457**, 1012–1014 (2009).
27. Culotta, A. Towards detecting influenza epidemics by analyzing Twitter messages. In *Proc. First Workshop on Social Media Analytics* 115–122 (ACM, 2010).
28. Yang, S., Santillana, M. & Kou, S. C. Accurate estimation of influenza epidemics using Google search data via ARGO. *Proc. Natl Acad. Sci. USA* **112**, 14473–14478 (2015).
29. Ning, S., Yang, S. & Kou, S. C. Accurate regional influenza epidemics tracking using Internet search data. *Sci. Rep.* **9**, 5238 (2019).
30. Ray, E. L. & Reich, N. G. Prediction of infectious disease epidemics via weighted density ensembles. *PLoS Comput. Biol.* **14**, e1005910 (2018).
31. Chakraborty, P. et al. Forecasting a moving target: ensemble models for ILI case count predictions. In *Proc. 2014 SIAM International Conference on Data Mining* (eds Zaki, M. et al.) 262–270 (SIAM, 2014).
32. Zou, B., Lampos, V. & Cox, I. Multi-task learning improves disease models from web search. In *Proc. 2018 World Wide Web Conference* 87–96 (IW3C2, 2018).
33. Matsubara, Y., Sakurai, Y., Van Panhuis, W. G. & Faloutsos, C. FUNNEL: automatic mining of spatially coevolving epidemics. In *Proc. 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 105–114 (ACM, 2014).
34. Wang, Z. et al. Dynamic poisson autoregression for influenza-like-illness case count prediction. In *Proc. 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1285–1294 (ACM, 2015).
35. Lamb, A., Paul, M. J. & Dredze, M. Separating fact from fear: tracking flu infections on Twitter. In *Proc. 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (eds Vanderwende, L. et al.) 789–795 (ACL, 2013).
36. Zou, B., Lampos, V. & Cox, I. Transfer learning for unsupervised influenza-like illness models from online search data. In *Proc. 2019 World Wide Web Conference* (eds Liu, L. & White, R.) 2505–2516 (ACM, 2019).
37. Paul, M. J. & Dredze, M. A model for mining public health topics from Twitter. *Johns Hopkins University* https://www.cs.jhu.edu/~mdredze/publications/2011.tech.twitter_health.pdf (2011).
38. Chen, L., Tozammel Hossain, K. S. M., Butler, P., Ramakrishnan, N. & Prakash, B. A. Flu gone viral: syndromic surveillance of flu on Twitter using temporal topic models. In *2014 IEEE International Conference on Data Mining* 755–760 (IEEE, 2014).
39. Rekatsinas, T. et al. SourceSeer: forecasting rare disease outbreaks using multiple data sources. In *Proc. 2015 SIAM International Conference on Data Mining* (eds Venkatasubramanian, S. & Ye, J.) 379–387 (SIAM, 2015).
40. Brooks, L. C., Farrow, D. C., Hyun, S., Tibshirani, R. J. & Rosenfeld, R. Flexible modeling of epidemics with an empirical Bayes framework. *PLoS Comput. Biol.* **11**, e1004382 (2015).
41. Ray, E. L., Sakrejda, K., Lauer, S. A., Johansson, M. A. & Reich, N. G. Infectious disease prediction with kernel conditional density estimation. *Stat. Med.* **36**, 4908–4929 (2017).
42. Brooks, L. C., Farrow, D. C., Hyun, S., Tibshirani, R. J. & Rosenfeld, R. Nonmechanistic forecasts of seasonal influenza with iterative one-week-ahead distributions. *PLoS Comput. Biol.* **14**, e1006134 (2018).
43. Zimmer, C. & Yaesoubi, R. Influenza forecasting framework based on Gaussian processes. In *Proc. 37th International Conference on Machine Learning* (eds Daumé, H. III & Singh, A.) 11671–11679 (PMLR, 2020).
44. Senanayake, R., O'Callaghan, S. & Ramos, F. Predicting spatio–temporal propagation of seasonal influenza using variational Gaussian process regression. In *Proc. 30th AAAI Conference on Artificial Intelligence* 3901–3907 (AAAI, 2016).
45. Volkova, S., Ayton, E., Porterfield, K. & Corley, C. D. Forecasting influenza-like illness dynamics for military populations using neural networks and social media. *PLoS ONE* **12**, e0188941 (2017).
46. Ayyoubzadeh, S. M., Ayyoubzadeh, S. M., Zahedi, H., Ahmadi, M. & Kalhori, S. R. N. Predicting COVID-19 incidence through analysis of google trends data in Iran: data mining and deep learning pilot study. *JMIR Publ. Health Surveill.* **6**, e18828 (2020).
47. Venna, S. R. et al. A novel data-driven model for real-time influenza forecasting. *IEEE Access* **7**, 7691–7701 (2018).
48. Adhikari, B., Xu, X., Ramakrishnan, N. & Prakash, B. A. EpiDeep: exploiting embeddings for epidemic forecasting. In *Proc. 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 577–586 (ACM, 2019).
49. Wang, L. et al. Examining deep learning models with multiple data sources for COVID-19 forecasting. In *2020 IEEE International Conference on Big Data (Big Data)* 3846–3855 (IEEE, 2020).
50. Jin, X., Wang, Y.-X. & Yan, X. Inter-series attention model for COVID-19 forecasting. In *Proc. 2021 SIAM International Conference on Data Mining* (eds Demeniconi, C. et al.) 495–503 (SIAM, 2021).
51. Wu, Y., Yang, Y., Nishiura, H. & Saitoh, M. Deep learning for epidemiological predictions. In *41st International ACM SIGIR Conference on Research & Development in Information Retrieval* 1085–1088 (ACM, 2018).
52. 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).
53. Roy, P. et al. Deep diffusion-based forecasting of COVID-19 by incorporating network-level mobility information. In *Proc. 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (eds Coscia, M. et al.) 168–175 (ACM, 2021).
54. Ibrahim, M. R. et al. Variational-LSTM autoencoder to forecast the spread of coronavirus across the globe. *PLoS ONE* **16**, e0246120 (2021).
55. Ramchandani, A., Fan, C. & Mostafavi, A. DeepCOVIDNet: an interpretable deep learning model for predictive surveillance of COVID-19 using heterogeneous features and their interactions. *IEEE Access* **8**, 159915–159930 (2020).
56. Rodríguez, A. et al. Steering a historical disease forecasting model under a pandemic: case of flu and COVID-19. In *Proc. 35th AAAI Conference on Artificial Intelligence* 4855–4863 (AAAI, 2021).
57. Panagopoulos, G., Nikolentzos, G. & Vazirgiannis, M. Transfer graph neural networks for pandemic forecasting. In *Proc. 35th AAAI Conference on Artificial Intelligence* 4838–4845 (AAAI, 2021).
58. Kamarthi, H., Kong, L., Rodríguez, A., Zhang, C. & Prakash, B. A. When in doubt: neural non-parametric uncertainty quantification for epidemic forecasting. In *Proc. 35th Conference on Neural Information Processing Systems* (eds Ranzato, M. et al.) 19796–19807 (NeurIPS, 2021).
59. Kamarthi, H., Kong, L., Rodríguez, A., Zhang, C. & Prakash, B. A. CAMul: calibrated and accurate multi-view time-series forecasting. In *Proc. ACM Web Conference 2022* (eds Laforest, F. et al.) 3174–3185 (ACM, 2022).
60. Shaman, J. & Karspeck, A. Forecasting seasonal outbreaks of influenza. *Proc. Natl Acad. Sci. USA* **109**, 20425–20430 (2012).

61. Kandula, S., Pei, S. & Shaman, J. Improved forecasts of influenza-associated hospitalization rates with Google Search Trends. *J. R. Soc. Interf.* **16**, 20190080 (2019).
62. Pei, S. & Shaman, J. Aggregating forecasts of multiple respiratory pathogens supports more accurate forecasting of influenza-like illness. *PLoS Comput. Biol.* **16**, e1008301 (2020).
63. Tabataba, F. S. et al. Epidemic forecasting framework combining agent-based models and smart beam particle filtering. In *2017 IEEE International Conference on Data Mining* 1099–1104 (IEEE, 2017).
64. Zhang, Q. et al. Forecasting seasonal influenza fusing digital indicators and a mechanistic disease model. In *Proc. 26th International Conference on World Wide Web* 311–319 (ACM, 2017).
65. Wang, R., Maddix, D., Faloutsos, C., Wang, Y. & Yu, R. Bridging physics-based and data-driven modeling for learning dynamical systems. In *Proc. 3rd Conference on Learning for Dynamics and Control* (eds Jadbabaie, A. et al.) 385–398 (PMLR, 2021).
66. Arık, S. Ö. et al. A prospective evaluation of AI-augmented epidemiology to forecast COVID-19 in the USA and Japan. *npj Digit. Med.* **4**, 1–18 (2021).
67. Qian, Z., Alaa, A. M. & van der Schaar, M. When and how to lift the lockdown? Global COVID-19 scenario analysis and policy assessment using compartmental Gaussian processes. In *Proc. 34th Conference on Neural Information Processing Systems* 10729–10740 (NeurIPS, 2020).
68. Chopra, A. et al. Differentiable agent-based epidemiology. In *Proc. 2023 International Conference on Autonomous Agents and Multiagent Systems* 1848–1857 (International Foundation for Autonomous Agents and Multiagent Systems, 2023).
69. Osthus, D., Gattiker, J., Priedhorsky, R. & Del Valle, S. Y. et al. Dynamic Bayesian influenza forecasting in the United States with hierarchical discrepancy (with discussion). *Bayes. Anal.* **14**, 261–312 (2019).
70. Osthus, D. & Moran, K. R. Multiscale influenza forecasting. *Nat. Commun.* **12**, 2991 (2021).
71. Wu, D. et al. DeepGLEAM: a hybrid mechanistic and deep learning model for COVID-19 forecasting. Preprint at https://arxiv.org/abs/2102.06684 (2021).
72. Kamarthi, H., Rodríguez, A. & Prakash, B. A. Back2Future: leveraging backfill dynamics for improving real-time predictions in future. In *Proc. Tenth International Conference on Learning Representations* (ICLR, 2022).
73. Wang, L., Chen, J. & Marathe, M. DEFSI: deep learning based epidemic forecasting with synthetic information. In *Proc. 33rd AAAI Conference on Artificial Intelligence* 9607–9612 (AAAI, 2019).
74. Rodríguez, A., Cui, J., Ramakrishnan, N., Adhikari, B. & Prakash, B. A. EINNs: epidemiologically-informed neural networks. In *Proc. 37th AAAI Conference on Artificial Intelligence* (eds Williams, B. et al.) 14453–14460 (AAAI, 2023).
75. Kargas, N. et al. STELAR: spatio-temporal tensor factorization with latent epidemiological regularization. In *Proc. 35th AAAI Conference on Artificial Intelligence* 4830–4837 (AAAI, 2021).
76. Recchia, G., Freeman, A. L. & Spiegelhalter, D. How well did experts and laypeople forecast the size of the COVID-19 pandemic? *PLoS ONE* **16**, e0250935 (2021).
77. Shea, K. et al. Harnessing multiple models for outbreak management. *Science* **368**, 577–579 (2020).
78. Polgreen, P. M., Nelson, F. D., Neumann, G. R. & Weinstein, R. A. Use of prediction markets to forecast infectious disease activity. *Clin. Infect. Dis.* **44**, 272–279 (2007).
79. Farrow, D. C. et al. A human judgment approach to epidemiological forecasting. *PLoS Comput. Biol.* **13**, e1005248 (2017).
80. McAndrew, T., Cambeiro, J. & Besiroglu, T. Aggregating human judgment probabilistic predictions of the safety, efficacy, and timing of a COVID-19 vaccine. *Vaccine* **40**, 2331–2341 (2022).
81. Reich, N. G. et al. Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the US. *PLoS Comput. Biol.* **15**, e1007486 (2019).
82. Adiga, A. et al. All models are useful: Bayesian ensembling for robust high resolution COVID-19 forecasting. In *Proc. 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* 2505–2513 (ACM, 2021).
83. McAndrew, T. & Reich, N. G. Adaptively stacking ensembles for influenza forecasting. *Stat. Med.* **40**, 6931–6952 (2021).
84. Kim, J.-S., Kavak, H., Züfle, A. & Anderson, T. COVID-19 ensemble models using representative clustering. *SIGSPATIAL Special* **12**, 33–41 (2020).
85. Altieri, N. et al. Curating a COVID-19 data repository and forecasting county-level death counts in the United States. *Harv. Data Sci. Rev.* https://doi.org/10.1162/99608f92.1d4e0dae (2021).
86. Rodríguez, A. et al. DeepCOVID: an operational deep learning-driven framework for explainable real-time COVID-19 forecasting. In *Proc. 35th AAAI Conference on Artificial Intelligence* 15393–15400 (AAAI, 2021).
87. Ferguson, N. M. et al. *Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand* (Imperial College COVID-19 Response Team, 2020).
88. Reich, N., Tibshirani, R., Ray, E. & Rosenfeld, R. On the predictability of COVID-19. *IIF Blog* https://forecasters.org/blog/2021/09/28/on-the-predictability-of-covid-19 (2021).
89. Probert, W. J. et al. Real-time decision-making during emergency disease outbreaks. *PLoS Comput. Biol.* **14**, e1006202 (2018).
90. Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C. & Vasilakis, C. Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions. *Euro. J. Oper. Res.* **290**, 99–115 (2021).
91. Atkins, B. D. et al. Anticipating future learning affects current control decisions: a comparison between passive and active adaptive management in an epidemiological setting. *J. Theor. Biol.* **506**, 110380 (2020).
92. Shea, K., Tildesley, M. J., Runge, M. C., Fonnesbeck, C. J. & Ferrari, M. J. Adaptive management and the value of information: learning via intervention in epidemiology. *PLoS Biol.* **12**, e1001970 (2014).
93. Mhasawade, V., Zhao, Y. & Chunara, R. Machine learning and algorithmic fairness in public and population health. *Nat. Mach. Intell.* **3**, 659–666 (2021).
94. Zhang, C. et al. A survey on federated learning. *Knowl. Based Syst.* **216**, 106775 (2021).
95. Reinhart, A. et al. An open repository of real-time COVID-19 indicators. *Proc. Natl Acad. Sci. USA* **118**, e2111452118 (2021).
96. Scarpino, S. V. & Petri, G. On the predictability of infectious disease outbreaks. *Nat. Commun.* **10**, 898 (2019).
97. Rosenkrantz, D. J. et al. Fundamental limitations on efficiently forecasting certain epidemic measures in network models. *Proc. Natl Acad. Sci. USA* **119**, e2109228119 (2022).
98. Karniadakis, G. E. et al. Physics-informed machine learning. *Nat. Rev. Phys.* **3**, 422–440 (2021).
99. Mastakouri, A. & Schölkopf, B. Causal analysis of COVID-19 Spread in Germany. *Adv. Neur. Inf. Process. Syst.* **33**, 3153–3163 (2020).
100. Kraemer, M. U. et al. Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence. *Science* **373**, 889–895 (2021).
101. Ray, E. L. et al. Ensemble forecasts of coronavirus disease 2019 (COVID-19) in the U.S. Preprint at *medRxiv* https://doi.org/10.1101/2020.08.19.20177493 (2020).
102. Riquelme, C. et al. Scaling vision with sparse mixture of experts. In *Proc. 35th Conference on Neural Information Processing Systems* (eds Ranzato, M. et al.) 8583–8595 (NeurIPS, 2021).

103. Angelini, G., De Angelis, L. & Singleton, C. Informational efficiency and behaviour within in-play prediction markets. *Int. J. Forecast.* **38**, 282–299 (2022).
104. Lutz, C. S. et al. Applying infectious disease forecasting to public health: a path forward using influenza forecasting examples. *BMC Public Health* **19**, 1659 (2019).
105. Pollett, S. et al. Recommended reporting items for epidemic forecasting and prediction research: the EPIFORGE 2020 guidelines. *PLoS Med.* **18**, e1003793 (2021).
106. Gibson, G. C., Reich, N. G. & Sheldon, D. Real-time mechanistic Bayesian forecasts of COVID-19 mortality. *Ann. Appl. Stat.* **17**, 1801–1819 (2023).
107. Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* **395**, 689–697 (2020).
108. Edeling, W. et al. The impact of uncertainty on predictions of the CovidSim epidemiological code. *Nat. Comput. Sci.* **1**, 128–135 (2021).
109. Balcan, D. et al. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl Acad. Sci. USA* **106**, 21484–21489 (2009).
110. Pei, S., Kandula, S., Yang, W. & Shaman, J. Forecasting the spatial transmission of influenza in the United States. *Proc. Natl Acad. Sci. USA* **115**, 2752–2757 (2018).
111. Chang, S. et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2021).
112. Gopalakrishnan, V. et al. Globally local: hyper-local modeling for accurate forecast of COVID-19. *Epidemics* **37**, 100510 (2021).
113. Geng, X. et al. A kernel-modulated SIR model for COVID-19 contagious spread from county to continent. *Proc. Natl Acad. Sci. USA* **118**, e2023321118 (2021).
114. Santillana, M., Nsoesie, E. O., Mekaru, S. R., Scales, D. & Brownstein, J. S. Using clinicians' search query data to monitor influenza epidemics. *Clin. Infect. Dis.* **59**, 1446–1450 (2014).
115. Soebiyanto, R. P., Adimi, F. & Kiang, R. K. Modeling and predicting seasonal influenza transmission in warm regions using climatological parameters. *PLoS ONE* **5**, e9450 (2010).
116. Paul, M. & Dredze, M. You are what you tweet: analyzing Twitter for public health. In *Proc. 5th International AAAI Conference on Weblogs and Social Media* 265–272 (AAAI, 2011).
117. Ghamizi, S. et al. Data-driven simulation and optimization for COVID-19 exit strategies. In *Proc. 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 3434–3442 (ACM, 2020).
118. Osthus, D. Fast and accurate influenza forecasting in the United States with Inferno. *PLoS Comput. Biol.* **18**, e1008651 (2022).
119. Nadella, P., Swaminathan, A. & Subramanian, S. Forecasting efforts from prior epidemics and COVID-19 predictions. *Euro. J. Epidemiol.* **35**, 727–729 (2020).
120. Hemming, V., Burgman, M. A., Hanea, A. M., McBride, M. F. & Wintle, B. C. A practical guide to structured expert elicitation using the IDEA protocol. *Methods Ecol. Evol.* **9**, 169–180 (2018).
121. Viboud, C., Boëlle, P.-Y., Carrat, F., Valleron, A.-J. & Flahault, A. Prediction of the spread of influenza epidemics by the method of analogues. *Am. J. Epidemiol.* **158**, 996–1006 (2003).


## Acknowledgements

This work was supported in part by the National Science Foundation (grant numbers Expeditions CCF-1918770, CAREER IIS-2028586, RAPID IIS-2027862, Medium IIS-1955883, Medium IIS-2106961, CCF-2115126 and PIPP CCF-2200269), the CDC MInD programme, the ORNL, faculty research awards from Facebook and funds/computing resources from Georgia Tech.

## Author contributions

A.R., H.K. and B.A.P. contributed to the conceptualization of the manuscript. All authors contributed to gathering, analysing and interpreting the literature. P.A., J.H., M.P. and S.S. contributed to the development of Figs. 1 and 2. A.R., H.K. and B.A.P. contributed to the writing of all sections.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s42256-024-00895-7.

**Correspondence and requests for materials** should be addressed to Alexander Rodríguez, Harshavardhan Kamarthi or B. Aditya Prakash.

**Peer review information** *Nature Machine Intelligence* thanks Sen Pei and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2024