Journal of Progressive Medical Sciences
ISSN: 3052-8518
J.P.M.S
DOI: https://doi.org/10.63939/q6smc190

Vol. 2 , No. 6 , August 2026

RESEARCH Open Access

The Algorithmic Virologist: Integrating Artificial Intelligence Across Viral Detection, Diagnosis, Surveillance, Prediction, and Therapeutics

Saif Jabbar Yasir

Department of Medical Microbiology, College of Medicine, University of Kufa, Najaf, Iraq. saif.alshehmani@uokufa.edu.iq

Abstract

Artificial intelligence is rapidly revolutionizing medical virology, yet its applications remain fragmented, ranging from virus detection and clinical diagnosis to epidemiological prediction, variant monitoring, and treatment development. This review provides a comprehensive overview of AI principles, from basic genomic biology and drug discovery to clinical physiology, and proposes the generation of integrated knowledge, under human supervision, through an "algorithmic virologist" as a conceptual framework to bridge these currently disparate fields. The research assesses deep learning, machine learning, hybrid computational approaches and recent multimodal and generative AI systems with explicit emphasis on their biological validity, diagnostic performance, generalizability, interpretability and translational ability. Current evidence shows that AI can improve reference-free detection of viral sequences, facilitate and inform multimodal clinical diagnosis, combine genomic and epidemiological signals for outbreak and variant prediction, speed up anti-viral target identification and virtual screening and help optimize therapeutics. Yet the review also points out an enduring divide – between computational prowess and real-world clinical utility. Heterogeneous and geographically non-representative datasets, small reference databases of viruses, lack of external and prospective validation, algorithmic black-boxing, potential model drift due to viral and epidemiological changes over time, inadequate interoperability with clinical technologies or other data in use at the bedside or by public health practitioners (and inadequate access at scale) are major barriers to

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

implementation. Consequently, the study proposes that the progress that follows cannot be based on predictive accuracy alone but also needs to keep in mind biological plausibility, external and temporal validity, assessment of uncertainty, explainability; reproducibility; equity and measurable clinical or public-health benefit. Accordingly, it is proposed that the algorithmic virologist is not just a medicine for human knowledge, but an integrated computer ecosystem that links detection, diagnosis, characterization, prediction, monitoring, therapeutic discovery, therapeutic improvement, and continuous verification. Responsibly integrated AI, the review concludes, could facilitate a transition from largely reactive to predictive and anticipatory virology, if they are biologically grounded algorithmic innovations that are experimentally validated, clinically accountable and meaningfully overseen by humans.

Keywords: Artificial Intelligence; Algorithmic Virologist; Machine Learning; Viral Detection; Clinical Diagnosis; Variant Surveillance; Epidemic Forecasting; Antiviral Discovery; Precision Virology; Predictive Virology.

Introduction

Viral diseases are still some of the most active and greatest global health hazards. The exceptional genetic diversity, mutation and recombination rates, interspecies transmission potential, immune evasion capacity, and ability to adapt to host and environmental changes of these pathogens continuously challenge traditional strategies for the detection of pathogens in clinical diagnosis, epidemiological surveillance as well as therapeutic development. As evident from recent outbreaks, control of viral diseases increasingly rests not just on the observation of pathogens after disease has occurred, but also on defining and predicting which viruses will gain prominence in causing widespread clinical and public-health problems. Meanwhile, contemporary virology is in an unprecedented era of data generation. Data assimilation from high-throughput sequencing, metagenomics, proteomics, structural biology, electronic health records and medical imaging along with new microbiome data types such as genomic surveillance of community distributed pathogens, mobility data and digital epidemiology generate multidimensional datasets whose scale and complexity render the incisiveness of traditional population level analytical approaches incapable. The central challenge thus is moving away from biological data capturing into the rapid integration of it and resulting clinically meaningful interpretation in this heterogeneous environment.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

Descriptive models: This challenge can leverage powerful computational approaches provided by artificial intelligence (AI). Abstract: V genomic data (Ge)Big C Pathomics Machine learning neural networks multimodal G machine, Deep learning models and emerging multimodal machine learning (MML) frameworks are increasingly designed to integrate and analyze abundant, heterogeneous, and high-dimensional multimodal data generated across massive datasets, enabling the identification of complex, hidden, and non-linear patterns that remain challenging to detect using conventional analytical approaches. Deep-learning methods overcome traditional sequence-alignment limitations to enable recognition of genetically-distinct pathogens [1], and the inclusion of these tools may enhance categorization of exhaustively sampled viral genomic sequences in viral genomics. In some cases, AI models extract biologically informative relations even from genomic data itself such as those of probable virus–host associations [2]. In addition to sequence analysis, AI is being adopted for multimodal clinical diagnosis, genomic and epidemiological surveillance, outbreak forecasting, mutation-effect estimation, variant prioritization, immune-escape prediction, virtual screening and drug repurposing. Drug discovery Molecular design Machine learning (ML), a subclass of artificial intelligence (AI) that can learn from experience without explicit programming [2], has made invaluable contributions to addressing COVID19 disease in the past 3 years. Together, these applications suggest that AI is suited to aspects of viral disease across the full trajectory rather than isolated computational tasks. Broadening this role is essential because viral disease itself is inherently multidimensional. Host immunity, clinical characteristics/environmental conditions, population behavior/epidemiological dynamics and therapeutic pressures all interact with viral genetics and protein structure. Thus, a naive algorithm that is constrained to using data from just a single modality may only extract one factor of the biological system it ultimately intends on replicating. Coupling genomic, structural, clinical, epidemiological and therapeutic data allows us to reproduce the disconnected prediction tasks as a single but much more comprehensive computational representation of viral disease.

But computational sophistication does not equal clinical readiness at all Most AI studies report high accuracy, sensitivity, specificity or area under the receiver operating characteristic curve in retrospective or internally validated settings, while significantly fewer show similar performance across independent institutions, diverse geography, sequencing platforms, viral lineages and changing epidemiological conditions. This

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

distinction is particularly relevant in virology as both pathogens and host populations evolve through time. Thus, dataset imbalance (37), incomplete viral reference repositories (38), algorithmic opacity (39), insufficient external validation (40) and model drift could easily render what appeared to be high-performing systems into clinically brittle ones. These limitations also highlight a more general issue regarding methodology: predictive performance is not synonymous with being biologically or clinically valid. A model can predict a viral sequence, phenotype, rank an emerging variant or therapeutic candidate without explaining the biological mechanism driving its prediction and/or providing evidence of clinical benefit. Therefore, we propose that any AI system developed for medical virology should now be judged on the basis of predictive accuracy but also biological plausibility, external and temporal validity, uncertainty quantification, interpretability, reproducibility, equity and clinical or public-health utility. Since viral evolution, immunity, interventions and population behavior continually change the environment in which algorithms operate, validation should also be viewed as a lifecycle process entailing continuous monitoring, recalibrating and evaluation of model drift.

While there have been rapid advances, the literature on AI in virology remains primarily siloed among viral detection, clinical diagnosis, epidemic forecasting, genomic and variant surveillance and antiviral discovery. Such compartmentalization gives important insights about specific applications, but fails to account for how these capabilities may function together across the viral disease continuum. Furthermore, an ongoing translational burden persists between exemplar computational prowess and proven biological authenticity, external generalizability and tangible clinical or public-health utility. In response to such fragmentation, this review offers the “Algorithmic Virologist” (AlgVi), an integrative human-supervised conceptual framework that connects: Detection → Diagnosis → Characterization → Prediction → Surveillance →Therapeutic Discovery →Therapeutic Optimization → Continuous Validation.

Instead of an autonomous artificial system, and instead of a replacement for virologists or clinicians, the Algorithmic Virologist imagines AI as an interdependent computational ecosystem where molecular, clinical, epidemiological and therapeutic algorithms interact in the context of continual experimental & human feedback loop.

This review evaluates artificial intelligence through four interrelated aspects of the viral disease continuum—detection, diagnosis, prediction and therapeutics—and critically assesses the leading computational paradigms along with their biological and clinical

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

functionalities as well as methodological barriers to translation. We further focus on dataset heterogeneity, generalizability, external and prospective validation, explainability, uncertainty, interoperability, privacy and regulatory oversight as well as model drift. This review reframes the question of what AI can or cannot do with virology to one of under what scientific, clinical, methodological and ethical conditions (as opposed to merely technical limits) it becomes a trusted part of virological practice by integrating these traditionally distinct domains into a single analytical framework. It thus suggests four interrelated metrics to monitor progress toward a predictive and anticipatory virology—biological validity, clinical utility, algorithmic transparency, and human accountability—and offers an intellectual framework to facilitate the responsible development of next-generation AI-supported virology.

Research Gap

Despite the rapid development of virology applications, the existing literature remains overwhelmingly fragmented, application-specific, and technology-oriented. However, most studies assess artificial intelligence in the context of discrete domains—be it viral sequence classification, clinical diagnosis, epidemic modelling and forecasting, variant prediction or antiviral drug discovery—separately and do not consider synergistic effects of these applications through the entire viral disease continuum. As a result, the conceptual and translation gap remains large between developing stand-alone high performing algorithms for individual tasks of virus detection and therapy. A second major gap is the divergence between computational performance and clinical validity. Although many of the models reported show high accuracy, sensitivity, specificity or predictive performance in retrospective or internally validated settings, far fewer studies assess whether such performance can be sustained across independent populations, healthcare systems, geographic regions, sequencing platforms (where applicable), viral variants and changing epidemiological conditions. As such, it is unclear whether algorithmic accuracy can yet be consistently equated with clinical utility.

The third gap concerns the data representativeness and algorithmic generalizability. AI models in virology are often trained upon datasets that are under- or over-sampled based on population, geography, pathogen, and health care setting. Another reason for systemic bias in model predictions could be the unrepresentative populations from low-resource settings and insufficiently characterized viral populations. This shortcoming is especially important in infectious diseases, since the areas most needed to monitor

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

novel pathogens may also be regions with low genomic and digital surveillance capacity.

The fourth gap is that integration of multimodal biological and clinical information has been far from adequate. Abstract: Viral infections arise from dynamic networks of interactions between viral genomes, protein structures in competing viruses, the host's immune repertoire and clinical features, the environment (higher copy number), population mobility and therapeutic pressures. However most existing AI model is still limited to the unimodal data. Current Systems Does Not Yet Capable to Recreate This Multidimensional Nature of Viral Disease Due To Lack of Integrated Genomic, Proteomic, Immunological, Clinical, Epidemiological and Pharmacogenomics Model [49].

Fifth, there remains the unresolved issue of explainability/uncertainty and biological interpretability. The most sophisticated deep-learning and generative models might pick up cross-sectional predictive patterns without providing a biological explanatory basis. This highlights an important boundary between statistical prediction and mechanistic explanation. Even with a very high accuracy, unless clinicians and researchers can ascertain the evidence that drives the prediction, estimate the uncertainty of those predictions and either prove or disprove biological plausibility in clinically consequential applications such performance may be inadequate.

The sixth gap is streamlining the development of continuous and adaptive validation frameworks. In summary, factors like viral evolution, population immunity, therapeutic pressure, human behavior and epidemiological dynamics vary across time and thus represent a considerable risk of model drift. However, most of the existing literature regards validation as a one-off exercise rather than something that will need to be performed continuously throughout the lifecycle of an AI system.

Lastly, there is still a large conceptual void about the use of artificial intelligence in virology. Previous research has the attribute to mostly consider AI as a set of domain-specific computational tools. In contrast, less attention has been directed towards a cohesive, human-supervised framework integrating detection, diagnosis, characterization, prediction and surveillance with therapeutic discovery and optimization. This review fills these gaps by conceptualizing the Algorithmic Virologist as a unified platform that extends across the continuum of viral disease. Instead of assessing the performance of an algorithm alone, the study critiques the link between technology readiness – or how close it is to a clinically usable solution – and

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

biological validity as well as clinical translation, generalizability, explainability and data governance and human oversight. In doing so, it seeks to shift the discussion away from the technical question of what artificial intelligence can do in virology, towards the more important question of what scientific, clinical, and ethical conditions under which artificial intelligence can be a trustworthy partner in contemporary and future virology.

Scope and Rationale

Current Advances in Artificial Intelligence in Virology: This review summarizes the latest developments in artificial intelligence in virology and assesses how machine learning frameworks are accelerating progress in diagnostic accuracy, epidemiological surveillance, and treatment discovery [1]. These computational frameworks leverage deep learning architectures that overcome the limitations of traditional similarity-based detection, enabling the precise identification of highly diverse pathogens and accurate prediction of epidemic trajectories [2, 3]. These frameworks provide fundamental tools for interpreting the spectrum of viral diseases, including high-throughput automated screening in diagnosis and rapid small-molecule identification in antiviral drug discovery [4]. Furthermore, the combination of generative models and predictive algorithms helps overcome the major obstacle of modeling host-pathogen interaction, from inferring complex protein configurations to optimizing numerous drug candidates [5, 6]. Furthermore, integrating machine learning into all aspects of viral diseases enables real-time risk classification and supports clinical decision-making [7, 8], allowing us to move from reactive containment to proactive precision medicine. These systems can integrate multi-omics datasets with electronic health records, leveraging real-time data integration to improve prediction accuracy, enabling clinicians to take proactive measures against patient deterioration and provide targeted antimicrobial interventions [9].

However, significant challenges remain in this area, including data heterogeneity, algorithmic bias, and the eventual need for more diverse and representative databases to ensure equitable clinical outcomes [10]. Some of these limitations call for a cross-disciplinary platform to connect computational knowledge and clinical practice so that algorithm-based interventions are backed by this kind of rigorous empirical evidence rather than just the observational literature [11–13]. For example, VIBRANT and VirSorter2 are now able to automatically recover and detect viral sequences with short lengths at high precision [14]. Not only identification, but also on knowing host-

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

pathogen crosstalk, where supervised classifiers and unsupervised language models were applied to facilitate antigen determination for rational design of next-gen vaccine candidates [15, 16].

We believe that AI's role in modern virology is significant not only because of how it automates discrete laboratory or computational functions, but because of its integration of disparate layers along the viral disease continuum into a cohesive analytical environment. Evidence indicates that the greatest impact will come from integrating genomic, clinical, epidemiological, structural and therapeutic components into a quasi-systematic analysis by machine learning rather than exploring these components through disjoint computational pipelines. The authors argue, however, that the scientific merit of such systems should not be judged based on predictive performance. Any true contribution will need to be evaluated based on whether such approaches yield reproducible biological insight, contribute to improved diagnostic or therapeutic decision-making, and stand the test of time across populations with heterogenous profiles and viral diversity. When combined, these pieces of evidence support the emergence of an integrative form of algorithmic virology with a potential ability to interrelate viral detection, diagnosis, surveillance, prediction and therapeutic development. But moving computational promise into practice with clinical impact will require substantial validation, broad datasets, open-source model development and continued collaboration between virologists, clinicians, computational scientists and public-health agencies.

Review Objectives

This review proposes a framework for systematically studying these computational developments [17], based on diagnostic accuracy, generalizability, and clinical interpretability. This assessment focuses on exploring the application gap between model performance on training datasets and its actual effectiveness in resource-limited clinical settings [18, 19]. Furthermore, this analysis questions the methodological soundness of current studies, advocating for evidence-based verification as a means of separating computational capabilities from actual, verifiable outcomes in patient populations [20, 21]. We also believe that supporting comparisons of current AI-based decision support systems with their medical counterparts in terms of transparent and understandable methodologies are important as part of this message (as this represents a trend toward generating interpretability in AI architectures that link this cognitive workflow to clinical inference [22, 23]). Thus, this review discusses whether current

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

algorithmic frameworks adequately accommodate microbial diversity and the challenges posed by existing genomic databases [24, 25]. This research also explores the effectiveness of using synthetic evaluation criteria and standardized validation lines—as proposed in some recent standardization efforts—to assess the plausibility of mechanisms beyond predictive performance [26]. Finally, the robustness of these tools is evaluated to compare their generalizability and identify critical gaps that hinder their clinical adoption across different genomic sequencing platforms and patient populations [27, 28]. Furthermore, the research explores the methodological trade-offs between model complexity and the diagnostic response time requirements at the point of care, where timely intervention is crucial to reducing viral transmission [29]. By focusing on standardized reporting and interoperability of data, this assessment aims to bridge the gap between high-dimensional computing outputs and the interpretable clinical insights needed for effective public health policies [30, 31].

The current literature might also be at fault for conflating high internal validation scores with clinically meaningful efficacy; and the evidence available to date suggests that this is a serious limitation. Hence, when evidenced-based approaches are required for applications of AI in virology, this review takes the position that consideration should be given to an expanded set of criteria including diagnostic sensitivity, generalizability, interpretability, biological plausibility and translational potential. Since there is always the possibility that algorithms may not faithfully transferring in these new settings, special attention should be paid to what sequencer platform, patient and viral lineage and resource environment shape especially durable performance of the algorithm. This means, the authors argue, that translational validity should become a paramount endpoint of future studies rather than an ancillary measurement. As such, the primary aim of this review is not only to list AI methods but to ask whether any current algorithmic systems can be sufficiently methodologically sound, generalizable, explainable and clinically relevant, given that they may provide valuable insights into medical virology. This evaluative lens is an essential precursor to subsequent analysis on detection, diagnosis, forecasting, surveillance and therapeutics.

Algorithmic Methods

Although machine learning frameworks in virology are diverse [32], they are generally classified based on the learning model and the level of data detail. Diagnostic processes primarily rely on supervised learning models, where classified datasets are used to train classifiers to accurately identify pathogens, but this is usually done by

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

pooling data to reduce the weaknesses of individual classifiers [33]. On the other hand, unsupervised learning is widely utilized for identifying hidden patterns of data and to describe an unknown viral agent by not depending on already available labels and consequently discovering a new characteristic genomic; motifs [15]. On the other hand, deep leaning architectures utilize multiple layers of neural networks to automatically learn hierarchical features from raw sequence data that have been shown to be more effective than manual feature engineering in identifying complex classification tasks [27, 34]. Specifically, support vector machines and random forests can reveal interpretable mappings between k-mer frequencies and viral taxa in supervised environments, while convolutional and recurrent neural networks are being increasingly applied to extract global correlations from long metagenomic reads or contigs [35, 36]. These architectures are advanced enough to allow bypassing handcrafted features as they use multiple non-linear transformations to learn mappings from primary sequences of nucleotides or amino acids to the target labels [37]. In addition, unclassified positive learning is being incorporated as an effective strategy to address the problem of incomplete viral genome databases, and to facilitate the construction of robust models when only some host-pathogen interactions are known [38].

The effectiveness of these approaches depends heavily on the quality and diversity of input data sources, which include high-throughput next-generation metagenomic sequencing [39], host gene transcription profiles [40], and clinical electronic health records. These rich multi-source inputs enable us to use the multi-omics, including gene expression or protein-protein interaction, and evolutionary traits between organisms - host-pathogen interactions - so as to improve taxonomy research and enhance our understandings of viral regulation [41]. Reinforcement learning (RL) is becoming increasing popular for representing dynamic viral transmission systems with its policies being optimized through iterative feedback originating from the environmental state-action rewards [42].

At present study assessment, there is no universal best machine-learning application for virology. Model choice should rather depend on the biological question, dataset structure, sample size, interpretability and clinical or research goal. Whereas, simple supervised models would be preferred over deep-learning architectures when its small sample size or interpretability that is important, the latter doing much better as genomic/protein data dimensionality increases. Among these, joint and semi-

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

supervised methods that can deal with incomplete viral databases, few labels, and poorly characterized host–pathogen interactions hold particular promise. Hence, the authors urge us not to embrace more complex architectures simply for their novelty value. However, methods used are algorithmic and thus, in virology they need to be viewed as a complementary system rather than one opposed to the other. They are used best when a balance among methodological complexity, biological purpose, and robustness, transparency, and reproducibility is preserved. Advances over the future may also originate from hybrid fashions that synergize organic facts, device gaining knowledge of, and proper experiments for validation.

Detection and Diagnosis

These high-performance models are applied in present diagnostic strategies to overcome the tension between taxa identification and the clinical demand of immediate, actionable classification. Although classical alignment tools depend on sequence homology, taxonomy-specific breadth-of-coverage metrics are now included in higher level IDs to augment reliability at the species-level [43]. In addition to these homology based approaches, machine learning methods can be used to fingerprint viral from non-viral sequences via k-mer frequency signatures [44] or long short-term memory models [45] when the sequence similarity to reference databases is low. Recent tools such as Vira Miner and Deep Vir Finder have showed its advantage of automatically extracting intrinsic genomic features from metagenomic contigs by using convolutional neural networks, which is quite limited in conventional hand-designed feature engineering [46, 47]. On the other hand, there is still a need for prospective validation of these high-dimensional analytical tools against large validated clinical datasets to prove that their accuracy in diagnosis remains invariant across populations with diverse characteristics [48, 49]. Furthermore, interpretability of these black-box models is crucial for clinical acceptance, in that without an understanding of the biological features driving a classification, practitioners are less likely to accept diagnostic outputs made by the model given its high stakes nature [50]. As a result, utilizing explainable artificial intelligence frameworks (e.g., visualizations based on SHAP or LIME) represents a critical step to convert complex underlying neural network activations into clinically meaningful, biologically motivated metrics for decision support [51]. However, efficient deployment pipelines still depend heavily on the current absence of independent validation across multi-institutional cohorts, an ongoing problem for many existing diagnostic models and the reason they are not more widely generalizable [52].

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

We argue that AI-assisted viral diagnosis is of greatest benefit when it extends (rather than replicates) conventional diagnostic capabilities. Such reference-independent algorithms also play a crucial role in identifying genetically divergent or previously uncharacterized pathogens that may be missed by homology-based pipelines. However, it is important to note that computational classification should not be considered equivalent to clinical diagnosis of infection. The authors stress that algorithmic identification needs to be interpreted in conjunction with laboratory findings, clinical context, prevalence and biological plausibility. Consequently, for diagnostic applications in particular, explainability is crucial as clinicians need to assess how and why a model has produced a high-stakes classification. By spotting intricate genomic and clinical patterns that standard methods could miss, AI can significantly elevate viral identification and diagnosis. Nonetheless, its diagnostic utility will eventually rely on prospective validation, clinical integration, interpretability as well as the stability of performance in diverse samples and structures.

Viral Detection

The computational identification of viral sequences has recently moved towards high-throughput methods such as rapid estimation of pairwise similarities (HMM and composition-based) followed by deep learning classifiers to detect pathogens in complex metagenomic backgrounds [53]. Some approaches aim only at fast classifications, such as Kraken2 using discriminatory k-mer hash tables [54], while others are more sophisticated, such as Vir Finder and Deep Vir Finder that use convolutional neural networks (CNN) to improve recovery rate of shorter viral sequences [55]. These models mostly take advantage of features like gene density and particular protein families to surpass the shortcomings of traditional alignment, but are still susceptible to biases due to fragmented metagenomic datasets [56]. Moreover, Vir Rep – hybrid language representation framework, which is an integrated major advancement of alignment-based approaches and deep learning to quantitative monitoring of low-abundance viruses in clinical samples [57]. In addition to these computational improvements, the introduction of camouflage modules and non-directional environmental genomics has become essential for distinguishing genuine viral signals from background noise and reducing misclassification rates [58, 59]. However, the application of these high-throughput pathways is often limited by the lack of standardized workflows in bioinformatics, which can introduce selection biases depending on the viral reference databases used [60]. These inconsistencies can be

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

partially addressed by standardized learning strategies, which have been explored as a means of training robust diagnostic models that maintain specificity and do not rely on the centralized collection of sensitive raw sequence data across disparate institutional groups [61]. Furthermore, more advanced approaches using hierarchical deep learning architectures are employed to modify the taxonomic structure and genomic feature informatics for biologically relevant multilevel classification [62].

We propose that this re-orientation away from systems based almost exclusively on identity-detection is the most significant maturation in viral genomics methodology since the advent of next-generation sequencing itself. This shift is especially important for emerging and poorly diverged viruses, where reference databases may not fully represent the diversity present in sampled populations. But increasing sensitivity should not come at the cost of specificity. Although some of the false-positive viral classifications in metagenomic datasets stem from contamination, short sequences, or imbalances in databases of genomic elements that incorrectly deposit before classification as viruses. Because of this, the authors appear to advocate hybrid systems comprising reference-based evidence, learned sequence representations, biological annotation and suitable negative controls as more scientifically-grounded than reliance on single classification strategy. High-throughput algorithmic viral detection enhances the ability to detect both known and uncharacterized viral genetic material. However, for reproducible use pipelines must be standardized, reference databases carefully curated, false-positive signals controlled and biological validation conducted independently. Consequently, AI should be thought of as an improved exploration and prioritization tool rather than a replacement for virological validation.

Clinical Diagnosis

Clinical diagnostic models increasingly rely on the integration of multimodal data—such as traditional laboratory measurements, patient symptom profiles, and diagnostic imaging—to provide comprehensive support for decision-making in infectious disease management [63]. Clinicians are trained to interpret these statistical learning algorithms on high-dimensional omics data and longitudinal clinical records, which enhances their ability to identify clinically relevant, subtle disease indicators that may not be accurately inferred from inferential methods [64, 65]. The use of extreme gradient enhancement algorithms and random forest classifiers on symptom-based triage data has proven effective in the diagnosis of respiratory viruses [66]. Moreover, linking structured clinical parameters with unstructured content including physician

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

notes provides a more comprehensive view of disease state that closes the loop between high-throughput binary diagnostic output and bedside interpretation [67]. The integration of multiple heterogeneous data streams together allows logic-based algorithms to emulate clinical reasoning in order to offer rational real-time diagnostic support even at the point of care where laboratory infrastructure may be lacking [68]. More recent frameworks highlight the need for host-response signatures—e.g., peripheral blood transcriptomics—to distinguish viral vs. bacterial infection and actionable prognosis to ensure that diagnostic models remain robust across clinical settings [69]. Continuing to standardize integrated frameworks, future studies need to focus on effective standardization in sample preparation and validation of biomarkers which may address the impact of pre-analytical variables on diagnostic performance [70].

Argue that the future of AI-assisted viral diagnosis is necessarily multimodal. Clinical decisions about diagnosing an illness invariably involve integrating information from different sources including contemporary symptoms, laboratory tests, imaging, host-response markers, medical history and epidemiological context; therefore, a clinically meaningful diagnostic model will not be associated with any single modality of reference. These systems with multimodal features and classifiers may thus come closer to the clinical reasoning than single-modality classifiers. However, Standardization of sample acquisition, biomarker measurement and clinical data annotation is viewed as essential by the authors since algorithmic sophistication cannot compensate for heterogeneous or biologically unreliable input data. The field of AI-based clinical diagnosis is transitioning, step-wise, from discrete classification to integrated decision support. Integration of molecular data, clinical features, imaging, and host-response information might provide improved diagnostic discrimination and prognostic assessment in complex or resource-limited settings. Standardization of data generation, external validation and the preservation of clinician oversight will be necessary to allow clinical adoption.

Prediction and Surveillance

Applying machine learning to predict disease outbreaks requires a shift from traditional, narrow-scale epidemiological models to more dynamic systems capable of integrating genomic, environmental, and patient data for real-time monitoring [71]. By using high-dimensional mobility patterns in conjunction with real-time population-level comorbidity statistics via recurrent neural networks, these systems can produce

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

highly accurate projections of transmission dynamics and rapidly identify emerging transmission hotspots [72, 73]. Furthermore, the development of tree-based algorithms and deep learning techniques has enhanced this predictive capability for biological preparedness by processing large, multiscale infectious disease datasets [74]. However, while these advancements can elevate our models to higher levels, the nature of pathogen evolution, along with deviation mechanisms that may occur on temporal and spatial scales, weakens their predictive power [75], and thus challenges remain in translating these models into adaptable policies. These challenges cannot be separated from the current limitations of reconciling genomic surveillance and timely epidemiological monitoring, given the need for stable prediction models limited by time variations in viral strains that often exceed their training horizons [76, 77].

These qualities make predictive surveillance one of the most strategically important applications of AI in virology because it has the potential to transform public-health practice from retrospective identification towards real-time prediction and anticipation. However, infectious-disease forecasting is a fundamentally dynamic problem: viral evolution, human mobility, behavioral adaptations to infection, policy interventions, and immunity all modify the data-generating environment over time. As a result, models that perform well in one epidemiological phase may fail really quickly in another. To this end, the authors therefore argue for surveillance algorithms that are adaptive systems (as opposed to predictive products) that get regularly updated over time. This is extremely powerful, the combination of genomic, environmental, mobility, clinical and epidemiological data for viral surveillance can be significantly enhanced with machine Learning. However, a fundamental element of reliable forecasting requires continuous recalibration and time-based verification, estimates of uncertainty, and the ability to adapt to the changing dynamics of pathogens and hosts. Consequently, forecasting should serve to prepare, not be taken as an inevitable truth.

Epidemic Forecasting

Predictive frameworks have evolved beyond static epidemiological forecasting by integrating multi-source data to create dynamic simulations of rapid viral spread pathways [78, 79]. These hybrid models achieve significant improvements in predicting disease outbreak pathways at the regional level compared to traditional zoning-based approaches by incorporating real-time mobility patterns with socioeconomic indicators [80, 81]. Furthermore, the ability to input spatial data using geographic information systems (GIS) enables highly accurate predictions at the local level, allowing for

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

targeted public health interventions [82]. However, most of these predictive efforts suffer from significant data gaps, such as geographic reporting bias that skews datasets toward resource-rich urban centers and underrepresents rural populations [83]. Moreover, the inherent reliance on irregular and fragmented reporting flows often leads to significant data collection delays, hindering these models from accurately monitoring nonlinear and multi-level transmission processes in real time [84, 85]. To fill these gaps, more researchers have begun using graphical techniques in conjunction with recurrent neural networks, contributing to the modeling of nonlinear dependencies between disease spread at multiple levels [86]. More specifically, graphical neural networks provide an efficient mechanism for modeling these interconnected dynamics by considering geographic areas or social clusters as nodes, and have proven highly effective compared to traditional machine learning methods in characterizing nonlinear spread patterns [87].

In light of these findings, the primary advantage of AI-based epidemiological predictions lies in their ability to model complex, nonlinear relationships that are often oversimplified by traditional models. However, increased complexity does not necessarily translate into more useful public health predictions. At best, models can represent the underlying surveillance infrastructure, with a disproportionate focus on urban or resource-rich populations, systematically underestimating transmission patterns elsewhere. Therefore, researchers consider geographic representation and clear uncertainty reporting to be just as important as prediction accuracy. By integrating mobility, geographic, socioeconomic, and epidemiological data, advanced prediction systems can improve outbreak forecasting at both regional and local levels. However, these systems remain effective only if the data are up-to-date, representative, and interoperable. Future systems should combine computational sophistication with uncertainty measurement and explicit geographic and socioeconomic equity in surveillance.

Variant Surveillance

Emerging computational paradigms have been trained on deep sequencing data to infer evolutionary trajectories and predict mutations in viral proteins with high potential for generating variants of concern, a priori to their emergence [88]. These models utilize transformer-based architectures to study viral fitness landscapes and protein structural stability in order to predict antigenic properties that may allow for immune escape [89]. Additionally, curating these structural facts with longitudinal genomic surveillance can

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

track mutations that improve binding affinity to human cell receptors (which in SARS-CoV-2 is the ACE2 protein), suggesting increasing viral transmissibility or virulence [90, 91]. These predictive abilities are enhanced by deep-learning approaches which assess the potential functional impact of amino-acid-altering mutations, effectively serving as an alarm bell for pathogens prone to escape from vaccine-induced protection [92]. In this way, these platforms support a proactive analysis of how certain genomic changes might impact clinical severity or epidemiological fitness [93], by systematically generating maps of the fitness landscapes for any emerging lineages. Current models that attempt to connect genomic data with epidemiological investigation are hampered by ongoing developments in examining these genomic-epidemiological linkages, where spatial graphs either need to be defined a priori or may fail to model the adaptive inter-regional dynamics of evolving viral populations [94]. To help overcome these limitations, causal graph neural networks should enable more accurate modeling of how collective individual behaviors and changes in policy and viral transmission interact [95]. These frameworks also utilize self-supervised message-passing to recognize important changes in disease progression that could lead to real-time simulations of dynamic, intelligent quarantining policy. Moreover, applying lightweight Tiny ML and Edge AI on wearable sensors would enable them to monitor physiological vitals in real time so they could serve immediate decentralized thorough early detection of the infection clusters [96].

They believe that AI-supported variant surveillance can develop into a real early-warning system, particularly if genomic sequence analysis is combined with protein structure, antigenicity, receptor binding, epidemiological growth and clinical phenotype. Sequence frequency of a mutation alone does not mean much biologically. Viral fitness is context-specific, shaped by epistasis, host immunity, transmission opportunity and population behavior. Thus, this makes the authors caution against calling computationally predicted mutations functionally valuable without functional or epidemiological validation to support it. It has the ability to improve the predicted identification of and delimitation between viral mutations and lineages that may have consequences for immune escape, transmissibility or severity [15]. However, its power resides in risk stratification and early prioritization rather than definitive prediction. Computational predictions should be complemented with laboratory experiments and real-time epidemiological data.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

Therapeutic Applications

Using unsupervised representational learning to decode viral fitness environments, computational strategies for identifying important conserved residues have begun to focus on key proteins that may have potent antiviral targets [97, 98]. By combining graph neural networks with supervised autolearning, researchers have recently been able to accelerate the identification of therapeutic agents through more comprehensive analysis of complex virus-drug association networks and high-resolution protein structures [99]. In addition to designing from scratch, drug reuse frameworks also employ techniques such as autoencoders and causal network discovery to find compounds that can mirror the specific viral infection signature by targeting host-pathogen pathways present in both the host and the pathogen [100]. Generative models—for example, the Anti-Viral DL framework—synthesize two-part virus-drug graphs and predict therapeutic interactions while addressing data scarcity issues through differential learning [101, 102].

Besides these structural analyses, reinforcement learning models go a step further beyond this and not only refine these predictions based on previously learnt molecular scaffolds to optimize binding affinity but also simulate the iterative process of candidate selection (103, 104). In addition to the rational design of molecules, algorithmic systems for decision support are beginning to be used to combine individual omics profiles with viral genomic information in order to recommend precision antiviral therapies tailored to subject- and strain-specific differences in metabolism (which may be associated with drug resistance) [1].

Outlook The most obvious impact of AI on antiviral development will be in the study view by narrowing down the enormous search space-plagued landscape of traditional drug discovery. It is primarily valuable in guiding the selection of biologically actionable targets, compounds, combinations, and probable resistance pathways for testing. However, the computational prioritization should not be confused with therapeutic effectiveness. Identified molecular interactions, host–pathogen associations and drug-repurposing candidates represent hypotheses that need biochemical, cellular, animal and ultimately clinical validation. The authors thus advocate a closed-loop model in which computational predictions continue to be exposed to experimental evidence. However, AI provides robust capabilities for target identification and drug repurposing, molecular design and precision antiviral selection. Its most significant impact is not the eradication of experimental development as part

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

of a therapeutic discovery paradigm, but rather, its acceleration and rationalization. Close integration of computational models and iterative laboratory and clinical validation will be required for successful translation.

Antiviral Discovery

Recent work in this area focuses on large-scale virtual screening of millions of chemical libraries using pre-trained deep learning models and transfer learning [105]. Along with PDA methods, these target prioritization workflows exploit predictive models to approximate the therapeutic viability of the candidate leading to a significant acceleration of the early stages of medicinal chemistry [106, 107]. Moreover, these strategies also aid in informed design and lead compound optimization by simulating pathogen evolutionary paths and predicting overall binding kinetics making them an excellent solution for the high attrition rates observed in classical development pipelines [108]. More than merely providing estimates of binding, high-fidelity docking simulations paired with machine learning models can allow rapid evaluation of large numbers of FDA-approved compounds to discover ligands that can interfere in host-virus interactions [109]. Modeling can be performed on the predictive profiles of fractions, allowing researchers to evaluate both pharmacokinetic and toxicological profiles at once so that candidate safety is guaranteed in parallel with lead compound optimization against highly mutable viral proteins [110,111]. Further, the coupling of molecular dynamics to quantitative structure-activity relationship modeling can be employed to fine-tune binding stability in specific pockets of proteins so that repurposed or de novo identified ligands retain potency against variable viral strains [112]

Virtual screening and AI-aided compound design are particularly appealing as these paradigms have the potential to drastically reduce the number of compounds from thousands or millions needing experimental activity data. Binding affinity does not directly allow to infer antiviral efficacy, however. Whether any computationally interesting molecule actually ends up being a useful clinical drug depends on pharmacokinetics, toxicity, cellular penetration, off-target effects, and how well the virus can resist it and how the host's biological pathways respond. Consequently, the authors prefer multi-objective AI frameworks that optimize molecular activity and ADMET properties concurrently alongside evolutionary robustness and manufacturability. This includes a comprehensive approach to vaccine development, supported by artificial intelligence for its ability to prioritize targets/virtual screening,

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

repurpose drugs, and optimize leading compounds in antiviral discovery. However, the development of clinically successful antivirals must move beyond single-criteria molecular predictions to integrated models of efficacy, safety, resistance, and pharmacovigilance.

Therapeutic Optimization

The clinical application of these computational approaches requires linking predictive models with real-time patient response data to tailor drug dosages and avoid/mitigate side effects [113]. However, these improved models still suffer from a significant discrepancy between the ideal molecular docking simulation and the realistic human pathophysiology that determines their clinical pharmacokinetic properties [114]. Furthermore, accurate prediction of absorption, distribution, metabolism, excretion, and toxicity (ADMET) characteristics and synergistic drug combinations to address interindividual metabolic variability [115] will depend on concordance to achieve accuracy in personalized therapy. A systematic effort to standardize these computations, based on community-level performance benchmarks and repeated feedback loops, will be crucial to bridging the gap between initial correlation predictions and reproducible translational success (116).

Furthermore, the process-scale chemistry of novel chemotypes has a limited scalability which hampers in silico hit identification and clinical validation [117]. These barriers can only be overcome through the incorporation of high-throughput experimental validation platforms like automated microfluidic screening to provide exposure data and other key biochemical details to close in a feedback loop for refining computational collections [118]. Therapeutic optimization (too) is more complex than drug discovery; the study argues because it relies on computational predictions operating in the full complexity of individual human physiology. Therapeutic effectiveness for viruses is likely to be substantially modulated by patient factors such as metabolic pathology, also comorbidities driven and virus-related like viral-genotype & immune status also concurrent therapies given, treatment response over time. Thus, the authors approach considers clinical and experimental data in real-time indispensable for an effective precision antiviral therapy. An Inherent Limitation of Static Predictive Models for the Continuously Evolving Host–Virus–Drug Interactions One potential advantage of this approach would be its ability to use patient-specific biological data, viral information, and effective treatments in an adaptive system as a way to optimize therapy moving forward. These systems could provide benefit in the areas of dosing,

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

combination therapy and resistance management but their clinical utility must be established through prospectively validating experiments and feedback from closed-loop experimental and clinical data.

Limitations and Future Directions

Although these methodologies have progressed rapidly, their application at scale in virology is currently limited by issues ensuring data quality (e.g. [118]), demographic biases within training cohorts and the fundamental scarcity of high-quality data for rare viral pathogens [119]. Furthermore, reliance on single datasets containing virally encoded proteins, which often fail to represent the structural variation in viral protein compositions or even host-pathogen interactions in diverse populations, frequently hinders the generalization of models [120]. The “black box” nature of some constructs also limits clinical confidence and necessitates integration with interpretable AI frameworks to better understand the biological motivations behind algorithmic recommendations [14, 121]. However, employing interpretive approaches requires balancing the benefit of understanding decision-making processes with the risk of disclosing detailed patient data [122]. Additionally, to curb the proliferation of outdated, unreliable, or scientifically unsound but poorly supported research, the reproducibility crisis must be addressed by standardizing databases, thereby reducing the need for validation of potential therapeutic candidates [123, 124]. Rigorous ethical and regulatory frameworks remain crucial for addressing algorithmic bias, algorithm liability, and the need for informed consent in AI-assisted healthcare delivery [125]. To advance clinically responsible algorithmic virology, stakeholders must adopt open and reusable data (FAIR) principles and unified learning architectures that integrate fragmented datasets while maintaining patient privacy [126]. Furthermore, a “human intervention” approach, which ensures repeated clinical validation of algorithm outputs, will help align deep learning predictions with current physiological references [127]. Ultimately, this integration must be aligned with existing regulatory guidelines for trusted AI-based medical products to ensure a successful transition from development to clinical practice [128].

In a study assessment, the main problem with algorithmic virology is not an abundance of complex computational models, but rather the quality, representativeness and transparency—and governance—of data via which such models are built. Improving model complexity without addressing biases in the data, limited external validation, poor interoperability and insufficient biological verification risks generating

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

technologically impressive but clinically weak systems. As a consequence, the authors understand future AI research to be built on ethical foundations underpinned by FAIR data practices, privacy preserving collaboration, explainability and human oversight with prospective validation as core elements rather than optional ethical extras.

The authors also suggest that Human-supervised by design responsible algorithmic virology should be the norm. The pattern discovered by algorithms may go beyond even the highest levels of human analytical capacity, but pattern interpretation should always remain anchored to virological expertise, clinical acumen as well as overall responsible institutional understanding and decision-making. AI's future success in virology will be less about constructing increasingly larger models and more about building trustworthy ecosystems of high-quality data, transparent algorithms, experimental validation, ethical governance, and human expertise. Responsible algorithmic virology, therefore, requires both technological innovation and institutional development.

Translation Challenges

Ultrasound imaging in robot-controlled techniques is a significant area of focus in peer-reviewed publications; however, a gap remains between experimental models and clinical application. This gap is characterized by the limited number of prospective clinical studies and the dominance of non-clinical research environments lacking representative patient-level data [129, 130]. Many aspects of this deficiency are exacerbated by the selection of study measures that prioritize statistical accuracy at the expense of clinical efficacy and mask the limited generalizability of heterogeneous populations in practice [131]. To mitigate the limitations of dataset bias, concomitant variable variation, and concept skew [132], researchers are compelled to conduct large-scale external validation studies across multiple centers. Finally, bridging the gap between algorithm outputs and clinician practice requires heterogeneous platforms seamlessly integrated into workflows, enabling continuous monitoring and updating of models over time to adapt to evolving clinical contexts [133]. These platforms should include interpretive tools, such as discriminatory maps or feature significance analysis, to enable clinicians to critically analyze the biological basis of algorithmic decisions [134]. Furthermore, future trials are still needed to measure the actual gains from these technologies and enhance clinical confidence, given the social and technical complexities inherent in real-world medical settings [135]. Finally, the maturation of this field will require additional rigorous yet flexible regulatory pathways that consider AI-enabled clinical decision support systems as an advanced medical device, while

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

consistently ensuring safety and accountability [136]. Moreover, while technology is a given in the development of useful tools for healthcare practice, data management must adapt so that ethics remain central to the rapid deployment of applied innovations.[137] According to the study, the greatest discrepancy in current AI and virology research lies between past computational performance and future clinical efficacy. After deployment, models trained under controlled research conditions face common challenges, including dataset changes, workflow disruptions, sample diversity, and the dynamics of viral biology. Therefore, lifecycle validation of the model should replace the traditional one-off assessment, which includes independent, multi-center studies, prospective validation, post-deployment monitoring, and periodic recalibration. The researchers also argue that regulatory bodies should view clinically used AI as a dynamic, not a static, technology. Translating research into routine medical practice goes far beyond technical correctness. AI systems should show reproducible clinical benefit, compatibility with workflows in healthcare facilities, external validity, interpretability, safety and performance over time. Indeed, these approaches require multicenter prospective evaluation and regulatory oversight that is adaptive to validate reliability in clinical practice.

Future Research

So, future studies should focus on developing foundation models that are trained on multimodal datasets which can encapsulate and approximate the complex and non-linear interactions of viral pathogenesis and host immune responses [138]. These large-scale architectures provide the ability to incorporate genomic, proteomic and longitudinal clinical data, allowing for the move beyond static unimodal analysis [9, 139]. These models, when combined with self-supervised learning on the bountiful untapped raw unlabeled biological treasure, can produce representation feature embeddings that contribute to better predictive robustness in low-data rate regimes [140,141]. Lastly, the need for collaboration and harmonization of international standards for sharing and validating genotype–phenotype correlations across populations is vital to prevent the emergent "accuracy paradox" in predictive models, generalizing across different ethnicities [22,142]. These types of collaborative frameworks should be predicated on specific requirements for clear descriptions of model limitations and uncertainty quantification so that internal validation metrics are always validated via independent, prospective analysis [143]. To achieve equitable progress, these frameworks should pave the way for inclusive, collaborative approaches

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

to eliminate bias, enabling organizations with limited digital capabilities to focus on and leverage shared global AI models [144]. Consensus-based data management protocols, such as those promoted by projects like Standing Together, are essential for standardizing interoperability across global healthcare systems and facilitating the ethical use of datasets from diverse sources [145]. Furthermore, prioritizing standardized learning and privacy-preserving analytical methods will offer an alternative solution to the problem of data segregation across borders, while maintaining robust regulatory transparency.[146 ,23]

Researchers anticipate that the next generation of algorithmic virology will move from task-specific models to multimodal platforms capable of integrating viral genomes, protein structures, host responses, clinical pathways, therapeutic inputs, and epidemiological signals. However, as researchers caution, scalability should not be the end goal. A basic model is only scientifically useful if its predictions are biologically relevant, externally validated, transparent regarding uncertainties, and applicable across all healthcare settings. The study also suggests that future research should move toward a proactive framework for virology, where AI not only retrospectively describes viral events but also continuously integrates molecular and population-level signals. This would facilitate the detection of clinically significant threats before widespread disease outbreaks. Achieving this goal requires international data standards, equitable sharing, standardized infrastructure, proactive validation, and ongoing dialogue between computational prediction and empirical science. Future research should be driven by multimodal, continuously learning systems that are biologically informed, privacy-preserving, externally validated, and based on globally representative metrics. It should be an integrated intelligence framework to facilitate earlier detection, better forecasting and prioritization of therapeutics and equitable preparedness for emerging viral threats — not autonomous algorithmic decision-making.

Conclusions

This review shows that AI is already changing the viral disease continuum from viral detection and clinical diagnosis to epidemic prediction, variant surveillance, antiviral discovery and therapeutic optimization. In these domains, machine learning, deep learning, generative models and multimodal analytical systems offer capabilities that can improve on conventional virological approaches through pattern detection in complex data types, multi-dimensional dataset integration and speed-ups of processes

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

that would otherwise have required significant experimental or computational resources.

Checked evidence suggests that among the most important contributions from AI lies in its potential to promote a shift from reactive towards predictive and anticipatory virology. A possible improvement in viral identification could be in recognizing new or unknown pathogens. Multimodal algorithms can combine molecular and clinical information for diagnostic or prognostic modeling in clinical diagnosis. Conversely, computational systems can integrate genomic, epidemiological, geographic and mobility data in surveillance that allows for the identification of new transmission dynamics and notable variants. In the field of therapeutics, AI can narrow down the search range (for molecular targets / antiviral candidates) and help with drug repurposing / optimization (target identification + validation).

However, the review further illustrates that high computational performance alone does not provide evidence of clinical readiness. Multiple surrounding issues such as dataset bias, limited external validation, incomplete reference viral databases, population under-representation, interoperability challenges, algorithmic opacity, model drift and privacy concerns and lack of prospective evaluation are continuing constraints to clinical or public-health application of experimental performance.

One of the more critical conclusions is that future iterations of AI in virology should not be evaluated merely by comparing increasingly sophisticated architectures or improvements on benchmark performance. Indicators of algorithmic quality such as biological plausibility, external validity, uncertainty quantification, interpretability, reproducibility, clinical utility and equity should be treated equally.

The results also point to isolated applications of AI slowly gravitating towards an integrated computational ecosystem. THIS BACKS UP THE USA OF THE ALGORITHMICAL VIROLOGIST AS CONCEPTUALIZED IN THIS REVIEW: NOT A STANDALONE ARTIFICAL ENTITY THAT IS INTENDED TO REPLACE VIROLOGISTS OR CLINICIAN, BUT AN END USER SUPERVISED, HUMAN-DRIVEN ANALYTICAL BRIDGE BETWEEN:

Detection → Diagnosis → Characterization → Prediction → Surveillance Therapeutic Discovery Therapeutic Optimization Continuous Validation. In this context, artificial intelligence enhances analytical capabilities; experimental science validates biological findings; clinicians provide contextually appropriate interpretations; and responsible human institutions bear the consequences of critical decisions. Therefore, the study

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

concludes that the future of AI in virology cannot be categorized as simply a matter of human intelligence versus machine intelligence. The most scientifically logical path is that of augmented virology, where computational intelligence and human expertise work together seamlessly within a transparent system of continuous verification and ethical governance. Ultimately, for the current generation of algorithmic virology efforts, the question is no longer whether AI can generate accurate predictions, but rather whether these predictions can maintain their biological validity, be clinically usable, generalizable across different environments and populations, be interpretable/interpretive in real-world conditions, and provide immediate, practical recommendations.

Recommendations

This review presents a synthesis of evidence pointing to future directions for the design of AI-assisted virology by identifying six interconnected strategic priorities, based on identified methodological, application, and governance gaps. Representative and Interoperable Data: The emergence of reliable algorithmic virology requires datasets that are broadly representative of human viral diversity, patient populations, healthcare settings, and the epidemiological contexts that will aid in prediction.. While retaining the need for harmonized data collection, annotation, curation and reporting to be advanced in an open & standardized manner; relevant International initiatives should focus on inclusion of underrepresented populations as well as low & middle-resource settings too. Increased uptake of FAIR data principles and coupled genomic, clinical, epidemiological, and therapeutic data infrastructures would enhance reproducibility and support cross-institutional validation while decreasing concern that AI systems predominantly represent well-sampled populations. Federated learning (as appropriate) would also mean that these private-data-preserving approaches could allow for international collaboration without having to centralize sensitive patient-level data in unrestricted ways.

External, Prospective and Lifecycle Validation: High performance for the model internal validation should not be deemed as sufficient evidence that a model is ready for clinical use. All AI systems targeting virological, clinical, or public-health applications are in urgent need of independent external validation across populations, institutions, sequencing platforms, viral lineages and resource environments. The evaluation of predictive algorithms should increasingly rely on prospective multicenter

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

contexts rather than retrospective benchmarking and focus on clinically meaningful outcomes in addition to predictive accuracy. Since viral evolution, population immunity, clinical practice and epidemiological conditions are constantly evolving phenomena, validation should extend throughout the algorithmic lifecycle through post-deployment monitoring, temporal reassessment of model performance over time periods with temporal drift detection as well as recalibration and revalidation.

This combined approach provides an avenue for developing more biologically informed multimodal AI models, which would be able to integrate viral genomics information with potential multidimensional aspects of viral disease biology (such as proteomic, structural, immunological, host-response, clinical epidemiologic exposure response and therapeutic data) Most importantly, where scientifically warranted, multiple modalities studied or invoked will similarly benefit from the integration of experimentally validated information pertaining to viable virological mechanisms and their underlying protein structures; host–pathogen interactions and how they evolve; or other critical knowledge from existing pharmacological literature. Increase computational complexity, should not be a goal to pursue per se, but rather the development of architectures and outputs that closely correspond more biologically—to how stored computational systems might perform those processes they claim to represent. This kind of integration might help pave the way for developing from independent, task-specific algorithms to an integrated framework of the Algorithmic Virologist type.

Justification, Anticipation of Application, and Repeatability: The most clinically consequential AI systems need to provide enough interpretability for researchers and healthcare professionals to effectively assess the evidence justifying algorithmic outputs. This explainability must be accompanied by explicit uncertainty quantification, reporting of model limitations and circumstances under which its performance may decline. These model-agnostic factors in addition to reproducibility, as the third one, can also be used to report on datasets and preprocessing layout or strategy followed through transparent reporting of pre-processing steps, validation strategies and performance metrics. This is particularly touchy in virology, where statistically tight predictions do not necessarily correspond to biological mechanisms—nor hold as viral populations evolve.

Human Supervision, Privacy and Dynamic Regulation: AI must be deployed under human supervision in high-stakes; diagnostic, therapeutic, epidemiological or public-

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

health deployments. Algorithmic recommendations should only augment the recommendations of an appropriate virological, clinical, and public-health judgment that is still made by a qualified person with accountability for important decisions remaining in human institutions. Data governance, transparent accountability mechanisms, and proper safeguards against algorithmic discrimination (for instance) should be built into the design. Moreover, regulatory frameworks need to acknowledge that clinically deployed AI may change via retraining, recalibration, changes in input distributions, and model updates. The regulatory assessment should therefore not be limited to static medical technology, instead focusing on the entire algorithmic lifecycle of AI.

Experimental Validation and Anticipation Virology: The computational prediction should be taken as guidance for the aerosol of experiments to validate without replacing biological validation. This includes predictions about novel viruses, functionally significant mutations, host–pathogen interactions, therapeutic targets, antiviral candidates or resistance pathways that should be validated experimentally in the laboratory setting or clinical population where possible. Therefore, closed loop systems linking computational prediction and experimental/clinical feedback should become an important research focus. Ultimately, these capabilities should facilitate a transition away from largely reactive virology towards predicting and anticipating virology—where molecular, clinical and population-level signals are continuously fused to identify emerging threats earlier, target interventions more efficiently and increase preparedness while maintaining explicit uncertainty quantification with human oversight. Collectively, these 6 priorities provide a translational road map for the Algorithmic Virologist. Progress is not a function of bringing an ever more sophisticated algorithm or benchmark: progress is the ability to generate biologically plausible, externally validated, interpretable, reproducible and equitable clinically-actionable conclusions in a system that monitors itself relentlessly and retains human accountability.

References

1. Ibrahim MM. Artificial Intelligence as a Catalyst for Advancements in Medical Virology. *Clinical Laboratory*; 72. Epub ahead of print 1 January 2026. DOI: 10.7754/clin.lab.2025.250474.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

2. Sinno A, Baghdadi R, Narch R, et al. Charting the virosphere: computational synergies of AI and bioinformatics in viral discovery and evolution. *Journal of Virology*; 99. Epub ahead of print 12 November 2025. DOI: 10.1128/jvi.01554-25.
3. Garg P, Singhal G, Singhal SS. Next-generation viral detection through AI-enhanced nanotechnology: advances, challenges, and future directions. *Frontiers in Molecular Biosciences*; 13. Epub ahead of print 18 March 2026. DOI: 10.3389/fmolb.2026.1750124.
4. Tian T, Zhang X, Zhang F, et al. Harnessing AI for advancing pathogenic microbiology: a bibliometric and topic modeling approach. *Frontiers in Microbiology* 2024; 15: 1510139.
5. Yakimovich A. Toward the novel AI tasks in infection biology. *PubMed Central*; 9. Epub ahead of print 9 February 2024. DOI: 10.1128/msphere.00591-23.
6. Bowyer S, Allen DJ, Furnham N. Unveiling the ghost: machine learning's impact on the landscape of virology. *PubMed Central*; 106. Epub ahead of print 13 January 2025. DOI: 10.1099/jgv.0.002067.
7. Wang Y, Zhong J, Kumar R. A Systematic Review of Machine Learning Applications in Infectious Disease Prediction, Diagnosis, and Outbreak Forecasting. *Preprints.org*. Epub ahead of print 15 April 2025. DOI: 10.20944/preprints202504.1250.v1.
8. Zhou H-Y, Li Y, Li J, et al. Harnessing the power of artificial intelligence to combat infectious diseases: Progress, challenges, and future outlook. *The Innovation Medicine* 2024; 100091.
9. Abbara S, Crabol Y, Bouillé JG de, et al. Artificial intelligence and infectious diseases: Scope and perspectives. *Infectious Diseases Now* 2025; 55: 105131.
10. Sarantopoulos A, Kourmpani CM, Yokarasa AL, et al. Artificial Intelligence in Infectious Disease Clinical Practice: An Overview of Gaps, Opportunities, and Limitations. *Tropical Medicine and Infectious Disease* 2024; 9: 228.
11. Chatterjee JM, Saxena SK. *Artificial Intelligence in Medical Virology*. Springer Nature. Epub ahead of print 1 January 2023. DOI: 10.1007/978-981-99-0369-6.
12. Singh T. Deep Learning Endowed Generative Models: Revolutionizing Viral Disease Identification and Analysis with AI-Driven Decision

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

Support. *International Scientific Journal of Engineering and Management* 2025; 4: 1–9.

13. Srivastava V, Kumar R, Wani MY, et al. Role of artificial intelligence in early diagnosis and treatment of infectious diseases. *Infectious Diseases* 2024; 57: 1–26.

14. Liu G, Yu D, Fan M-M, et al. Antimicrobial resistance crisis: could artificial intelligence be the solution? *Military Medical Research* 2024; 11: 7.

15. Wong F, Fuente-Núñez C de la, Collins JJ. Leveraging artificial intelligence in the fight against infectious diseases. *Science* 2023; 381: 164–170.

16. Meslamani AZA, Sobrino I, Fuente J de la. Machine learning in infectious diseases: potential applications and limitations. *Annals of Medicine* 2024; 56: 2362869.

17. Arab RAE, Almoosa Z, Alkhunaizi M, et al. Artificial intelligence in hospital infection prevention: an integrative review. *Frontiers in Public Health* 2025; 13: 1547450.

18. Aşkın HŞ, Şahin A, Aşkın L. Exploring the Potential of Artificial Intelligence in Infectious Disease. *Gaziantep Islam Science and Technology University* 2024; 5: 168–181.

19. Omar M, Brin D, Glicksberg BS, et al. Utilizing Natural Language Processing and Large Language Models in the Diagnosis and Prediction of Infectious Diseases: A Systematic Review. *medRxiv*. Epub ahead of print 17 January 2024. DOI: 10.1101/2024.01.14.24301289.

20. Theodosiou AA, Read RC. Artificial intelligence, machine learning and deep learning: Potential resources for the infection clinician. *Journal of Infection* 2023; 87: 287–294.

21. Algain S, Marra AR, Kobayashi T, et al. Can we rely on artificial intelligence to guide antimicrobial therapy? A systematic literature review. *Antimicrobial Stewardship & Healthcare Epidemiology*; 5. Epub ahead of print 1 January 2025. DOI: 10.1017/ash.2025.47.

22. Giacobbe DR, Marelli C, Muccio M, et al. From Klebsiella and Candida to artificial intelligence: a perspective from infectious diseases doctors and researchers. *Frontiers in Medicine* 2025; 12: 1676920.

23. Mairi A, Hamza L, Touati A. Artificial intelligence and its application in clinical microbiology. *Expert Review of Anti-infective Therapy* 2025; 23: 469–490.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

24. Cardona D, Cardona-Acevedo S, Valencia-Arias A, et al. Bibliometric analysis of the application of artificial intelligence techniques in bacteriology: a PRISMA-guided research agenda. *Frontiers in Microbiology* 2025; 16: 1641967.

25. Jumaah AS, Hassan A, Jasim FS. *Precision Microbiome Medicine and AI-Driven Infectious Disease Diagnostics*. Epub ahead of print 1 January 2025. DOI: 10.62906/bs.book.456.

26. Sufi F. Editorial: Generative AI and large language models in microbial evolution, resistance mechanisms, and antimicrobial drug discovery. *Frontiers in Microbiology* 2026; 17: 1848209.

27. Suster CJE, Pham D, Kok J, et al. Emerging applications of artificial intelligence in pathogen genomics. *Frontiers in Bacteriology*; 3. Epub ahead of print 6 March 2024. DOI: 10.3389/fbrio.2024.1326958.

28. Ibiam VA, Omale LE, Taiwo O. The role of Artificial Intelligence models in clinical decision support for infectious disease diagnosis and personalized treatment planning. *International Journal of Science and Research Archive* 2025; 14: 1337–1347.

29. Assudani PJ, Bhurgy AS, Kollem S, et al. Artificial intelligence and machine learning in infectious disease diagnostics: a comprehensive review of applications, challenges, and future directions. *Microchemical Journal* 2025; 218: 115802.

30. (BTK) B tyrosine kinase, signaling a key component of B cell receptor (BCR), from plays a crucial role in regulating cell proliferation and survival in various B cell malignancies. T study employs a comprehensive computational approach to evaluate phytochemicals, et al. Abstracts. *FEBS Open Bio* 2025; 15: 14–58.

31. Omar M, Brin D, Glicksberg BS, et al. Utilizing natural language processing and large language models in the diagnosis and prediction of infectious diseases: A systematic review. *American Journal of Infection Control* 2024; 52: 992–1001.

32. Coxe T, Azad RK. Silicon versus Superbug: Assessing Machine Learning's Role in the Fight against Antimicrobial Resistance. *Antibiotics* 2023; 12: 1604.

33. Mishra S, Kumar R, Tiwari S, et al. Machine learning approaches in the diagnosis of infectious diseases: a review. *Bulletin of Electrical Engineering and Informatics* 2022; 11: 3509–3520.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

34. Sheppard SK, Arning N, Eyre DW, et al. Machine learning and statistical inference in microbial population genomics. *Genome biology* 2025; 26: 313.

35. Lucaci AG, Shaikh H, Chong LC, et al. Consensus statement from the second RdRp Summit: towards a unified framework for RNA virus biology. *Peer Community Journal*; 6. Epub ahead of print 28 May 2026. DOI: 10.24072/pcjournal.727.

36. Peng X, Wei Y, Zhou X. Enhancing pathogen identification through AI-assisted metagenomic sequencing. *Frontiers in Microbiology* 2025; 16: 1634194.

37. Awe OI, Obura HO, Charles S, et al. Enhanced deep Convolutional Neural Network for SARS-CoV-2 variants classification. *Frontiers in Artificial Intelligence* 2025; 8: 1512003.

38. Iuchi H, Kawasaki J, Kubo K, et al. Bioinformatics approaches for unveiling virus-host interactions. *Computational and Structural Biotechnology Journal* 2023; 21: 1774–1784.

39. Chisompola D, Luwaya E, Nzobokela J, et al. AI-powered analysis of viral metagenomic sequencing data for rapid outbreak investigation and novel pathogen discovery. *Frontiers in Microbiology* 2026; 16: 1717859.

40. Whisnant AW, Dölken L. Viruses, Vectors, and Villains: Governing the Risks and Rewards of Artificial Intelligence in Virology. *AI* 2026; 7: 93.

41. Jiang Y, Luo J, Huang D, et al. Machine Learning Advances in Microbiology: A Review of Methods and Applications. *Frontiers in Microbiology* 2022; 13: 925454.

42. Li L, Yang L, Li Z, et al. From immunology to artificial intelligence: revolutionizing latent tuberculosis infection diagnosis with machine learning. *Military Medical Research* 2023; 10: 58.

43. Ma H, Tan TW, Ban K. A multi-task CNN learning model for taxonomic assignment of human viruses. *BMC Bioinformatics* 2021; 22: 194.

44. Guo J, Bolduc B, Zayed AA, et al. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* 2021; 9: 37.

45. Wu L, Wijesekara Y, Piedade GJ, et al. Benchmarking bioinformatic virus identification tools using real-world metagenomic data across biomes. *Genome biology* 2024; 25: 97.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

46. Miao Y, Bian J, Dong G, et al. DETIRE: a hybrid deep learning model for identifying viral sequences from metagenomes. *Frontiers in Microbiology* 2023; 14: 1169791.

47. Azevedo KS, Souza LC de, Coutinho MGF, et al. Deepvirusclassifier: a deep learning tool for classifying SARS-CoV-2 based on viral subtypes within the coronaviridae family. *BMC Bioinformatics* 2024; 25: 231.

48. Dong C, Liu Y, Nie J, et al. Artificial Intelligence in Infectious Disease Diagnostic Technologies. *Diagnostics* 2025; 15: 2602.

49. Cheah BCJ, Vicente CR, Chan KR. Machine Learning and Artificial Intelligence for Infectious Disease Surveillance, Diagnosis, and Prognosis. *Viruses* 2025; 17: 882.

50. Sahragard R, Arabfard M, Najafi A. Predicting host-pathogen interactions with machine learning algorithms: A scoping review. *Infection Genetics and Evolution* 2025; 130: 105751.

51. Marcos-Zambrano LJ, López-Molina VM, Bakır-Güngör B, et al. A toolbox of machine learning software to support microbiome analysis. *Frontiers in Microbiology* 2023; 14: 1250806.

52. Dubey S, Dubey S, Raghuwanshi K, et al. GenoDense-Net: unraveling the genomic puzzle of the global pathogen. *Tropical Diseases Travel Medicine and Vaccines* 2025; 11: 32.

53. Galeeva JS, Kuzmichenko P, Manolov AI, et al. Bioinformatics Tools and Approaches for Virus Discovery in Genomic Data: A Systematic Review. *Viruses* 2025; 17: 1538.

54. Glickman C, Hendrix J, Strong M. Simulation study and comparative evaluation of viral contiguous sequence identification tools. *BMC Bioinformatics* 2021; 22: 329.

55. Ho SFS, Wheeler NE, Millard A, et al. Gauge your phage: benchmarking of bacteriophage identification tools in metagenomic sequencing data. *Microbiome* 2023; 11: 84.

56. Schackart K, Graham JB, Ponsero AJ, et al. Evaluation of computational phage detection tools for metagenomic datasets. *Frontiers in Microbiology* 2023; 14: 1078760.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

57. Dong Y, Chen W, Zhao X. VirRep: a hybrid language representation learning framework for identifying viruses from human gut metagenomes. *Genome biology* 2024; 25: 177.

58. Ghorbani A, Rostami M, Guzzi PH. AI-enabled pipeline for virus detection, validation, and SNP discovery from next-generation sequencing data. *Frontiers in Genetics* 2024; 15: 1492752.

59. Santos JD, Sobral D, Pinheiro M, et al. INSaFLU-TELEVIR: an open web-based bioinformatics suite for viral metagenomic detection and routine genomic surveillance. *Genome Medicine* 2024; 16: 61.

60. Vries JJC de, Brown JR, Fischer N, et al. Benchmark of thirteen bioinformatic pipelines for metagenomic virus diagnostics using datasets from clinical samples. *Journal of Clinical Virology* 2021; 141: 104908.

61. Tian Z, Zhao F. AI-empowered human microbiome research. *Gut* 2025; 75: 1432–1446.

62. Nawaz MS, Fournier-Viger P, Nawaz S, et al. HieRMVir: Interpretable Viral Classification via Hierarchical Deep Learning. *IEEE Journal of Biomedical and Health Informatics* 2025; 1–8.

63. Hudu SA, Alshrari AS, Abu-Shoura EJI, et al. A Critical Review of the Prospect of Integrating Artificial Intelligence in Infectious Disease Diagnosis and Prognosis. *Interdisciplinary Perspectives on Infectious Diseases* 2025; 2025: 6816002.

64. Sahoo A, Singh J, Alam K, et al. Medical Artificial Intelligence: Opportunities and Challenges In Infectious Disease Management. *Current Medicinal Chemistry* 2025; 33: 1616–1634.

65. Cesaro A, Hoffman SC, Das P, et al. Challenges and applications of artificial intelligence in infectious diseases and antimicrobial resistance. *npj Antimicrobials and Resistance* 2025; 3: 2.

66. Livieratos A, Kagadis GC, Gogos C, et al. AI Methods Tailored to Influenza, RSV, HIV, and SARS-CoV-2: A Focused Review. *Pathogens* 2025; 14: 748.

67. Shorten C, Khoshgoftaar TM, Furht B. Deep Learning applications for COVID-19. *Journal Of Big Data* 2021; 8: 18.

68. Chakraborty S. Democratizing nucleic acid-based molecular diagnostic tests for infectious diseases at resource-limited settings – from point of care to extreme point of care. *Sensors & Diagnostics* 2024; 3: 536–561.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

69. Er AG, Gülten E, Temoçin F, et al. Artificial Intelligence in Infectious Diseases and Clinical Microbiology: Current and Future Directions. *Klimik Dergisi/Klimik Journal* 2026; 1.

70. Alsulimani A, Akhter N, Jameela F, et al. The Impact of Artificial Intelligence on Microbial Diagnosis. *Microorganisms* 2024; 12: 1051.

71. Cha S, Elkanzi M. A Review on the Role of Machine Learning in Predicting the Spread of Infectious Diseases. *Metaheuristic Optimization Review* 2024; 2: 14–27.

72. Syrowatka A, Kuznetsova M, Alsubai A, et al. Leveraging artificial intelligence for pandemic preparedness and response: a scoping review to identify key use cases. *npj Digital Medicine* 2021; 4: 96.

73. Kumar M. Leveraging AI to Predict and Manage Infectious Disease Outbreaks: Insights Gained from COVID-19. *Journal of Artificial Intelligence & Cloud Computing* 2023; 1–6.

74. Rahman SZ, Senthil R, Ramalingam V, et al. Predicting Infectious Disease Outbreaks with Machine Learning and Epidemiological Data. *Journal Of Advanced Zoology* 2023; 44: 110–121.

75. Numérique LL École supérieure en Sciences et Technologies de l'Informatique et du, Melchane S, Elmir Y, et al. Artificial Intelligence for Infectious DiseasePrediction and Prevention: A ComprehensiveReview. *Acta Universitatis Sapientiae Informatica* 2025; 16: 160–197.

76. Hassan A. AI for Pandemic Preparedness and Infectious Disease Surveillance: Predicting Outbreaks, Modeling Transmission, and Optimizing Public Health Interventions. *International Journal of Research Publication and Reviews* 2025; 6: 4605–4619.

77. Isiaka AB, Anakwenze VN, Ilodinso CR, et al. Harnessing Artificial Intelligence for Early Detection and Management of Infectious Disease Outbreaks. *International Journal of Innovative Research and Development*. Epub ahead of print 14 March 2024. DOI: 10.24940/ijird/2024/v13/i2/feb24016.

78. Zhao AP, Li S, Cao Z, et al. AI for science: Predicting infectious diseases. *Journal of Safety Science and Resilience* 2024; 5: 130–146.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

79. Farhat F, Sohail SS, Alam MT, et al. COVID-19 and beyond: leveraging artificial intelligence for enhanced outbreak control. *Frontiers in Artificial Intelligence* 2023; 6: 1266560.

80. Fazil AW. Artificial Intelligence in Infectious Disease Detection and Control: A Systematic Literature Review of Advancements, Challenges, and Future Directions. *Afghanistan Journal of Infectious Diseases* 2025; 3: 113–126.

81. Zhong J, Wang Y, Kumar R. A Systematic Review of Machine Learning Applications in Infectious Disease Prediction, Diagnosis, and Outbreak Forecasting. *Preprints.org*. Epub ahead of print 1 July 2025. DOI: 10.20944/preprints202504.1250.v2.

82. Olawade DB, Wada OJ, David-Olawade AC, et al. Using artificial intelligence to improve public health: a narrative review. *Frontiers in Public Health* 2023; 11: 1196397.

83. Dhanda SS, Panwar D, Lin C, et al. Advancement in public health through machine learning: a narrative review of opportunities and ethical considerations. *Journal Of Big Data*; 12. Epub ahead of print 4 July 2025. DOI: 10.1186/s40537-025-01201-x.

84. Mayaki LD. Enhancing real-time infectious disease surveillance through AI-driven early warning and predictive outbreak detection systems. *GSC Biological and Pharmaceutical Sciences* 2025; 33: 1–14.

85. Michael B, Veerasami H, Jayaprakash N. Artificial intelligence and big data for decoding infectious disease transmission dynamics and outbreak prediction. *Decoding Infection and Transmission* 2026; 4: 100079.

86. Anand LG, Lawrence A. Predictive Modeling of Infectious Disease Outbreaks Through AI and Big Data Analytics. Epub ahead of print 4 December 2025. DOI: 10.14293/pr2199.002523.v1.

87. GulMohamed RB, Hetany WA, Almaimani HA, et al. DengueGNN: Graph-based deep learning for modeling disease spread dynamics and prediction. *Scientific Reports*; 16. Epub ahead of print 30 March 2026. DOI: 10.1038/s41598-026-43073-y.

88. Zvyagin M, Brace A, Hippe K, et al. GenSLMs: Genome-scale language models reveal SARS-CoV-2 evolutionary dynamics. *OSTI OAI (US Department of Energy Office of Scientific and Technical Information)* 2023; 37: 683–705.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

89. Kamyshnyi OM, Halabitska I, Oksenych V, et al. Forecasting Influenza Epidemics and Pandemics in the Age of AI and Machine Learning. *Reviews in Medical Virology*; 36. Epub ahead of print 1 January 2026. DOI: 10.1002/rmv.70107.

90. Ali H. Integrating multimodal health datasets with AI-driven epidemiological modeling to accurately forecast infectious disease transmission across populations. *Magna Scientia Advanced Biology and Pharmacy* 2026; 17: 33–51.

91. David M Lucky. Enhancing real-time infectious disease surveillance through AI-driven early warning and predictive outbreak detection systems. *Zenodo (CERN European Organization for Nuclear Research)*. Epub ahead of print 31 December 2025. DOI: 10.5281/zenodo.17877222.

92. David M Lucky. Enhancing real-time infectious disease surveillance through AI-driven early warning and predictive outbreak detection systems. *Zenodo (CERN European Organization for Nuclear Research)*. Epub ahead of print 31 December 2025. DOI: 10.5281/zenodo.17877221.

93. Ghosh BM. Artificial Intelligence in Predictive Epidemiology and Public Health Surveillance. In: *Zenodo (CERN European Organization for Nuclear Research)*. European Organization for Nuclear Research. Epub ahead of print 15 March 2026. DOI: 10.5281/zenodo.19038107.

94. Gao J, Heintz J, Mack C, et al. Evidence-driven spatiotemporal COVID-19 hospitalization prediction with Ising dynamics. *Nature Communications* 2023; 14: 3093.

95. Wang L, Adiga A, Chen J, et al. CausalGNN: Causal-Based Graph Neural Networks for Spatio-Temporal Epidemic Forecasting. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Association for the Advancement of Artificial Intelligence, pp. 12191–12199.

96. Ghorbian M, Ghorbian Saeid, Ghobaei-Arani M. AI-driven techniques for detection and mitigation of SARS-CoV-2 spread: a review, taxonomy, and trends. *Clinical and Experimental Medicine* 2025; 25: 204.

97. Boniolo F, Dorigatti E, Ohnmacht AJ, et al. Artificial intelligence in early drug discovery enabling precision medicine. *Expert Opinion on Drug Discovery* 2021; 16: 991–1007.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

98. Han W, Chen N, Xu X, et al. Predicting the antigenic evolution of SARS-COV-2 with deep learning. *Nature Communications* 2023; 14: 3478.

99. Prasad K, Kumar V. Artificial intelligence-driven drug repurposing and structural biology for SARS-CoV-2. *Current Research in Pharmacology and Drug Discovery* 2021; 2: 100042.

100. Galindez G, Matschinske J, Rose T, et al. Lessons from the COVID-19 pandemic for advancing computational drug repurposing strategies. *Nature Computational Science* 2021; 1: 33–41.

101. Sabrol H, Madamala R, Arora S, et al. AI-Enhanced Drug Repurposing: A Graph Neural Network and Self-Supervised Learning Framework. 2025; 877–881.

102. Kasaraneni CK. AI-Driven Drug Repurposing: A Graph Neural Network and Self-Supervised Learning Approach. 2025; 1–6.

103. Murata T. COVID-19 and Networks. *New Generation Computing* 2021; 39: 469–481.

104. Atitallah SB, Rabah CB, Driss M, et al. Self-supervised learning for graph-structured data in healthcare applications: A comprehensive review. *Computers in Biology and Medicine* 2025; 188: 109874.

105. Cao L, Liu Q. COVID-19 Modeling: A Review. *medRxiv* 2022; 57: 1–42.

106. Kolarič A, Jukič M, Bren U. Machine learning in antiviral drug design. *Bioorganic & Medicinal Chemistry* 2025; 132: 118426.

107. Serafim MSM, Santos VS dos, Gertrudes JC, et al. Machine learning techniques applied to the drug design and discovery of new antivirals: a brief look over the past decade. *Expert Opinion on Drug Discovery* 2021; 16: 961–975.

108. Li K, Yuan Q, Qi C, et al. Artificial Intelligence Revolutionizes Anti-Infective Drug Discovery: From Target Identification to Lead Optimization. *iMetaMed*; 1. Epub ahead of print 16 October 2025. DOI: 10.1002/imm3.70011.

109. Keith JA, Vassilev-Galindo V, Cheng B, et al. Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems. *arXiv (Cornell University)* 2021; 121: 9816–9872.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

110. Ameer N, Razzak S, Fatima I, et al. Role of Artificial Intelligence in Drug Discovery and Design: From Foundational Principles to Emerging Applications in Antiviral Therapeutics. *Currents in pharmaceutical research* 2025; 3: 56–97.

111. V NL, K D. Computational Frameworks for Antiviral Drug Discovery: A Comparative Review of Integrative Models. 2025; 707–713.

112. Krishnateja D, Aruna R, Vyshnavi DV, et al. Computational Approaches for the Design and Development of Antiviral Compounds: A Review. *International Journal of All Research Education & Scientific Methods* 2025; 13: 3342–3354.

113. Sadybekov A, Katritch V. Computational approaches streamlining drug discovery. *Nature* 2023; 616: 673–685.

114. Muratov E, Amaro RE, Andrade CH, et al. A critical overview of computational approaches employed for COVID-19 drug discovery. *Chemical Society Reviews* 2021; 50: 9121–9151.

115. Singh N, Villoutreix BO. Resources and computational strategies to advance small molecule SARS-CoV-2 discovery: Lessons from the pandemic and preparing for future health crises. *Computational and Structural Biotechnology Journal* 2021; 19: 2537–2548.

116. Al-Nema M, Rendeiro MM, Liao D, et al. Computational approaches for flavivirus drug discovery: targeting structural and non-structural proteins. *RSC Medicinal Chemistry*. Epub ahead of print 1 January 2026. DOI: 10.1039/d6md00108d.

117. Delft A von, Hall MD, Kwong AD, et al. Accelerating antiviral drug discovery: lessons from COVID-19. *Nature Reviews Drug Discovery* 2023; 22: 585–603.

118. Namba-Nzanguim CT, Turón G, Simoben CV, et al. Artificial intelligence for antiviral drug discovery in low resourced settings: A perspective. *Frontiers in Drug Discovery*; 2. Epub ahead of print 2 November 2022. DOI: 10.3389/fddsv.2022.1013285.

119. Liu W, Pang P, Wu CA, et al. New approach methodologies for drug discovery. *Cell* 2026; 189: 1877–1903.

120. Winkler DA. Computational repurposing of drugs for viral diseases and current and future pandemics. *Journal of Mathematical Chemistry* 2024; 62: 2844–2879.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

121. Udegbe FC, Ebulue OR, Ebulue CC, et al. MACHINE LEARNING IN DRUG DISCOVERY: A CRITICAL REVIEW OF APPLICATIONS AND CHALLENGES. *Computer Science & IT Research Journal* 2024; 5: 892–902.

122. Pushkaran AC, Arabi AA. From understanding diseases to drug design: can artificial intelligence bridge the gap? *Artificial Intelligence Review*; 57. Epub ahead of print 11 March 2024. DOI: 10.1007/s10462-024-10714-5.

123. Askr H, Elgeldawi E, Ella HA, et al. Deep learning in drug discovery: an integrative review and future challenges. *Artificial Intelligence Review* 2022; 56: 5975–6037.

124. Hasselgren C, Oprea TI. Artificial Intelligence for Drug Discovery: Are We There Yet? *arXiv (Cornell University)* 2023; 64: 527–550.

125. Carini C, Seyhan AA. Tribulations and future opportunities for artificial intelligence in precision medicine. *Journal of Translational Medicine* 2024; 22: 411.

126. Bae H, Ji H, Konstantinov K, et al. Artificial Intelligence-Driven Nanoarchitectonics for Smart Targeted Drug Delivery. *Advanced Materials*; 37. Epub ahead of print 7 August 2025. DOI: 10.1002/adma.202510239.

127. Tade RS, Jain SN, Satyavijay JT, et al. Artificial Intelligence in the Paradigm Shift of Pharmaceutical Sciences: A Review. *Nano Biomedicine and Engineering* 2023; 16: 64–77.

128. Cascarano A, Mur-Petit J, Hernández-González J, et al. Machine and deep learning for longitudinal biomedical data: a review of methods and applications. *Artificial Intelligence Review* 2023; 56: 1711–1771.

129. Zhang X, Zhang D, Zhang X, et al. Artificial intelligence applications in the diagnosis and treatment of bacterial infections. *Frontiers in Microbiology* 2024; 15: 1449844.

130. Krishnan G, Singh S, Pathania M, et al. Artificial intelligence in clinical medicine: catalyzing a sustainable global healthcare paradigm. *Frontiers in Artificial Intelligence* 2023; 6: 1227091.

131. Sounderajah V, Guni A, Liu X, et al. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. *Nature Medicine* 2025; 31: 3283–3289.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

132. Araf I, Idri A, Chairi I. Cost-sensitive learning for imbalanced medical data: a review. *Artificial Intelligence Review*; 57. Epub ahead of print 1 March 2024. DOI: 10.1007/s10462-023-10652-8.

133. McFadden B, Reynolds M, Inglis TJJ. Developing machine learning systems worthy of trust for infection science: a requirement for future implementation into clinical practice. *Frontiers in Digital Health* 2023; 5: 1260602.

134. Wang Y, Liu L, Wang C. Trends in using deep learning algorithms in biomedical prediction systems. *Frontiers in Neuroscience* 2023; 17: 1256351.

135. Esteva A, Chou K, Yeung S, et al. Deep learning-enabled medical computer vision. *npj Digital Medicine* 2021; 4: 5.

136. Çubukçu HC, Topçu Dİ, Yenice S. Machine learning-based clinical decision support using laboratory data. *Clinical Chemistry and Laboratory Medicine (CCLM)* 2023; 62: 793–823.

137. Zhou SK, Greenspan H, Davatzikos C, et al. A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies With Progress Highlights, and Future Promises. *arXiv (Cornell University)* 2021; 109: 820–838.

138. Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health: A scoping review. *npj Digital Medicine* 2022; 5: 171.

139. Allam A, Feuerriegel S, Rebhan M, et al. Analyzing Patient Trajectories With Artificial Intelligence. *Journal of Medical Internet Research*; 23. Epub ahead of print 3 December 2021. DOI: 10.2196/29812.

140. Lu M, Huang C, Hung C, et al. Artificial intelligence predicts direct-acting antivirals failure among hepatitis C virus patients: A nationwide hepatitis C virus registry program. *Clinical and Molecular Hepatology* 2023; 30: 64–79.

141. Pham M-K, Tan T, Crane M, et al. Explainable AI for infection prevention and control: modeling CPE acquisition and patient outcomes in an Irish hospital with transformers. *BMC Medical Informatics and Decision Making* 2025; 25: 391.

142. Lu M, Yu M. Correspondence on Letter regarding "Toward hepatitis C virus elimination using artificial intelligence". *Clinical and Molecular Hepatology* 2024; 30: 274–275.

143. Ajuwon BI, Awotundun ON, Richardson A, et al. Machine learning prediction models for clinical management of blood-borne viral infections: a

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin

systematic review of current applications and future impact. *International Journal of Medical Informatics* 2023; 179: 105244.

144. Yang J, Clifton L, Dung NT, et al. Mitigating machine learning bias between high income and low–middle income countries for enhanced model fairness and generalizability. *Scientific Reports* 2024; 14: 13318.

145. Arora A, Alderman J, Palmer J, et al. The value of standards for health datasets in artificial intelligence-based applications. *Nature Medicine* 2023; 29: 2929–2938.

146. Goetz L, Seedat N, Vandersluis R, et al. Generalization—a key challenge for responsible AI in patient-facing clinical applications. *npj Digital Medicine* 2024; 7: 126.

Published by DAC Democratic Arabic Center
DAC Democratic Arabic Center GmbH, Berlin