Clinical Microbiology and Infection 26 (2020) 584–595

Contents lists available at ScienceDirect

Clinical Microbiology and Infection

journal homepage: www.clinicalmicrobiologyandinfection.com

Narrative review

Machine learning for clinical decision support in infectious diseases: a narrative review of current applications

N. Peiffer-Smadja [1, 2, *], T.M. Rawson [1], R. Ahmad [1], A. Buchard [3], P. Georgiou [4], F.-X. Lescure [2, 5], G. Birgand [1], A.H. Holmes [1]

[1] National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance, Imperial College London, London, UK
[2] French Institute for Medical Research (Inserm), Infection Antimicrobials Modelling Evolution (IAME), UMR 1137, University Paris Diderot, Paris, France
[3] Babylon Health, London, UK
[4] Department of Electrical and Electronic Engineering, Imperial College, London, UK
[5] Infectious Diseases Department, Bichat-Claude Bernard Hospital, Assistance-Publique Hôpitaux de Paris, Paris, France

ARTICLE INFO

Article history:
Received 27 June 2019
Received in revised form
29 August 2019
Accepted 9 September 2019
Available online 17 September 2019

Editor: A. Kalil

Keywords:
Artificial intelligence
Clinical decision support system
Infectious diseases
Information technology
Machine learning

ABSTRACT

Background: Machine learning (ML) is a growing field in medicine. This narrative review describes the current body of literature on ML for clinical decision support in infectious diseases (ID).
Objectives: We aim to inform clinicians about the use of ML for diagnosis, classification, outcome prediction and antimicrobial management in ID.
Sources: References for this review were identified through searches of MEDLINE/PubMed, EMBASE, Google Scholar, biorXiv, ACM Digital Library, arXiV and IEEE Xplore Digital Library up to July 2019.
Content: We found 60 unique ML-clinical decision support systems (ML-CDSS) aiming to assist ID clinicians. Overall, 37 (62%) focused on bacterial infections, 10 (17%) on viral infections, nine (15%) on tuberculosis and four (7%) on any kind of infection. Among them, 20 (33%) addressed the diagnosis of infection, 18 (30%) the prediction, early detection or stratification of sepsis, 13 (22%) the prediction of treatment response, four (7%) the prediction of antibiotic resistance, three (5%) the choice of antibiotic regimen and two (3%) the choice of a combination antiretroviral therapy. The ML-CDSS were developed for intensive care units ($n = 24$, 40%), ID consultation ($n = 15$, 25%), medical or surgical wards ($n = 13$, 20%), emergency department ($n = 4$, 7%), primary care ($n = 3$, 5%) and antimicrobial stewardship ($n = 1$, 2%). Fifty-three ML-CDSS (88%) were developed using data from high-income countries and seven (12%) with data from low- and middle-income countries (LMIC). The evaluation of ML-CDSS was limited to measures of performance (e.g. sensitivity, specificity) for 57 ML-CDSS (95%) and included data in clinical practice for three (5%).
Implications: Considering comprehensive patient data from socioeconomically diverse healthcare settings, including primary care and LMICs, may improve the ability of ML-CDSS to suggest decisions adapted to various clinical contexts. Currents gaps identified in the evaluation of ML-CDSS must also be addressed in order to know the potential impact of such tools for clinicians and patients. **N. Peiffer-Smadja, Clin Microbiol Infect 2020;26:584**

Introduction

Artificial intelligence (AI), initiated in 1956, is often defined as the study of 'intelligent agents', devices that perceive their environment and take actions that maximize the likelihood of successfully achieving their goals [1]. AI in healthcare began with the development of expert systems, based on rules captured from interviews with medical experts, then translated and programmed [2]. With about 450 rules, MYCIN, the first expert system in medicine, was developed in 1976 and aimed to suggest antibiotic regimens for severe bacterial infections [3]. However, MYCIN was never actually used in practice because of the lack of system

* Corresponding author. Nathan Peiffer-Smadja, National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance, Imperial College London, London, UK.
E-mail address: n.peiffer-smadja@ic.ac.uk (N. Peiffer-Smadja).

https://doi.org/10.1016/j.cmi.2019.09.009
1198-743X/

integration into clinical work. Another major limitation of expert systems is the large volume of rules needed to capture experts' knowledge in difficult clinical decisions.

Machine learning (ML) was developed to overcome the constraints of expert systems [2]. In ML, engineers programme algorithms able to define their own rules from data. Thus, human hand-coded rules are replaced by the artificial finding of rules from data. This allows ML systems to learn from data and interpret unknown situations. Among the panel of ML techniques developed, deep learning based on artificial neural networks is the most famous [4]. The performance and ability of the machine to learn is driven by the volume and quality of data provided, that is why ML systems are sometimes called data-intensive systems (Fig. 1).

The availability of healthcare data dramatically increased with electronic health record systems and the rise of connected devices [5]. The optimal analysis and interpretation of this important volume of data, called 'big data', needs the computing power of modern machines. In this context, ML tools integrating and making sense of huge amounts of complex data are becoming popular. Clinical decision support systems (CDSS) can be defined as software in which the characteristics of an individual patient are used to present patient-specific assessments or recommendations to the clinician towards a decision [6].Most current CDSS or computer-aided diagnosis or therapy are expert systems [7], but Machine Learning-Clinical Decision Support Systems (ML-CDSS) are drawing increased interest (Fig. 2) [8–10].

ML systems have been developed in many fields of medicine, including radiology, with the interpretation of images from chest X-ray or magnetic resonance imaging for diagnostic purposes [11–15]. The first FDA approval for an autonomous AI system took place in 2018 with IDx, a ML system used to detect diabetic retinopathy in retinal fundus photographs [16]. In infectious diseases (IDs), most ML work focuses on research, drug development or clinical microbiology. Systems have been developed to analyse bacterial genome and improve the prediction of resistance [17,18], HIV genotype and predict susceptibility to antiretroviral drugs [19], patterns of epidemics for surveillance purpose [20,21] or to discover new antibacterial drugs or vaccines [22–24] (Table 1).

This article focuses on ML adapted to clinicians' decision (ML-CDSS) in ID. More precisely, we describe the objectives, characteristics, development and assessment of ML systems that may directly help clinicians with their work to diagnose infectious diseases, predict severity, decide whether they should use an antimicrobial and which antimicrobial to choose or what dosage.

Material and methods

Search strategy

References for this review were identified through searches of MEDLINE/PubMed, EMBASE, Google Scholar, biorXiv, ACM Digital Library, arXiV and IEEE Xplore Digital Library for articles by use of a combination of ML keywords ('deep learning', 'artificial intelligence', 'artificial learning', 'machine learning', 'machine intelligence', 'neural networks', 'probabilistic networks', 'knowledge representation', 'statistical learning', 'bayesian learning'), decision-making keywords ('medical decision', 'decision tool', 'support tool', 'clinical decision', 'physician decision', 'decision algorithm', 'CDSS', 'clinical management', 'decision making'), antimicrobial keywords ('antimicrobial', 'antibiotic', 'anti-infective', 'antifungal', 'antiparasitic', 'antiviral') or sepsis keywords ('sepsis', 'septic shock'). We included articles resulting from these searches and relevant references cited in those articles up to July 2019.

Study selection

Prospective and retrospective articles in English that reported original research on ML-CDSS for ID were included. We included development reports, implementation studies, clinical trials or qualitative studies in primary, secondary, tertiary care including intensive care and paediatrics. We excluded studies that describe expert system CDSS, as defined by the use of manually programmed rules, studies with ML systems that use data not currently available in routine clinical care (e.g. bacterial genome) or with outcomes irrelevant to clinicians' use for usual care (e.g. development of a new biomarker).

Definitions of learning methods

ML methods can be divided into supervised learning, unsupervised learning and reinforcement learning [2]. Supervised learning refers to algorithms using labelled data as a training dataset. Labelled data are data in which the outcome of interest is defined; for example, to train an algorithm for sepsis prediction, we use a dataset in which patients are already defined as having sepsis or not. The algorithm will then choose the best model to predict the outcome of interest. In unsupervised learning, data are used without a predefined outcome of interest. Algorithms are left to their own to find patterns and to extract hidden structure from data without any expert labelling. Unsupervised learning is mainly used in medicine to do clustering, aiming to discover inherent grouping in the data, such as similar groups of patients based on clinical data [25]. Reinforcement learning involves algorithms discovering actions that yield the greatest rewards through trial and error. In this category, the algorithm is programmed to consider survival or a reduced hospital length of stay as a reward. A training dataset is used by the algorithm to conduct multiple tests in order to develop the model that obtains the highest reward [2,26].

Analysis of data in the selected articles

In each article, we extracted the variables that were analysed by the ML-CDSS. We split these variables into structured and unstructured variables. Structured variables are variables in which the possible values are predefined, e.g. for wound aspect, 0 = no necrosis, 1 = non-viable tissue, etc. Unstructured variables refer to free text such as a description of the wound by a clinician. We further split structured variables into demographic data (e.g. age, ethnicity), medical or surgical history (e.g. comorbidities, date of admission), vitals (e.g. temperature), symptoms or physical examination findings (e.g. cough, lung auscultation), laboratory workup (e.g. creatinine), microbiology workup (e.g. antibiotic susceptibility testing), therapeutic history (e.g. type of surgery, administered drugs) or other workup (biopsy results, medical imaging). Medical imaging can be considered as a structured variable, for example 'presence of lung infiltrates' can be mapped a priori to a uniquely defined input value, or as an unstructured variable when the image itself is fed as input to the CDSS.

Results

General characteristics of ML-CDSS

Among the 126 abstracts identified and assessed for eligibility, 14 were excluded because the variables were not available in routine clinical practice (e.g. bacterial genome), 29 because the CDSS was an expert system and eight because the outcome of the CDSS was not clinical (e.g. development of a novel biomarker).

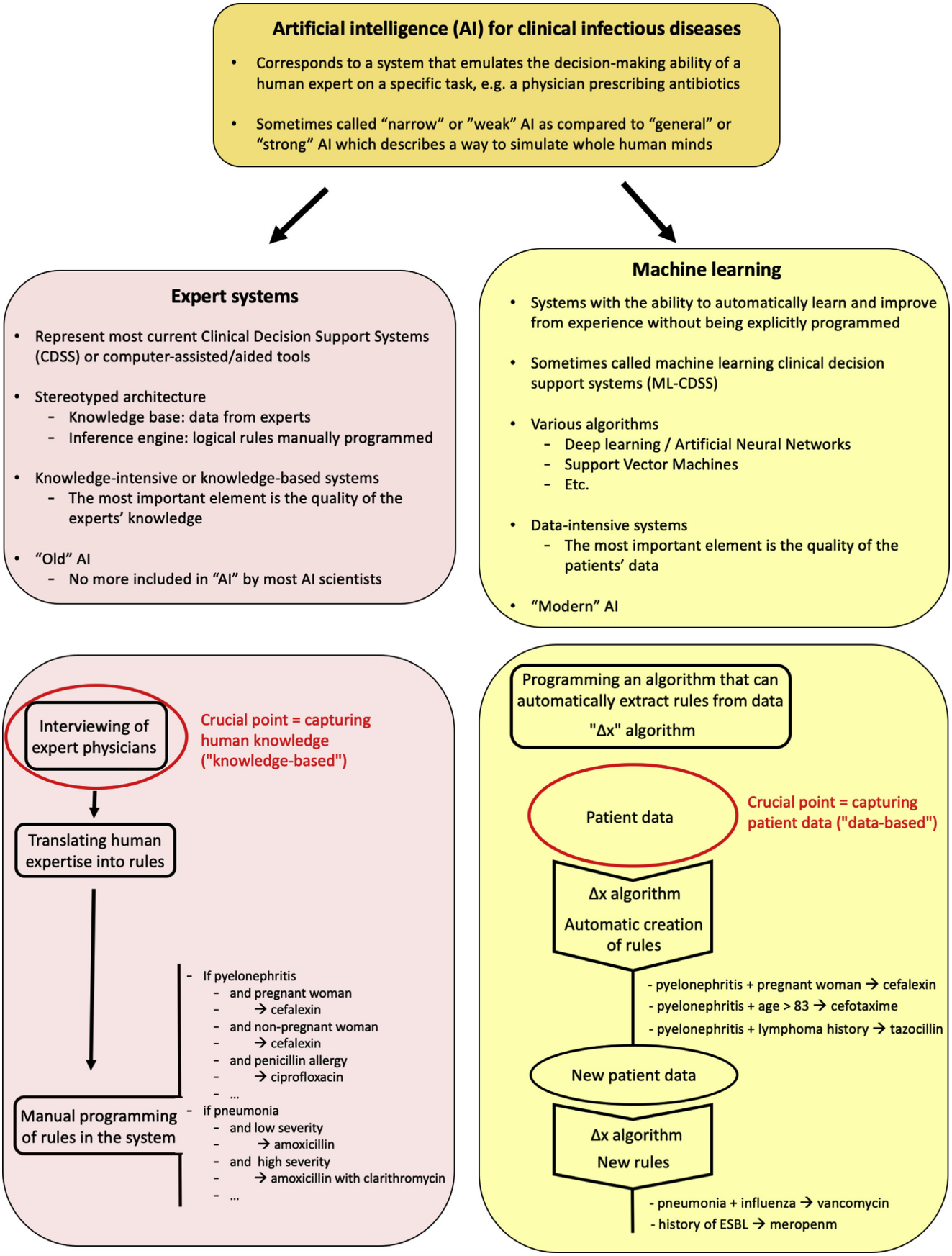

Fig. 1. Artificial intelligence for clinical infectious diseases.

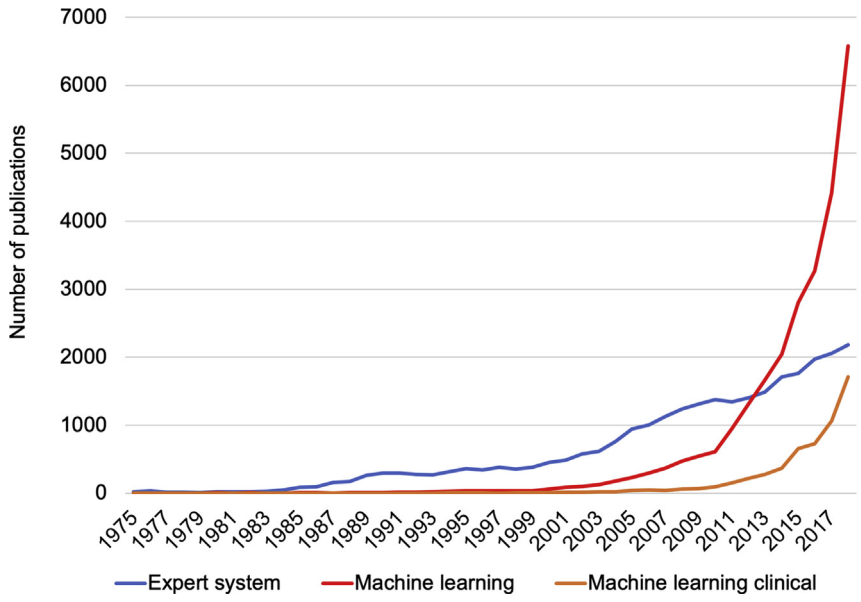

Fig. 2. Trends in the number of publications using selected search terms in PubMed in August 2019.

In total, 75 papers with 60 unique ML-CDSS addressing clinical ID decisions were included. Thirty-one articles were found in MEDLINE/PubMed, 12 in free repository of electronic preprints such as arXiv ($n = 8$) or bioRxiv ($n = 4$), ten in Google Scholar and ten in IEEE Xplore Digital Library. The main characteristics of the ML-CDSS are summarized in Table 2 and detailed in the Supplementary material.

Overall, 37 (62%) ML-CDSS focused on bacterial infections, 10 (17%) on viral infections, nine (15%) on tuberculosis and four (7%) on any kind of infection. Among them, 20 (33%) addressed the diagnosis of infection, 18 (30%) the prediction, early detection or stratification of sepsis, 13 (22%) the prediction of treatment response, four (7%) the prediction of antibiotic resistance, three (5%) the choice of antibiotic regimen and two (3%) the choice of a combination antiretroviral therapy (cART). Most of the ML-CDSS were developed with data from adult patients ($n = 54$, 90%) but some targeted paediatric patients ($n = 3$, 5%) [27–29] or neonates ($n = 3$, 5%) [30–32].

The vast majority of ML-CDSS ($n = 58$, 97%) used supervised learning whereas two (3%) used reinforcement learning [33,34].

Prediction, early detection or stratification of sepsis

Eighteen ML-CDSS (30%) addressed the prediction, early detection or stratification of sepsis in the Intensive Care Unit (ICU). Among these ML-CDSS, 16 exclusively analysed structured patient variables: vitals ($n = 15$), laboratory data ($n = 12$), basic demographic information ($n = 7$), medical history limited to main comorbidities and date of admission ($n = 7$), therapeutic data ($n = 5$) and electrocardiogram waveform ($n = 1$). Two ML-CDSS added unstructured clinical data to their model, one specifically looked for antibiotic prescription in nursing notes to predict sepsis [35], but the other did not give details [36]. No CDSS for sepsis prediction used symptoms, physical examination findings nor microbiology data.

All the 18 ML-CDSS were developed with data from high-income countries (HICs), with six ML-CDSS using the MIMIC (Multiparameter Intelligent Monitoring in Intensive Care) dataset. This dataset contains openly available data from ~50 000 critical care patients admitted to a Medical Center in Boston [37].

Seventeen ML-CDSS were evaluated with measures of performance such as sensitivity or specificity and one ML-CDSS was evaluated in the clinical setting. This ML-CDSS was developed by a company that published a set of papers describing its development, performance and impact in quasi-experimental studies [38]. They reported a relative reduction of in-hospital mortality between 30% and 60% after implementation of their ML-CDSS in ICUs [39–41] but do not use statistical methods adapted to quasi-experimental designs [42]. Indeed, they directly compare in-hospital mortality before and after the use of the ML-CDSS, a measure that is susceptible to biases such as the history bias, i.e. the possibility that the change in the outcome may be linked to other changes in the environment rather than to the addition of the ML-CDSS [43]. They also published a randomized clinical trial in which 67 patients were randomized to a group monitored by a machine learning algorithm and 75 patients to a control group [44]. The trial reported a relative reduction of 58% of mortality in the group monitored by the machine learning algorithm in which patients had blood cultures drawn and antibiotics administered approximately 2.8 hours before the control group.

Diagnosis of infection

Twenty ML-CDSS (33%) focused on the diagnosis of infection, to help clinicians decide if a patient has an infection, which infection (e.g. viral or bacterial) or an alternative non-infectious diagnosis. More precisely, six ML-CDSS address the diagnosis of tuberculosis (TB) in outpatient settings [45–50], five the diagnosis of bacterial infection in hospitalized patients [27,51–54], four the diagnosis of surgical site infection (SSI) [30,55–57], three the diagnosis of infection in emergency department [58–60] and two the distinction between bacterial and viral meningitis [61,62]. Among the five ML-CDSS to diagnose bacterial infection in hospitalized patients, two ML-CDSS included any kind of infection [27,51], two focused on the prediction of positive blood cultures [53,54] and one on MRSA infection [52]. The ML-CDSS for patients hospitalized in surgical wards aimed to diagnose SSIs following open abdominal surgery [55,57] surgery for head or neck cancer [56] or following any intervention on neonates [30].

Among the 20 ML-CDSS, 18 exclusively analysed structured patient variables: demographics data ($n = 15$), medical history

Table 1
List of reviews with a focus on machine learning in the field of infectious diseases

Topic	Title	Year of publication	Review type	Detailed outcome	Settings	Including CDSS
Bacterial resistance	Genome-based prediction of bacterial antibiotic resistance [18]	2018	Narrative	Analysis of bacterial genome to predict resistance	Clinical microbiology lab	No
	Machine learning: novel bioinformatics approaches for combating antimicrobial resistance [17]	2017	Narrative	Analysis of bacterial genome to predict resistance	Clinical microbiology lab	No
Sepsis	Emerging technologies for molecular diagnosis of sepsis [99]	2018	Narrative	Microbiological diagnosis of sepsis	Clinical microbiology lab	No
				Prediction of sepsis	ICU	Yes
HIV	A survey of machine learning applications in HIV clinical research and care [100]	2017	Systematic	HIV/AIDS clinical research and medical care studies that utilize machine learning methodology	Multiple settings	Yes
	Computer-aided optimization of combined anti-retroviral therapy for HIV: new drugs, new drug targets and drug resistance [101]	2016	Narrative	Analysis of HIV genotype to predict drug susceptibility in vitro or response to combination antiretroviral therapy in vivo	Virology lab	Yes
Infection control	Using online social networks to track a pandemic: A systematic review [102]	2016	Systematic	Analysis of online social network data to track pandemics	Health authorities	No
	Automated surveillance of healthcare-associated infections: state of the art [21]	2017	Narrative	Surveillance of healthcare-associated infections	Infection control unit	No
	Introduction to machine learning in digital healthcare epidemiology [20]	2018	Narrative	Prediction, detection of trends and patterns for surveillance purposes	Infection control unit	No
Molecular biology	Computational approaches for prediction of pathogen-host protein-protein interactions [103]	2015	Narrative	Prediction of pathogen-host protein-protein interaction	Research lab	No
	Progress in computational studies of host-pathogen interactions [104]	2013	Narrative	Prediction of pathogen-host protein-protein interaction	Research lab	No
	Identification of legionella effectors using bioinformatic approaches [105]	2012	Narrative	Identification of *L. pneumophila* effectors	Research lab	No
Vaccine development	Comparative pathogenesis and systems biology for biodefense virus vaccine development [24]	2010	Narrative	Analysis of Virus-Host Interactions	Research lab/Research and development	No
	Systems serology for evaluation of HIV vaccine trials [106]	2017	Narrative	Defining humoral signatures in response to vaccines	Research lab/Research and development	No
Drug discovery	Machine-learning techniques applied to antibacterial drug discovery [23]	2014	Narrative	Antibiotic drug discovery	Research lab/Research and development	No
Medical imaging	Computer-assisted detection of infectious lung diseases: a review [107]	2013	Narrative	Analysis of medical imaging of respiratory tract infections	Radiology department	No
Microscopy	Microscopy in infectious disease research-imaging across scales [108]	2018	Narrative	Analysis of microscopic images in ID research	Research lab	No

($n = 14$), laboratory results ($n = 13$), vitals ($n = 9$), symptoms and physical examination findings ($n = 8$), therapeutic history ($n = 7$), chest X-ray for TB diagnosis ($n = 2$) and microbiology data ($n = 2$). Two ML-CDSS added free clinical text; nursing notes, clinical narrative or chief complaint [51,59]. Seventeen ML-CDSS were developed with data from HICs and three with data from low- and middle-income countries (LMICs), addressing the screening of TB in South Africa [48] or the diagnosis of TB in Turkey [49] or Iran [63].

Nineteen ML-CDSS were evaluated by measures of sensitivity, specificity and receiver ROC curves. One study described a ML-CDSS trained to diagnose bacterial infection using six routinely available blood parameters with data from 160 203 individuals and its use in a prospective observational cohort. Among 104 patients included, the ML-CDSS predicted a bacterial infection in three individuals who were not identified by clinicians as having an infection on admission but were diagnosed later with a bacterial infection [58]. No study assessed the clinical or microbiological impact of the use of ML-CDSS in clinical settings.

Determinants of treatment outcomes

Ten ML-CDSS (17%) addressed the prediction of treatment success for outpatients. Among them, five ML-CDSS predicted the virological response to HIV therapy [64–68] and three to HCV therapy [69–71]. They all analysed therapeutic history but only four used viral genotype or medical history [65–67,71]. Therapeutic history was limited to previous HIV or HCV therapies but not to other drugs that the patient might have been taking. Two ML-CDSS predicted treatment outcome in TB using demographics data (e.g. education level, homelessness), TB history including constitutional symptoms, TB treatment and structured data from the chest radiograph (e.g. size of cavity) [72,73].

Three ML-CDSS (5%) addressed the prediction of *C. difficile* colitis complications [74] or recurrence [75,76] in hospital settings. We found four ML-CDSS (7%) for the prediction of drug resistance: one predicted the risk of developing multidrug-resistant TB in Chinese patients with TB [77] one combined demographic data and medical history for personalized prediction of baseline antibiotic resistance in urinary tract infection [78] one analysed demographic data, living conditions and treatment history of patients with positive blood cultures to predict baseline susceptibility to ampicillin, ceftriaxone and gentamicin [28] and one used ML to measure the impact of antibiotic exposure on the acquisition of colonization with extended-spectrum β-lactamase-producing Gram-negative bacteria [79].

Thirteen ML-CDSS in this category were developed with patient data from HICs and two with patient data from LMICs to predict TB outcomes, one in Pakistan [72] and the other in Eastern Europe [73]

Table 2
Summary of machine learning clinical decision support systems

ML-CDSS characteristics	$n = 60$ (%)
Medical settings[a]	
Intensive Care Unit	24 (40)
Infectious diseases consultation	15 (25)
Medical ward	9 (15)
Surgical ward	5 (8)
Emergency department	4 (7)
Primary care	3 (5)
Antimicrobial stewardship team	1 (2)
Geographical settings[a]	
High-income countries	54 (90)
Low- and middle-income countries	7 (12)
Population	
Adults	53 (88)
Neonates	3 (5)
Paediatric patients	3 (5)
Retirement-home	1 (2)
Types of decision support	
Diagnostic	
Diagnosis of infection	20 (33)
Prediction of sepsis	18 (30)
Prediction of antibiotic resistance	4 (7)
Therapeutic	
Prediction of treatment response	13 (22)
Antibiotic selection	3 (5)
HIV therapy selection	2 (3)
Type of infection	
Bacterial infection	37 (62)
Viral infection	10 (17)
Mycobacterial infection	9 (15)
Any kind of infection	4 (7)
Types of learning	
Supervised learning	58 (97)
Reinforcement learning	2 (3)
Data	
Clinical data (e.g. demographics, vitals)	52 (87)
Laboratory	38 (63)
Therapy	28 (47)
Microbiology	15 (25)
Other workup (e.g. ECG, imaging)	6 (10)
Unstructured clinical data (free text)	5 (8)
Evaluation[b]	
Performance	57 (95)
Use	3 (5)
Adoption	0

ML, machine learning; CDSS, clinical decision support systems; HIV, human immunodeficiency virus.
[a] One article included patients in medical wards, surgical wards and intensive care units in high and middle-income countries.
[b] We separated studies that describe the performance (e.g. receiver operating characteristic curves) of the ML-CDSS, studies that describe the use of the CDSS in real-life settings and studies that describe the adoption of the CDSS in routine clinical practice.

In addition, one ML-CDSS used data both from HICs and middle-income countries [79]. Another ML-CDSS was developed in HICs and then specifically adapted to the specificities of human immunodeficiency virus (HIV) care in LMICs such as limited access to genotype data, infrequent visits to clinics and restricted list of available drugs [67]. It used the HIV Drug Resistance Database that includes large datasets from patients around the world including sub-Saharan countries.

The 17 ML-CDSS predicting treatment outcomes or antibiotic susceptibilities were evaluated by measures of sensitivity, specificity and receiver operating characteristic (ROC) curves.

Treatment selection

Five ML-CDSS (8%) addressed the choice of antimicrobials: three ML-CDSS aimed to find the optimal antibiotic regimen [33,80,81] and two the optimal cART for HIV [34,82]. All the CDSS focused on the choice of an agent but did not give individual advice on the dose or duration.

The two ML-CDSS to guide cART focused on outpatient settings: one analysed treatment history and treatment objectives but did not consider clinical or biological data [82], whereas the other used demographics, medical history, CD4+ cell count, viral load, genotypic data and treatment history [34]. Among the three ML-CDSS for antibiotic management, one was adapted to ICUs [33]. another one to primary care [81] and the third one targeted antimicrobial stewardship teams [80]. Their use of patient data was heterogeneous; only one ML-CDSS took into account the identification of the pathogen in blood culture results but not the antimicrobial susceptibility testing [33], medical history was not used in two ML-CDSS [33,80] and limited to breastfeeding, pregnancy, allergy and kidney failure in the third [81]. They did not use the therapeutic history of patients, neither data from the physical examination nor local antibiotic resistance rates and were not adapted for a dynamic use over time.

All of these ML-CDSS concerned HICs, using data from North American or European patients and only one ML-CDSS was evaluated in a clinical setting. It was a ML-CDSS focusing on inappropriate prescriptions for review by antimicrobial stewardship pharmacists. The addition of a ML module to an expert system was shown to identify inappropriate prescriptions of piperacillin-tazobactam that were missed by the expert system [80]. However, this study does not report the clinical or microbiological outcomes of this intervention.

Discussion

Healthcare settings

Most ML-CDSS found in the literature were developed for secondary and tertiary care ($n = 57$, 95%) whereas only three ML-CDSS (5%) targeted primary care [48,78,81]. Four ML-CDSS were specifically tailored for emergency care [46,58–60], compared with more than 20 ML-CDSS for ICUs. Moreover, only seven ML-CDSS (12%) were adapted to LMICs [48,49,63,67,72,73,79], six of them concerning HIV or TB.

ML techniques usually need a large amount of data (e.g. more than 128 000 retinal images in a recent ophthalmologic ML system) [83]. Thus, it is no surprise that most ML-CDSS in ID currently should focus on domains generating high-quality databases including ICU and HIV patients from HICs. The MIMIC dataset was a recurrent data source for ML systems predicting sepsis [37]. In the same way, ML-CDSS targeting HIV therapy selection or the prediction of virological response were also based on large open-access databases [66]. The development of ML-CDSS is more difficult in primary care where available patient data for the clinician is limited and databases are scarce. The same statement can be made for LMICs where data extraction is difficult due to the frequent lack of clinical information system. The availability of data is also correlated with healthcare access: vulnerable populations tend to have a lower access to the healthcare system with a more fragmented care. This could undermine the ability of ML-CDSS to make adequate predictions for patients from vulnerable populations and lead to a potential increase in healthcare inequalities [84,85].

The unequal spreading of ML-CDSS across healthcare settings reflects the unequal availability of high-quality and large clinical databases. This is why a special effort should be made to target the development of such databases in underfunded settings such as primary care or LMICs and to include data from diverse settings and populations in training and validation datasets. Global open access

to these databases should be a requirement for future funding in order to increase the opportunities for ML-CDSS development everywhere in the world. Moreover, we have not found any study focusing on fungal or parasitic diseases in this review. These areas may represent interesting opportunities for developing CDSS using machine learning for diagnosis, prediction of resistance or choice of antifungal or antiparasitic therapy. The analysis of the outputs of ML-CDSS at the population level should also be encouraged as it may be useful for public health and research institutes to regularly monitor trends in diagnosis and prescribing and correct potential biases.

If these constraints and potential issues are considered, ML-CDSS appear as credible options for supporting infection management in countries where ID specialists and antimicrobial teams may not be readily available [86].

Choice of patient variables

In the present review, many ML-CDSS did not include some clinically relevant data. Medical history analysed in ML-CDSS was often limited to the presence of two or three predefined diseases (kidney failure, cirrhosis, etc.) [56,87]. No ML-CDSS for the diagnosis of infection considered patients exposed to infected individuals or high-risk environment (e.g. travel history). We did not find any ML-CDSS predicting sepsis and taking into account the results from blood cultures or from any other microbiology sample. Among the 60 ML-CDSS, chest X-Ray was the only medical imaging analysed [46,48,72]. No ML-CDSS to predict antibiotic resistance or to select an antibiotic regimen included data from local antibiotic resistance rates. Paradoxically, ML-CDSS often use less data than a human clinician would do. In some articles, the authors analysed the performance of ML-CDSS when using a reduced set of variables and they found that the sensitivity and specificity of ML-CDSS were systematically better when they used a larger panel of variables [60,87,48]. especially when adding unstructured data [59,36]. Ensuring integration of physical examination or detailed medical history in these systems is of critical importance. Progress in Natural Language Processing (NLP, i.e. the ability of computers to analyse human language) may help the integration of free medical text in future decision systems [88]. Nonetheless, the development of ML-CDSS using minimal variables may be of interest when data are not available across some areas or in resource limited settings. A particular attention should be paid to which variables are used by the ML-CDSS to predict their outcome: e.g. we found a ML-CDSS that used the prescription of antibiotics in ICU to predict sepsis [35], which could provide good performance but seems clinically irrelevant.

ML-CDSS are constrained by the quality and availability of the clinical data used for their development and validation. To ensure that future ML tools are useful for clinicians, efforts should be made to provide comprehensive databases that include relevant clinical data.

Evaluation and use of ML-CDSS

The evaluation of ML-CDSS in clinical ID is still lacking with 57 ML-CDSS (95%) reporting measures of performance such as sensitivity and specificity and only three (5%) evaluated in clinical practice. The validation of the performance in the original database should preferably be done in a prospective database collected for validation purpose in another location. Then, pilot studies testing the usability and usefulness of ML-CDSS in clinical settings should be conducted, followed by clinical trials, preferably randomized, evaluating patients' outcomes, process improvement or cost-effectiveness.

Special attention should be paid to the integration and implementation of systems into clinical practice, and their adoption and utilization by clinicians [89]. Co-design including clinicians, pharmacists engineers, data managers and service users may be key for success and may allow to study the interaction between CDSS and healthcare professionals [90]. The place of ML-CDSS in the workflow of clinicians has to be considered as different clinical tasks are likely to require different degrees of human involvement (Fig. 3) [10]. Data should be easily entered into the CDSS, ideally by automatic extraction from the EHR, and daily refreshed with new patient information. The CDSS need to fit into the complex medical decision-making pathway and support clinicians at different stages sometimes by presenting different

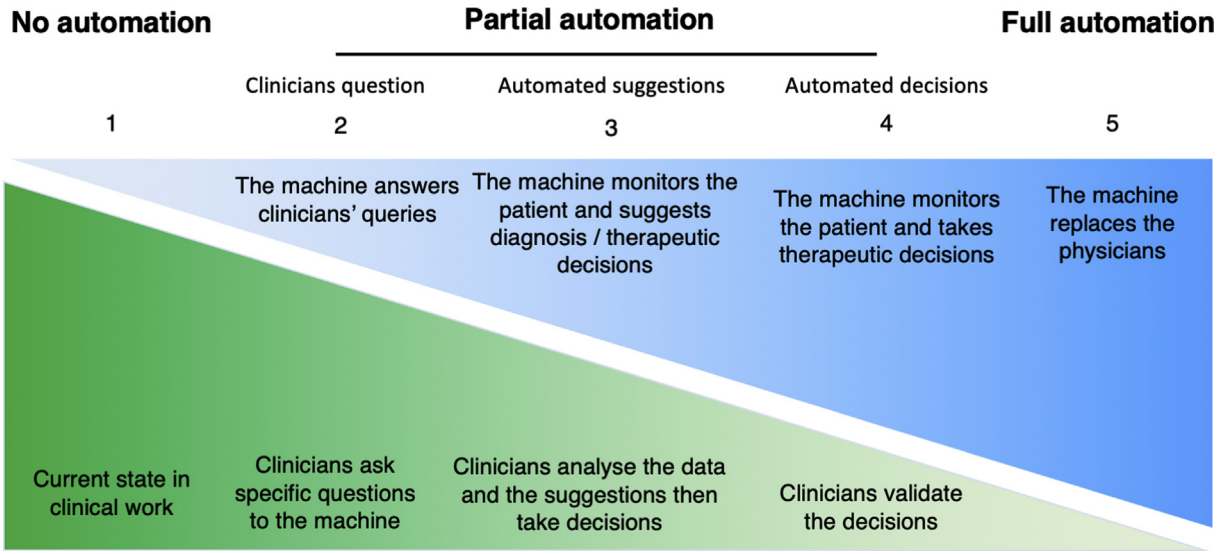

Fig. 3. Interaction between clinicians and machine-learning systems according to automation level.

aspects of data [91]. According to the numerous limits that we found concerning ML-CDSS in ID, the replacement of physicians by ML tools seems highly unlikely. AI is best suited to augmenting human intelligence, bringing big-data into focus to support human decision-making and developers should aim to emphasize the synergy and interaction between physicians and algorithms instead of reporting less interesting clinical-versus-algorithm comparison [8,92,93] Clinicians' expectations are that ML-CDSS will relieve them of repetitive tasks that need to be performed without errors and allow them to focus on tasks that require human skills: analysis of the broader context, abstraction capability but also exercising empathy and building relationships. However, we have not found any in-depth describing of the interaction of physicians with ML-CDSS in real-world settings. Implementation research may help to take into account the global context in which ML-CDSS are used [7].

Unexpected consequences of the use of ML also have to be included in the research agenda. Consequences commonly cited are the risk of 'deskilling' clinicians who use ML-CDSS routinely, the risk of a blind obedience to the CDSS or the risk of decreased interaction between the clinician and the patient. Studies have shown that the development of the electronic health record (EHR) shifted clinicians' time from the patient to the computer and ML systems may represent an additional risk to do so [94,95]. By definition, unexpected consequences cannot be fully anticipated and research on ML-CDSS in real-world settings should aim to collect information on consequences both at the clinician and patient -levels.

Perspectives

In Table 3 we summarize some essential criteria for future ML-CDSS adapted for clinicians' use. Clinicians must know when and why errors by ML in clinical-decision predictions might occur and when they cannot rely on an automated system whose intrinsic mechanisms they do not understand (the so-called 'black box') [8]. Visualization tools highlighting the variable or combination of variables that determined the output ought to be available for the prescriber and are an important parameter in future ML-CDSS [28,96]. ML-CDSS should also detail the trade-off that they apply between risk-taking and caution and show the confidence interval of their decisions and suggestions [9]. The impact of under- and overdiagnosis or treatment is not the same when considering upper respiratory tract infection or sepsis therapy. Besides system improvement, we need to encourage and fund the education of clinicians about AI and ML in order to develop an AI-literate workforce in healthcare worldwide.

In addition, global challenges for the development of artificial intelligence and ML for clinical decisions are to be considered. Table 4 describes the main challenges for ML-CDSS that have been identified in the literature. For example, controls have to ensure that ML-CDSS are registered on public databases, thoroughly evaluated independently of economic interest and validated using standardized frameworks. The privacy protection of data [97], cost-effectiveness research at the organizational and global level [89,98], and ethics discussion about accountability and liability when using ML-CDSS are other potential challenges to deal with.

Table 3
Necessary requirements for clinical decision support systems using machine learning

Target	Type of decision	A clinically meaningful decision should be targeted, for example the diagnosis of a bacterial infection requiring antibiotics or the choice of an antibiotic regimen
	Justification of ML	The use of machine learning instead of expert systems should be justified, for example by the analysis of big data that could not be captured by expert systems.
	settings	ML-CDSS should not only be developed for secondary or tertiary care but should also target such primary care prescribers as GPs or nurse prescribers
		The ML-CDSS should offer a good performance in diverse settings including low- and middle-income countries
	Patients	The ML-CDSS should be performant for different populations of patients like patients with diabetes, chronic kidney disease, cirrhosis, etc.
		If some comorbidities prevent the use of the ML-CDSS, it should be clearly stated and explained
	Users	Different practitioners may not need the same tools even for the same decision. For example, surgeons do not have the same need for ML-CDSS addressing antimicrobial management as antimicrobial stewardship specialists or nurse prescribers in primary care
Data	Comprehensiveness	ML-CDSS should be able to analyse data from different sources such as structured clinical data, vitals, laboratory data, therapeutic history but also unstructured clinical data (e.g. free medical text, medical imaging)
	Availability	ML-CDSS should not require data that are not available for most patients
	Extraction	ML-CDSS should be able to automatically extract data from the electronic health record (EHR)
Decision	Uncertainty	A way to present uncertainty should be found by the developers: e.g. it could give a confidence interval for the diagnosis of bacterial infection or a number of suggestions for antimicrobial therapy with associated probabilities of success according to its calculation
	Clarity	ML-CDSS should display what parameters were the most important for the suggested diagnosis or therapeutic decision
	Flexibility	Clinicians should be able to take out a specific clinical or laboratory element that they do not think reliable and thus have access to modified suggestions. They should also be able to add or emphasize elements that they find particularly important or reliable.
	Memory	Clinicians need ML-CDSS that can keep track of patients and for whom the data can be updated to modify the suggestions of the system
	Limits	ML-CDSS should display very clearly what parameters have been used or not, e.g. medical history, medical imaging, etc. and other potential limits of its decision
Interface	Integration	The ML-CDSS should ideally be integrated in the EHR used by the prescriber. A special attention should be paid to the interoperability of ML-CDSS across different EHR systems
	Speed	The ML-CDSS should be quick to use, should automatically extract pertinent data from the EHR and give its output immediately
	User-friendly	Clinicians should be able to use ML-CDSS without specific training
	Standardization	As multiple ML-CDSS may be available for different diseases and conditions, interfaces for ML-CDSS should be standardized to a common format to increase usability. We should avoid having a software for every decision, e.g. diagnosis, antimicrobial prescribing, dosing, prevention of venous thromboembolic disease, etc.

ML-CDSS, machine learning clinical decision support systems.

Table 4
Global challenges for machine learning clinical decision support systems

Challenges	Description	Potential solutions
Data quantity	There is a need for a huge quantity of clinical data to develop a machine learning (ML) system, especially for unsupervised ML	Sharable/open-source database such as MIMIC-III, a freely accessible critical care database that has helped to develop many ML-CDSS for sepsis in intensive care units
Data quality	ML-clinical decision support systems (CDSS) developers need to have access to high-quality data for training and validation dataset, especially for supervised ML	Electronic health record data standardization Pre-processing of data should be detailed in the description of ML-CCDSS
Comprehensiveness	ML-CDSS must be able to analyse data from different sources such as structured clinical data, vitals, laboratory data but also unstructured free text (e.g. physical examination, medical history, etc.).	Developers should try to include all the sources of data that are used by clinicians to take clinical decision Clinicians should be included during development to make sure that unstructured data are not left out in ML-CDSS (co-design)
Availability	The data that are analysed by the system should be available for most patients (e.g. a CDSS analysing genetic data could not be used in usual care)	Importance of co-design with users, physicians, other prescribers and clinical managers to know what data is available for patients Possibility of adapting the ML-CDSS to different settings in which available data could be different
Interoperability	A multitude of technology platforms are involved in the generation, collection and mobilization of health data. ML-CDSS should be able to have good performance with different platforms and electronic health record systems.	Initiatives as FHIR (Fast Healthcare Interoperability Resources) that standardize healthcare data to common format
Equity	Most ML-CDSS are data-driven. As patients with low socioeconomic status usually have a limited access to healthcare, less data from these patients are available and ML systems could create new inequalities or increase health disparities. Patients with comorbidities should also be included in ML-CDSS.	Developing ML-CDSS with training and validation datasets from socioeconomically diverse health care systems Ensuring that patients with comorbidities (e.g. chronic kidney disease, cirrhosis) are included in the training dataset Ensuring that social determinants of health are being considered in the support system Test ML-CDSS for potential discriminatory behaviour
Adaptability	ML-CDSS should be usable in different clinical settings including primary, secondary and tertiary care. A special attention should be paid to low- and middle-income countries that should not be left out by progress in the development of AI.	Diverse datasets coming from different countries Specificities of LMICs should be considered early in the development of ML-CDSS Availability of antimicrobials in every healthcare setting should be considered
Interpretability	Clinicians should be able to understand the output of ML-CDSS and their determinants for an intelligent use and interpretation	Education of an AI-literate clinician workforce (e.g. training in informatics and data science by the National Library of Medicine) Highlighting the variable or combination of variables that determined the output (e.g. 'saliency map') Reporting of confidence level in the prediction
Security	Alternatives should be developed to make sure that healthcare systems could function normally in case of an electronic problem. ML-CDSS should also be protected against hacking and other ill-willed practices	Systematic debugging Security software Anticipation and preparedness
Validation	ML-CDSS should be validated by an independent organization. Both the risk of under-regulation and over-regulation should be considered	Quality standards, precertification and certification Such regulatory framework as the Software as Medical Device (SaMD) framework developed by the International Medical Device Regulators Forum (IMDRF) working group chaired by the Food and Drug Administration (FDA)
Confidence	ML-CDSS developers should pay special attention to the confidence of clinicians and patients in artificial intelligence. Limited confidence or fear of the users would limit the use of ML and blind confidence also called automation bias could be dangerous	Behavioural science, sociological and anthropological research on artificial intelligence in medicine Training and education of clinicians about the benefits and limits of artificial intelligence and machine learning Studies on the interaction between machines and specifically human skills Co-design between engineers and clinicians
Evaluation	Evidence of the utility and benefits of ML-CDSS should not be limited to analysis of performance (e.g. ROC curves) but should consider clinical outcomes. Ideally, ML-CDSS should be evaluated in real-world settings from pilot study to clinical impact in real-world settings	Studies should not limit their outcomes to the evaluation of performance but should include clinical and microbiological outcomes Randomized controlled trial should try to determine the improvement of medical decision with the addition of ML-CDSS instead of comparing support systems with human clinicians
Accountability	The accountability of each actor (clinicians, ML-CDSS developers, certification organisms) should be defined beforehand	Research and forums on the accountability of AI and ML systems Inspiration from the work that has been done in other fields as driverless cars
Cost	Cost of ML-CDSS including development, data processing, implementation and impact on healthcare should be systematically studied, at the organizational and global level	Cost-effectiveness analysis and research
Privacy	ML-CDSS should respect the privacy of health data and ensure proper consent and patient governance in data collection	Global regulation as the European Union General Data Protection Regulation or United States Health Insurance Portability and Accountability Act Use of anonymized or pseudonymized data Privacy audits
Independence	Regulation agencies should make sure that ML-CDSS are independent from drug and diagnostic test companies and do not favour any treatments or tests for economic purposes	Regulatory frameworks should consider the need for independence Public funding for the development of AI and ML in healthcare
Unexpected consequences	ML-CDSS have the potential to deeply modify health systems and unexpected consequences should be systematically considered and studied	Systematic research on the limits of the ML-CDSS Studying potentially unexpected consequences on clinicians like 'deskilling' (loss of clinicians' skill due to the repeated use of ML-CDSS) or burn-out

Table 4 (*continued*)

Challenges	Description	Potential solutions
Implementation	The implementation of ML-CDSS should be thought of and organized as early as possible in the development	ML-CDSS developing teams should include specialists from implementation science Evaluation should include research on implementation, adoption, use and sustainability Specific task force committees to deal with AI implementation issues may be useful for developing a common vision at a specialty-wide level
Sustainability	ML-CDSS should be sustainable and updated according to new research, data or regulation	Specific and sustainable funding should be given to developers and adopters of ML-CDSS

Conclusion

Current ML tools cover a range of clinical outcomes such as the prediction of sepsis in ICUs, the diagnosis of TB or SSI, the prediction of virological success of cART or the selection of an antibiotic regimen. However, 57 ML-CDSS out of 60 only reported performance measures such as sensitivity or specificity and evidence is lacking regarding their use and impact in real-life clinical settings. Moreover, current ML-CDSS mainly use patient data from HICs, relying on large available open-access datasets and generally analyse a limited number of structured patient variables. Future ML-CDSS in ID should be developed in diverse health settings, including primary care and LMICs that are currently underrepresented, and be embedded in a structured process of integration into clinical settings. Future studies should aim to report clinical outcomes following sustainable use in routine clinical care.

Transparency Declaration

Dr Buchard reports that he is employed by Babylon Health. Dr Buchard has provided an expertise on machine learning methods exclusively on selected articles sent by the corresponding author. No tools developed by Babylon Health are discussed in the review.

The authors have no other relevant conflicts of interest to disclose.

Funding

The research was funded by a PhD studentship from Île-de-France Regional Health Agency, Paris, France to Nathan Peiffer-Smadja, by a grant from the French Association de Chimiothérapie Anti-Infectieuse (ACAI) to Nathan Peiffer-Smadja and by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Healthcare Associated Infection and Antimicrobial Resistance at Imperial College London in partnership with Public Health England (PHE). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health or Public Health England.

Acknowledgements

The support of the Economic and Social Research Council (ESRC) as part of the Antimicrobial Cross Council initiative supported by the seven UK research councils, and also the support of the Global Challenges Research Fund, are gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cmi.2019.09.009.

References

[1] Crevier D. A.I: The tumultuous history of the search for artificial intelligence. New York, NY, USA: Basic Books, Inc.; 1993.
[2] Mitchell TM. Machine learning. 1st ed. New York, NY, USA: McGraw-Hill, Inc.; 1997.
[3] Shortliffe EH. Computer-based medical consultations: mycin. Elsevier; 1976. https://doi.org/10.1016/B978-0-444-00179-5.50008-1.
[4] Hinton G. Deep learning - a technology with the potential to transform health care. JAMA 2018;320:1101–2. https://doi.org/10.1001/jama.2018.11100.
[5] Beam AL, Kohane IS. Big data and machine learning in health care. JAMA 2018;319:1317–8. https://doi.org/10.1001/jama.2017.18391.
[6] Sim I, Gorman P, Greenes RA, Haynes RB, Kaplan B, Lehmann H, et al. Clinical decision support systems for the practice of evidence-based medicine. J Am Med Inform Assoc JAMIA 2001;8:527–34.
[7] Rawson TM, Moore LSP, Hernandez B, Charani E, Castro-Sanchez E, Herrero P, et al. A systematic review of clinical decision support systems for antimicrobial management: are we failing to investigate these interventions appropriately? Clin Microbiol Infect 2017;23:524–32. https://doi.org/10.1016/j.cmi.2017.02.028.
[8] Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. JAMA 2017;318:517. https://doi.org/10.1001/jama.2017.7797.
[9] Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf 2019. https://doi.org/10.1136/bmjqs-2018-008370. bmjqs-2018-008370.
[10] Rawson TM, Ahmad R, Toumazou C, Georgiou P, Holmes AH. Artificial intelligence can improve decision-making in infection management. Nat Hum Behav 2019;3:543–5. https://doi.org/10.1038/s41562-019-0583-9.
[11] Choy G, Khalilzadeh O, Michalski M, Do S, Samir AE, Pianykh OS, et al. Current applications and future impact of machine learning in radiology. Radiology 2018;288:318–28. https://doi.org/10.1148/radiol.2018171820.
[12] Lindsey R, Daluiski A, Chopra S, Lachapelle A, Mozer M, Sicular S, et al. Deep neural network improves fracture detection by clinicians. Proc Natl Acad Sci 2018;115:11591–6. https://doi.org/10.1073/pnas.1806905115.
[13] Jha S, Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists. JAMA 2016;316:2353. https://doi.org/10.1001/jama.2016.17438.
[14] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115–8. https://doi.org/10.1038/nature21056.
[15] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44–56. https://doi.org/10.1038/s41591-018-0300-7.
[16] Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. Npj Digit Med 2018;1:39. https://doi.org/10.1038/s41746-018-0040-6.
[17] Macesic N, Polubriaginof F, Tatonetti NP. Machine learning: novel bioinformatics approaches for combating antimicrobial resistance. Curr Opin Infect Dis 2017;30:511–7. https://doi.org/10.1097/QCO.0000000000000406.
[18] Su M, Satola SW, Read TD. Genome-based prediction of bacterial antibiotic resistance. J Clin Microbiol 2018. https://doi.org/10.1128/JCM.01405-18.
[19] Zazzi M, Cozzi-Lepri A, Prosperi MCF. Computer-aided optimization of combined anti-retroviral therapy for HIV: new drugs, new drug targets and drug resistance. Curr HIV Res 2016;14:101–9. https://doi.org/10.2174/1570162X13666151029102254.
[20] Roth JA, Battegay M, Juchler F, Vogt JE, Widmer AF. Introduction to machine learning in digital healthcare epidemiology. Infect Control Hosp Epidemiol 2018;39:1457–62. https://doi.org/10.1017/ice.2018.265.
[21] Sips ME, Bonten MJM, van Mourik MSM. Automated surveillance of healthcare-associated infections: state of the art. Curr Opin Infect Dis 2017;30:425–31. https://doi.org/10.1097/QCO.0000000000000376.
[22] Lee EY, Lee MW, Fulan BM, Ferguson AL, Wong GCL. What can machine learning do for antimicrobial peptides, and what can antimicrobial peptides do for machine learning? Interface Focus 2017;7:20160153. https://doi.org/10.1098/rsfs.2016.0153.
[23] Durrant JD, Amaro RE. Machine-learning techniques applied to antibacterial drug discovery. Chem Biol Drug Des 2015;85:14–21. https://doi.org/10.1111/cbdd.12423.

[24] Bowick GC, Barrett ADT. Comparative pathogenesis and systems biology for biodefense virus vaccine development. J Biomed Biotechnol 2010;2010:236528. https://doi.org/10.1155/2010/236528.
[25] Deo RC. Machine learning in medicine. Circulation 2015;132:1920–30. https://doi.org/10.1161/CIRCULATIONAHA.115.001593.
[26] Russell S, Norvig P. Artificial intelligence: a modern approach. Third. Pearson; 2016.
[27] Lamping F, Jack T, Rübsamen N, Sasse M, Beerbaum P, Mikolajczyk RT, et al. Development and validation of a diagnostic model for early differentiation of sepsis and non-infectious SIRS in critically ill children - a data-driven approach using machine-learning algorithms. BMC Pediatr 2018;18:112. https://doi.org/10.1186/s12887-018-1082-2.
[28] Oonsivilai M, Mo Y, Luangasanatip N, Lubell Y, Miliya T, Tan P, et al. Using machine learning to guide targeted and locally-tailored empiric antibiotic prescribing in a children's hospital in Cambodia. Wellcome Open Res 2018;3. https://doi.org/10.12688/wellcomeopenres.14847.1.
[29] Desautels T, Hoffman J, Barton C, Mao Q, Jay M, Calvert J, et al. Pediatric severe sepsis prediction using machine learning. bioRxiv 2017:223289. https://doi.org/10.1101/223289.
[30] Bartz-Kurycki MA, Green C, Anderson KT, Alder AC, Bucher BT, Cina RA, et al. Enhanced neonatal surgical site infection prediction model utilizing statistically and clinically significant variables in combination with a machine learning algorithm. Am J Surg 2018;216:764–77. https://doi.org/10.1016/j.amjsurg.2018.07.041.
[31] Hu Y, Lee VCS, Tan K. Prediction of clinicians' treatment in preterm infants with suspected late-onset sepsis - an ML approach. In: 2018 13th IEEE conference on industrial electronics and applications (ICIEA); 2018. p. 1177–82. https://doi.org/10.1109/ICIEA.2018.8397888.
[32] Mani S, Ozdas A, Aliferis C, Varol HA, Chen Q, Carnevale R, et al. Medical decision support using machine learning for early detection of late-onset neonatal sepsis. J Am Med Inform Assoc JAMIA 2014;21:326–36. https://doi.org/10.1136/amiajnl-2013-001854.
[33] Tsoukalas A, Albertson T, Tagkopoulos I. From data to optimal decision making: a data-driven, probabilistic machine learning approach to decision support for patients with sepsis. JMIR Med Inform 2015;3:e11. https://doi.org/10.2196/medinform.3445.
[34] Parbhoo S, Bogojeska J, Zazzi M, Roth V, Doshi-Velez F. Combining kernel and model based learning for HIV therapy selection. AMIA Summits Transl Sci Proc 2017;2017:239–48.
[35] Apostolova E, Velez T. Toward automated early sepsis alerting: identifying infection patients from nursing notes. ArXiv180903995 Cs. 2018. http://arxiv.org/abs/1809.03995. [Accessed 20 February 2019].
[36] Culliton P, Levinson M, Ehresman A, Wherry J, Steingrub JS, Gallant SI. Predicting severe sepsis using text from the electronic health record. ArXiv171111536 Cs. 2017. http://arxiv.org/abs/1711.11536. [Accessed 30 January 2019].
[37] Johnson AEW, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016;3:160035. https://doi.org/10.1038/sdata.2016.35.
[38] Calvert JS, Price DA, Chettipally UK, Barton CW, Feldman MD, Hoffman JL, et al. A computational approach to early sepsis detection. Comput Biol Med 2016;74:69–73. https://doi.org/10.1016/j.compbiomed.2016.05.003.
[39] McCoy A, Das R. Reducing patient mortality, length of stay and readmissions through machine learning-based sepsis prediction in the emergency department, intensive care unit and hospital floor units. BMJ Open Qual 2017;6:e000158. https://doi.org/10.1136/bmjoq-2017-000158.
[40] Burdick H, Pino E, Gabel-Comeau D, Gu C, Huang H, Lynn-Palevsky A, et al. Evaluating a sepsis prediction machine learning algorithm in the emergency department and intensive care unit: a before and after comparative study. bioRxiv 2018. https://doi.org/10.1101/224014.
[41] Burdick H, Pino E, Gabel-Comeau D, McCoy A, Gu C, Roberts J, et al. Effect of a sepsis prediction algorithm on patient mortality, length of stay, and readmission. bioRxiv 2018. https://doi.org/10.1101/457465.
[42] Bärnighausen T, Oldenburg C, Tugwell P, Bommer C, Ebert C, Barreto M, et al. Quasi-experimental study designs series-paper 7: assessing the assumptions. J Clin Epidemiol 2017;89:53–66. https://doi.org/10.1016/j.jclinepi.2017.02.017.
[43] Ho AMH, Phelan R, Mizubuti GB, Murdoch JAC, Wickett S, Ho AK, et al. Bias in before-after studies: narrative overview for anesthesiologists. Anesth Analg 2018;126:1755–62. https://doi.org/10.1213/ANE.0000000000002705.
[44] Shimabukuro DW, Barton CW, Feldman MD, Mataraso SJ, Das R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. BMJ Open Respir Res 2017;4:e000234. https://doi.org/10.1136/bmjresp-2017-000234.
[45] Saybani MR, Shamshirband S, Golzari Hormozi S, Wah TY, Aghabozorgi S, Pourhoseingholi MA, et al. Diagnosing tuberculosis with a novel support vector machine-based artificial immune recognition system. Iran Red Crescent Med J 2015;17. https://doi.org/10.5812/ircmj.17(4)2015.24557.
[46] El-Solh AA, Hsiao C-B, Goodnough S, Serghani J, Grant BJB. Predicting active pulmonary tuberculosis using an artificial neural network. Chest 1999;116:968–73. https://doi.org/10.1378/chest.116.4.968.
[47] Ansari AQ, Gupta NK, Ekata. Adaptive neurofuzzy system for tuberculosis. In: 2012 2nd IEEE international conference on parallel, distributed and grid computing; 2012. p. 568–73. https://doi.org/10.1109/PDGC.2012.6449883.
[48] Melendez J, Sánchez CI, Philipsen RHHM, Maduskar P, Dawson R, Theron G, et al. An automated tuberculosis screening strategy combining X-ray-based computer-aided detection and clinical information. Sci Rep 2016;6. https://doi.org/10.1038/srep25265.
[49] Er O, Temurtas F, Tanrıkulu AÇ. Tuberculosis disease diagnosis using artificial neural networks. J Med Syst 2010;34:299–302. https://doi.org/10.1007/s10916-008-9241-x.
[50] Santos AM, Pereira BB, Seixas JM, Mello FCQ, Kritski AL. Neural networks: an application for predicting smear negative pulmonary tuberculosis. In: Auget J-L, Balakrishnan N, Mesbah M, Molenberghs G, editors. Advances in statistical methods for the health sciences. Boston, MA: Birkhäuser Boston; 2007. p. 275–87. https://doi.org/10.1007/978-0-8176-4542-7_18.
[51] Revuelta-Zamorano P, Sánchez A, Rojo-Álvarez JL, Álvarez-Rodríguez J, Ramos-López J, Soguero-Ruiz C. Prediction of healthcare associated infections in an intensive care unit using machine learning and big data tools. In: Kyriacou E, Christofides S, Pattichis CS, editors. XIV mediterranean conference on medical and biological engineering and computing 2016, vol. 57. Cham: Springer International Publishing; 2016. p. 840–5. https://doi.org/10.1007/978-3-319-32703-7_163.
[52] Hartvigsen T, Sen C, Brownell S, Teeple E, Kong X, Rundensteiner E. Early prediction of MRSA infections using electronic health records. In: Proceedings of the 11th international joint conference on biomedical engineering systems and technologies; 2018. p. 156–67. https://doi.org/10.5220/0006599601560167.
[53] Hernandez B, Herrero P, Rawson TM, Moore LSP, Evans B, Toumazou C, et al. Supervised learning for infection risk inference using pathology data. BMC Med Inform Decis Mak 2017;17. https://doi.org/10.1186/s12911-017-0550-1.
[54] Van Steenkiste T, Ruyssinck J, De Baets L, Decruyenaere J, De Turck F, Ongenae F, et al. Accurate prediction of blood culture outcome in the intensive care unit using long short-term memory neural networks. Artif Intell Med 2018. https://doi.org/10.1016/j.artmed.2018.10.008.
[55] Sanger PC, van Ramshorst GH, Mercan E, Huang S, Hartzler AL, Armstrong CAL, et al. A prognostic model of surgical site infection using daily clinical wound assessment. J Am Coll Surg 2016;223:259–70. https://doi.org/10.1016/j.jamcollsurg.2016.04.046. e2.
[56] Kuo P-J, Wu S-C, Chien P-C, Chang S-S, Rau C-S, Tai H-L, et al. Artificial neural network approach to predict surgical site infection after free-flap reconstruction in patients receiving surgery for head and neck cancer. Oncotarget 2018;9:13768–82. https://doi.org/10.18632/oncotarget.24468.
[57] Ke C, Jin Y, Evans H, Lober B, Qian X, Liu J, et al. Prognostics of surgical site infections using dynamic health data. J Biomed Inform 2017;65:22–33. https://doi.org/10.1016/j.jbi.2016.10.021.
[58] Rawson TM, Hernandez B, Moore LSP, Blandy O, Herrero P, Gilchrist M, et al. Supervised machine learning for the prediction of infection on admission to hospital: a prospective observational cohort study. J Antimicrob Chemother 2019;74:1108–15. https://doi.org/10.1093/jac/dky514.
[59] Horng S, Sontag DA, Halpern Y, Jernite Y, Shapiro NI, Nathanson LA. Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning. Groza T, ed. PLoS One 2017;12:e0174708. https://doi.org/10.1371/journal.pone.0174708.
[60] Taylor RA, Moore CL, Cheung K-H, Brandt C. Predicting urinary tract infections in the emergency department with machine learning. PLoS One 2018;13:e0194085. https://doi.org/10.1371/journal.pone.0194085.
[61] Revett K, Gorunescu F, Ene M. A machine learning approach to differentiating bacterial from viral meningitis. IEEE; 2006. p. 155–62. https://doi.org/10.1109/JVA.2006.2.
[62] D'Angelo G, Pilla R, Tascini C, Rampone S. A proposal for distinguishing between bacterial and viral meningitis using genetic programming and decision trees. Soft Comput 2019. https://doi.org/10.1007/s00500-018-03729-y.
[63] Shamshirband S, Hessam S, Javidnia H, Amiribesheli M, Vahdat S, Petković D, et al. Tuberculosis disease diagnosis using artificial immune recognition system. Int J Med Sci 2014;11:508–14. https://doi.org/10.7150/ijms.8249.
[64] Hatzakis GE, Mathur M, Gilbert L, Panos G, Wanchu A, Patel AK, et al. Neural network-longitudinal assessment of the Electronic Anti-Retroviral THerapy (EARTH) cohort to follow response to HIV-treatment. AMIA Annu Symp Proc AMIA Symp 2005:301–5.
[65] Beerenwinkel N, Lengauer T, Däumer M, Kaiser R, Walter H, Korn K, et al. Methods for optimizing antiviral combination therapies. Bioinforma Oxf Engl 2003;19:i16–25.
[66] Zazzi M, Incardona F, Rosen-Zvi M, Prosperi M, Lengauer T, Altmann A, et al. Predicting response to antiretroviral treatment by machine learning: the EuResist project. Intervirology 2012;55:123–7. https://doi.org/10.1159/000332008.
[67] Revell AD, Wang D, Perez-Elias M-J, Wood R, Cogill D, Tempelman H, et al. 2018 update to the HIV-TRePS system: the development of new computational models to predict HIV treatment outcomes, with or without a genotype, with enhanced usability for low-income settings. J Antimicrob Chemother 2018;73:2186–96. https://doi.org/10.1093/jac/dky179.
[68] Petersen ML, LeDell E, Schwab J, Sarovar V, Gross R, Reynolds N, et al. Super learner analysis of electronic adherence data improves viral prediction and may provide strategies for selective HIV RNA monitoring. J Acquir Immune Defic Syndr 1999 2015;69:109–18. https://doi.org/10.1097/QAI.0000000000000548.

[69] Maiellaro PA, Cozzolongo R, Marino P. Artificial neural networks for the prediction of response to interferon plus ribavirin treatment in patients with chronic hepatitis C. Curr Pharm Des 2004;10:2101–9.

[70] Takayama T, Ebinuma H, Tada S, Yamagishi Y, Wakabayashi K, Ojiro K, et al. Prediction of effect of pegylated interferon alpha-2b plus ribavirin combination therapy in patients with chronic hepatitis C infection. PLoS One 2011;6:e27223. https://doi.org/10.1371/journal.pone.0027223.

[71] Sargolzaee Aval F, Behnaz N, Raoufy MR, Alavian SM. Predicting the outcomes of combination therapy in patients with chronic hepatitis C using artificial neural network. Hepat Mon 2014;14:e17028. https://doi.org/10.5812/hepatmon.17028.

[72] Hussain OA, Junejo KN. Predicting treatment outcome of drug-susceptible tuberculosis patients using machine-learning models. Inform Health Soc Care 2018:1–17. https://doi.org/10.1080/17538157.2018.1433676.

[73] Sauer CM, Sasson D, Paik KE, McCague N, Celi LA, Sánchez Fernández I, et al. Feature selection and prediction of treatment failure in tuberculosis. Huang Z, ed. PLoS One 2018;13:e0207491. https://doi.org/10.1371/journal.pone.0207491.

[74] Li BY, Oh J, Young VB, Rao K, Wiens J. Using machine learning and the electronic health record to predict complicated Clostridium difficile infection. Open Forum Infect Dis 2019;6:ofz186. https://doi.org/10.1093/ofid/ofz186.

[75] LaBarbera FD, Nikiforov I, Parvathenani A, Pramil V, Gorrepati S. A prediction model for Clostridium difficile recurrence. J Community Hosp Intern Med Perspect 2015;5:26033. https://doi.org/10.3402/jchimp.v5.26033.

[76] Escobar GJ, Baker JM, Kipnis P, Greene JD, Mast TC, Gupta SB, et al. Prediction of recurrent Clostridium difficile infection using comprehensive electronic medical records in an integrated healthcare delivery system. Infect Control Hosp Epidemiol 2017;38:1196–203. https://doi.org/10.1017/ice.2017.176.

[77] Li S, Tang B, He H. An imbalanced learning based MDR-TB early warning system. J Med Syst 2016;40. https://doi.org/10.1007/s10916-016-0517-2.

[78] Yelin I, Snitser O, Novich G, Katz R, Tal O, Parizade M, et al. Personal clinical history predicts antibiotic resistance of urinary tract infections. Nat Med 2019;25:1143–52. https://doi.org/10.1038/s41591-019-0503-6.

[79] Tacconelli E, Górska A, De Angelis G, Lammens C, Restuccia G, Schrenzel J, et al. Estimating the association between antibiotic exposure and colonization with extended-spectrum β-lactamase-producing Gram-negative bacteria using machine learning methods: a multicentre, prospective cohort study. Clin Microbiol Infect 2020;26:87–94.

[80] Beaudoin M, Kabanza F, Nault V, Valiquette L. Evaluation of a machine learning capability for a clinical decision support system to enhance antimicrobial stewardship programs. Artif Intell Med 2016;68:29–36. https://doi.org/10.1016/j.artmed.2016.02.001.

[81] Shen Y, Yuan K, Chen D, Colloc J, Yang M, Li Y, et al. An ontology-driven clinical decision support system (IDDAP) for infectious disease diagnosis and antibiotic prescription. Artif Intell Med 2018;86:20–32. https://doi.org/10.1016/j.artmed.2018.01.003.

[82] Ying Hao, Lin Feng, MacArthur Rodger D, Cohn Jonathan A, Barth-Jones Daniel C, Ye Hong, et al. A self-learning fuzzy discrete event system for HIV/AIDS treatment regimen selection. IEEE Trans Syst MAN Cybern B Cybern 2007;37. https://doi.org/10.1109/TSMCB.2007.895360.

[83] Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402. https://doi.org/10.1001/jama.2016.17216.

[84] Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med 2018;169:866. https://doi.org/10.7326/M18-1990.

[85] Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. JAMA Intern Med 2018;178:1544–7. https://doi.org/10.1001/jamainternmed.2018.3763.

[86] Osakunor DNM, Sengeh DM, Mutapi F. Coinfections and comorbidities in African health systems: at the interface of infectious and noninfectious diseases. PLoS Negl Trop Dis 2018;12. https://doi.org/10.1371/journal.pntd.0006711.

[87] Taneja I, Reddy B, Damhorst G, Dave Zhao S, Hassan U, Price Z, et al. Combining biomarkers with EMR data to identify patients in different phases of sepsis. Sci Rep 2017;7:10800. https://doi.org/10.1038/s41598-017-09766-1.

[88] Ford E, Carroll JA, Smith HE, Scott D, Cassell JA. Extracting information from the text of electronic medical records to improve case detection: a systematic review. J Am Med Inform Assoc 2016;23:1007–15. https://doi.org/10.1093/jamia/ocv180.

[89] Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017;19:e367. https://doi.org/10.2196/jmir.8775.

[90] Stead WW. Clinical implications and challenges of artificial intelligence and deep learning. JAMA 2018;320:1107–8. https://doi.org/10.1001/jama.2018.11029.

[91] Rawson TM, Charani E, Moore LSP, Hernandez B, Castro-Sánchez E, Herrero P, et al. Mapping the decision pathways of acute infection management in secondary care among UK medical physicians: a qualitative study. BMC Med 2016;14. https://doi.org/10.1186/s12916-016-0751-y0751-y.

[92] Maddox TM, Rumsfeld JS, Payne PRO. Questions for artificial intelligence in health care. JAMA 2018. https://doi.org/10.1001/jama.2018.18932.

[93] He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019;25:30–6. https://doi.org/10.1038/s41591-018-0307-0.

[94] Asan O, Smith P, Montague E. More screen time, less face time – implications for EHR design. J Eval Clin Pract 2014;20:896–901. https://doi.org/10.1111/jep.12182.

[95] Margalit RS, Roter D, Dunevant MA, Larson S, Reis S. Electronic medical record use and physician–patient communication: an observational study of Israeli primary care encounters. Patient Educ Couns 2006;61:134–41. https://doi.org/10.1016/j.pec.2005.03.004.

[96] Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An interpretable machine learning model for accurate prediction of sepsis in the ICU. Crit Care Med 2018;46:547–53. https://doi.org/10.1097/CCM.0000000000002936.

[97] Price WN, Cohen IG. Privacy in the age of medical big data. Nat Med 2019;25:37–43. https://doi.org/10.1038/s41591-018-0272-7.

[98] Calvert J, Hoffman J, Barton C, Shimabukuro D, Ries M, Chettipally U, et al. Cost and mortality impact of an algorithm-driven sepsis prediction system. J Med Econ 2017;20:646–51. https://doi.org/10.1080/13696998.2017.1307203.

[99] Sinha M, Jupe J, Mack H, Coleman TP, Lawrence SM, Fraley SI. Emerging technologies for molecular diagnosis of sepsis. Clin Microbiol Rev 2018;31. https://doi.org/10.1128/CMR.00089-17. e00089-17.

[100] Bisaso KR, Anguzu GT, Karungi SA, Kiragga A, Castelnuovo B. A survey of machine learning applications in HIV clinical research and care. Comput Biol Med 2017;91:366–71. https://doi.org/10.1016/j.compbiomed.2017.11.001.

[101] Zazzi M, Cozzi-Lepri A, Prosperi MCF. Computer-aided optimization of combined anti-retroviral therapy for HIV: new drugs, new drug targets and drug resistance. Curr HIV Res 2016;14:101–9.

[102] Al-Garadi MA, Khan MS, Varathan KD, Mujtaba G, Al-Kabsi AM. Using online social networks to track a pandemic: a systematic review. J Biomed Inform 2016;62:1–11. https://doi.org/10.1016/j.jbi.2016.05.005.

[103] Nourani E, Khunjush F, Durmuş S. Computational approaches for prediction of pathogen-host protein-protein interactions. Front Microbiol 2015;6:94. https://doi.org/10.3389/fmicb.2015.00094.

[104] Zhou H, Jin J, Wong L. Progress in computational studies of host-pathogen interactions. J Bioinform Comput Biol 2013;11:1230001. https://doi.org/10.1142/S0219720012300018.

[105] Segal G. Identification of legionella effectors using bioinformatic approaches. Methods Mol Biol Clifton NJ 2013;954:595–602. https://doi.org/10.1007/978-1-62703-161-5_37.

[106] Ackerman ME, Barouch DH, Alter G. Systems serology for evaluation of HIV vaccine trials. Immunol Rev 2017;275:262–70. https://doi.org/10.1111/imr.12503.

[107] Bağcı U, Bray M, Caban J, Yao J, Mollura DJ. Computer-assisted detection of infectious lung diseases: a review. Comput Med Imaging Graph Off J Comput Med Imaging Soc 2012;36:72–84. https://doi.org/10.1016/j.compmedimag.2011.06.002.

[108] Laketa V. Microscopy in infectious disease research - imaging across scales. J Mol Biol 2018;430:2612–25. https://doi.org/10.1016/j.jmb.2018.06.018.