Title

Evolving Applications of Artificial Intelligence and Machine Learning in Infectious Diseases Testing

Journal

Clinical Chemistry, 68(1)

ISSN

0009-9147

Authors

Tran, Nam K
Albahra, Samer
May, Larissa
et al.

Publication Date

2021-12-30

DOI

10.1093/clinchem/hvab239

Peer reviewed

Clinical Chemistry 68:1
125–133 (2022)

Review

Evolving Applications of Artificial Intelligence and Machine Learning in Infectious Diseases Testing

Nam K. Tran,[a,*] Samer Albahra,[a] Larissa May,[b] Sarah Waldman,[c] Scott Crabtree,[c] Scott Bainbridge,[a] and Hooman Rashidi[a,*]

BACKGROUND: Artificial intelligence (AI) and machine learning (ML) are poised to transform infectious disease testing. Uniquely, infectious disease testing is technologically diverse spaces in laboratory medicine, where multiple platforms and approaches may be required to support clinical decision-making. Despite advances in laboratory informatics, the vast array of infectious disease data is constrained by human analytical limitations. Machine learning can exploit multiple data streams, including but not limited to laboratory information and overcome human limitations to provide physicians with predictive and actionable results. As a quickly evolving area of computer science, laboratory professionals should become aware of AI/ML applications for infectious disease testing as more platforms are become commercially available.

CONTENT: In this review we: (*a*) define both AI/ML, (*b*) provide an overview of common ML approaches used in laboratory medicine, (*c*) describe the current AI/ML landscape as it relates infectious disease testing, and (*d*) discuss the future evolution AI/ML for infectious disease testing in both laboratory and point-of-care applications.

SUMMARY: The review provides an important educational overview of AI/ML technique in the context of infectious disease testing. This includes supervised ML approaches, which are frequently used in laboratory medicine applications including infectious diseases, such as COVID-19, sepsis, hepatitis, malaria, meningitis, Lyme disease, and tuberculosis. We also apply the concept of "data fusion" describing the future of laboratory testing where multiple data streams are integrated by AI/ML to provide actionable clinical knowledge.

[a]Department of Pathology and Laboratory Medicine, UC Davis School of Medicine, CA; [b]Department of Emergency Medicine, UC Davis School of Medicine, CA; [c]Department of Internal Medicine, Division of Infectious Diseases, UC Davis School of Medicine, CA.
*Address correspondence to: N.K.T. at Department of Pathology and Laboratory Medicine, University of California Davis, 4400 V St. Sacramento, CA 95817, USA. E-mail nktran@ucdavis.edu. H.H.R. at Department of Pathology and Laboratory Medicine, University of California Davis, 4400 V St., Sacramento, CA 95817, USA. Fax 916-734-0431; e-mail hrashidi@ucdavis.edu.
Received June 1, 2021; accepted October 15, 2021.
https://doi.org/10.1093/clinchem/hvab239

Introduction

Infectious disease testing is one of the most technologically diverse spaces in laboratory medicine. Beginning with Koch's postulates published in 1890, infectious disease testing has evolved from simple microscopy and microbiological culture to modern techniques ranging from immunoassays (e.g., direct pathogen detection, serology) and MALDI–TOF–MS, to molecular diagnostics (Fig. 1) (1–3). In some cases, several tests may be used in combination to produce definitive results (e.g., blood culture → MALDI–TOF–MS → antimicrobial susceptibility testing) (4).

In the 21st century, infectious disease testing has heavily leveraged information technology (5). Both laboratory and point-of-care (POC) platforms quickly transmit results to electronic medical record (EMR) systems to integrating other data streams (e.g., vitals monitors, imaging systems, home testing devices) to help physicians determine the best course of action. By having all medical data stored in the EMR, it was hoped clinical decision-making would become more efficient (6). Unfortunately, this digitization of medical data has created a state of "information overload" for healthcare professionals (7) — causing data to be unintentionally ignored, misinterpreted, and/or potentially masking clinically significant patterns.

On average, humans can comfortably interpret and integrate up to 4 variables at a time (8) — limiting the value of EMR data. For example, the number of variables used for sepsis prediction, can range from a few simple parameters (e.g., heart rate, white blood cell [WBC] count, respiratory rate [RR], and body temperature) to hundreds of time-series measurements which could quickly overwhelm an individual (9–12). Computers do not share these limitations since they can handle a wider range of variables simultaneously and recognize patterns that are not apparent to the human eye. Thus, the use of predictive analytics via artificial intelligence (AI)/machine learning (ML) could

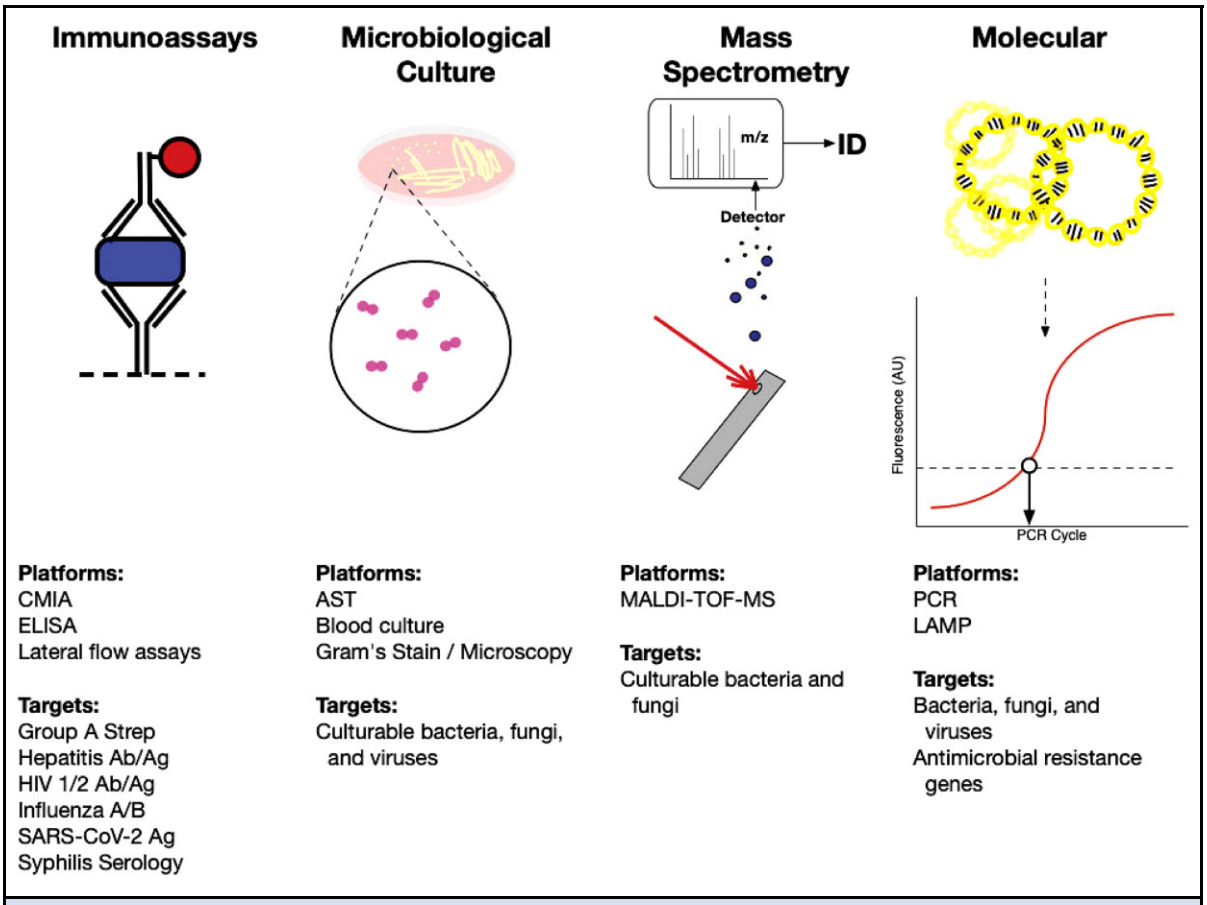

Fig. 1. Infectious disease testing domains. Four common techniques used for infectious disease testing. Examples of testing platforms, and common infectious targets are identified below the visual depiction of each testing method. Abbreviations: AST, antimicrobial susceptibility testing; AU, arbitrary units; CMIA, chemiluminescent microparticle immunoassay; ELISA, enzyme-linked immunosorbent assay; HIV, human immunodeficiency virus; ID, identification; LAMP, loop-mediated amplification; m/z, mass/charge ratio; SARS-CoV-2, severe acute respiratory syndrome-coronavirus-2.

enhance our ability to identify clinically significant patterns, including those for infectious diseases (13).

Artificial Intelligence and Machine Learning

Artificial intelligence is a rapidly evolving field of computer science and statistics aimed at producing systems that mimic human behavior (14, 15). In contrast, ML is an application of AI that enables systems to automatically learn from new data without explicit programming. Machine learning algorithms ultimately enable one to improve the predictive analytic performance for a given task and/or acquire new skills over time when trained with more data. Classically, ML is grouped into 3 major categories discussed in this review (14–24): (*a*) supervised learning, including classification and regression approaches, (*b*) unsupervised learning, and (*c*) reinforcement learning (Fig. 2). Here we provide a general overview of ML techniques with emphasis on applications for infectious disease testing. For more detailed discussion on AI/ML theory and techniques, we refer the reader to other publications (16–24).

MACHINE LEARNING TECHNIQUES

Laboratory medicine data often favor the use of supervised learning for ML applications largely due to the large amount of information and types of data (e.g., image versus numerical values versus text) available in this discipline (11, 14, 16, 25, 26). It must be noted, however, that as more studies use AI/ML, and computing power and portability increases, these predictive analytics will quickly evolve over time.

The type of data also influences which ML methods are used (14). Example of ML methods include

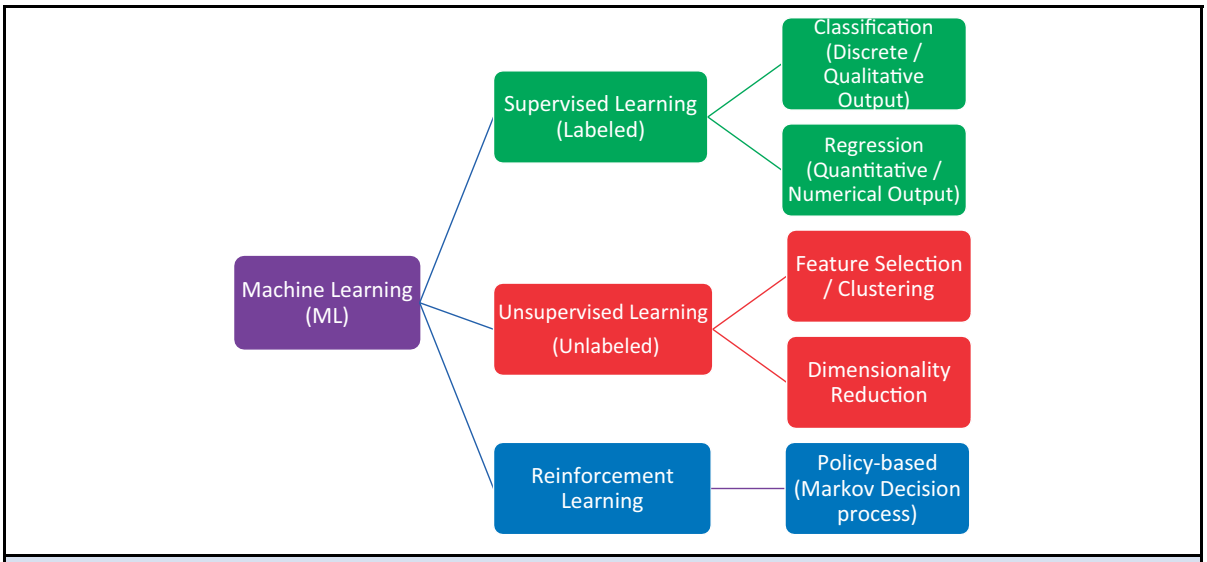

Fig. 2. Machine learning categories. Machine learning is subdivided into supervised, unsupervised, and reinforcement learning categories. These categories are further divided into classification, regression, feature selection/clustering, dimensionality reduction, and policy-based methods.

parametric techniques such as logistic regression, while nonparametric methods include neural network and non-neural network approaches such as linear regression, logistic regression, naïve Bayes, gradient boosting machine/decision tree, *k*-nearest neighbor (*k*-NN), support vector machine (SVM), and random forest (RF) (Fig. 3).

Machine Development Tools

Traditional AI/ML development requires substantial expertise and time. Programming languages used for various AI/ML tools include Python and R (1, 11, 14, 25, 26). Often, developers may take months to train and test a single supervised algorithm. This process could be further prolonged if there is a need to determine the best feature combinations to include or exclude in their models. Additionally, given the rapidly moving AI/ML landscape, not all data scientists may be comfortable with all methods. For example, some data scientists may prefer one technique versus another (e.g., SVM vs neural networks)—potentially creating bias and/or limiting the potential for discovery of better performing ML algorithms. Alternately, a priori determination of feature set combinations impacts ML model performance.

Due to these potential limitations, automated ML platforms have gained popularity (11, 25, 26). These "Auto-ML" platforms allow users to import train and test data, which are then automatically run through the whole process of feature selection, model building, and validation. An AI/ML programming task that could take a single data scientist months now takes a few hours. Examples of automated ML platforms include H2OAI AutoML, Microsoft ML.NET, Tree-based Pipeline Optimization Tool, Machine Intelligence Learning Optimizer, and AutoSKLearn (11, 17, 25).

Machine Learning Applications for Infectious Disease Testing

AI/ML is poised to transform not only bedside medicine, but also the field of infectious disease testing. In particular, infectious disease testing has seen a surge in miniaturization, automation, and increasing computing power, creating a unique opportunity to exploit AI/ML. We can categorize the applications into 3 categories: (*a*) laboratory diagnostics, (*b*) clinical prognosis, and (*c*) clinical diagnostics. Application examples are discussed next.

LABORATORY DIAGNOSTICS

Novel 2019 coronavirus infectious disease (COVID-19). The COVID-19 pandemic resulted in a substantial demand for molecular testing in the USA (27, 28). Early in the pandemic, molecular capacity was not sustainable and impacted the US response in controlling the spread of severe acute respiratory syndrome (SARS)-coronavirus (CoV)-2 infections. Asymptomatic spread of COVID-19 further increased demand for testing (29).

Alternative testing approaches using MALDI–TOF–MS have been proposed as a low-cost, rapid, and

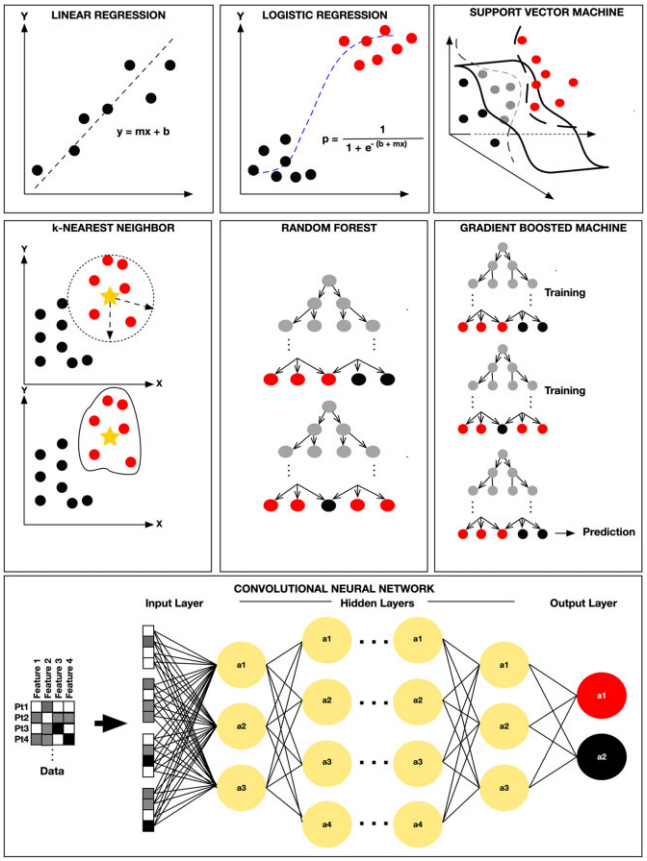

Fig. 3. Machine learning techniques. Machine learning methods discussed in this review. Data points are shown in different colors to describe their classification (e.g., red, with disease; black, without disease; gray, classification not defined). In the last row, the convolutional neural network example is analyzing a matrix of data.

high-throughput solution to alleviate demand on molecular testing (26, 30). Briefly, MALDI–TOF–MS detects ionizable proteins from respiratory samples such as anterior nares swabs, producing hundreds or even thousands of MS spectra peaks per samples. The use of AI/ML provides means to identify MALDI–TOF–MS profiles specific to COVID-19. Recent studies have used a neural network approach to analyze MALDI–TOF–MS spectra that achieved a sensitivity and specificity of 100% and 96%, respectively, with an area under the receiver operating characteristic (ROC) curve of 0.99 when using 487 peaks that span 1993.91 to 199 590.89 *m/z* (26). It must be also noted that these published MALDI–TOF–MS methods currently detect the host response to COVID-19, rather than the virus itself.

Antimicrobial susceptibility testing (AST). Antimicrobial resistance is one of the top 10 global public health threats defined by the World Health Organization (31). The inappropriate use of antimicrobial therapy fuels the persistence of existing drug resistant organisms, and the development of multidrug resistant strains. Currently, detection of resistant organisms relies on microdilution in vitro AST or molecular approaches (32). Microdilution AST involves exposing a known concentration of cultured microbes to antimicrobials to determine a minimum inhibitory concentration. Molecular approaches target specific genes to rapidly determine genotypic resistance rather than determining in vitro susceptibility to antimicrobials.

The application of AI/ML for AST has been reported for MALDI–TOF–MS techniques (32). For MALDI–TOF–MS, ML (e.g., neural networks, SVM, RF, *k*-NN) is used to analyze spectra representing ionizable proteins (mostly ribosomal proteins) from cultured bacteria. In one study, SVM was used to differentiate methicillin resistant *Staphylococcus aureus* versus methicillin sensitive *S. aureus* with an accuracy of 85% (33).

In another study, 5 different ML algorithms (*k*-NN, RF, SVM, naïve Bayes, and logistic regression) were used to identify carbapenem resistant versus sensitive *Klebsiella pneumonia* MALDI–TOF–MS spectra, achieving an accuracy of over 93% (34).

Malaria. Detection of *Plasmodium* species remains challenging due to the unique nature of the parasite (35). Microscopic examination with thin and thick smears requires experienced personnel. Rapid antigen testing is available but may not detect or easily differentiate between *Plasmodium* subspecies. Machine learning techniques have been proposed to microscopically analyze blood smears, including at the point of care via smartphone-based applications. In a study using neural network, SVM, and *k*-NN approaches, the investigators were able to produce models that achieved a sensitivity and specificity as high as 99.5% and 99.1% with an accuracy of 99.2% using a smartphone based slide imaging software (35).

Lyme disease. Lyme disease is the most common vector-borne infectious disease in North America and Europe (36). Timely diagnosis is necessary to prevent progression, especially if the early disease presentation is missed. Current Lyme disease tests, however, exhibit poor sensitivity (<50%) for early presenters. One group has proposed a multiplex POC sero-diagnostic test targeting antigens OspC, BmpA, P41, ErpD, Crasp1, OspA, DbpB, VlsE, P35, and Mod-C6 enhanced by ML to improve clinical performance, with a reported sensitivity of 90.5% and specificity of 87.0% (37).

CLINICAL PROGNOSIS

Hepatitis B virus infection. Hepatitis B virus has resulted in over 250 million chronically infected individuals worldwide (38). At present, hepatitis B surface antigen and hepatitis B core antigen are the primary biomarkers for predicting virological relapse. Unfortunately, these methods still exhibit poor performance. The use of ML to predict early virological relapse following discontinuation of therapy has shown promise. A supervised ML approach has been used to predict early virological relapse using soluble immune markers profiles (39). Optimal ML models using a combination of interleukin-2, monocyte induced interferon gamma, regulation on activation normal T-cell expressed and secreted, stem cell factor, and tumor necrosis factor related apoptosis-inducing ligand produced the highest predictive values for early virological relapse following treatment cessation.

Hepatitis C virus (HCV) infection. Serology and molecular approaches remain the gold standard for HCV screening. Detection of strains resistant to direct-acting antivirals, however, as well as predicting chronic HCV infection progression, has proven challenging (40). In one study, SVM was used to analyze HCV genomes to determine which patients could produce sustained viral response, producing an accuracy 95.4% (40). Another study employed longitudinal ML models incorporating laboratory (i.e., albumin, platelet, aspartate aminotransferase, alanine transaminase, alkaline phosphatase, alpha fetoprotein), and existing scoring (i.e., MELD, Ishak) data to predict chronic HCV infection progression (41). The area under the ROC curve for optimal RF models predicting fibrosis progression was 0.79, while models predicting clinical progression achieved an area under the ROC curve of 0.86.

CLINICAL DIAGNOSTICS

Meningitis. Meningitis results in 36 000 hospitalizations annually in the USA, costing between $234 and $310 million per year (42). Detection and differentiation of viral versus bacterial meningitis has been augmented by molecular platforms. Unfortunately, molecular tests still rely on invasive collection of cerebrospinal fluid (CSF). More traditional tests including CSF Gram stain and measurement of CSF WBC, glucose, and protein concentrations can provide rapid results, but are not always sensitive or specific. Several investigators have studied the use of AI/ML to enhance detection of meningitis (43, 44). One of these studies described a neural network-based approach based on 6 features including lymphocyte count, blood glucose, and age that produced an accuracy of 86.3% (44). The second study reported improved performance using naïve Bayes, neural network, and genetic programming techniques with age, race, sex, WBC, blood glucose, CSF glucose/protein/leukocytes (if available) as features (43). Naïve Bayes and neural network techniques exhibited sensitivity/specificity of 98%/98% and 99%/100%, respectively. The AI/ML approach produced an algorithm that achieved a sensitivity of 100% and 99%.

Sepsis. Sepsis is defined as life-threatening organ dysfunction caused by a dysregulated host response to infection (45). Early recognition of sepsis is critical to survival (46). Unfortunately, parameters for recognizing sepsis are not always sensitive nor specific. The systemic inflammatory response syndrome criteria introduced in 1992 is a major foundation of both past and present sepsis criteria (47). Although updated in more recent times (48, 49), the systemic inflammatory response syndrome criteria remain the same and focus on abnormal body temperature, RR, HR, and WBC count: parameters that are hardly specific for infection. Subsequent sepsis criteria have attempted to improve specificity by including other elements such as indicators of organ

dysfunction, but resulted in far too many variables to account for by a single human.

Electronic Health Record sepsis ML algorithms: Since EMR systems serve as the central repository of data, the use of AI/ML and statistical algorithms have been used to predict sepsis (49, 50). Sensitivity and specificity are reported to be 87% for statistical (non-AI/ML) algorithms with an area under the ROC curve of 0.94 when using 15 features (i.e., age, gender, blood pressure, HR, temperature, oxygen saturation, RR, WBC count, microbiological culture results, lactate, high sensitivity C-reactive protein, procalcitonin, arterial blood gas, use of vasopressors, and use of antibiotics). Application of AI/ML in sepsis has improved prediction performance (51, 52) including its use in more challenging populations such as burn patients (53, 54). Using heart rate, body temperature, hemoglobin, blood urea nitrogen, and total CO_2 as features with *k*-NN, sensitivity, and specificity was observed to be 95.8% and 87.8%, respectively, with an area under the ROC curve of 0.96 for predicting burn sepsis (53).

Sepsis host-response ML approaches: Expanding from traditional indicators of sepsis, other investigators have used ML to predict sepsis using a multi-RNA host-response approach. One study described a 29-host-mRNA based blood test that combined ML with a less than 30-min turnaround time POC test (55). This technique used a neural network based that achieved an area under the ROC curve of 0.92 and 0.91 for identifying individuals enrolled within 36 h of admission with bacterial or viral infections, respectively (56).

Tuberculosis. Tuberculosis remains a global healthcare problem. The use of AI/ML techniques has been studied for over a decade to aid in the diagnosis of tuberculosis, including the use of neural networks and SVM (57, 58). Support vector machines incorporating CD4 counts, human immunodeficiency virus status, purified protein derivative status, chest pain, weight loss, coughs, night sweats, fever, shortness of breath, total hemoglobin, platelet count, WBC count, neutrophil count, lymphocyte count, erythrocyte sedimentation rate, alanine aminotransferase, alkaline phosphatase, and lactate dehydrogenase concentration as features have exhibited 100% sensitivity and specificity (58). The investigators leading these studies, however, concede that future investigations are needed to determine the generalizability of this technique given the sample size.

Machine Learning Best Practices

Machine learning algorithm performance is dependent on the quality of data used for training and testing. Challenges include the ability for laboratories and/or clinical services to generate, collect, standardize, and quality control data that can be used in AI/ML. Overcoming these challenges will pave the way for optimal training and testing of candidate ML algorithms with datasets to increase overall performance. Developers and investigators also should be not tied to a single ML method or even particular features since this can bias analyses. A stepwise approach to develop ML for laboratory medicine applications has been discussed (16). Step 1 involves assessing the quality and accessibility of the data, followed by Step 2, which requires method validation to identify optimal AI/ML model(s). Once optimal AI/ML models have been identified, Step 3 involves determining their ability to work for other secondary and tertiary datasets (generalizability). Finally, Step 4 involves evaluating the data in more "real world" conditions to further assess performance and refine (go back to Step 2) as needed to achieve a desirable outcome.

Future Applications

The future of infectious disease diagnostics will involve the fusion of multiple data streams (i.e., vital signs, host-response biomarkers, traditional laboratory tests, etc.) with AI/ML to produce high quality clinically actionable results. This concept of "data fusion" is not new and is already prevalent in normal everyday life. Data fusion (in contrast to data integration) is the process of gathering data from multiple sources to produce more sophisticated models and better understand a disease or problem (59, 60). For example, data fusion of data from geofenced and connected smart devices (e.g., watches, tablets, cellular phones, etc.) provide a means to enhance the user experience while helping them to predict customer needs (60). Semiautonomous and autonomous vehicles also leverage this concept (61). In the context of infectious disease testing, AI/ML could enhance clinician and patient experience through integrating in vitro diagnostic testing with information from the EMR, in the form of data fusion, to facilitate predictive analytics (Fig. 4).

POC infectious disease testing represents another potential area that could be transformed by AI/ML By design, POC devices reside in a decentralized ecosystem—necessitating operation in a network-centric manner to gather a substantial amount of medical information (62). Wearable POC devices (e.g., smart watches), in particular, can now monitor oxygen saturation, and provide one-lead electrocardiogram and heart rate. Diabetics are now able to directly record discrete or continuous blood glucose concentrations, body weight, and blood pressure into their smart phones (63), stored in a centralized application (e.g., Apple HealthKit,

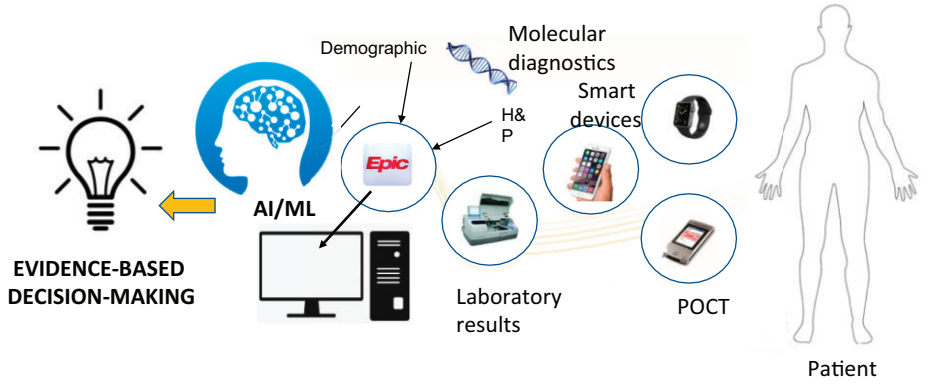

Fig. 4. Diagnostic sensor and data fusion concept. Conceptual model of "diagnostic sensor and data fusion." Patients using wearable smart devices and point-of-care testing feed data toward the electronic medical record. This is augmented by laboratory results, patient history/physical examination (H&P) data, as well as outside molecular results and demographics. The electronic medical record data is analyzed by artificial intelligence/machine learning (AI/ML) to facilitate evidence-based clinical decision-making.

Google Fit), and upload their data to compatible EMR, producing a valuable collection of persistent patient data that could be analyzed by AI/ML to predict disease. Such an example has been employed through the recent FDA emergency use authorization of the COVID Plus Monitor (Tiger Tech Solutions) that identifies patients who may exhibit signs of SARS-CoV-2 infection (64).

Conclusion

Artificial intelligence and ML will transform infectious disease testing. The use of AI/ML for sepsis management highlights the power of this technology for mining EMR data and prompting clinical action with better sensitivity and specificity compared to traditional approaches. Machine learning shows promise for a range of infectious disease applications such as COVID-19, hepatitis, malaria, Lyme disease, and tuberculosis. However, AI/ML should be implemented in a systematic and rational way to ensure data quality is not compromised, and model development is performed with minimal bias. The use of automated ML platforms is also an exciting development where thousands of candidate models could be automatically trained and tested across a range of feature combinations and allowing developers to quickly identify optimal algorithms for further development. Ultimately, the future of AI/ML infectious disease testing may revolve around concepts of "data fusion" to not only integrate multiple data streams, but also convert vast amounts of heterogenous data into actionable knowledge.

Nonstandard Abbreviations: AI, artificial intelligence; ML, machine learning; POC, point-of-care; EMR, electronic medical record; WBC, white blood cell; RR, respiratory rate; *k*-NN, *k*-nearest neighbor; SVM, support vector machine; RF, random forest; COVID-19, novel 2019 coronavirus infectious disease; SARS-CoV-2, severe acute respiratory syndrome-coronavirus-2; ROC, receiver operating characteristic; AST, antimicrobial susceptibility testing; HCV, hepatitis C virus; CSF, cerebrospinal fluid.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.*

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:*

Employment or Leadership: N.K. Tran, S. Albahra, and H. Rashidi are co-owners of MILO-ML, LLC.
Consultant or Advisory Role: Nam Tran is a consultant for Roche Diagnostics/Roche Molecular Systems. UC Davis Health is a Roche Diagnostics Molecular Center of Excellence.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.
Patents: N.K Tran, S. Albahra, and H. Rashidi are co-inventors of the Machine Intelligence Learning Optimizer platform (patent pending). N.K. Tran, H. Rashidi, and L. May are investigators developing the MALDI–TOF–MS COVID-19 testing platform.
Other Remuneration: S. Albahra, Machine Intelligence Learning Optimizer; L. May, Roche Diagnostics, Inflammatix; H. Rashidi, Machine Intelligence Learning Optimizer; N.K. Tran, Machine Intelligence Learning Optimizer, Roche Diagnostics, Roche Molecular Systems.

References

1. Falkow S. Molecular Koch's postulates applied to bacterial pathogenicity – a personal recollection 15 years later. Nat Rev Micro 2004;2:67–72.
2. O'Farrell B. Evolution in lateral flow-based immunoassay systems. Lateral Flow Immunoassay 2008;31:1–31.
3. Hansen G, Marino J, Wang ZX, Beavis KG, Rodrigo J, Labog K, et al. Clinical performance of the point-of-care cobas Liat for detection of SARS-CoV-2 in 20 minutes: a multicenter study. J Clin Microbiol 2021;59: e02811–20.
4. Fournier PE, Drancourt M, Colson P, Rolain JM, La Scola B, Raoult D. Modern clinical microbiology: new challenges and solutions. Nat Rev Microbiol 2013;11: 574–85.
5. Hota B, Jones RC, Schwartz DN. Informatics and infectious diseases: what is the connection and efficacy of information technology tools for therapy and health care epidemiology. Am J Infect Control 2007;36:S47–S56.
6. Petrides AK, Bixho I, Goonan EM, Bates DW, Shaykevich S, Lipsitz SR, et al. The benefits and challenges of an interfaced electronic health record and laboratory information system: effects on laboratory processes. Arch Pathol Lab Med 2017;141:410–7.
7. Furlow B. Information overload and unsustainable workloads in the era of electronic health records. Lancet Respir Med 2020;8:243–4.
8. Halford GS, Baker R, McCredden JE, Bain JD. How many variables can human process. Psychol Sci 2005;16:70–7.
9. Levy MM, Artigas A, Phillips GS, Rhodes A, Beale R, Osborn T, et al. Outcomes of the surviving sepsis campaign in intensive care units in the USA and Europe: a prospective cohort study. Lancet Infect Dis 2012;12: 919–24.
10. Sakr Y, Jaschinski U, Wittebole X, Szakmany T, Lipman J, Ñamendys-Silva SA, et al. Sepsis in intensive care unit patients: worldwide data from the intensive care over nations audit. Open Forum Infect Dis 2018;19:313.
11. Tran NK, Albahra S, Pham TN, Holmes JH 4th, Greenhalgh D, Palmieri TL, et al. Novel application of an automated-machine learning development tool for predicting burn sepsis: proof of concept. Sci Rep 2020;10: 12354.
12. Fleuren LM, Klausch TL, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. Intensive Care Med 2020;46: 383–400.
13. Gauglitz G. Artificial vs. human intelligence in analytics. Anal Bioanal Chem 2019;411:5631–2.
14. Thair S, Mewes C, Hinz J, Bergmann I, Büttner B, Sehmisch S, et al. Gene expression-based diagnosis of infections in critically ill patients-prospective validation of the SepsisMetaScore in a longitudinal severe trauma cohort. Crit Care Med 2021;49:e751–60.
15. Tsalik EL, Henao R, Montgomery JL, Nawrocki JW, Aydin M, Lydon EC, et al.; Antibacterial Resistance Leadership Group. Discriminating bacterial and viral infection using a rapid host gene expression test. Crit Care Med 2021; 49:1651–3.
16. Rashidi HH, Tran NK, Vali Betts EV, Howell LP, Green R. Artificial intelligence and machine learning in pathology: the present landscape of supervised methods. Acad Pathol 2019;6:2374289519873088.
17. Rashidi HH, Tran NK, Albahra S, Dang L. Machine learning in healthcare and laboratory medicine: general overview of supervised learning and auto-ML. Int J Lab Hematol 2021;43 Suppl 1:15–22.
18. Caruana R, Nicelescu-Mizil A. An empirical comparison of supervised learning algorithms. In: Proceedings of the 23rd International Conference on Machine Learning. 1-59593-383-2, Pittsburgh, PA, 2006. pp. 161–68. https://dl.acm.org/doi/abs/10.1145/1143844.1143865.
19. Culhane A, Buehler LK, Rashidi HH. Bioinformatics basics: Applications in biological science and medicine. BioMed Eng OnLine 2006;5:41. https: //doi.org/10.1186/1475-925X-5-41.
20. Aloise D, Deshpande A, Hansen P, Popat P. NP-hardness of Euclidean sum-of-squares clustering. Mach Learn 2009;75:245–8.
21. Cordeiro A, Boris MR. Minkowski metric, feature weighting and anomalous cluster initialization in k-means clustering. Pattern Recogn 2012;45:1061–75.
22. Celebi ME, Kingravi HA, Vela PA. A comparative study of efficient initialization methods for the k-means clustering algorithm. Expert Syst Appl 2013;40:200–10.
23. Greenmeier L. 20 Years after Deep Blue: how AI has advanced since conquering chess. https://www.scientificamerican.com/article/20-years-after-deep-blue-how-ai-has-advanced-since-conquering-chess/. *Scientific American* 2021 (Accessed May 31, 2021).
24. Murphy M. Google's AI just cracked the game that supposedly no computer could beat. *Quartz* 2016. https:// qz.com/603313/googles-ai-just-cracked-the-game-that-supposedly-no-computer-could-beat/ (Accessed on May 31, 2021).
25. Tran NK, Sen S, Palmieri TL, Lima K, Falwell S, Wajda J, Rashidi HH. Artificial intelligence and machine learning for predicting acute kidney injury in severely burned patients: a proof of concept. Burns 2019;18:S305–311.
26. Tran NK, Howard T, Walsh R, Pepper J, Loegering J, Phinney B, et al. Novel application of automated machine learning with MALDI-TOF-MS for rapid high-throughput screening of COVID-19: a proof of concept. Sci Rep 2021;11:8219.
27. Medical Laboratory Observer (MLO). Data collection tool shows shortages of COVID-19 and other testing supplies. https://www.mlo-online.com/disease/infectious-disease/ article/21159031/data-collection-tool-shows-shortages-of-covid19-and-other-testing-supplies (Accessed May 20, 2021).
28. Vandendberg O, Martiny D, Rochas O, van Belkum A, Kozlakidis Z. Considerations for diagnostic COVID-19 tests. Nat Rev Microbiol 2020;14:1–13.
29. Johansson MA, Quandelacy TM, Kada S. SARS-CoV-2 transmission from people without COVID-19 symptoms. Jama Netw Open 2021;4:e2035057.
30. Nachtigall FM, Pereira A, Trofymchuk OS, Santos LS. Detection of SARS-CoV-2 in nasal swabs using MALDI-MS. Nat Biotechnol 2020;38:1168–73.
31. World Health Organization (WHO). Key facts: Antimicrobial resistance. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance (Accessed August 5, 2021).
32. Weis CV, Jutzeler CR, Borgwardt K. Machine learning for microbial identification and antimicrobial susceptibility testing on MALDI-TOF-mass spectra: a systematic review. Clin Microbiol Infect 2020;26:1310–7.
33. Sogowa K, Watanabe M, Ishige T, Segawa S, Miyabe A, Murata S, et al. Rapid Staphylococcus aureus discrimination between methicillin-sensitive and methicillin-resistant using MALDI-TOF mass spectrometry. Biocontrol Sci 2017;22:163–9.
34. Huang TS, Lee SS, Lee CC, Chang FC. Detection of carbapenem-resistant Klebsiella pneumoniae on the basis of matrix-assisted laser desorption ionization time-of-flight mass spectrometry by using supervised machine learning approach. PLoS ONE 2020;15:e0228459.
35. Brozan RN, McMorrow ML, Kachur SP. Diagnosis of malaria: challenges for clinicians in endemic and non-endemic regions. Mol Diagn Ther 2008;12:299–306.
36. Mead PS. Epidemiology of Lyme disease. Infect Dis Clin North Am 2015;29:187–210.
37. Joung HA, Ballard ZS, Wu J, Tseng DK, Teshome H, Zhang L, et al. Point-of-care serodiagnostic test for early-stage Lyme disease using a paper-based immunoassay and machine learning. ACS Nano 2020;229–40.
38. Testoni B, Lebossé F, Scholtes C, Berby F, Miaglia C, Subic M, et al. Serum hepatitis B core-related antigen (HBcrAg) correlates with covalently closed circular DNA transcriptional activity in chronic hepatitis B patients. J Hepatol 2019;70:615–25.
39. Wubbolding M, Alfonso JC, Lin CY, Binder S, Falk C, Debarry J, et al. Pilot study using machine learning to identify immune profiles for the prediction of early virological relapse after stopping nucleos(t)ide analogues in HbeAg negative CHB. Hepatol Commun 2020;5:97–11.
40. Haga H, Sata H, Koseki A, Saito T, Okumoto K, Hoshikawa K, et al. A machine learning-based treatment prediction model using whole genome variants of hepatitis C virus. PLoS ONE 2020;15:e0242028.
41. Konerman MA, Zhang Y, Zhu J, Higgins PDR, Lok ASF, Waljee AK. Improvement of predictive models of risk of disease progression in chronic hepatitis C by incorporating longitudinal data. Hapatology 2015;61:1832–41.
42. Khetsuriani N, Quiroz ES, Holman RC, Anderson LJ. Viral meningitis-associated hospitalizations in the United States, 1988-1999. Neuroepidemiology 2003;22: 345–52.
43. 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;23:11775–91.
44. Revett K, Gorunescu F, Ene M. A machine learning approach to differentiating bacterial from viral meningitis. In: John Vincent Atanasoff, editor. Proceedings of the IEEE, International Symposium on Modern Computing (JVA'06). Los Alamitos: IEEE, 2006. p. 155–162. https:// westminsterresearch.westminster.ac.uk/item/92718/a-machine-learning-approach-to-differentiating-bacterial-from-viral-meningitis.
45. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA 2016;315:801–10.
46. Husabo G, Nilsen RM, Flaatten H, Solligard E, Frich JC, Bondevik GT, et al. Early diagnosis of sepsis in emergency departments, time to treatment, and association with mortality: an observational study. PLoS ONE 2020; 15:e0227652.
47. Bone RC, Balk RA, Cerra FB, Dellinger RP, Fein AM, Knaus WA, et al. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. Chest 1992;17:1644–55.

48. Dellinger RP, Levy MM, Carlet JM, Bion J, Parker MM, Jaeschke R, et al. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock. Crit Care Med 2008;2008:17:296–327.

49. Downing NL, Rolnick J, Poole SF, Hall E, Wessels AJ, Heidenreich P, Shieh L. Electronic health record-based clinical decision support alert for severe sepsis: a randomized evaluation. BMJ Qual Saf 2019;28:762–8.

50. 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: a prospective multicenter clinical outcomes evaluation of real-world patient data from US hospitals. BMJ Health Care Inform 2020;27:e100109.

51. Shimabukuro DW, Barton CW, Feldman MD, Mataro SJ, Das R. Effect of machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomized clinical trial. BMJ Open Respir Res 2017;4:e000234.

52. Goh KM, Wang L, Yeow AY, Poh H, Li K, Yeow JL, Tan GYH. Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. Nat Commun 2021;12:711.

53. Greenhalgh DG, Saffle JR, Holmes JH, Gamelli RL, Palmieri TL, Horton JW, et al.; American Burn Association Consensus Conference on Burn Sepsis and Infection Group. American Burn Association Consensus Conference to define sepsis and infection in burns. J Burn Care Res 2007;28:776–90.

54. Mann-Salinas EA, Baun MM, Meininger JC, Murray CK, Aden JK, Wolf SE, Wade CE. Novel predictors of sepsis outperform the American Burn Association sepsis criteria in the burn intensive care unit patient. J Burn Care Res 2013;34:31–43.

55. Ducharme J, Self WH, Osborn TM, Ledeboer NA, Romanowsky J, Sweeney TE, et al. A multi-mRNA host-response molecular blood test for the diagnosis and prognosis of acute infections and sepsis: proceedings from a clinical advisory panel. J Pers Med 2020;10:266.

56. Mayhew MB, Buturovic L, Luethy R, Midic U, Moore AR, Roque JA, et al. A generalizable 29-mRNA neural-network classifier for acute bacterial and viral infections. Nat Commun 2020;11:1177.

57. Orhan ER, Temurtas F, Tanrikulu AC. Tuberculosis disease diagnosis using artificial neural networks. J Med Syst 2010;34:299–302.

58. Saybani MR, Shamshirband S, Hormozi SG, Wah TY, Aghabozorgi S, Pourhoseingholi MA, et al. Diagnosing tuberculosis with a novel support vector machine-based artificial intelligence recognition system. Iran Red Cresent Med J 2015;17:e24557.

59. Hall DL, Llinas J. An introduction to multi-sensor data fusion. Proc IEEE 1997;85:6–23.

60. Guiry JJ, Van de Ven P, Nelson J. Multi-sensor fusion for enhanced contextual awareness of everyday activities with ubiquitous devices. Sensors (Basel) 2014;14:5687–701.

61. Fayyad J, Jaradat MA, Gruyer D, Najjaran H. Deep learning sensor fusion for autonomous vehicle perception and localized: a review. Sensors (Basel) 2020;20:4220.

62. Kost GJ. Geospatial science and point-of-care testing: creating solutions for population access, emergencies, outbreaks and disasters. Front Public Health 2019;7:1–31.

63. Doyle-Delgado K, Chamberlain JJ. Use of diabetes-related applications and digital health tools by people with diabetes and their health care providers. Clin Diabetes 2020;38:449–61.

64. Food and Drug Administration (FDA). Coronavirus (COVID-19) update: FDA Authorizes first machine learning-based screening device to identify certain biomarkers that may indicate COVID-19 infection. https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-first-machine-learning-based-screening-device-identify (Accessed May 31, 2021).