Review

Artificial Intelligence and Mapping a New Direction in Laboratory Medicine: A Review

Daniel S. Herman,[a] Daniel D. Rhoads (iD) ,[b,c] Wade L. Schulz,[d] and Thomas J.S. Durant[d,*]

BACKGROUND: Modern artificial intelligence (AI) and machine learning (ML) methods are now capable of completing tasks with performance characteristics that are comparable to those of expert human operators. As a result, many areas throughout healthcare are incorporating these technologies, including in vitro diagnostics and, more broadly, laboratory medicine. However, there are limited literature reviews of the landscape, likely future, and challenges of the application of AI/ML in laboratory medicine.

CONTENT: In this review, we begin with a brief introduction to AI and its subfield of ML. The ensuing sections describe ML systems that are currently in clinical laboratory practice or are being proposed for such use in recent literature, ML systems that use laboratory data outside the clinical laboratory, challenges to the adoption of ML, and future opportunities for ML in laboratory medicine.

SUMMARY: AI and ML have and will continue to influence the practice and scope of laboratory medicine dramatically. This has been made possible by advancements in modern computing and the widespread digitization of health information. These technologies are being rapidly developed and described, but in comparison, their implementation thus far has been modest. To spur the implementation of reliable and sophisticated ML-based technologies, we need to establish best practices further and improve our information system and communication infrastructure. The participation of the clinical laboratory community is essential to ensure that laboratory data are sufficiently available and incorporated conscientiously into robust, safe, and clinically effective ML-supported clinical diagnostics.

Introduction

In the past 10 evolution of artificial intelligence (AI) and machine learning (ML) technologies. AI-based systems are now capable of partially or fully automating well-defined tasks with performance metrics that are comparable to those of expert human operators. As a result, many areas throughout the healthcare industry are incorporating AI/ML-based technologies into amenable workflows. Laboratory medicine and in vitro diagnostics (IVD) are on the leading edge of this wave, as AI is increasingly implemented within business intelligence tools, instrumentation cleared by the Food and Drug Administration, and laboratory-developed tests. In addition, much of the data generated by the clinical laboratory is naturally suited for AI as these data are often structured, discrete, and high-fidelity and produced in large volumes. Accordingly, AI and clinical decision support systems (CDSSs) often rely heavily on laboratory data.

As the use of AI continues to increase, laboratorians will benefit from improved AI literacy and technological awareness so that we can become effective stewards of the technology and laboratory data alike. To this end, this review begins with a brief introduction to AI and ML and then proceeds to survey the breadth of ML approaches and relevant applications. Subsequent topics include an overview of AI/ML-based technologies currently implemented in the clinical laboratory or having demonstrated the potential for such use, AI-based systems outside the clinical laboratory and how they leverage laboratory data, barriers to the adoption of AI, and future directions of AI and laboratory medicine.

Machine Learning Overview

In 1950, Alan Turing authored a seminal paper in the philosophical journal *Mind*, wherein he proposed the question: "Can machines think?" ([1]). A short time later, McCarthy et al. assembled a group of computer scientists and mathematicians for a 2-month summer research workshop at Dartmouth College to collectively pursue the formative study of a new field termed "artificial intelligence" ([2]). AI encompasses a wide spectrum ranging from general AI, referring to the creation of intelligent thinking machines, to narrow AI, or computational algorithms, that can learn specific tasks. Early

[a]Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA, USA; [b]Department of Laboratory Medicine, Cleveland Clinic, Cleveland, OH, USA; [c]Department of Pathology, Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, Cleveland, OH, USA; [d]Department of Laboratory Medicine, Yale University, New Haven, CT, USA.
*Address correspondence to this author at: 20 York St/PS 535, New Haven, CT 06510, USA. E-mail thomas.durant@yale.edu.

implementations of AI were primarily comprised of "if, then" rules used to automate protocol-driven processes. Such rule-based algorithms are used throughout clinical practice today and can be classified as nonadaptive AI (e.g., autoverification and antibiotic susceptibility reporting cascades). ML, which later emerged as a sub-discipline of AI, leverages techniques that are considered adaptive in that they can learn patterns from among data within a defined framework. The primary focus of this review is to consider contemporary IVD and laboratory management technologies that leverage ML techniques. While the boundaries between AI, ML, and even traditional statistics are fuzzy, in this discussion of ML, we broadly included algorithms in which the components and/or the relationships between those components were learned by a machine rather than either hard-coded by a human or derived using simple summary statistics.

The discipline of ML can be further subcategorized into supervised, semisupervised, and unsupervised ML. Most contemporary applications in the clinical laboratory use supervised ML, which will be the primary focus of this review. For the interested reader, further information on basic ML definitions and concepts can be found in recent literature (3, 4). Additional definitions of ML vernacular, referred to in this report, can be found in Table 1.

Supervision is synonymous with the use of training labels (i.e., annotations) ascribed to each of the samples within a data set. These labels are frequently assigned a priori by a human subject matter expert. The labels serve as the ground truth during the learning process. The machinery that is used for rote learning of informative patterns involves a multitude of mathematical approaches (e.g., neural networks, support vector machines, and regression algorithms) and reviews of these techniques are abundant in the literature. Importantly, most of these algorithms use an error function to monitor and direct iterative refinements to the patterns that are identified from the data. The optimization of such an error (or loss) function is arguably the raison d'être of ML.

In supervised ML, a data set is comprised of 2 components: (i) raw data (input features) and (ii) labels (training outcome). Raw data are commonly categorized as either structured or unstructured. Structured data can be intuitively visualized as a data set of fields and values, such as a data table with column headers and rows that are described by a data model. Conversely, unstructured data such as digital images have more complex encodings and require more manipulation. Training labels can take many forms, including continuous variables (e.g., concentrations for a calibration curve), dichotomous variables (e.g., benign or malignant), or categorical variables (e.g., artifact, benign, premalignant, or malignant).

Taken together, raw data and labels are used to train an ML model, which can then be applied to partially or completely automate a laboratory process (Fig. 1; Table 1). Over the last 2 decades, ML has been applied to an ever-expanding set of use cases. In the following section, we will review salient examples of ML applications in the clinical laboratory, including those in clinical practice or described in the emerging literature.

Applications of Machine Learning in Laboratory Medicine

While we are currently in a resurgence of the application of AI to healthcare, ML is ubiquitous in laboratory medicine and can be easily overlooked. The most common application of ML in the laboratory is the conversion of raw measurement signals into analyte concentrations, typically achieved by constructing calibration curves that model signal-to-concentration relationships using linear regression (5). ML also assists with data interpretation as it can distill multivariate data into a more intuitive form. A representative and widely adopted example of this is prenatal open neural tube and aneuploidy screening, wherein the predicted disease risk is generally estimated using a combination of regression and related discriminant analyses of protein biomarkers (6, 7). A calculated disease risk based on a set of laboratory results is a natural extension beyond simply reporting the individual results of many individual assays. Indeed, the current resurgence of ML has begun to reveal the much broader potential for this historically accepted but underutilized practice, now leveraging larger data sets with more variables and patients and more sophisticated methods.

CLINICAL CHEMISTRY AND IMMUNOLOGY

Chemistry and Immunology laboratories are particularly well suited for leveraging ML because they generate large and highly structured data that can be input into computational methods. These tools can help to translate raw data (e.g., spectra, profiles) into discrete results, review data for autoverification, and suggest clinical interpretations of multivariate results.

For the interpretation and quality control of electrophoresis traces and mass spectra, laboratories generally rely on labor-intensive procedures. Accordingly, there is growing interest in tools that can automate these processes to improve throughput and quality (8, 9). Electrophoresis of serum proteins is a long-standing and widely used diagnostic modality for screening and monitoring of monoclonal gammopathy-related disease (10, 11). However, result interpretation or verification generally involves labor-intensive review performed by highly trained individuals, which makes scaling this approach to other use cases time- and cost-prohibitive (12). To this end, there are several published reports that describe various approaches to

Review

Table 1. Definitions of common machine learning vernacular.		
Term	**Definition**	**Example**
Artificial intelligence (AI)	Referring to a broad set of technologies which are capable of automated decision making, or similarly intelligent behavior, via the analysis of data. Fundamentally, AI can be categorized as being either nonadaptive (e.g., rule-based) or adaptive (e.g., machine learning).	Nonadaptive AI: Autoverification of laboratory results Adaptive AI: Automated classification of leukocytes from digital image of peripheral blood smear using a trained machine learning algorithm
Machine learning (ML)	A subdiscipline of AI which leverages mathematical functions to analyze input data and, without explicit instruction, provide inferences from the data.	See Fig. 1
ML algorithm	A collection of mathematical functions that are used for a machine learning task.	See Fig. 1
ML model	The resulting file, or artifact, created from training an ML algorithm.	Logistic regression: vectors of coefficients Neural network: computational graph with associated matrix of weights.
Supervised ML	An approach to training ML algorithms in which the provided input data is associated with an outcome label. Commonly, the intent is for the model to learn how to map input data to the appropriate output data (label).	A model used to classify leukocytes may have initially been provided digital images of individual leukocytes with an associated label (e.g., lymphocyte, monocyte, etc.). Following algorithm training, the model could then be shown an image of a leukocyte without a label and infer a morphologic classification, within the scope of the original set of training labels.
Unsupervised ML	An approach to ML in which the ML algorithm analyzes input data without an associated label, to infer patterns, structure, or clusters within the data set.	Principal component analysis (i.e., dimensionality reduction) T-Stochastic Neighbor Embedding (tSNE)
		Continued

Term	Definition	Example
Semisupervised ML	An approach that uses a combination of supervised and unsupervised ML. This approach is commonly employed when there is a large amount of training data, with a limited amount of labeled data.	Photographs of person A, taken on a smartphone, may be initially collated into a separate "person A album." This may be done initially without end-user input (unsupervised). The photo application may then prompt the user to provide explicit labeling (supervised) of photos the unsupervised algorithm was unsure of, to better classify pictures of person A.
Underfitting	A model that fails to learn the inherent structure of the data provided to it during training.	A model that demonstrates suboptimal performance on the training data set.
Overfitting	A model that learns the inherent structure of the training data set so well, that when it is confronted with new data, that was not well represented in the training data set, performance is consequently poor.	A model that demonstrates optimal performance on the training data set, but suboptimally on the test data set, may be suffering from overfitting.

automating gel or capillary electrophoresis pattern interpretation, mostly using artificial neural networks (12–15). While these techniques currently demonstrate suboptimal specificity and are not yet widely adopted in clinical practice, integrating them within commercially available electrophoresis platforms as interpretation support tools would likely dramatically improve the efficiency and consistency of these evaluations.

For the analysis of mass spectra, ML approaches can support the translation of spectra into concentrations and quality review by enabling or increasing autoverification. Commercial products have recently become available for automated peak quantitation and quality analysis that have been shown to improve workflows (16, 17). ASCENT (Indigo BioAutomation) uses an exponentially modified Gaussian model to estimate peak areas and has shown good performance even in the setting of low-concentration analytes and modest signal-to-noise ratios (17). ML approaches have also been proposed for the automated and more sensitive quality review of spectrometry data and discrete results of general chemistries (8, 9).

Downstream of result generation, ML-based decision support tools show considerable promise in assisting pathologists' interpretation of complex multianalyte data. At the most basic level, ML approaches can help to jointly interpret discrete laboratory results with concrete relationships, like thyroid-stimulating hormone and free thyroxine (18). Extending further, ML can support and potentially improve upon manual interpretation for larger multianalyte panels such as steroid and amino acid profiles (12, 19–21).

Clinical chemistry laboratories also generate digital images (i.e., unstructured data), which given the recent advancements in computer vision, are highly conducive to semiautomated analysis. Urine sediment analysis is a labor-intensive process in which sedimentary particles are identified and quantitated by a skilled medical laboratory scientist. With ML, this analysis can be automated using one of the many highly performant image classification techniques that are currently available. Currently, there are many commercial products that offer digital imaging and subsequent classification of urine sedimentary particles, which can be presented to the end-user for review and verification (22–26). However, as these tools are young, manual review of pathological samples is still generally recommended (26). In immunology, ML-based image analysis is commonly combined with immunofluorescent assays for the detection and classification of antineutrophil cytoplasmic antibodies (27, 28). While there are currently only a few instances of digital imaging in the clinical chemistry

Table 2. Representative examples of how machine learning applications are being applied in laboratory medicine using structured and unstructured data.

Raw data	Label	ML algorithm	Clinical purpose
Structured data			
Basic demographic and clinical information, and CBC/differential results	PBFC results classified as negative or positive	DT or GLM	Predict PBFC results as negative or positive, proposed as an approach to triage PBFC utilization
Urine steroid metabolites quantified by GCMS and demographic data	Normal or abnormal; if abnormal, then classify by disease category.	RF, WSRF, or XGBT	Map data to comment templates to generate semi- or fully automated interpretive comments
Unstructured data			
Bounding-box coordinates and cropped images of individual intestinal protozoa, yeast, and PBCs	Species-level classification (e.g., *Giardia duodenalis* cyst, *Blastocystis* sp., etc.)	Deep CNN	Detection and classification of potential intestinal protozoa, yeast, or PBCs. Classifications reviewed by user prior to result verification.
Images of leukocytes from Romanowsky stained peripheral blood smears	Leukocyte differential; 17-cell types	ANN	Automatically classify leukocytes, subject to expert human operator review prior to result release

Abbreviations: ANN = artificial neural network; AANN = auto-associative neural network; DT = decision tree; GLM = generalized linear model; PBC = peripheral blood cell; PBFC = peripheral blood flow cytometry; RF = random forest; WSRF = weighted-subspace random forest; XGBT = extreme gradient boosted tree.
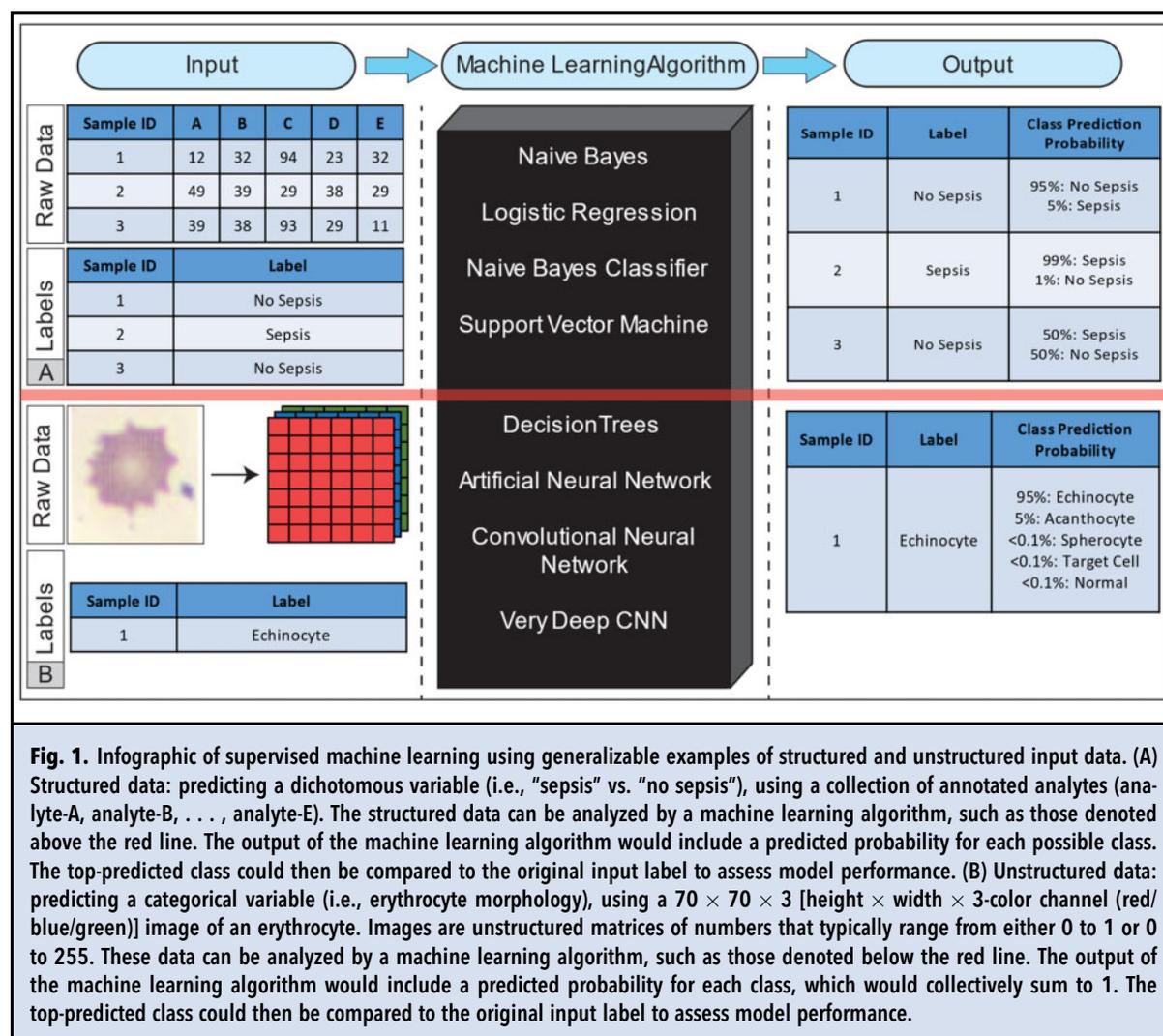
laboratory, there are many potential opportunities (e.g., sample tube imaging on automated preprocessing units) for which automated analysis by ML tools will likely be essential.

There are also many emerging laboratory testing approaches for which sophisticated and rapid computational tools are foundational. For instance, there is considerable interest in bringing mass spectrometry into the operating room for real-time biochemical profiling of surgical tissue samples (29, 30). In recent proof-of-concept applications, tissue-based samples (e.g., gas-phase ionic species or water droplet) are collected by a handheld surgical device and input to a mass spectrometer via a tube conduit. The resulting mass spectra are then analyzed in real time to enable rapid, in situ biochemical profiling. While this emerging IVD technology remains in the development stages, current approaches incorporate ML methods to distinguish between benign and malignant tissue spectra. Recent publications have described this approach for the identification of malignancies of various tissue types, including ovarian, thyroid, and lung (29, 31).

HEMATOPATHOLOGY

While ML has seen many forms of progress in recent years, one of the primordial advances, with respect to the current ML resurgence, was in the context of digital image analysis (32). In 2012, Krizhevsky et al. successfully integrated convolutional neural networks within a deep neural network, which resulted in a significant improvement in automated image classification performance relative to the state-of-the-art predicate method (32, 33). Subsequently, the computer vision field saw

Fig. 1. Infographic of supervised machine learning using generalizable examples of structured and unstructured input data. (A) Structured data: predicting a dichotomous variable (i.e., "sepsis" vs. "no sepsis"), using a collection of annotated analytes (analyte-A, analyte-B, . . . , analyte-E). The structured data can be analyzed by a machine learning algorithm, such as those denoted above the red line. The output of the machine learning algorithm would include a predicted probability for each possible class. The top-predicted class could then be compared to the original input label to assess model performance. (B) Unstructured data: predicting a categorical variable (i.e., erythrocyte morphology), using a $70 \times 70 \times 3$ [height $\times$ width $\times$ 3-color channel (red/blue/green)] image of an erythrocyte. Images are unstructured matrices of numbers that typically range from either 0 to 1 or 0 to 255. These data can be analyzed by a machine learning algorithm, such as those denoted below the red line. The output of the machine learning algorithm would include a predicted probability for each class, which would collectively sum to 1. The top-predicted class could then be compared to the original input label to assess model performance.

widespread adoption and expansion of this approach, which has led to highly performant image classification algorithms (34). Not surprisingly, there has been substantial interest in the adoption of image classification technologies in medical specialties such as radiology and anatomic pathology, where the collection of digital images continues to accelerate (35–37). Laboratory medicine will similarly be image-rich, particularly in the context of hematopathology.

Hematopathology as a specialty puts significant consideration on the visual interpretation of patterns, such as the morphologic characterization of cells (38, 39). Accordingly, there is a long history of researchers and manufacturers combining digital imaging and computer vision technologies to automate some facets of the hematopathology workflow with efforts dating back as far as the 1970s. For example, Bacus et al. used digital imaging and feature engineering to model erythrocyte morphology classification based on

cellular features (e.g., size and spicularity) (40). Since this time, there have been successful reports of automated classification of peripheral blood cells using a variety of modern ML technologies. Core lab commercial applications have been available since the early 2000s and, more recently, an increasing number of point-of-care devices (38, 39, 41–44).

One widely adopted, Food and Drug Administration (FDA)-cleared image analysis system enabled by ML is the Cellavision DM96 (45). The Cellavision system uses a static (i.e., locked weights) artificial neural network–based approach to precharacterize leukocytes and erythrocytes for an automated differential count and morphologic analysis, respectively. Like other image classification systems in the laboratory, substantial equivalence for this analyzer was established in the setting where a certified laboratory technologist was required to verify or modify suggested classifications for each cell prior to result release (45). Despite the need for expert review, there are numerous reports that

detail high classification accuracy and good correlation with manual differential counts ([46], [47]). More recently, there are emerging applications that do not require expert operator review, particularly for point-of-care testing. Some examples include the HemoScreen™ (PixCell Medical) and the Sight OLO® (Sight Diagnostics), both of which offer a 5-part cell differential using ML-based image analysis but without the requirement for expert operator review ([48]).

Malaria diagnosis by digital imaging of peripheral blood has received significant attention from international research communities ([49]). A recent review by Poostchi et al. offers a thorough overview of this topic and nicely organizes the considerable heterogeneity across published studies ([49]). These tests face particularly daunting challenges, such as conducting multiplanar focusing, handling noisy interfering objects (e.g., parasite-like inclusions), and enumerating objects of interest at high magnification from a digital microscopic image in the setting of overlapping cells ([50]). Recent reports are now pursuing approaches that use deep learning that may be able to overcome some of these difficulties ([50]–[52]). However, general performance remains suboptimal, and it is likely that more progress will be needed before we see widespread clinical implementation of these tools ([49]).

Beyond digital image analysis, other areas of ML-based laboratory support have been described in hematopathology. Optimizing test utilization is an area of rising interest, particularly in the context of emerging value-based reimbursement models and stewardship strategies for labor-intensive test methods. To this end, there are many publications that investigate the use of routine hematologic data to predict the likelihood of more complex test result abnormalities to guide ordering practices. As an example, Turbett et al. described a statistically driven method of triaging testing for anaplasmosis using complete blood count measurements ([53]). While their approach did not formally leverage an ML framework, this workflow is amenable to ML modeling and has been described elsewhere in published conference proceedings ([54]). Similarly, Zhang et al. investigated the use of decision trees and regression models to identify peripheral blood flow cytometry specimens, which were more likely to be abnormal. Using this approach, a generalized linear model achieved 100% sensitivity and 54% specificity (area under the curve = 0.919) ([55]).

Lastly, multiparameter flow cytometry is a partially structured, high-dimensional data set, which by those 2 criteria is inherently well suited as an input for ML modeling. Consequently, it has received increasing attention in ML-based literature over recent years ([56]–[58]). Recently, Gaidano et al. describe a decision tree model to classify B-cell non-Hodgkin lymphoma with an overall accuracy of 92.68% ([59]). Similarly, Ng et al. demonstrated the use of a random forest classifier that could be used to screen out negative cases with 100%

sensitivity ([60]). Conference proceedings have described the use of neural network–based approaches to decision support tools for multiparameter flow cytometry interpretation ([61], [62]). Unsupervised clustering and dimensionality reduction techniques are increasingly applied to flow cytometry data, addressing some of the potential ambiguity of manual gating, which impacts numerical data, and is of particular relevance to investigative projects ([63]). However, to date, these approaches have had limited impact on clinical practice where multidimensional data are generally interpreted by inspection of bivariate plots. Dimensionality reduction can be combined with ML classifiers, which may be collectively leveraged to assist in diagnostic interpretation ([64]). A persistent challenge is that clinical use often relies on recognition of subtle change to variable expression patterns, as well as understanding nuances of the specific instrumentation and reagents being used which are aspects not readily characterized by automated clustering methods. While there is great potential for leveraging ML-based technologies to analyze multiparameter flow cytometry data, efficient and clinically focused applications that incorporate ML remain limited at this time.

CLINICAL MICROBIOLOGY

Clinical microbiology is historically a manually intensive section of the clinical laboratory. In recent years, however, there has been increasing interest in employing automated workflows to improve efficiency and help to mitigate the ongoing nationwide shortage of medical laboratory scientists ([65]–[67]). With this shift in practice, there are opportunities for the integration of ML-based tools in these new workflows.

BACTERIOLOGY CULTURES

Total laboratory automation systems that enable culture-based microbiology testing are generally comprised of three main components: an inoculation unit, an incubation system, and a high-resolution imaging system. Digital plate imaging offers many advantages by creating a retrievable record of bacterial growth, allowing remote plate reading, and enhancing visual inspection using various lighting conditions ([68]). In addition, opportunities to employ ML to aid in analyzing these complex images, which have traditionally been interpreted manually, are also emerging. A convergence of maturing technologies including back-end automation of bacteriology cultures (digital imaging of petri plates using Kiestra or WaspLAB), affordable storage of large digital image files, and affordable computational power are creating these opportunities. Computer algorithms are currently being developed and employed as decision support tools for bacteriology culture interpretation ([69]–[71]). It is reasonable and feasible for software to

interpret at least some Petri plate image results, such as "no growth" bacterial cultures, and to autoverify these interpretations without human intervention, although this is not yet practiced in laboratory medicine.

In addition to automated interpretation of traditional growth media, more advanced interpretations of bacterial cultures are likely to be the next step in development and clinical implementation. The Accelerate Pheno® (Accelerated Diagnostics) is a recently FDA-cleared IVD device that uses digital dark-field microscopy images of single-cell bacterial cell growth in the presence of varying concentrations of antimicrobial agents to determine minimum inhibitory concentrations (MIC) and susceptibility interpretations. This device relies on a hierarchical system that combines multivariate logarithmic regression and computer vision (72). Using this approach, the Accelerate Pheno® system can provide rapid (e.g., hours) phenotypic MICs using direct samples from positive blood culture bottles. While further research is needed to evaluate its impact on patient care, it remains a timely example of ML-based tools being used in concert with unique, culture-based technology.

MICROSCOPIC ANALYSIS OF PRIMARY SPECIMENS

Beginning with Leeuwenhoek, microbiology's cornerstone has been the visual examination of primary samples. In clinical microbiology, these analyses include microscopic evaluation for bacteria (e.g., Gram stain), fungi, acid-fast bacilli including mycobacteria, ova and parasite exam, and blood parasites (e.g., apicomplexa). These microscopic analyses are used for screening or to achieve a definitive clinical diagnosis. Inherent strengths of ML tools align well with many image analysis use cases in clinical microbiology. These use cases include rare-event detection (e.g., acid-fast bacilli) and segmentation/classification (e.g., Nugent-scored Gram stain). ML tools have demonstrated utility in early studies for these applications: Nugent score Gram stain interpretation for bacterial vaginosis diagnosis (73), classification of bacteria in positive blood cultures (74), mycobacteria detection in sputa (75–77), ova and parasite detection and classification from stool samples (78), and blood parasite detection and quantification (79–81).

In clinical microbiology practice, classification algorithms can be tuned based on the desired role in clinical practice. Generally, rare event detection algorithms (e.g., acid-fast bacilli detection) are most useful when weighted toward high sensitivity at the expense of specificity, wherein the machine can screen an image for suspect events and a human expert can make the final decision as to whether the rare event is artifact or microbe. Similarly, high-sensitivity algorithms could be used to screen and autoverify no-growth urine cultures by analyzing nutrient agar Petri dishes. The utility of

these tools is to improve the efficiency of the time employed by the expert medical laboratory scientist reviewing the specimen and to potentially improve the sensitivity or reproducibility of the test (78, 82).

A recent and representative example of high-sensitivity algorithms was recently examined in the context of enteric parasite diagnostics. Mathison et al. described their clinical validation of convolutional neural network–based software, developed as a decision support tool to detect and classify intestinal protozoa found within digital images of trichrome-stained fecal specimens (78). The method presented in this report achieved high concordance with manual microscopic review and demonstrated superior sensitivity relative to the predicate method, as shown with serial dilution experiments. Similarly, Smith et al. described the application of convolutional neural networks to the interpretation of blood culture Gram stains, demonstrating sensitivity above 90% for Gram-positive cocci in clusters and chains and Gram-negative rods (74). These works demonstrate that ML-based technologies can achieve high-performance metrics and, in conjunction with expert review, can be helpful in a clinical microbiology setting. As digital imaging of microscopic analysis expands in microbiology, we are likely to see more applications such as these (5, 82–84).

INTERPRETATION OF MICROBIAL SEQUENCING DATA AND PCR

Next-generation sequencing (NGS) and real-time PCR interpretation can both benefit from computer-aided interpretation. Algorithms can be used in place of cycle thresholds for quantifying the initial target of a PCR assay (85). Whole-genome sequencing data may one day be used in place of phenotypic test results for predicting clinical efficacy of antibiotics (86), and early work has demonstrated that machines can predict antimicrobial resistance and susceptibility in certain instances (87–90). Relying on ML algorithms to interpret real-time PCR and to predict phenotypes or taxonomic identities using trained software instead of mechanistic decision trees could change how clinical microbiology is practiced. Nontargeted sequencing techniques, such as shotgun metagenomics, may rely on these techniques and offer important potential to enhance pathogen discovery, strain typing, and resistance prediction from clinical samples (91).

MOLECULAR DIAGNOSTICS

The development of high-throughput and high-multiplexity nucleic acid technologies has transformed the field of molecular diagnostics. These methods have been enabled in part by advances in ML. For example, at their core, many NGS methods analyze thousands of images of millions of microscopic clusters of labeled nucleic acids (92). These tools generate massive amounts

of data, necessitating robust pipelines for big data management that would be impossible for humans to interpret alone (93, 94). In addition, massive multiplexity requires sophisticated approaches to identify analytically valid results and interpretations amidst the sea of data (93). As previously discussed, ML is well suited for assisting in the analysis of large, well-structured data. Over the last decade, publications involving genomics and ML abound (94).

Modern NGS assays generate high-dimensional, structured data sets that can provide highly useful diagnostic and prognostic insights. However, given the size and complexity of these data sets, analysis of NGS data sets is labor-intensive and time-consuming. Accordingly, many software products use ML to streamline various aspects of the NGS data analysis pipeline. Like other areas of ML application, this technology can be used to make human interpretations more efficient or to provide new diagnostic capacity. These platforms can assist with variant calls, curation, and clinical interpretations (95, 96). These approaches may provide particular benefit in the annotation of variants of uncertain significance identified in the clinical setting. Early methods that relied on similarity of sequence or structure often had limited performance to predict clinical impact (97, 98), but newer methods are being developed to provide interpretations from functional analysis (99) to clinical impact (100). ML techniques have also been used to generate more complex interpretive results from broad genomic assays (e.g., generating polygenic risk scores for complex diseases), which are now available via both clinician-ordered and direct-to-consumer pathways (95, 101–103).

While molecular diagnostics in laboratory medicine predominantly interrogate nucleic acid sequences and the concentration of specific molecules, similar methods can be applied to other kinds of biological variation. These -omics–oriented tests can encompass studies including epigenomics, transcriptomics, proteomics, metabolomics, and microbiomics (104). These tests often include an ML component in the processing or analysis of the raw data being generated, often at a large scale. However, as a new clinical diagnostic area, the ability to combine multiple sets of -omics data (i.e., multiomics) and integrate high-fidelity phenotypic data represents a challenging yet promising, data-driven direction for molecular diagnostics. Much research is aimed at identifying clinically useful biomarkers using 1 or more testing modality (105–107). With the cost of these assays rapidly decreasing, our ability to identify meaningful information from such complex data is ever-improving. While the breadth of multiomic diagnostic testing strategies is beyond the scope of this review, they collectively represent the rapidly growing next frontier for ML and laboratory medicine (108).

GENERAL LABORATORY MEDICINE PRACTICE

The total testing process is a commonly referenced framework used for evaluating laboratory testing from quality assurance and quality control standpoints. Errors in the preanalytical phase of testing are thought to account for the highest frequency of laboratory errors but are challenging to prevent as many of the relevant processes are often beyond laboratory oversight (109). To this end, there are several recent publications that investigate the utility of ML techniques for identifying preanalytical errors. Recent and representative examples include the use of optical character recognition to identify mislabeled samples (110), the detection of spuriously increased glucose results due to intravenous fluid contamination(111), and logistic regression and support vector machines to detect "wrong blood in tube" errors (112). Both approaches demonstrated how a combination of commonly available laboratory results can be utilized for preanalytical quality assurance purposes. While such approaches are not yet widely adopted, they are representative of potentially value-added components, which may be integrated with commercially available laboratory information systems or middleware.

Regarding the analytic phase, intermittent testing of quality control material is the gold standard for evaluating analytic methods for instability (113). Moving averages are an alternative approach for identifying method drifts or shifts whereby the mean of consecutive patients' results are compared to control limits established for a specific patient population (114). This approach could identify problems in between the testing of external quality control samples, which can be as infrequent as daily. However, choosing the parameters of a moving averages protocol to maximize sensitivity and minimize false alarms is challenging. To this end, Ng et al. described an ML-based approach to detect systematic error among patient results using a simulated annealing model to optimize moving average protocols (e.g., control and truncation limits) and were able to implement these protocols in their production environment (115). Similarly, as part of quality assurance in the analytic phase, result verification processes aim to identify test errors prior to result release, which is commonly implemented using a rule-based system (i.e., nonadaptive AI) to accept or reject results. Demirci et al. recently described an adaptive, ML-based approach using artificial neural networks to develop a model that could be used for autoverification purposes. The model was found to be 91% sensitive and 100% specific as compared to result verification decisions by 7 clinical chemists (116).

Bordering between analytical and postanalytical, abnormal result flagging remains a critical aspect of laboratory services and a major driver of clinical decision-making (117, 118). Determination of clinically meaningful reference intervals is challenging, particularly when considering subpopulations of community testing

cohorts (e.g., sex- or age-specific reference intervals) and poorly standardized assays. Several methods have been developed to leverage in-practice clinical data for defining or refining reference intervals (119–121). Poole et al. describe an unsupervised ML approach to identifying clinical diagnoses associated with extreme results and then excluding patients with these diagnoses from reference interval calculation, as typically seen with the a posteriori approach (121). Such approaches show great potential for ensuring reference intervals used are most appropriate for local patient populations. Indeed, as laboratory software continues to expand in functionality, these examples demonstrate the potential of leveraging data generated as part of routine clinical practice to improve that clinical practice by using ML methods. One final example of ML applications in general laboratory practice is resource stewardship, including blood banking and transfusion, wherein predictive analytics can guide the prospective utilization of blood products (122).

Laboratory Data and Machine Learning Outside of the Laboratory

Laboratory data are a critical component of the diagnostic process. From chronic disease to acute infection, almost all clinical decision-making will leverage the results of a laboratory test to guide the diagnosis, treatment, or prediction of the outcome of disease. Within the clinical laboratory, basic mathematical calculations, rule-based engines, and ML-based tools are all routinely used. But additional applications of these data exist in tools developed external to the clinical laboratory, spanning from real-time clinical predictive models to health system business intelligence.

The interest in leveraging ML to drive electronic health record (EHR)-based tools, such as CDSSs, is an area of increasing interest (123). CDSSs create the links between data, algorithm, and end-user to be able to deliver information at the point of care and allow for changes in clinical decision-making. Like other applications of AI and ML, CDSSs can range from rule-based triggers through black-box ML. The goal of these systems is to provide the user with suggested clinical pathways or interventions based on the existing, high-dimensional data within the clinical record. Clinical laboratory data are often a key component of clinical predictive models and CDSSs, as these data are often structured and are typically available as soon as a test result is verified, unlike manually entered data, such as flow sheet and clinical notes, which may not be finalized until the end of a provider's shift. The goals of applying more advanced ML-based approaches to CDSS include the reduction of alert fatigue by providing more appropriate alerts at more precise times, by leveraging and distilling more EHR data as compared to manually developed rule-based approaches. Such models have been commonly implemented to identify and provide treatment guidance, including for detecting acute kidney injury (124–126) and oncology treatment recommendation (127), as well as to predict clinical outcomes, including acute deterioration (128, 129) and postoperative outcomes (130, 131).

Within the laboratory, calculations and predictive models, regardless of how basic they are, are regulated by the Centers for Medicare and Medicaid Services via accreditation by deemed organizations such as the College of American Pathologists. For example, the College of American Pathologists provides validation and documentation requirements for calculations of the international normalized ratio and estimated glomerular filtration rate, as well as for the validation of autoverification rules. Outside of the laboratory, the regulation of AI and CDSS is less clear and is an area of active discussion and evolving guidance from the FDA (132–135). While laboratory data provide a rich, structured data set to drive these algorithms, active integration of laboratorians in the development, implementation, and management of these tools is critical as common laboratory decisions, such as the deployment of a new assay or changing a reference interval, could have unintended consequences on both rule-based and AI-driven CDSSs. In addition to the data and algorithmic validity of these approaches, it is also crucial to assess the impact of such tools on the relevant clinical outcomes. Even in situations when experts would expect an ML-triggered intervention would be likely to improve clinical outcomes and very unlikely to cause harm, recent studies have found that this may not always be the case (136).

Key Barriers

While the potential for the application of ML in laboratory medicine is massive, the progress to date has been modest. This gap is due to a combination of many factors, including the complexity of pathophysiology, risks of automation, limited access to data and tools, and inadequate data quality.

THE COMPLEXITY OF PATHOPHYSIOLOGY

A fundamental challenge is that medicine is complicated. Laboratory tests provide snapshots of bits of descriptive information. However, there are many ways in which normal physiology can be disrupted, and within each pathology, there is considerable between-patient variability. Whether ML models can be built with sufficient sophistication to achieve the desired clinical performance for complex problems, such as screening for asymptomatic disease, remains an open question. Additional testing comes at higher costs and runs the risk of false-positive findings and overmedicalization.

Thus, when considering whether to implement ML for a particular clinical use case, we need to understand the full implementation context and evaluate ML approaches as we would any other clinical practice or laboratory test by performing analytical validations and verifying or establishing clinical effectiveness (137).

RISKS OF IMPLEMENTING ML IN PRACTICE

The decision to implement a new process in clinical practice requires a comprehensive assessment of risks and benefits. There is considerable uncertainty and potential risk in adopting any innovative technology, including ML for clinical care. Because of the novelty of some ML approaches, the guidelines and regulations for developing, assessing, and monitoring ML tools are under active development. This lack of best practices has allowed premature or poor applications of ML in clinical practice and contributed to the overall slow implementation of useful tools (136, 138, 139).

One way of mitigating the risks of implementing ML is for models to make recommendations to laboratorians or providers rather than automatically determining clinical actions. A step in this direction that can mitigate risk while enabling more automation is to use ML models that are interpretable whenever possible. Interpretability is an enigmatic and heterogeneously defined concept. As specified in the FDA's draft guidance for clinical decision support (September 2019), one formulation is whether the end-user can "independently review the basis for [a model's] recommendations." An example of this concept is that for imaging models, it is easier to validate models, verify individual results, and monitor performance of a model that can highlight specific fields of relevance to the model's prediction, such as red blood cells suspicious for a parasite (Fig. 2). Such a model is not fully interpretable because we cannot understand the model's thinking, but it is considerably more interpretable than a completely black-box model. Ideally, if the model's user can understand precisely why an ML model is calling a cell, or patient, positive or negative, they can develop trust in the model and, when appropriate, second-guess a model's prediction.

UPCOMING REGULATORY FRAMEWORK

AI- and ML-based software that can continuously learn presents a novel challenge for existing premarket approval pathways by the FDA. While the FDA has cleared AI/ML-based software in the past, until recently, these applications have only leveraged algorithms that are locked and are not continuously updating via analysis of newly acquired data (133). Due to the locked nature of these models, the output, or answer, will be the same for a given set of input data throughout the life span of the product. Accordingly, novel regulatory frameworks are needed for algorithms designed to continuously analyze new data, learn, and behave adaptively. To this end, in recent years, the FDA has incrementally progressed toward the development of a formal regulatory framework for the approval and oversight of AI/ML-based software (i.e., Software as a Medical Device), with particular focus on how to manage their adaptive nature. Principal components of this newly proposed framework include predetermined change control plans, which specify "what" aspects of the algorithm will change and "how" the manufacturer will effectuate that change. In addition, the FDA will seek to ensure good ML practices are used in product development and oversight of ML algorithms, promote transparency regarding device function, minimize algorithm bias and optimize robustness, and monitor real-world performance. Similarly, the FDA has also issued draft guidance on the separate but related topic of CDSSs, which, in some instances, may also be classified as Software as a Medical Device (140).

As medical devices leveraging AI/ML continue to proliferate in the clinical laboratory and EHR, understanding pre- and postmarket approval pathways will be essential. It remains unclear whether the initial validation and ongoing monitoring of ML algorithms would be the responsibility of the vendor, end-user, or a combination thereof. Concepts outlined in draft regulatory guidance, regarding good ML practice and methods for algorithm validation, are generally analogous to good laboratory medicine practice. Validation and monitoring of algorithms will reduce the risk of bias, overfitting, and performance degradation over time and, much like a laboratory test, can limit the impact of such issues on patient care and outcomes (4, 134, 141). Furthering ML education in laboratory medicine will likely assist laboratory professionals in understanding the fundamental purpose of why AI/ML technologies require close monitoring, thereby improving the adoption and implementation of these practices. Lastly, validation and monitoring ML algorithms will likely come with difficulties in collating the necessary data and presenting it to end-users for efficient review.

LIMITED ACCESS TO DATA AND TOOLS

Training models to grasp medical complexity requires a lot of data. The magnitude of information needed often necessitates learning from data collected as part of clinical practice. Unfortunately, these data exist in silos, and there are many obstacles to aggregating said data.

Clinical data are siloed in part because of concern for patient privacy and ambiguity regarding who owns patients' data. The use of patient data for research requires explicit consent or a waiver of consent because of the risk of reidentification and unauthorized access to sensitive data (142, 143). While the potential long-term benefits to the practice of medicine are substantial, the
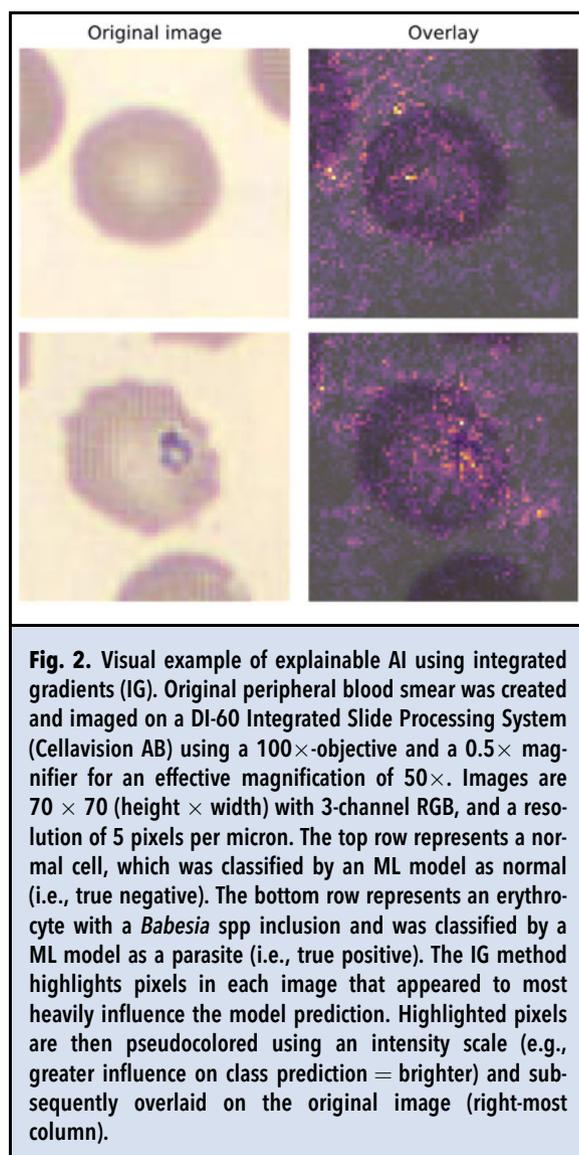
Fig. 2. Visual example of explainable AI using integrated gradients (IG). Original peripheral blood smear was created and imaged on a DI-60 Integrated Slide Processing System (Cellavision AB) using a 100×-objective and a 0.5× magnifier for an effective magnification of 50×. Images are 70 × 70 (height × width) with 3-channel RGB, and a resolution of 5 pixels per micron. The top row represents a normal cell, which was classified by an ML model as normal (i.e., true negative). The bottom row represents an erythrocyte with a *Babesia* spp inclusion and was classified by a ML model as a parasite (i.e., true positive). The IG method highlights pixels in each image that appeared to most heavily influence the model prediction. Highlighted pixels are then pseudocolored using an intensity scale (e.g., greater influence on class prediction = brighter) and subsequently overlaid on the original image (right-most column).

concern for the individual has been held as paramount. Movements toward more broadly consenting patients for research are now empowering patients to decide whether to share their data and contribute to these efforts (144).

The usefulness of patients' data has also made such data a valuable commodity that health systems have been reticent to share. However, this has been changing rapidly as health systems and patient groups are making data available in federated research networks (145–147), contributing to health information exchanges (148), publishing deidentified data sets for proof-of-concept research (149), and sharing data with interested companies. In addition, within clinical practice, EHR applications have begun enabling sharing of data between different health systems.

A related challenge to data access and aggregation is limited interoperability, which refers to the inconsistency of the processes and practices for transmitting data between devices and software applications (150, 151). There are specifications for how to transmit laboratory data between devices, such as an instrument and the laboratory information system (152). However, the specifications that are in common use are so broad that enormous effort is needed to program each interface, and substantial amounts of information can be lost between the instrument and the EHR. As a result, aggregating data across clinical practices requires enormous harmonization efforts. To harmonize data, each aggregation effort invests inordinate time in the formulation of and mapping to a common data model with standard vocabularies (153). However, to our knowledge, none of these aggregation efforts contain the metadata needed to sufficiently interpret and harmonize laboratory test results, likely due to the complexity of the data, poor access to the pertinent metadata, and limited participation from and engagement with clinical laboratory experts.

Improvement to the interoperability of laboratory data is possible as there are new and evolving specifications. The Fast Healthcare Interoperability Resources specification is richer, more detailed, and built atop computer industry standard conventions (154). More detailed and precise specifications have also been developed specifically for communication within clinical laboratories, published as the Laboratory Analytical Workflows by the IVD Industry Connective Consortium and the Integrating the Healthcare Enterprise and further codified in Clinical and Laboratory Standards Institute AUTO16. These specifications spell out how to transmit critical information regarding laboratory test results, such as the unique device identifiers that precisely label a test result's performing instrument and method. These achievements bring the potential to considerably streamline laboratory data exchange but have been minimally adopted thus far. If laboratory community members were more engaged in and advocating for interoperability, it could catalyze the implementation of these new standards.

These same barriers also exist within individual clinical laboratories. Our information systems and interfaces are not designed to flexibly enable the implementation of new computational tools. Poor interoperability between devices and information systems makes it extremely challenging to access data or to connect external tools that use ML. Moreover, our information systems' rigidity severely limits their ability to implement more sophisticated computational methods internally. There is progress being made on these fronts, but it has been glacial. This slow pace is likely due to a combination of factors, including the absence of coordinated demand from the clinical laboratory community, immense inertia to change, limited competition among information

system vendors, and the misalignment of incentives across various stakeholders.

Beyond the challenges of accessing big data, the quality of the available data has also been a substantial barrier to progress. Even if we had interoperable devices and complete metadata, the limited standardization and harmonization of laboratory tests add considerable complexity to learning ML models that can generalize across clinical practices. In addition, as clinical data are not collected for the purpose of secondary research, the data include considerable bias, including the consequences of inequities of care associated with patients' race (155). This bias can manifest as differential data missingness or can surreptitiously affect predictors or outcomes. Naïve ML approaches to learning predictive models that do not account for the data ascertainment schema or inappropriately make use of biased surrogate outcomes risk maintaining or exacerbating these biases (156). The ethical use of ML requires conscientious consideration of the clinical use case how the ML application will affect diverse patients. The importance of such efforts is illustrated by the controversy of the inclusion of race in the calculation of estimated glomerular filtration rate (157). Race was included in the estimated glomerular filtration rate prediction model because it improved predictive performance in the available training data, but its inclusion reinforces race-based medicine, only modestly improves performance, and likely exacerbates underlying inequities in the care of patients with chronic kidney disease (158).

Future Directions and Opportunities

Medical device manufacturers are actively exploring diagnostic approaches that are newly feasible with AI. We can likely expect the further expansion of IVD technologies into important and potentially unanticipated directions. A salient and recent example is how the Accelerate Pheno® is used to provide rapid antimicrobial susceptibility test results from real-time digital image analysis of bacterial cell growth. Similarly, Chen et al. recently described the application of AI to quantify protein biomarkers within a microbubbling digital assay format using bright-field smartphone microscopy (159). Further, researchers are continuing to investigate whether AI can enable measurements of nontraditional analytes such as those found in the analysis of breath, the pupillary light reflex, or vocal patterns (160–162). While the latter examples may fall beyond the scope of traditional laboratory medicine, it remains an open question how such tests would fit into existing point-of-care management strategies and regulatory structures.

Computational pathology is an emerging field that incorporates AI and extends beyond digital pathology and whole-slide imaging. Computational pathology is a pathologist-led diagnostic approach involving the integration and analysis of raw clinical data, including multiple data sources (e.g., laboratory information system, EHR, imaging, etc.) (163). Luo et al. recently published a straightforward example of this that involved the combination of patient demographic information and routinely available laboratory measurements to predict disease or future laboratory results. They demonstrated high accuracy with their method and highlighted the potential value of multianalyte analyses (164). While the results of preliminary efforts in computational pathology are promising, there are a multitude of challenges to be considered when using real-world data. One of the major challenges confronted when deploying prediction models in production environments is the sparsity of data. Researchers have investigated technical solutions to ameliorate the issues associated with sparse matrices, and ML-based methods to impute missing data have shown promise in improving the accuracy of laboratory data-based predictions (165). Further progression of computational pathology will undoubtedly require the close coordination between computational specialists, pathologists, and clinicians to ensure high-quality and clinical useful results.

As discussed in previous sections, there are many open questions regarding the best practices and regulatory frameworks for ML in clinical practice. Research on ML has demonstrated repeatedly that the technology is susceptible to many kinds of errors such as overfitting. This may limit the generalizability of a model to future data and result in unexpected deviations from previous apparent performance. These types of errors intuitively resemble the random and systematic error we observe with IVD assays. Although the precise error mechanisms are different, much can be learned from well-established clinical laboratory frameworks for ensuring quality. Laboratory practices of quality control and quality assurance are excellent correlates for appropriate postdeployment monitoring of ML models in production environments. While such monitoring is not yet required, draft guidance on these issues from the FDA and other regulatory bodies suggests that mandates in this area are forthcoming. However, delineations of responsibility for such provisions remain to be defined (134). Furthermore, ML-driven approaches can maintain or exacerbate healthcare disparities due to biases present in the data or within existing care delivery systems. It is critical that best practices and regulatory frameworks consider how to evaluate for and mitigate these effects.

Beyond the technically oriented future of AI, socialization of AI technology with laboratorians and clinicians is

an area of active discussion among professional organizations and researchers. A recent survey of laboratory professionals indicated that a quarter of respondents were concerned about potential job loss and quality issues with AI implementation. In addition, 72% of respondents were unsure or indicated they have never been in contact with an AI application in their daily activities. These results may suggest that, as the presence of AI continues to increase in the laboratory, there will be a need to promote education on technological awareness, ML literacy, and the scope and purpose of AI among laboratorians (166). Historically, computer science and AI have progressed much faster than clinical medicine. While there is a perceived benefit of implementing newly developed AI technologies in real time as they emerge, there remains a need to resolve the discrepant pace of the 2 fields and emphasize the need for evidence-based implementations of AI/ML-based models (167). Lastly, as many of the models being implemented, both in and outside the laboratory, rely on data generated by the laboratory, laboratorians are uniquely qualified to be stewards of these technologies, which offers a potential area of growth opportunity for the profession (168).

Finally, additional opportunities exist for business intelligence applications in laboratory medicine, particularly in how the clinical laboratory demonstrates value to healthcare delivery organizations amidst shifting reimbursement paradigms. While laboratory services have historically been reimbursed in a fee-for-service structure, value-driven healthcare initiatives are poised to change the utilization and management of laboratory resources. Recently, the Laboratory 2.0 concept has been introduced to encourage the application of laboratory practice principles and the analysis of laboratory data to optimize clinical care practices traditionally outside of the laboratory's domain (169). While there are limited examples of business intelligence platforms assisting in laboratory management in this regard, this is an area of big potential and where we are likely to see expansion soon.

AI and ML have and will continue to dramatically alter the way in which laboratory data are analyzed and drive clinical care decisions. The ongoing development of more sophisticated ML methods, coupled with emerging laboratory measurement technologies, should lead to further improvements in clinical efficiency and patient outcomes. It is paramount that these approaches be rigorously designed, evaluated, and monitored to ensure quality, achieve effectiveness, and minimize harm. Laboratorians have an important role to play in the development and stewardship of ML in laboratory medicine to enable us to realize the full potential of data-driven healthcare.

References

1. Turing AM. Computing machinery and intelligence. In: R Epstein, G Roberts, G Beber, editors. Parsing the Turing test. Dordrecht: Springer Netherlands; 2009. p. 23–65.

2. McCarthy J, Minsky ML, Rochester N, Shannon CE. A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. AI Magazine 2006;27:12.

3. Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial intelligence and machine learning in pathology: the present landscape of supervised methods. Acad Pathol 2019;6:2374289519873088.

4. Harrison JH, Gilbertson JR, Hanna MG, Olson NH, Seheult JN, Sorace JM, et al. Introduction to artificial intelligence and machine learning for pathology. [Epub

ahead of print] Arch Pathol Lab Med January 25, 2021 as doi:10.5858/arpa.2020-0541-CP.

5. Rhoads DD. Computer vision and artificial intelligence are emerging diagnostic tools for the clinical microbiologist. J Clin Microbiol 2020;58:e00511–20.

6. Wald NJ, Cuckle HS, Densem JW, Nanchahal K, Royston P, Chard T, et al. Maternal serum screening for Down's syndrome in early pregnancy. BMJ 1988;297:883–7.

7. Williams CJ, Lee SS, Fisher RA, Dickerman LH. A comparison of statistical methods for prenatal screening for Down syndrome. Appl Stochastic Models Bus Ind 1999;15:89–101.

8. Wang H, Wang H, Zhang J, Li X, Sun C, Zhang Y. Using machine learning to develop an autoverification

system in a clinical biochemistry laboratory. Clin Chem Lab Med 2020;59:883–91.

9. Yu M, Bazydlo LAL, Bruns DE, Harrison JH. Streamlining quality review of mass spectrometry data in the clinical laboratory by use of machine learning. Arch Pathol Lab Med 2019;143:990–8.

10. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos M-V, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol 2014;15:e538–48.

11. Milani P, Murray DL, Barnidge DR, Kohlhagen MC, Mills JR, Merlini G, et al. The utility of MASS-FIX to detect and monitor monoclonal proteins in the clinic. Am J Hematol 2017;92:772–9.

12. Altinier S, Sarti L, Varagnolo M, Zaninotto M, Maggini M, Plebani M. An expert system for the classification of serum protein electrophoresis patterns. Clin Chem Lab Med 2008;46:1458–63.

13. Männer GA, Schweiger CR, Söregi G, Pohl AL. Detection of monoclonal gammopathies in serum electrophoresis by neural networks. Clin Chem 1993;39:1984–85.

14. Kratzer MA, Ivandic B, Fateh-Moghadam A. Neuronal network analysis of serum electrophoresis. J Clin Pathol 1992;45:612–5.

15. Ognibene A, Motta R, Caldini A, Terreni A, Dea ED, Fabris M, Messeri G. Artificial neural network-based algorithm for the evaluation of serum protein capillary electrophoresis. Clin Chem Lab Med 2004;42:1451–2.

16. Vicente FB, Lin DC, Haymond S. Automation of chromatographic peak review and order to result data transfer in a clinical mass spectrometry laboratory. Clin Chim Acta 2019;498:84–9.

17. Zabell APR, Foxworthy T, Eaton KN, Julian RK. Diagnostic application of the exponentially modified Gaussian model for peak quality and quantitation in high-throughput liquid chromatography-tandem mass spectrometry. J Chromatogr A 2014;1369:92–7.

18. Hadlow NC, Rothacker KM, Wardrop R, Brown SJ, Lim EM, Walsh JP. The relationship between TSH and free T_4 in a large population is complex and nonlinear and differs by age and sex. J Clin Endocrinol Metab 2013;98:2936–43.

19. Wilkes EH, Rumsby G, Woodward GM. Using machine learning to aid the interpretation of urine steroid profiles. Clin Chem 2018;64:1586–95.

20. Wilkes EH, Emmett E, Beltran L, Woodward GM, Carling RS. A machine learning approach for the automated interpretation of plasma amino acid profiles. Clin Chem 2020;66:1210–18.

21. Eisenhofer G, Durán C, Cannistraci CV, Peitzsch M, Williams TA, Riester A, et al. Use of steroid profiling combined with machine learning for identification and subtype classification in primary aldosteronism. JAMA Netw Open 2020;3:e2016209.

22. Enko D, Stelzer I, Böckl M, Derler B, Schnedl WJ, Anderssohn P, et al. Comparison of the diagnostic performance of two automated urine sediment analyzers with manual phase-contrast microscopy. Clin Chem Lab Med 2020;58:268–73.

23. Laiwejpithaya S, Wongkrajang P, Reesukumal K, Bucha C, Meepanya S, Pattanavin C, et al. UriSed 3 and UX-2000 automated urine sediment analyzers vs manual microscopic method: a comparative performance analysis. J Clin Lab Anal 2018;32:e22249.

24. Oyaert M, Delanghe J. Progress in automated urinalysis. Ann Lab Med 2019;39:15–22.

25. Liang Y, Kang R, Lian C, Mao Y. An end-to-end system for automatic urinary particle recognition with convolutional neural network. J Med Syst 2018;42:165.

26. İnce FD, Ellidağ HY, Koseoğlu M, Şimşek N, Yalçın H, Zengin MO. The comparison of automated urine analyzers with manual microscopic examination for urinalysis automated urine analyzers and manual urinalysis. Pract Lab Med 2016;5:14–20.

27. Nagy G, Csípő I, Tarr T, Szűcs G, Szántó A, Bubán T, et al. Anti-neutrophil cytoplasmic antibody testing by indirect immunofluorescence: computer-aided versus conventional microscopic evaluation of routine diagnostic samples from patients with vasculitis or other inflammatory diseases. Clin Chim Acta 2020;511:117–24.

28. De Bruyne S, Speeckaert MM, Van Biesen W, Delanghe JR. Recent evolutions of machine learning applications in clinical laboratory medicine. Crit Rev Clin Lab Sci 2020;58:131–152.

29. Zhang J, Rector J, Lin JQ, Young JH, Sans M, Katta N, et al. Nondestructive tissue analysis for ex vivo and in vivo cancer diagnosis using a handheld mass spectrometry system. Sci Transl Med 2017;9:eaan3968.

30. Balog J, Sasi-Szabó L, Kinross J, Lewis MR, Muirhead LJ, Veselkov K, et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. Sci Transl Med 2013;5:194ra93.

31. Sans M, Zhang J, Lin JQ, Feider CL, Giese N, Breen MT, et al. Performance of the masspec pen for rapid diagnosis of ovarian cancer. Clin Chem 2019;65:674–83.

32. Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. Adv Neural Inform Process Syst 2012;25:1097–105.

33. Angelova A, Krizhevsky A, Vanhoucke V, Ogale A, Ferguson D. Real-time pedestrian detection with deep network cascades. Proceedings of the British Machine Vision Conference 2015. British Machine Vision Association; 2015. p. 32.1–12.

34. Khan A, Sohail A, Zahoora U, Qureshi AS. A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 2020;53:5455–516.

35. 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.

36. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inform 2016;7:29.

37. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: challenges and opportunities. Med Image Anal 2016;33:170–75.

38. Mohammed EA, Mohamed MMA, Far BH, Naugler C. Peripheral blood smear image analysis: a comprehensive review. J Pathol Inform 2014;5:9.

39. Acevedo A, Alférez S, Merino A, Puigví L, Rodellar J. Recognition of peripheral blood cell images using convolutional neural networks. Comput Methods Programs Biomed 2019;180:105020.

40. Bacus JW, Belanger MG, Aggarwal RK, Trobaugh FE. Image processing for automated erythrocyte classification. J Histochem Cytochem 1976;24:195–201.

41. Ramesh N, Dangott B, Salama ME, Tasdizen T. Isolation and two-step classification of normal white blood cells in peripheral blood smears. J Pathol Inform 2012;3:13.

42. Egelé A, Stouten K, van der Heul-Nieuwenhuijsen L, de Bruin L, Teuns R, van Gelder W, et al. Classification of several morphological red blood cell abnormalities by DM96 digital imaging. Int J Lab Hematol 2016;38:e98–e101.

43. Prinyakupt J, Pluempitiwiriyawej C. Segmentation of white blood cells and comparison of cell morphology by linear and naïve Bayes classifiers. Biomed Eng Online 2015;14:63.

44. Durant TJS, Olson EM, Schulz WL, Torres R. Very deep convolutional neural networks for morphologic classification of erythrocytes. Clin Chem 2017;63:1847–55.

45. 510(K) Summary: DiffMaster Octavia. 2018. https://www.accessdata.fda.gov/cdrh_docs/pdf/K003301.pdf (Accessed September 2021).

46. Briggs C, Longair I, Slavik M, Thwaite K, Mills R, Thavaraja V, et al. Can automated blood film analysis replace the manual differential? An evaluation of the CellaVision DM96 automated image analysis system. Int J Lab Hematol 2009;31:48–60.

47. Kratz A, Bengtsson H-I, Casey JE, Keefe JM, Beatrice GH, Grzybek DY, et al. Performance evaluation of the CellaVision DM96 system: WBC differentials by automated digital image analysis supported by an artificial neural network. Am J Clin Pathol 2005;124:770–81.

48. Ben-Yosef Y, Marom B, Hirshberg G, D'Souza C, Larsson A, Bransky A. The HemoScreen, a novel haematology analyser for the point of care. J Clin Pathol 2016;69:720–25.

49. Poostchi M, Silamut K, Maude RJ, Jaeger S, Thoma G. Image analysis and machine learning for detecting malaria. Transl Res 2018;194:36–55.

50. Abbas N, Saba T, Rehman A, Mehmood Z, Kolivand H, Uddin M, Anjum A. Plasmodium life cycle stage classification based quantification of malaria parasitaemia in thin blood smears. Microsc Res Tech 2019;82:283–95.

51. Liang Z, Powell A, Ersoy I, Poostchi M, Silamut K, Palaniappan K. CNN-based image analysis for malaria diagnosis. 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Shenzhen, China: IEEE; 2016. p. 493–496.

52. Bibin D, Nair MS, Punitha P. Malaria parasite detection from peripheral blood smear images using deep belief networks. IEEE Access 2017;5:9099–108.

53. Turbett SE, Anahtar MN, Pattanayak V, Azar MM, Coffey KC, Eng G, et al. Use of routine complete blood count results to rule out anaplasmosis without the need for specific diagnostic testing. Clin Infect Dis 2020;70:1215–21.

54. Durant T, Peaper D. Logistic regression modeling: statistically driven stewardship of anaplasma polymerase chain reaction testing by complete blood count and basic metabolic profile. Am J Clin Pathol 2016;146(suppl 1):S89–93.

55. Zhang ML, Guo AX, Kadauke S, Dighe AS, Baron JM, Sohani AR. Machine learning models improve the diagnostic yield of peripheral blood flow cytometry. Am J Clin Pathol 2020;153:235–42.

56. Lakoumentas J, Drakos J, Karakantza M, Nikiforidis GC, Sakellaropoulos GC. Bayesian clustering of flow cytometry data for the diagnosis of B-chronic lymphocytic leukemia. J Biomed Inform 2009;42:251–61.

57. Biehl M, Bunte K, Schneider P. Analysis of flow cytometry data by matrix relevance learning vector quantization. PLoS One 2013;8:e59401.

58. Manninen T, Huttunen H, Ruusuvuori P, Nykter M. Leukemia prediction using sparse logistic regression. PLoS One 2013;8:e72932.

59. Gaidano V, Tenace V, Santoro N, Varvello S, Cignetti A, Prato G, et al. A clinically applicable approach to the classification of B-cell non-Hodgkin lymphomas with flow cytometry and machine learning. Cancers (Basel) 2020;12:12.

60. Ng DP, Zuromski LM. Augmented human intelligence and automated diagnosis in flow cytometry for hematologic malignancies. Am J Clin Pathol 2021;155:597–605.

61. Kern W, Elsner F, Zhao M, Mallesh N, Schabath R, Haferlach C, et al. An artificial neural network providing highly reliable decision support in a routine setting for classification of B-cell neoplasms based on flow cytometric raw data. Blood 2019;134:886–6.

62. Höllein A, Zhao M, Schabath R, Haferlach T, Haferlach C, Krawitz P, et al. An artificial intelligence (AI) approach for automated flow cytometric diagnosis of B-cell lymphoma. Blood 2018;132:2856–56.

63. Duetz C, Bachas C, Westers TM, van de Loosdrecht AA. Computational analysis of flow cytometry data in hematological malignancies: future clinical practice? Curr Opin Oncol 2020;32:162–9.

64. Flores-Montero J, Grigore G, Fluxá R, Hernández J, Fernandez P, Almeida J, et al. EuroFlow Lymphoid

Screening Tube (LST) data base for automated identification of blood lymphocyte subsets. J Immunol Methods 2019;475:112662.

65. Garcia E, Kundu I, Kelly M, Soles R. The American Society for Clinical Pathology's 2018 Vacancy Survey of Medical Laboratories in the United States. Am J Clin Pathol 2019;152:155–68.

66. Garcia E, Kundu I, Ali A, Soles R. The American Society for Clinical Pathology's 2016–2017 Vacancy Survey of Medical Laboratories in the United States. Am J Clin Pathol 2018;149:387–400.

67. Williams RE, Trotman RE. Automation in diagnostic bacteriology. J Clin Pathol Suppl Coll Pathol 1969;3: 8–13.

68. Bailey AL, Ledeboer N, Burnham C-AD. Clinical microbiology is growing up: the total laboratory automation revolution. Clin Chem 2019;65:634–43.

69. Dauwalder O, Michel A, Eymard C, Santos K, Chanel L, Luzzati A, et al. Use of artificial intelligence for tailored routine urine analyses. Clin Microbiol Infect 2020;27: 1168.e1–6.

70. Faron ML, Buchan BW, Relich RF, Clark J, Ledeboer NA. Evaluation of the WASPLab segregation software to automatically analyze urine cultures using routine blood and MacConkey agars. J Clin Microbiol 2020;58: e01683–19.

71. Yarbrough ML, Lainhart W, McMullen AR, Anderson NW, Burnham C-AD. Impact of total laboratory automation on workflow and specimen processing time for culture of urine specimens. Eur J Clin Microbiol Infect Dis 2018;37:2405–11.

72. Pancholi P, Carroll KC, Buchan BW, Chan RC, Dhiman N, Ford B, et al. Multicenter evaluation of the accelerate phenotest BC kit for rapid identification and phenotypic antimicrobial susceptibility testing using morphokinetic cellular analysis. J Clin Microbiol 2018;56: e01329–17.

73. Wang Z, Zhang L, Zhao M, Wang Y, Bai H, Wang Y, et al. Deep neural networks offer morphologic classification and diagnosis of bacterial vaginosis. J Clin Microbiol 2021;59(2):e02236–20.

74. Smith KP, Kang AD, Kirby JE. Automated interpretation of blood culture gram stains by use of a deep convolutional neural network. J Clin Microbiol 2018;56: e01521–17.

75. Panicker RO, Soman B, Saini G, Rajan J. A review of automatic methods based on image processing techniques for tuberculosis detection from microscopic sputum smear images. J Med Syst 2016;40(1):17.

76. Horvath L, Hänselmann S, Mannsperger H, Degenhardt S, Last K, Zimmermann S, et al. Machine-assisted interpretation of auramine stains substantially increases through-put and sensitivity of microscopic tuberculosis diagnosis. Tuberculosis (Edinb) 2020;125: 101993.

77. Yang M, Nurzynska K, Walts AE, Gertych A. A CNN-based active learning framework to identify mycobacteria in digitized Ziehl-Neelsen stained human tissues. Comput Med Imaging Graph 2020;84:101752.

78. Mathison BA, Kohan JL, Walker JF, Smith RB, Ardon O, Couturier MR. Detection of intestinal protozoa in trichrome-stained stool specimens by use of a deep convolutional neural network. J Clin Microbiol 2020; 58:e02053–19.

79. Rajaraman S, Antani SK, Poostchi M, Silamut K, Hossain MA, Maude RJ, et al. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. PeerJ 2018;6:e4568.

80. Uc-Cetina V, Brito-Loeza C, Ruiz-Piña H. Chagas parasite detection in blood images using AdaBoost. Comput Math Methods Med 2015;2015:139681.

81. Das DK, Ghosh M, Pal M, Maiti AK, Chakraborty C. Machine learning approach for automated screening of malaria parasite using light microscopic images. Micron 2013;45:97–106.

82. Smith KP, Wang H, Durant TJS, Mathison BA, Sharp SE, Kirby JE, et al. Applications of artificial intelligence in clinical microbiology diagnostic testing. Clin Microbiol Newsl 2020;42:61–70.

83. Kim H, Ganslandt T, Miethke T, Neumaier M, Kittel M. Deep learning frameworks for rapid gram stain image data interpretation: protocol for a retrospective data analysis. JMIR Res Protoc 2020;9:e16843.

84. Rhoads DD, Sintchenko V, Rauch CA, Pantanowitz L. Clinical microbiology informatics. Clin Microbiol Rev 2014;27:1025–47.

85. Buchan BW, Hoff JS, Gmehlin CG, Perez A, Faron ML, Munoz-Price LS, et al. Distribution of SARS-CoV-2 PCR cycle threshold values provide practical insight into overall and target-specific sensitivity among symptomatic patients. Am J Clin Pathol 2020;154:479–85.

86. Ellington MJ, Ekelund O, Aarestrup FM, Canton R, Doumith M, Giske C, et al. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. Clin Microbiol Infect 2017;23:2–22.

87. Ferreira I, Beisken S, Lueftinger L, Weinmaier T, Klein M, Bacher J, et al. Species identification and antibiotic resistance prediction by analysis of whole-genome sequence data by use of ARESdb: an analysis of isolates from the Unyvero Lower Respiratory Tract Infection Trial. J Clin Microbiol 2020;58:e00273–20.

88. Nguyen M, Long SW, McDermott PF, Olsen RJ, Olson R, Stevens RL, et al. Using machine learning to predict antimicrobial MICs and associated genomic features for Nontyphoidal Salmonella. J Clin Microbiol 2019; 57:e01260–18.

89. Davis JJ, Boisvert S, Brettin T, Kenyon RW, Mao C, Olson R, et al. Antimicrobial resistance prediction in PATRIC and RAST. Sci Rep 2016;6:27930.

90. Pesesky MW, Hussain T, Wallace M, Patel S, Andleeb S, Burnham C-AD, et al. Evaluation of machine learning and rules-based approaches for predicting antimicrobial resistance profiles in gram-negative bacilli from whole genome sequence data. Front Microbiol 2016; 7:1887.

91. Rahman SF, Olm MR, Morowitz MJ, Banfield JF. Machine learning leveraging genomes from metagenomes identifies influential antibiotic resistance genes in the infant gut microbiome. mSystems 2018;3: e00123–17.

92. Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, et al. Accurate multiplex polony sequencing of an evolved bacterial genome. Science 2005;309:1728–32.

93. Tolan NV, Parnas ML, Baudhuin LM, Cervinski MA, Chan AS, Holmes DT, et al. "Big data" in laboratory medicine. Clin Chem 2015;61:1433–40.

94. Telenti A. Machine learning to decode genomics. Clin Chem 2020;66:45–7.

95. Kearney E, Wojcik A, Babu D. Artificial intelligence in genetic services delivery: utopia or apocalypse? J Genet Couns 2020;29:8–17.

96. Technology | Fabric Genomics. 2021. https://fabricgenomics.com/products/technology/ (Accessed September 2021).

97. Oulas A, Minadakis G, Zachariou M, Spyrou GM. Selecting variants of unknown significance through network-based gene-association significantly improves risk prediction for disease-control cohorts. Sci Rep 2019;9:3266.

98. Schulz WL, Tormey CA, Torres R. Computational approach to annotating variants of unknown significance in clinical next generation sequencing. Lab Med 2015; 46:285–289.

99. Zimmerman L, Zelichov O, Aizenmann A, Barbash Z, Vidne M, Tarcic G. A novel system for functional determination of variants of uncertain significance using deep convolutional neural networks. Sci Rep 2020;10: 4192.

100. Lai C, Zimmer AD, O'Connor R, Kim S, Chan R, van den Akker J, et al. LEAP: using machine learning to support variant classification in a clinical setting. Hum Mutat 2020;41:1079–90.

101. Ho DSW, Schierding W, Wake M, Saffery R, O'Sullivan J. Machine learning SNP based prediction for precision medicine. Front Genet 2019;10:267.

102. Dudbridge F. Power and predictive accuracy of polygenic risk scores. PLoS Genet 2013;9:e1003348.

103. 23andMe offers new genetic report on type 2 diabetes. 23andMe Blog. 2019. https://blog.23andme.com/health-traits/type-2-diabetes/ (Accessed September 2021).

104. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. Genome Biol 2017;18:83.

105. Yang Z, LaRiviere MJ, Ko J, Till JE, Christensen T, Yee SS, et al. A multianalyte panel consisting of extracellular vesicle miRNAs and mRNAs, cfDNA, and CA19-9 shows utility for diagnosis and staging of pancreatic ductal adenocarcinoma. Clin Cancer Res 2020;26: 3248–58.

106. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med 2008;14:985–90.

107. Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP, et al. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med 2014; 370:1287–97.

108. Nicora G, Vitali F, Dagliati A, Geifman N, Bellazzi R. Integrated multi-omics analyses in Oncology: a review of machine learning methods and tools. Front Oncol 2020;10:1030.

109. Plebani M, Sciacovelli L, Aita A, Chiozza ML. Harmonization of pre-analytical quality indicators. Biochem Med (Zagreb) 2014;24:105–13.

110. Hawker CD, McCarthy W, Cleveland D, Messinger BL. Invention and validation of an automated camera system that uses optical character recognition to identify patient name mislabeled samples. Clin Chem 2014; 60:463–70.

111. Baron JM, Mermel CH, Lewandrowski KB, Dighe AS. Detection of preanalytic laboratory testing errors using a statistically guided protocol. Am J Clin Pathol 2012; 138:406–13.

112. Rosenbaum MW, Baron JM. Using Machine learning-based multianalyte delta checks to detect wrong blood in tube errors. Am J Clin Pathol 2018;150:555–66.

113. Rosenbaum MW, Flood JG, Melanson SEF, Baumann NA, Marzinke MA, Rai AJ, et al. Quality control practices for chemistry and immunochemistry in a cohort of 21 large academic medical centers. Am J Clin Pathol 2018;150:96–104.

114. Hoffmann RG, Waid ME. The "average of normals" method of quality control. Am J Clin Pathol 1965;43: 134–41.

115. Ng D, Polito FA, Cervinski MA. Optimization of a moving averages program using a simulated annealing

algorithm: the goal is to monitor the process not the patients. Clin Chem 2016;62:1361–71.

116. Demirci F, Akan P, Kume T, Sisman AR, Erbayraktar Z, Sevinc S. Artificial neural network approach in laboratory test reporting: learning algorithms. Am J Clin Pathol 2016;146:227–37.

117. Horn PS, Pesce AJ. Reference intervals: an update. Clin Chim Acta 2003;334:5–23.

118. Horowitz GL. The power of asterisks. Clin Chem 2015; 61:1009–11.

119. Zierk J, Arzideh F, Kapsner LA, Prokosch H-U, Metzler M, Rauh M. Reference interval estimation from mixed distributions using truncation points and the Kolmogorov-Smirnov distance (KOSMIC). Sci Rep 2020;10:1704.

120. Holmes DT, Buhr KA. Widespread incorrect implementation of the Hoffmann method, the correct approach, and modern alternatives. Am J Clin Pathol 2019;151:328–36.

121. Poole S, Schroeder LF, Shah N. An unsupervised learning method to identify reference intervals from a clinical database. J Biomed Inform 2016;59:276–84.

122. Mitterecker A, Hofmann A, Trentino KM, Lloyd A, Leahy MF, Schwarzbauer K, et al. Machine learning-based prediction of transfusion. Transfusion 2020;60:1977–86.

123. Sloane EB, J. Silva R. Artificial intelligence in medical devices and clinical decision support systems. In: Ernesto Iadanza, editor. Clinical engineering handbook. 2nd Ed. Elsevier; 2020. p. 556–68.

124. Gameiro J, Branco T, Lopes JA. Artificial intelligence in acute kidney injury risk prediction. J Clin Med 2020; 9(3):678.

125. Huang C, Li S-X, Mahajan S, Testani JM, Wilson FP, Mena CI, et al. Development and validation of a model for predicting the risk of acute kidney injury associated with contrast volume levels during percutaneous coronary intervention. JAMA Netw Open 2019;2: e1916021.

126. Sandokji I, Yamamoto Y, Biswas A, Arora T, Ugwuowo U, Simonov M, et al. A time-updated, parsimonious model to predict AKI in hospitalized children. J Am Soc Nephrol 2020;31:1348–57.

127. Choi GH, Yun J, Choi J, Lee D, Shim JH, Lee HC, et al. Development of machine learning-based clinical decision support system for hepatocellular carcinoma. Sci Rep 2020;10:14855.

128. Finlay GD, Rothman MJ, Smith RA. Measuring the modified early warning score and the Rothman index: advantages of utilizing the electronic medical record in an early warning system. J Hosp Med 2014;9:116–9.

129. Haimovich AD, Ravindra NG, Stoytchev S, Young HP, Wilson FP, van Dijk D, et al. Development and validation of the quick COVID-19 severity index: a prognostic tool for early clinical decompensation. Ann Emerg Med 2020;76:442–53.

130. Durant TJS, Jean R, Huang C, Coppi A, Schulz WL, Geirsson A, et al. Evaluation of a risk stratification model using preoperative and intraoperative data for major morbidity or mortality after cardiac surgical treatment. JAMA Netw Open 2020;3:e2028361.

131. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1–coronary artery bypass grafting surgery. Ann Thorac Surg 2009; 88:S2–22.

132. Collins GS, Moons KGM. Reporting of artificial intelligence prediction models. Lancet 2019;393:1577–9.

133. Food and Drug Administration. Proposed regulatory framework for modifications to artificial intelligence/ machine learning (AI/ML)-based software as a medical device (SaMD). 2021. https://www.fda.gov/media/122535/download (Accessed September 2021).

134. Schulz WL, Durant TJS, Krumholz HM. Validation and regulation of clinical artificial intelligence. Clin Chem 2019;65:1336–1.

135. Parikh RB, Obermeyer Z, Navathe AS. Regulation of predictive analytics in medicine. Science 2019;363: 810–2.

136. Wilson FP, Martin M, Yamamoto Y, Partridge C, Moreira E, Arora T, et al. Electronic health record alerts for acute kidney injury: multicenter, randomized clinical trial. BMJ 2021;372:m4786.

137. Waltz TJ, Powell BJ, Fernández ME, Abadie B, Damschroder LJ. Choosing implementation strategies to address contextual barriers: diversity in recommendations and future directions. Implement Sci 2019;14: 42.

138. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. JAMA 2017;318:517–8.

139. Baggerly KA, Coombes KR. Deriving chemosensitivity from cell lines: forensic bioinformatics and reproducible research in high-throughput biology. Ann Appl Stat 2009;3:1309–34.

140. Food and Drug Administration. Clinical Decision Support Software: Draft Guidance for Industry and Food and Drug Administration Staff. 2019. https://www.fda.gov/media/109618/download (Accessed September 2021).

141. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. NPJ Digit Med 2020;3:118.

142. El Emam K, Jonker E, Arbuckle L, Malin B. A systematic review of re-identification attacks on health data. PLoS One 2011;6:e28071.

143. Benitez K, Malin B. Evaluating re-identification risks with respect to the HIPAA privacy rule. J Am Med Inform Assoc 2010;17:169–77.

144. Grady C, Eckstein L, Berkman B, Brock D, Cook-Deegan R, Fullerton SM, et al. Broad consent for research with biological samples: workshop conclusions. Am J Bioeth 2015;15:34–42.

145. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. Stud Health Technol Inform 2015;216:574–8.

146. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. J Am Med Inform Assoc 2014;21:578–82.

147. Murphy SN, Mendis ME, Berkowitz DA, Kohane I, Chueh HC. Integration of clinical and genetic data in the i2b2 architecture. AMIA Annu Symp Proc 2006; 2006:1040.

148. Shapiro JS, Mostashari F, Hripcsak G, Soulakis N, Kuperman G. Using health information exchange to improve public health. Am J Public Health 2011;101: 616–23.

149. Johnson AEW, Pollard TJ, Shen L, Lehman L-WH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016;3:160035.

150. Kuperman GJ. Health-information exchange: why are we doing it, and what are we doing? J Am Med Inform Assoc 2011;18:678–82.

151. Greene DN, McClintock DS, Durant TJS. Interoperability: COVID-19 as an impetus for change. Clin Chem 2021;67:592–5.

152. McDonald CJ, Hammond WE. Standard formats for electronic transfer of clinical data. Ann Intern Med 1989;110:333–5.

153. Makadia R, Ryan PB. Transforming the premier perspective hospital database into the Observational Medical Outcomes Partnership (OMOP) common data model. EGEMS (Wash DC) 2014;2:1110.

154. Bender D, Sartipi K. HL7 FHIR: an agile and RESTful approach to healthcare information exchange. In: Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems. Piscataway (NJ): IEEE; 2013. p. 326–331.

155. 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.

156. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366:447–53.

157. Eneanya ND, Yang W, Reese PP. Reconsidering the consequences of using race to estimate kidney function. JAMA 2019;322:113–4.

158. Ahmed S, Nutt CT, Eneanya ND, Reese PP, Sivashanker K, Morse M, et al. Examining the potential impact of race multiplier utilization in estimated glomerular filtration rate calculation on African-American care outcomes. J Gen Intern Med 2021;36:464–71.

159. Chen H, Li Z, Zhang L, Sawaya P, Shi J, Wang P. Quantitation of femtomolar-level protein biomarkers using a simple microbubbling digital assay and bright-field smartphone imaging. Angew Chem 2019;131: 14060–6.

160. Master CL, Podolak OE, Ciuffreda KJ, Metzger KB, Joshi NR, McDonald CC, et al. Utility of pupillary light reflex metrics as a physiologic biomarker for adolescent sport-related concussion. JAMA Ophthalmol 2020; 138:1135–41.

161. Tracy JM, Özkanca Y, Atkins DC, Hosseini Ghomi R. Investigating voice as a biomarker: deep phenotyping methods for early detection of Parkinson's disease. J Biomed Inform 2020;104:103362.

162. Abdel-Aziz MI, Brinkman P, Vijverberg SJH, Neerincx AH, de Vries R, Dagelet YWF, et al.; Amsterdam UMC Breath Research Group. eNose breath prints as a surrogate biomarker for classifying patients with asthma by atopy. J Allergy Clin Immunol 2020;146:1045–55.

163. Louis DN, Gerber GK, Baron JM, Bry L, Dighe AS, Getz G, et al. Computational pathology: an emerging definition. Arch Pathol Lab Med 2014;138:1133–38.

164. Luo Y, Szolovits P, Dighe AS, Baron JM. Using machine learning to predict laboratory test results. Am J Clin Pathol 2016;145:778–88.

165. Luo Y, Szolovits P, Dighe AS, Baron JM. 3D-MICE: integration of cross-sectional and longitudinal imputation for multi-analyte longitudinal clinical data. J Am Med Inform Assoc 2018;25:645–53.

166. Ardon O, Schmidt RL. Clinical laboratory employees' attitudes toward artificial intelligence. Lab Med 2020; 51:649–54.

167. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44–56.

168. Henricks WH, Wilkerson ML, Castellani WJ, Whitsitt MS, Sinard JH. Pathologists as stewards of laboratory information. Arch Pathol Lab Med 2015;139:332–7.

169. Crawford JM, Shotorbani K, Sharma G, Crossey M, Kothari T, Lorey TS, et al. Improving American healthcare through "Clinical Lab 2.0": a Project Santa Fe report. Acad Pathol 2017;4:237428951770106.