Clinical applications of machine learning in pre-analytical, analytical and post-analytical phases of laboratory medicine: a narrative review

Lei Zhang, Zhi-De Hu

Department of Laboratory Medicine, the Affiliated Hospital of Inner Mongolia Medical University, Hohhot, China

Contributions: (I) Conception and design: ZD Hu; (II) Administrative support: Both authors; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: Both authors; (V) Data analysis and interpretation: L Zhang; (VI) Manuscript writing: Both authors; (VII) Final approval of manuscript: Both authors.

Correspondence to: Dr. Zhi-De Hu. Department of Laboratory Medicine, the Affiliated Hospital of Inner Mongolia Medical University, Hohhot, China. Email: hzdlj81@163.com.

Background and Objective: In an era of big data, various machine learning (ML) algorithms have increasingly been applied in clinical medicine. In particular, ML algorithms have greatly impacted the current concepts and developmental models of laboratory medicine. This narrative review aimed to address the clinical applications of ML algorithms in the pre-analytical, analytical, and post-analytical phases of laboratory medicine.

Methods: PubMed database were searched to identify studies published between the years 2000/1/1–2022/8/1. Articles investigating the clinical applications of ML algorithms in laboratory medicine were included. With typical examples, the clinical applications of ML algorithms were summarized in the pre-analytical, analytical, and post-analytical phases of laboratory medicine.

Key Content and Findings: In the pre-analytical phase, ML algorithms can be used to reduce the rate of pre-analytical error and verify specimens with low quality. In the analytical phase, ML algorithms can be used to reduce laboratory costs and optimize laboratory work procedures. In the post-analytical phase, ML algorithms can be used to integrate the existing test results to guide the diagnosis and treatment of diseases.

Conclusions: ML is a strong driver for the development of laboratory medicine.

Keywords: Machine learning (ML); laboratory medicine; artificial intelligence (AI); big data

Received: 06 November 2022; Accepted: 19 December 2022; Published: 30 December 2022.

doi: 10.21037/amj-22-92

View this article at: https://dx.doi.org/10.21037/amj-22-92

Introduction

Machine learning (ML) is a branch of artificial intelligence (AI) that aims to find general rules in complex data through pre-set algorithms and apply these rules to new data for classification and prediction (1). In recent years, thanks to the rapid advancement of computer software and hardware and the vigorous development of the internet, massive biomedical data can be obtained within a short period, which paves the way for AI applications in modern medical sciences (2). Numerous AI methods, represented by ML algorithms, are gradually changing modern medical models.

As an important part of modern medicine, laboratory medicine explores the mechanisms underlying the occurrence and development of diseases through laboratory testing, thus providing a scientific basis for risk assessment, diagnosis, stratification, prognosis assessment, and treatment monitoring (3). In general, a laboratory testing process is divided into three phases: pre-analytical, analytical, and post-analytical. The pre-analytical phase involves the selection of proper laboratory tests and the collection and transport of qualified specimens, during which the influence of specimen quality on laboratory tests should be avoided (4). In the analytical phase, the laboratory test procedure should

Table 1 The search strategy summary

Items	Specification
Date of search	2022/8/1
Databases and other sources searched	PubMed
Search terms used	"Machine learning", "laboratory medicine", "biomarker", "laboratory test"
Timeframe	2000/1/1–2022/8/1
Inclusion and exclusion criteria	None. The searched papers that provide new aspects of laboratory medicine and ML were read
Selection process	The authors read the articles together

ML, machine learning

be continuously optimized to ensure that the test results are timely and accurate; meanwhile, the cost of laboratory tests should be continuously reduced, to meet the clinical needs for disease diagnosis and treatment with the lowest resource consumption (5). The post-analytical phase requires a scientific and reasonable interpretation of the clinical relevance of the test results, to provide patients with better medical care (6).

In recent years, ML algorithms have greatly reshaped the landscape of laboratory medicine (7). Accumulating studies indicated that ML algorithms can be used to reduce laboratory costs and errors, and improve laboratory quality management. Here we summarize the application of ML in laboratory medicine by giving examples. We present the following article in accordance with the Narrative Review reporting checklist (available at https://amj.amegroups.com/article/view/10.21037/amj-22-92/rc).

Methods

The PubMed database was searched to identify studies published between the years 2000/1/1–2022/8/1 using the search terms "machine leaning", "laboratory medicine", "biomarker", and "laboratory test". A manual search was also performed using the references of the review articles retrieved and primary research. With no further inclusion or exclusion criteria, the searched papers that provide new aspects of laboratory medicine and ML were read. Two authors drafted the manuscript together with the typical examples in this field. *Table 1* lists the summary of the search strategy.

Key content and findings

Application of ML in the pre-analytical phase

As mentioned earlier, the purpose of the pre-analytical phase is to ensure specimen quality and minimize errors. The advances in laboratory testing methodologies have dramatically lowered the incidence of errors in analysis, and most errors in the testing process are seen in the pre-analytical phase (8), which may include misidentification, inappropriate container, insufficient volume, and clotting of an anticoagulated specimen (9).

Misidentification is a common error in the pre-analytical phase. In clinical practice, misidentification is recognized mainly by delta check (i.e., by comparing historical records) (10), which, however, is mainly based on human judgment and lacks uniform objective criteria. Different laboratory technicians in different laboratories may have different understandings of the delta check, resulting in large diversities in recognizing misidentification among different laboratories and individuals. In addition, manual judgment is time-consuming, which is not conducive to saving laboratory resources. Therefore, several studies have explored the value of ML in recognizing misidentification (11-14). In most of these studies, specific laboratory test data were first downloaded from the laboratory information system (LIS), and then the data that could be used for analysis (e.g., patients who have received duplicated testing within seven days.) were screened by using inclusion and exclusion criteria. Computer software was then used to randomly create misidentification in half of the specimens, and the accuracy in recognizing artificial misidentification was compared between ML algorithms and human judgments. All of these studies found that ML algorithms were much more accurate than human judgments (11-14). In one study, researchers used ML algorithms to analyze misidentification in electrolytes and renal function tests and found that the accuracy of manual identification was only about 77.8%. In contrast, even the simplest ML algorithm, the decision tree, achieved an accuracy of 86.5%, and the accuracy of the artificial neural network even reached

92.1% (14). More importantly, the accuracy of recognizing misidentification can be significantly improved if the ML results are presented to lab technicians to alert them to the risk of misidentification (15). Thus, the accuracy of ML alone in recognizing misidentification is much higher than that of manual identification, and the accuracy can be further improved if the ML results are presented to laboratory staff for comprehensive judgment.

Hemolysis, icterus, and lipemia (HIL) of blood samples are common pre-analytical errors that pose large challenges to laboratory tests (16,17). Traditionally, HIL is mainly observed by the visual inspection, which is time-consuming and can be affected by subjective factors, leading to low accuracy in clinical practice. Some newly-developed biochemical instruments can detect the HIL status of the specimen and describe the status of the specimen by using indicators such as the hemolysis index (H-index), icterus index (I-index), and/or lipemia index (L-index) (18,19). However, approximately 10 minutes are required for the biochemical instrument to describe the specimen status, which will affect the efficiency of the biochemical instrument and even the laboratory turnaround time. A recent study used deep learning to analyze sample images to determine whether HIL existed. It was found that all areas under the receiver operating characteristic curve (AUCs) of deep learning in recognizing HIL were above 0.98, showing significantly higher accuracy than biochemical instruments (20). Therefore, deep learning can dramatically increase the accuracy in identifying low-quality serum samples (20).

In addition to recognizing misidentification and low-quality samples, ML can also be used for identifying the clotting of specimens. In coagulation tests, the clotting of the samples will affect the accuracy of the test results. In clinical practice, the clotting of specimens is mainly judged by visual inspection, which, however, is not able to identify small clots in some coagulated blood specimens. Since clotting can cause changes in the results of a coagulation test, the likelihood of clotting can be predicted based on the results of the coagulation test. A recent real-world study used backpropagation (BP) neural networks to determine the likelihood of clotting in a blood sample (21). The results showed that the BP neural network method based on the coagulation test results was extremely accurate in predicting blood clotting, and the AUC reached 0.97.

Application of ML in the analytical phase

The analysis phase includes the entire process from the entrance of a specific sample into the laboratory to the reporting of the test results. In this process, ML can optimize laboratory work procedures, reduce laboratory costs, and increase laboratory efficiency. ML algorithms serve different purposes for different laboratory tests or test panels. Here, we illustrate the applications of ML algorithms in different clinical settings.

Since low-density lipoprotein cholesterol (LDL-C) is a key risk factor and therapeutic target for cardio-cerebrovascular diseases (CVDs), LDL-C testing is of great value for the prevention and treatment of CVDs. The reference method for LDL-C testing is beta quantification following ultracentrifugation, which, however, is time-consuming and labor-intensive and requires very expensive instrumentation, making it unsuitable for routine testing. Early in 1972, Friedewald discovered that LDL-C concentration was related to the concentrations of high-density lipoprotein cholesterol (HDL-C), total cholesterol (TC), and triglycerides (TG) and invented an LDL-C calculation formula, the famous Friedewald formula (22):

$$LDL\text{-}C = TC - HDL\text{-}C - \frac{TG}{5} \quad [1]$$

in which the difference between TC and HDL-C is also known as non-HDL-C. Many laboratories use the Friedewald formula to calculate the concentration of LDL-C, rather than directly testing it. Although the Friedewald formula has been widely used, it has some limitations. In particular, the prediction accuracy of the formula decreases as the TG concentration increases. This is mainly because the Friedewald formula assumes the triglyceride/cholesterol ratio in very-low-density lipoprotein (VLDL) to be 5:1. The mathematical basis for this assumption is linear regression, which does not take into account that the triglyceride/cholesterol ratio in VLDL is affected by a variety of factors. Unlike conventional linear regression, ML algorithms are more flexible and do not presuppose a linear relationship between the dependent and independent variables. For example, the random forest (RF) algorithm, in essence, is to build multiple decision trees through the training dataset, operate using these decision trees in the testing dataset, and calculate the probability of classification according to the operation results of multiple decision trees. Therefore, ML algorithms may be more advantageous in predicting LDL-C. So far, several studies have evaluated the accuracy of ML algorithms in predicting LDL-C, and all of these algorithms were based on TC, TG, and HDL-C (23-29). These studies have found that ML

algorithms had higher accuracies than Friedewald's formula and even the Martin formula, which was proposed more recently (30). ML algorithms are also quite accurate in individuals with higher and lower LDL-C concentrations. Notably, ML algorithms can be directly incorporated into the LIS and are easy to use.

The liver enzymes test is an important part of the liver function test. The common liver enzymes tested include aspartate aminotransferase (AST), alanine aminotransferase (ALT), alkaline phosphatase (AKP), and γ-glutamyl transferase (GGT). Although the clinical values of these enzymes are specific, they may overlap each other to some extent. Therefore, some enzymatic tests may be redundant from the perspective of saving laboratory testing costs. One study proposed that ALT and AKP results could be used to predict GGT measurements (31). Using ML algorithms, the researchers found that the ALT and AKP decision trees had an accuracy of up to 90% in predicting GGT. In other words, tests for GGT in 90% of liver function tests are not needed because GGT can be accurately predicted by ALT and AKP measurements. One of the most important roles of ML in the analytic phase is to use low-cost laboratory tests to predict high-cost laboratory tests. In addition to GGT, the level of ferritin can also be predicted based on the results of routine blood tests (32,33).

In addition to the prediction of laboratory results, ML has been widely used in auto-verification (34), establishing the rules for urine sediment examination (35), morphologic classification of erythrocytes (36), and data analyses in metabolomics (37).

Application of ML in the post-analytical phase

The mission of laboratory medicine in the post-analytical phase is to translate the test results into effective clinical information and provide scientific evidence for the diagnosis and evaluation of diseases. The role of ML in this process is to integrate the existing test results to guide the diagnosis and treatment of diseases. Here we use two samples to illustrate how to use ML algorithms to study the clinical value of laboratory tests.

Pleural fluid biochemistry is an important approach for diagnosing tuberculosis pleurisy. In particular, adenosine deaminase (ADA) has a diagnostic accuracy of about 90% for this disease (38). Other biomarkers in the pleural fluid, including lactate dehydrogenase (LDH) and leukocyte count, also have certain diagnostic values for tuberculous pleurisy. Therefore, the clarification of whether biomarkers (e.g., LDH) in pleural fluid can improve the diagnostic accuracy of ADA is necessary. In other words, do the combinations of multiple biomarkers (including ADA) have higher diagnostic performance than ADA alone? A study published in 2019 used ML algorithms such as support vector machine (SVM) and RF to explore the diagnostic value of the combination of these pleural fluid markers for tuberculous pleurisy; the AUC of ADA was found to be only 0.89 but reached 0.97 with the application of RF algorithm (39). Therefore, although ADA has a high diagnostic value for tuberculous pleural effusion (TPE), it can achieve higher diagnostic accuracy if it is used in combination with other biomarkers by using ML algorithms.

Assessing the prognosis of diabetic nephropathy is the basis for developing individualized treatment protocols and thus improving patient outcomes. At present, many markers and scoring systems can be used to predict the progression of diabetic nephropathy, with the most widely-used system being the chronic kidney disease classification system released by the Kidney Disease Improving Global Outcomes (KDIGO). However, the accuracy of this system in predicting the prognosis of chronic diabetic nephropathy is far from satisfactory. Therefore, new prognostic factors for diabetic nephropathy are urgently needed. A cohort study published in 2021 used the RF algorithm combined with multiple biomarkers (KIM-1, TNFR1, and TNFR2) to predict the prognosis of patients with diabetic nephropathy and found that the AUC of the RF algorithm was 0.77, whereas the AUC of the KDIGO grading system was only 0.62 (40). Therefore, ML algorithms have more advantages in predicting the prognosis of diabetic nephropathy.

Furthermore, ML algorithms are widely used in the screening of Downs syndrome (41) and the diagnosis of malignant pleural mesothelioma (42).

Conclusions

The past few years have witnessed the wider application of various ML algorithms in laboratory medicine. These advanced ML algorithms have brought more insights and addressed a variety of problems in this field. This article introduces the applications of ML in laboratory medicine by giving some typical examples, aiming to refresh our knowledge in this emerging interdisciplinary field. Laboratory technicians are encouraged to master this new technology and apply it in clinical practice, thus promoting the development of laboratory medicine. It is foreseeable that, with the optimization of ML algorithms and the advances in computer software and hardware performance,

ML will become a strong driver for the development of laboratory medicine.

Acknowledgments

Funding: This work was supported by the Foundation from the Commission of Health of Inner Mongolia (202201260).

Footnote

Reporting Checklist: The authors have completed the Narrative Review reporting checklist. Available at https://amj.amegroups.com/article/view/10.21037/amj-22-92/rc

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at https://amj.amegroups.com/article/view/10.21037/amj-22-92/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med 2019;380:1347-58.
2. He J, Baxter SL, Xu J, et al. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019;25:30-6.
3. Lippi G, Plebani M. A modern and pragmatic definition of Laboratory Medicine. Clin Chem Lab Med 2020;58:1171.
4. Cornes M. The preanalytical phase - Past, present and future. Ann Clin Biochem 2020;57:4-6.
5. Schmidt RL, Pearson LN. Estimating the cost of quality of errors in the analytical phase. Clin Chim Acta 2019;495:60-6.
6. Sciacovelli L, Aita A, Padoan A, et al. Performance criteria and quality indicators for the post-analytical phase. Clin Chem Lab Med 2016;54:1169-76.
7. Rabbani N, Kim GYE, Suarez CJ, et al. Applications of machine learning in routine laboratory medicine: Current state and future directions. Clin Biochem 2022;103:1-7.
8. Lippi G, Betsou F, Cadamuro J, et al. Preanalytical challenges - time for solutions. Clin Chem Lab Med 2019;57:974-81.
9. Cornes MP, Atherton J, Pourmahram G, et al. Monitoring and reporting of preanalytical errors in laboratory medicine: the UK situation. Ann Clin Biochem 2016;53:279-84.
10. Markus C, Tan RZ, Loh TP. Evidence-based approach to setting delta check rules. Crit Rev Clin Lab Sci 2021;58:49-59.
11. Zhou R, Liang YF, Cheng HL, et al. A highly accurate delta check method using deep learning for detection of sample mix-up in the clinical laboratory. Clin Chem Lab Med 2022;60:1984-92.
12. 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.
13. Farrell CL, Giannoutsos J. Machine learning models outperform manual result review for the identification of wrong blood in tube errors in complete blood count results. Int J Lab Hematol 2022;44:497-503.
14. Farrell CJ. Identifying mislabelled samples: Machine learning models exceed human performance. Ann Clin Biochem 2021;58:650-2.
15. Farrell CL. Decision support or autonomous artificial intelligence? The case of wrong blood in tube errors. Clin Chem Lab Med 2022;60:1993-7.
16. Simundic AM, Baird G, Cadamuro J, et al. Managing hemolyzed samples in clinical laboratories. Crit Rev Clin Lab Sci 2020;57:1-21.
17. Lippi G, von Meyer A, Cadamuro J, et al. Blood sample quality. Diagnosis (Berl) 2019;6:25-31.
18. Mondejar R, Mayor Reyes M, Melguizo Madrid E, et al. Utility of icteric index in clinical laboratories: more than a preanalytical indicator. Biochem Med (Zagreb) 2021;31:020703.
19. Cao Y, Branzell I, Vink M. Determination of clinically acceptable cut-offs for hemolysis index: An application of bootstrap method using real-world data. Clin Biochem 2021;94:74-9.
20. Yang C, Li D, Sun D, et al. A deep learning-based system for assessment of serum quality using sample images. Clin

Chim Acta 2022;531:254-60.
21. Fang K, Dong Z, Chen X, et al. Using machine learning to identify clotted specimens in coagulation testing. Clin Chem Lab Med 2021;59:1289-97.
22. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18:499-502.
23. P P A, Kumari S, Rajasimman AS, et al. Machine learning predictive models of LDL-C in the population of eastern India and its comparison with directly measured and calculated LDL-C. Ann Clin Biochem 2022;59:76-86.
24. Barakett-Hamade V, Ghayad JP, Mchantaf G, et al. Is Machine Learning-derived Low-Density Lipoprotein Cholesterol estimation more reliable than standard closed form equations? Insights from a laboratory database by comparison with a direct homogeneous assay. Clin Chim Acta 2021;519:220-6.
25. Singh G, Hussain Y, Xu Z, et al. Comparing a novel machine learning method to the Friedewald formula and Martin-Hopkins equation for low-density lipoprotein estimation. PLoS One 2020;15:e0239934.
26. Oh GC, Ko T, Kim JH, et al. Estimation of low-density lipoprotein cholesterol levels using machine learning. Int J Cardiol 2022;352:144-9.
27. Çubukçu HC, Topcu Dİ. Estimation of Low-Density Lipoprotein Cholesterol Concentration Using Machine Learning. Lab Med 2022;53:161-71.
28. Kwon YJ, Lee H, Baik SJ, et al. Comparison of a Machine Learning Method and Various Equations for Estimating Low-Density Lipoprotein Cholesterol in Korean Populations. Front Cardiovasc Med 2022;9:824574.
29. Tsigalou C, Panopoulou M, Papadopoulos C, et al. Estimation of low-density lipoprotein cholesterol by machine learning methods. Clin Chim Acta 2021;517:108-16.
30. Martin SS, Blaha MJ, Elshazly MB, et al. Comparison of a novel method vs the Friedewald equation for estimating low-density lipoprotein cholesterol levels from the standard lipid profile. JAMA 2013;310:2061-8.
31. Lidbury BA, Richardson AM, Badrick T. Assessment of machine-learning techniques on large pathology data sets to address assay redundancy in routine liver function test profiles. Diagnosis (Berl) 2015;2:41-51.
32. Kurstjens S, de Bel T, van der Horst A, et al. Automated prediction of low ferritin concentrations using a machine learning algorithm. Clin Chem Lab Med 2022;60:1921-8.
33. Luo Y, Szolovits P, Dighe AS, et al. Using Machine Learning to Predict Laboratory Test Results. Am J Clin Pathol 2016;145:778-88.
34. Wang H, Wang H, Zhang J, et al. Using machine learning to develop an autoverification system in a clinical biochemistry laboratory. Clin Chem Lab Med 2021;59:883-91.
35. Cao Y, Cheng M, Hu C. UrineCART, a machine learning method for establishment of review rules based on UF-1000i flow cytometry and dipstick or reflectance photometer. Clin Chem Lab Med 2012;50:2155-61.
36. Durant TJS, Olson EM, Schulz WL, et al. Very Deep Convolutional Neural Networks for Morphologic Classification of Erythrocytes. Clin Chem 2017;63:1847-55.
37. Streun GL, Steuer AE, Poetzsch SN, et al. Towards a New Qualitative Screening Assay for Synthetic Cannabinoids Using Metabolomics and Machine Learning. Clin Chem 2022;68:848-55.
38. Zhang M, Li D, Hu ZD, et al. The diagnostic utility of pleural markers for tuberculosis pleural effusion. Ann Transl Med 2020;8:607.
39. Ren Z, Hu Y, Xu L. Identifying tuberculous pleural effusion using artificial intelligence machine learning algorithms. Respir Res 2019;20:220.
40. Chan L, Nadkarni GN, Fleming F, et al. Derivation and validation of a machine learning risk score using biomarker and electronic patient data to predict progression of diabetic kidney disease. Diabetologia 2021;64:1504-15.
41. He F, Lin B, Mou K, et al. A machine learning model for the prediction of down syndrome in second trimester antenatal screening. Clin Chim Acta 2021;521:206-11.
42. Niu Y, Hu ZD. Diagnostic accuracy of pleural effusion biomarkers for malignant pleural mesothelioma: a machine learning analysis. J Lab Precis Med 2021;6:4.

doi: 10.21037/amj-22-92
Cite this article as: Zhang L, Hu ZD. Clinical applications of machine learning in pre-analytical, analytical and post-analytical phases of laboratory medicine: a narrative review. AME Med J 2022;7:37.