J Lipid Atheroscler. 2024 May;13(2):111-121
https://doi.org/10.12997/jla.2024.13.2.111
pISSN 2287-2892·eISSN 2288-2561

JLA Journal of Lipid and Atherosclerosis

Review

Check for updates

Applicability of Artificial Intelligence in the Field of Clinical Lipidology: A Narrative Review

Walter Masson,[1] **Pablo Corral**,[2] **Juan P Nogueira**,[3,4] **Augusto Lavalle-Cobo** [5]

[1]Department of Cardiology, Hospital Italiano de Buenos Aires, Buenos Aires, Argentina
[2]Faculty of Medicine, FASTA University, Mar del Plata, Argentina
[3]Endocrinology, Nutrition and Metabolism Research Center, Faculty of Health Sciences, Universidad Nacional de Formosa, Argentina
[4]International University of the Americas, San José, Costa Rica
[5]Department of Cardiology, Sanatorio Finochietto, Buenos Aires, Argentina

OPEN ACCESS

Received: Aug 9, 2023
Revised: Oct 10, 2023
Accepted: Dec 19, 2023
Published online: Feb 27, 2024

Correspondence to
Walter Masson
Department of Cardiology, Hospital Italiano de Buenos Aires, Perón 4190. Ciudad Autónoma de Buenos Aires, C1199ABB, Buenos Aires, Argentina.
Email: walter.masson@hospitalitaliano.org.ar

© 2024 The Korean Society of Lipid and Atherosclerosis.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs
Walter Masson
https://orcid.org/0000-0002-5620-6468
Pablo Corral
https://orcid.org/0000-0003-0017-8725
Juan P Nogueira
https://orcid.org/0000-0002-8764-4700
Augusto Lavalle-Cobo
https://orcid.org/0000-0002-1257-9211

Funding
None.

Conflict of Interest
Walter Masson, Pablo Corral, Juan P. Nogueira, Augusto Lavalle-Cobo are editors of

ABSTRACT

The development of advanced technologies in artificial intelligence (AI) has expanded its applications across various fields. Machine learning (ML), a subcategory of AI, enables computers to recognize patterns within extensive datasets. Furthermore, deep learning, a specialized form of ML, processes inputs through neural network architectures inspired by biological processes. The field of clinical lipidology has experienced significant growth over the past few years, and recently, it has begun to intersect with AI. Consequently, the purpose of this narrative review is to examine the applications of AI in clinical lipidology. This review evaluates various publications concerning the diagnosis of familial hypercholesterolemia, estimation of low-density lipoprotein cholesterol (LDL-C) levels, prediction of lipid goal attainment, challenges associated with statin use, and the influence of cardiometabolic and dietary factors on the discordance between apolipoprotein B and LDL-C. Given the concerns surrounding AI techniques, such as ethical dilemmas, opacity, limited reproducibility, and methodological constraints, it is prudent to establish a framework that enables the medical community to accurately interpret and utilize these emerging technological tools.

Keywords: Artificial intelligence; Deep learning; Dyslipidemias; Lipids; Machine learning

INTRODUCTION

Since epidemiological studies first established total cholesterol and low-density lipoprotein cholesterol (LDL-C) as causative factors in the development of atherosclerotic vascular disease, there has been an exponential increase in research in the field of lipidology.[1]

In recent years, artificial intelligence (AI) technologies have been transforming medicine and healthcare.[2] AI refers to the simulation of human intelligence processes by machines, particularly computer systems. Machine learning (ML), a subcategory of AI, enables computers to identify patterns in vast datasets and formulate responses to various questions across multiple disciplines through the use of algorithms. Furthermore, deep learning (DL) is a subset of ML distinguished by its use of biologically inspired neural network architectures to process inputs.

Generated by ximilinkpress

Journal of Lipid and Atherosclerosis. However, they were not involved in the peer reviewer selection, evaluation, or decision process of this article. No other potential conflicts of interest relevant to this article were reported.

Data Availability Statement

All data generated or analyzed during this study are included in this published article.

Author Contributions

Conceptualization: Masson W; Formal analysis: Masson W, Corral P, Nogueira JP; Project administration: Masson W; Supervision: Masson W, Corral P, Lavalle-Cobo A; Validation: Corral P, Nogueira JP, Lavalle-Cobo A; Writing - original draft: Masson W, Nogueira JP; Writing - review & editing: Masson W, Nogueira JP.

The advanced technologies developed for AI have fueled its applications in many areas, such as security, business administration, finance, commerce, tourism, meteorology, and biology, among others.[3] Similarly, AI's applications in the health sciences, especially in medicine, are extensive and span nearly all domains, such as genomics, epidemiology, diagnostics, prognosis, clinician workflow, telemedicine, and administrative management.[4] Furthermore, the remarkable expansion of clinical lipidology in recent years has seen numerous intersections with AI. The pursuit of precision medicine within certain contexts of contemporary lipidology may benefit from these innovative techniques.

Therefore, the aim of this narrative review was to explore some of the applications of AI in the field of clinical lipidology.

MATERIALS AND METHODS

A literature search was conducted to identify studies that established connections between various aspects of lipidology and AI. Two independent reviewers systematically searched the electronic databases of PubMed/MEDLINE, Embase, Science Direct, Scopus, and SciELO. The search terms used included “artificial intelligence,” “deep learning,” and “machine learning,” which were combined with the following lipid-related terms: “cholesterol,” “dyslipidemia,” “low-density lipoprotein cholesterol,” “high-density lipoprotein cholesterol,” “triglycerides,” “hypercholesterolemia,” “lipid-lowering treatment,” and “statins.” The search for relevant articles concluded on July 31, 2023.

Studies that analyzed the applicability of AI techniques for different diagnostic, prognostic, or therapeutic aspects of lipidology were included. There were no restrictions based on language, geography, or publication status. Studies that were excluded consisted of expert opinions, reviews, and those with a patient population of fewer than 100.

1. Ethical approval

This article is based on previously conducted studies and does not contain any studies with human participants or animals performed by any of the authors.

DIAGNOSIS OF FAMILIAL HYPERCHOLESTEROLEMIA (FH)

FH is the most common genetic disorder of lipid metabolism. FH results in life-long exposure to high LDL-C levels, which, if left untreated, significantly increase the risk of cardiovascular events.[5] A clinical diagnosis of FH is typically made by documenting characteristic clinical features alongside significantly raised LDL-C levels. Additionally, a genetic diagnosis can be confirmed through the detection of heterozygous or biallelic pathogenic variants, primarily in the *LDLR, APOB*, or *PCSK9* genes.[6] Nevertheless, AI techniques may have the potential to improve FH identification.

Using only a basic lipid profile, age, and sex, Hesse et al. demonstrated that an ML model more effectively identified genetically confirmed FH in a cohort of individuals suspected of having FH than did LDL-C cutoff values, and it performed comparably to the Dutch Lipid Clinic Network criteria.[7] The ML model achieved an area under the receiver operating characteristic curve (AUROC) of 0.711 in an external dataset with a high FH prevalence

(n=1,376; FH prevalence=64%), outperforming the LDL-C cutoff (AUROC=0.642) and matching the Dutch Lipid Clinic Network criteria (AUROC=0.705). Furthermore, the model demonstrated higher accuracy when evaluated on individuals with medium or lower FH prevalence, with AUROC values of 0.801 and 0.856, respectively.

Similarly, another study employed three ML algorithms to predict the presence of FH genetic mutations in two independent European cohorts with FH.[8] Pina et al.[8] discovered that the three ML algorithms outperformed the clinical Dutch Lipid Score in predicting carriers of FH-causative mutations, with AUROC values of 0.79, 0.83, and 0.83 for the Gothenburg cohort, and 0.70, 0.78, and 0.76 for the Milan cohort, compared to the Dutch Lipid Score's AUROC of 0.68 and 0.64 for the Gothenburg and Milan cohorts, respectively.

Nolde et al.[9] demonstrated that a model based on neural networks outperformed traditional clinical diagnostic criteria in predicting pathogenic mutations responsible for FH. The AUROC for the neural network model in predicting an FH gene variant was 0.87, significantly higher than that of other clinical criteria (Dutch Lipid Score: AUROC of 0.743; MEDPED: AUROC of 0.764; Simon Broome: AUROC of 0.635). The model's global accuracy in predicting mutations causing FH was 80.6%, with a specificity of 83.3%, sensitivity of 74.3%, negative predictive value of 88.5%, and positive predictive value of 66.0%.

Finally, Banda et al.[10] developed an ML-based classifier to identify potential FH patients using electronic health record data. The classifier achieved a positive predictive value of 0.88, sensitivity of 0.75, and specificity of 0.99 for detecting genetically confirmed FH cases. In addition, the AUROC was more informative for low-prevalence outcomes.

Underdiagnosis and undertreatment of FH are significant problems affecting the management of the condition. The cardiovascular consequences of FH often remain silent until they manifest as premature cardiovascular mortality and morbidity. Therefore, altering the natural history of FH depends on early detection and the initiation of lipid-lowering therapies. The studies analyzed in this review suggest that an ML-based approach could improve the identification rate of index FH cases. The potential applicability of these novel diagnostic tools for FH could be significant in the future.

LDL-C ESTIMATION

In clinical settings, LDL-C is typically estimated using the Friedewald equation. However, this method is known to be inaccurate in cases of high triglycerides, non-fasting states, or when patients exhibit very low LDL-C values.[11] To account for variations in triglyceride levels, alternative formulas have been developed to estimate LDL-C more accurately than traditional methods. These include the Martin-Hopkins formula and the Sampson equation.

In this context, several studies have evaluated various AI techniques for estimating LDL-C levels. Oh et al.[12] utilized a substantial single-center electronic health record database to develop an ML algorithm that estimates LDL-C from standard lipid profiles, analyzing 823,657 tests. The ML algorithms outperformed the traditional Friedewald and the more recent Martin-Hopkins equations in estimating LDL-C. ML algorithms employing gradient boosting (LDL-CX) and neural networks (LDL-CN) demonstrated a stronger correlation with directly measured LDL-C than the conventional methods, with correlation coefficients

https://doi.org/10.12997/jla.2024.13.2.111

of 0.9662 and 0.9668 for LDL-CX and LDL-CN, respectively, compared to 0.9563 and 0.9585 for the Friedewald and Martin-Hopkins equations. Furthermore, the overall bias was significantly lower for LDL-CX (−0.27 mg/dL) and LDL-CN (−0.01 mg/dL) than for the Friedewald formula (−3.80 mg/dL) and Martin-Hopkins equation (−2.00 mg/dL), particularly at elevated triglyceride levels.

A retrospective study was conducted to compare the performance of an ML algorithm with that of a direct homogeneous LDL-C assay.[13] The principal findings indicated that the ML algorithm demonstrated better agreement with the direct homogeneous LDL-C assay than other equations, particularly in cases of mild and severe hypertriglyceridemia. Overall, the intraclass correlation coefficients for the estimated LDL-C values compared to the directly measured LDL-C values were 0.894, 0.937, 0.935, 0.869, and 0.925 for the Friedewald formula, the Martin-Hopkins equation, the Sampson equation, the de Cordova equation, and the ML-based estimation, respectively (p<0.001).

Similarly, another study introduced a novel method for estimating LDL-C from the standard lipid profile using an ML approach, which was based on 17,500 lipid profiles from 10,936 individuals.[14] The authors found that the correlation coefficients between the estimated and measured LDL-C values were 0.982 for the AI model, surpassing the 0.950 achieved by the Friedewald equation and the 0.962 by the Martin-Hopkins method. Furthermore, the AI model demonstrated superior performance across various subgroups categorized by LDL-C and triglyceride levels. This included individuals with triglycerides greater than 500 mg/dL, where the mean difference in LDL-C estimation was −27.17 mg/dL compared to the Friedewald equation and −4.44 mg/dL compared to the Martin-Hopkins method. Similarly, for individuals with LDL-C levels below 70 mg/dL, the mean difference was −3.82 mg/dL when compared to the Friedewald equation and −1.84 mg/dL compared to the Martin-Hopkins method.

Ghayad et al.[15] evaluated an ML algorithm based on age, sex, healthcare setting, and triglyceride levels against a direct LDL-C assay. The ML algorithm demonstrated good agreement with direct LDL-C measurements in patients with normal triglyceride levels and mild hypertriglyceridemia, as indicated by an intraclass correlation coefficient greater than 0.9. However, its performance was less robust in patients with severe hypertriglyceridemia and in those with very low LDL-C levels, where the intraclass correlation coefficient was slightly below 0.9.

Another study applied ML techniques to three databases: health check-up participants at the Resource Center for Health Science (n=2,664), clinical patients at Gifu University Hospital (n=7,409), and clinical patients at Fujita Health University Hospital (n=14,842).[16] Subsequently, nine different ML models were developed. A separate test dataset (n=3,711) from Fujita Health University Hospital was used to compare and validate the models against the Friedewald formula and the Martin method. The coefficients of determination for the models using the health check-up dataset were equal to or lower than those of the Martin method. However, the coefficients of determination for several models using the clinical patient data surpassed those of the Martin method. Thus, the study's results indicate that the performance of ML models varies with the data source, which complicates the task of confirming whether new techniques outperform traditional methods.

Finally, Kwon et al.[17] aimed to develop a deep neural network model for estimating LDL-C levels and to compare its performance with that of previous LDL-C estimation equations.

https://doi.org/10.12997/jla.2024.13.2.111

JLA Journal of Lipid and Atherosclerosis

They used two large, independent datasets from Korean populations, comprising 176,400 individuals. The DL method exhibited lower bias and error than the Friedewald, Martin, and Sampson equations, demonstrating high agreement with LDL-C levels measured by a homogeneous assay.

Considering these reports, new AI-based techniques could offer a valuable approach for estimating LDL-C levels. These methods may not only surpass the traditional Friedewald formula but could also outperform alternatives such as the Martin-Hopkins formula or the Sampson equation. Nonetheless, these models need further refinement and validation across diverse populations. Enhanced LDL-C estimation would lead to more accurate cardiovascular risk stratification and optimize lipid-lowering treatments, ultimately helping to meet the lipid targets suggested by clinical guidelines and reduce patients' cardiovascular risk.

PREDICTING THE ACHIEVEMENT OF LIPID GOALS

Based on scientific evidence, current guidelines recommend using LDL-C levels as the main therapeutic goals.[18,19] However, despite advances in lipid-lowering treatments, many patients do not reach the goals recommended by the guidelines.[20]

In this context, a study evaluated various predictors of lipid goal achievement in outpatients with type 2 diabetes using ML.[21] The study conducted a real-world analysis of the lipid profiles of 11,252 patients, employing an ML model to identify the most relevant factors for predicting the achievement of an LDL-C level below 100 mg/dL within two years of initiating lipid-lowering therapy. The model exhibited strong predictive capabilities, with an AUROC of 0.79. The most significant predictors of achieving the treatment goal were baseline LDL-C values, their reduction after 6 months, and no therapy discontinuation. Other factors associated with a higher likelihood of success were elevated high-density lipoprotein cholesterol, presence of albuminuria, body mass index, younger age, male sex, lower blood glucose, hemoglobin A1c levels, and the use of antihypertensive medication. Notably, the ML model also determined the minimum LDL-C reduction required by the 6-month follow-up to improve the chances of meeting the therapeutic target within 2 years. For instance, to achieve the two-year goal, the model predicted that patients with baseline LDL-C levels ranging from 100 to 125 mg/dL, 150 to 175 mg/dL, and over 200 mg/dL would need to reduce their LDL-C by at least 14%, 33%, and 47%, respectively, by the 6-month mark. The authors suggested that these insights could be a valuable asset for guiding therapeutic decisions and promoting more comprehensive analysis and validation. However, they acknowledged limitations such as potential biases, the challenge of generalizing results to different populations, the exclusion of certain data such as statin dosages or some socioeconomic factors, and issues related to non-adherence. These caveats underscore the necessity for additional research, which should also encompass other high-cardiovascular-risk groups, including patients with a history of cardiovascular disease, chronic renal failure, or FH.

BARRIERS RELATED TO THE USE OF STATINS

Statin intolerance remains an important clinical challenge, although the prevalence of this clinical scenario might often be overestimated.[22] Importantly, the nocebo effect could play a key role in this significant public health problem.[23]

https://doi.org/10.12997/jla.2024.13.2.111

JLA Journal of Lipid and Atherosclerosis

A recent study demonstrated the potential of an AI approach to sift through extensive social media data, providing insights into public perceptions of statins.[24] This qualitative research focused on discussions related to statins on a social media platform. A total of 10,233 statin-related discussions authored by 5,188 individuals were identified. The three most prevalent topics within these discussions included concerns about elevated LDL-C levels while following a ketogenic diet, requests for advice and shared experiences regarding changes in lipid panels, and anecdotal views on the effectiveness and side effects of statins. Sentiment analysis conducted on the social media posts and discussions about statins indicated that the overall sentiment was mainly neutral to negative, with 30.8% of posts being negative, 66.6% neutral, and only 2.6% positive.

Sarraju et al.[25] developed a DL model to classify statin nonuse and the reasons for it, utilizing unstructured electronic health records from a cohort of 56,530 patients with atherosclerotic cardiovascular disease. The DL model classifiers were able to identify statin nonuse with an AUROC of 0.94 and the reasons for nonuse with a weighted-average AUROC of 0.88, as evaluated against manual expert chart review. This DL-based approach pinpointed key patient-level reasons, such as side effects and patient preference, as well as clinician-level reasons, including guideline-discordant practices, for statin nonuse. It also highlighted differences by patient race or ethnicity. Similarly, a novel DL approach accurately identified statin nonuse in patients with diabetes (n=33,461).[26] This method also classified reasons for statin nonuse from unstructured electronic health record data, which included patient reasons (side effects and statin hesitancy), clinician reasons (guideline-discordant practice), and system reasons (clinical inertia), with variations observed by patient race and ethnicity. Older individuals (>75 years of age) were more likely to experience statin-associated side effects or contraindications (23.4%) and were less likely to be affected by clinical inertia or guideline-discordant practice compared to younger individuals (p<0.05 for comparisons). Hispanic patients were the most likely to encounter guideline-discordant practice (24.7%, p<0.05), while Black patients were the most likely to be subject to clinical inertia (24.0%, p<0.05), in comparison to other racial groups.

Interestingly, the last two studies identified some common reasons for statin non-adherence, such as adverse effects or perceived inadequate lipid control/guideline-discordant practice, while other reasons varied, including patient preference, statin hesitancy, or clinical inertia. However, all these reasons have been previously documented in the medical literature.[27] Therefore, accurately identifying these barriers using AI techniques could enhance statin utilization among patients with high cardiovascular risk in real-world settings, offering a strategy to bridge significant gaps in dyslipidemia treatment. It is crucial that future research be tailored to the diverse geographical areas and cultural characteristics of each population.

ASSOCIATION BETWEEN CARDIOMETABOLIC AND DIETARY FACTORS AND APOLIPOPROTEIN B (ApoB)/ LDL-C DISCORDANCE

ApoB is an adequate representative of all atherogenic particles.[28] When there is a discrepancy between LDL-C and ApoB levels, elevated ApoB has been shown to be a more accurate predictor of atherosclerotic cardiovascular disease risk.[29] Webb et al.[30] examined the influence of adiposity, diet, and inflammation on the discordance between LDL-C and

https://doi.org/10.12997/jla.2024.13.2.111

ApoB by employing an ML model. In this instance, the ML model was applied to data from the National Health and Nutrition Examination Survey to explore cardiometabolic and dietary factors associated with concordance or discordance between LDL-C and ApoB. The ML analysis revealed that body mass index, dietary saturated fatty acids, dietary fiber, serum C-reactive protein, and uric acid were the variables most strongly associated (R^2=0.70) with the pattern of low LDL-C and high ApoB. Accurately identifying patients with lipid discordance is crucial, as their condition has been linked to a higher prevalence of atherosclerosis.[31,32]

A summary of the main characteristics of the studies analyzed in this review is presented in **Table 1**.

LIMITATIONS, CHALLENGES, AND OPPORTUNITIES RELATED TO AI

Adapting new AI-based techniques to meet patient needs presents a significant challenge. Nevertheless, it is essential to address certain concerns associated with AI techniques. Ethical issues, such as patient privacy, must be taken into consideration.[33] Additionally, when interpreting information, one should be mindful of publication-related issues, including a lack of transparency or methodological clarity that can impact the reproducibility of results.[34] Familiarizing oneself with the specific methodological aspects of new techniques, such as mathematical models and software, will pose a challenge for the medical community. A key distinction is that conventional statistics are model-driven, whereas AI and ML are data-driven, operating without a predefined understanding of the relationship between data and outcomes.[35] In AI and ML, the software identifies patterns and forms clusters of data with shared characteristics that may influence the outcome. Traditional statistical methods, on the other hand, presuppose knowledge of the model that generated the data, assuming a known relationship between input variables and the output. The issue here is that the relationship between input and output is selected by the user and may lead to a less than optimal model. In contrast, ML methods do not start with a presumed model; instead, they begin with the data, and a mathematical algorithm develops a model with prediction as the primary objective. Compared to traditional statistical methods, ML algorithms can manage more variables but also necessitate a larger sample size for analysis. Another challenge is integrating these technologies into existing workflows. AI is not intended to replace clinical judgment. Rather, the most effective use of AI in clinical settings occurs when it is integrated seamlessly into the clinical workflow.[36] Finally, it is crucial to consider the specific characteristics of different geographical regions, cultures, and ethnic groups. In this regard, some authors have suggested the creation of a framework to guide research groups in designing, conducting, and reporting their studies; to assist editors and peer reviewers in evaluating contributions to the literature; and to enable patients, clinicians, and policymakers to critically assess where new findings may offer patient benefits.[37]

The concerns and opportunities related to AI techniques in the area of lipidology are shown in **Fig. 1**.

CONCLUSION

In this literature review, we analyzed potential applications of AI, ML, and DL in lipidology. Diagnostic applications, such as screening for FH or calculating LDL-C, could significantly

https://doi.org/10.12997/jla.2024.13.2.111

Table 1. Characteristics of the studies included

Study (year)	Data origin	Country	Main results
Diagnosis of FH			
Banda et al.[10] (2019)	Internal data set: FH patients (n=197) and matched non-cases (n=6,590). External data set: FH patients (n=466) and matched non-cases (n=5,000).	USA	The ML-based classifier obtained a positive predictive value of 0.88, sensitivity of 0.75, and specificity of 0.99 for detecting genetically confirmed FH cases.
Pina et al.[9] (2020)	Patients diagnosed with FH	Sweden	The ML algorithms performed better (AUROC range between 0.70 and 0.83) than the clinical DLCNc (AUROC range between 0.64-0.68) in predicting carriers of FH-causative mutations.
Hesse et al.[7] (2022)	Clinically suspected of having FH. Internal dataset: n=678. External dataset: high FH prevalence, n=1376; medium FH prevalence, n=3,304; low FH prevalence: n=1,616.	South Africa	The ML model achieved an AUROC value of 0.711 on the dataset with high FH prevalence, which was superior to the LDL-C cutoff (AUROC=0.642) and comparable to the DLCNc (AUROC=0.705). The model achieved higher accuracy when tested on individuals with medium (AUROC=0.801) or lower (AUROC=0.856) FH prevalence.
Nolde et al.[9] (2023)	Adult patients referred to a specialist lipid clinic (n=885).	Australia	The model based on neural networks (AUROC=0.870) was superior to traditional clinical diagnostic criteria (AUROC range between 0.635 and 0.764) in predicting pathogenic mutations of FH.
LDL-C estimation			
Singh et al.[14] (2020)	Electronic health record (17,500 lipid profiles performed on 10,936 unique individuals).	USA	The ML model had a better correlation with direct LDL-C (r=0.982) than the Friedewald (r=0.95) or Martin-Hopkins (r=0.962) formulas, even in patients with high triglycerides (mean difference of −27.17 mg/dL compared to Friedewald and −4.44 mg/dL compared to Martin Hopkins) and very low LDL-C (mean difference of −3.82 mg/dL compared to Friedewald and −1.84 mg/dL compared to Martin-Hopkins).
Barakett-Hamade et al.[13] (2021)	A total of 31,922 observations from 19,279 subjects.	Lebanon	The ML algorithm agreed better with direct LDL measurement than other equations, especially in mild and severe hypertriglyceridemia.
Oh et al.[12] (2022)	Single-center electronic health record database (n=823,657 tests)	South Korea	ML algorithms (r=0.966 and r=0.967) showed better correlation with directly measured LDL-C than the Friedewald formula (r=0.956) and Martin-Hopkins equations (r=0.958).
Ghayad et al.[15] (2022)	The analysis comprised 31,853 retrospective and 6,599 prospective observations.	Lebanon	The ML algorithm was in satisfactory agreement with direct LDL-C in observations with normal triglyceridemia and mild hypertriglyceridemia (intraclass correlation coefficient >0.9)
Kwon et al.[17] (2022)	Participants from 2 independent population-based cohorts. Internal dataset: n=129,930; External dataset: n=46,470.	South Korea	The DL model had lower bias and root error than the Friedewald, Martin, and Sampson equations, showing a high agreement with LDL-C measured by a homogeneous assay.
Hidekazu et al.[16] (2023)	Participants from independent datasets. Three internal datasets: n=24,915. External dataset: n=3,711.	Japan	The coefficients of determination of the ML models on the health check-up dataset were equal to or inferior to those of the Martin method. The coefficients of determination of the ML models on clinical patients exceeded those of the Martin method.
Predicting the achievement of LDL-C goals			
Masi et al.[21] (2023)	Electronic medical records of patients with diabetes.	Italy	The ML model demonstrated good predictive performance (AUROC=0.79). The most significant predictors of achieving the LDL-C goal were baseline LDL-C levels and their reduction after 6 months.
Barriers related to the use of statins			
Somani et al.[24] (2022)	Reddit was used as the data source. A total of 10,233 unique statin-related discussions from 5,188 unique authors were included.	US	Anecdotal perspectives on statin efficacy and adverse effects were one of the most common topics of statin-related discussions. Neutral or negative sentiments were more frequent.
Sarraju et al.[25] (2022)	Adults with ASCVD from an electronic health record cohort (n=56,530).	US	The DL approach identified statin nonuse and potentially actionable reasons for statin nonuse in high-risk populations (AUROC=0.94).
Sarraju et al.[26] (2023)	Adults with diabetes and no statin prescriptions from an electronic health record cohort (n=33,461).	US	The DL approach identified statin nonuse (AUROC=0.99) and potentially actionable reasons for statin nonuse including key patient, clinician, and system factors. Reasons for nonuse varied by clinical and demographic characteristics.
Association between ApoB/LDL-C discordance and metabolic and diet factors			
Webb et al.[30] (2022)	Data derived from the US NHANES (n=14,265).	US	ML showed that BMI, dietary saturated fatty acids, dietary fiber, C-reactive protein and uric acid were the most strongly associated variables with the low LDL-C/high ApoB pattern (R^2=0.70),

FH, familial hypercholesterolemia; ML, machine learning; AUROC, area under the receiver operating characteristic curve ; DLCNc, Dutch Lipid Clinic Network criteria; LDL-C, low-density lipoprotein cholesterol; LDL, low-density lipoprotein; ASCVD, atherosclerotic cardiovascular disease; DL, deep learning; ApoB, apolipoprotein B; NHANES, National Health and Nutrition Examination Survey; BMI, body mass index.

impact clinical practice. Understanding the barriers patients face with lipid-lowering medications could enable physicians to address these obstacles effectively. Furthermore, the development of improved predictive models based on lipid profiles may be anticipated in the near future. Additionally, novel AI techniques could uncover previously unknown

https://doi.org/10.12997/jla.2024.13.2.111

JLA Journal of Lipid and Atherosclerosis

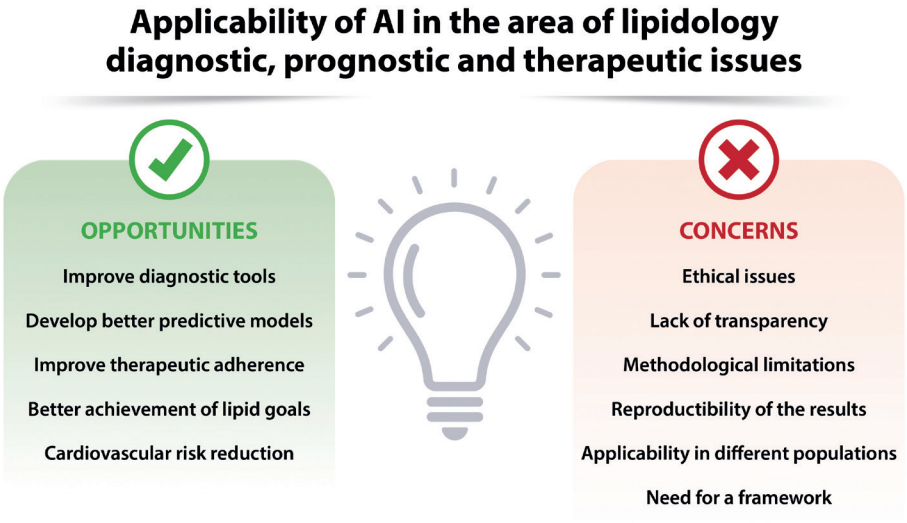

Fig. 1. Concerns and opportunities of using AI techniques in the field of lipidology.
AI, artificial intelligence.

disease correlations and facilitate the delivery of precision medicine. However, the social, methodological, and ethical complexities associated with these applications warrant further investigation and regulation.

REFERENCES

1. Parhofer KG. Update lipidology: evidence-based treatment of dyslipidemia. Inn Med (Heidelb) 2023;64:611-621. PUBMED | CROSSREF
2. Tang L, Li J, Fantus S. Medical artificial intelligence ethics: a systematic review of empirical studies. Digit Health 2023;9:20552076231186064. PUBMED | CROSSREF
3. Gomes MA, Kovaleski JL, Pagani RN, da Silva VL, Pasquini TC. Transforming healthcare with big data analytics: technologies, techniques and prospects. J Med Eng Technol 2023;47:1-11. PUBMED | CROSSREF
4. Hong L, Luo M, Wang R, Lu P, Lu W, Lu L. Big data in health care: applications and challenges. Data Inf Manag 2018;2:175-197. CROSSREF
5. Kallapur A, Sallam T. Pharmacotherapy in familial hypercholesterolemia- current state and emerging paradigms. Trends Cardiovasc Med 2023;33:170-179. PUBMED | CROSSREF
6. Berberich AJ, Hegele RA. The complex molecular genetics of familial hypercholesterolaemia. Nat Rev Cardiol 2019;16:9-20. PUBMED | CROSSREF
7. Hesse R, Raal FJ, Blom DJ, George JA. Familial hypercholesterolemia identification by machine learning using lipid profile data performs as well as clinical diagnostic criteria. Circ Genom Precis Med 2022;15:e003324. PUBMED | CROSSREF
8. Pina A, Helgadottir S, Mancina RM, Pavanello C, Pirazzi C, Montalcini T, et al. Virtual genetic diagnosis for familial hypercholesterolemia powered by machine learning. Eur J Prev Cardiol 2020;27:1639-1646. PUBMED | CROSSREF
9. Nolde JM, Pang J, Chan DC, Ward NC, Mian A, Schlaich MP, et al. Neural network modelling for predicting gene variants causative of familial hypercholesterolaemia in the clinic. Heart Lung Circ 2023;32:e44-e45. PUBMED | CROSSREF
10. Banda JM, Sarraju A, Abbasi F, Parizo J, Pariani M, Ison H, et al. Finding missed cases of familial hypercholesterolemia in health systems using machine learning. NPJ Digit Med 2019;2:23. PUBMED | CROSSREF
11. Rossouw HM, Nagel SE, Pillay TS. Comparability of 11 different equations for estimating LDL cholesterol on different analysers. Clin Chem Lab Med 2021;59:1930-1943. PUBMED | CROSSREF

https://doi.org/10.12997/jla.2024.13.2.111

12. Oh GC, Ko T, Kim JH, Lee MH, Choi SW, Bae YS, et al. Estimation of low-density lipoprotein cholesterol levels using machine learning. Int J Cardiol 2022;352:144-149. PUBMED | CROSSREF
13. Barakett-Hamade V, Ghayad JP, Mchantaf G, Sleilaty G. 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;519:220-226. PUBMED | CROSSREF
14. Singh G, Hussain Y, Xu Z, Sholle E, Michalak K, Dolan K, 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. PUBMED | CROSSREF
15. Ghayad JP, Barakett-Hamadé V, Sleilaty G. Prospective validation of a machine learning model for low-density lipoprotein cholesterol estimation. Lab Med 2022;53:629-635. PUBMED | CROSSREF
16. Hidekazu I, Nagasawa H, Yamamoto Y, Doi H, Saito M, Ishihara Y, et al. Dataset dependency of low-density lipoprotein-cholesterol estimation by machine learning. Ann Clin Biochem 2023;60:396-405. PUBMED | CROSSREF
17. Kwon YJ, Lee H, Baik SJ, Chang HJ, Lee JW. Comparison of a machine learning method and various equations for estimating low-density lipoprotein cholesterol in Korean populations. Front Cardiovasc Med 2022;9:824574. PUBMED | CROSSREF
18. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. Eur Heart J 2020;41:111-188. PUBMED | CROSSREF
19. Visseren FL, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. Eur J Prev Cardiol 2022;29:5-115. PUBMED | CROSSREF
20. Ersbøll AK, Kristensen MS, Nybo M, Hede SM, Mikkelsen KH, Gislason G, et al. Trends in low-density lipoprotein cholesterol goal achievement and changes in lipid-lowering therapy after incident atherosclerotic cardiovascular disease: Danish cohort study. PLoS One 2023;18:e0286376. PUBMED | CROSSREF
21. Masi D, Zilich R, Candido R, Giancaterini A, Guaita G, Muselli M, et al. Uncovering predictors of lipid goal attainment in type 2 diabetes outpatients using logic learning machine: insights from the AMD annals and AMD artificial intelligence study group. J Clin Med 2023;12:4095. PUBMED | CROSSREF
22. Bytyçi I, Penson PE, Mikhailidis DP, Wong ND, Hernandez AV, Sahebkar A, et al. Prevalence of statin intolerance: a meta-analysis. Eur Heart J 2022;43:3213-3223. PUBMED | CROSSREF
23. Penson PE, Mancini GB, Toth PP, Martin SS, Watts GF, Sahebkar A, et al. Introducing the 'Drucebo' effect in statin therapy: a systematic review of studies comparing reported rates of statin-associated muscle symptoms, under blinded and open-label conditions. J Cachexia Sarcopenia Muscle 2018;9:1023-1033. PUBMED | CROSSREF
24. Somani S, van Buchem MM, Sarraju A, Hernandez-Boussard T, Rodriguez F. Artificial intelligence-enabled analysis of statin-related topics and sentiments on social media. JAMA Netw Open 2023;6:e239747. PUBMED | CROSSREF
25. Sarraju A, Coquet J, Zammit A, Chan A, Ngo S, Hernandez-Boussard T, et al. Using deep learning-based natural language processing to identify reasons for statin nonuse in patients with atherosclerotic cardiovascular disease. Commun Med (Lond) 2022;2:88. PUBMED | CROSSREF
26. Sarraju A, Zammit A, Ngo S, Witting C, Hernandez-Boussard T, Rodriguez F. Identifying reasons for statin nonuse in patients with diabetes using deep learning of electronic health records. J Am Heart Assoc 2023;12:e028120. PUBMED | CROSSREF
27. Mostaza JM, Lahoz C. Main barriers in the management of dyslipidaemias: Intolerants. Clin Investig Arterioscler 2021;33 Suppl 1:40-45. PUBMED
28. Glavinovic T, Thanassoulis G, de Graaf J, Couture P, Hegele RA, Sniderman AD. Physiological bases for the superiority of apolipoprotein B over low-density lipoprotein cholesterol and non-high-density lipoprotein cholesterol as a marker of cardiovascular risk. J Am Heart Assoc 2022;11:e025858. PUBMED | CROSSREF
29. Sniderman AD, Thanassoulis G, Glavinovic T, Navar AM, Pencina M, Catapano A, et al. Apolipoprotein B particles and cardiovascular disease: a narrative review. JAMA Cardiol 2019;4:1287-1295. PUBMED | CROSSREF
30. Webb RJ, Mazidi M, Lip GY, Kengne AP, Banach M, Davies IG. The role of adiposity, diet and inflammation on the discordance between LDL-C and apolipoprotein B. Nutr Metab Cardiovasc Dis 2022;32:605-615. PUBMED | CROSSREF

31. Masson W, Lobo M, Molinero G, Siniawski D. Discordant lipid pattern and carotid atherosclerotic plaque. Importance of remnant cholesterol. Arq Bras Cardiol 2017;108:526-532. PUBMED | CROSSREF
32. Masson W, Siniawski D, Lobo M, Molinero G. Lipid discordance and carotid plaque in obese patients in primary prevention. Endocrinol Diabetes Nutr (Engl Ed) 2018;65:39-44. PUBMED | CROSSREF
33. Kayaalp M. Patient privacy in the era of big data. Balkan Med J 2018;35:8-17. PUBMED | CROSSREF
34. Wachter S, Mittelstadt B, Russell C. Counterfactual explanations without opening the black box: automated decisions and the GDPR. Harv J Law Technol 2018;31:841-887. CROSSREF
35. Ley C, Martin RK, Pareek A, Groll A, Seil R, Tischer T. Machine learning and conventional statistics: making sense of the differences. Knee Surg Sports Traumatol Arthrosc 2022;30:753-757. PUBMED | CROSSREF
36. Sahni NR, Carrus B. Artificial intelligence in U.S. health care delivery. N Engl J Med 2023;389:348-358. PUBMED | CROSSREF
37. Vollmer S, Mateen BA, Bohner G, Király FJ, Ghani R, Jonsson P, et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. BMJ 2020;368:l6927. PUBMED | CROSSREF