JACC: ADVANCES VOL. 3, NO. 12, 2024

EDITORIAL COMMENT

EHR-Based Screening of Familial Hypercholesterolemia

Finding the Lipid in the Haystack

Ramzi Ibrahim, MD, Jack Hartnett, MB, BCh, BAO

Familial hypercholesterolemia (FH) is a genetic disorder characterized by elevated low-density lipoprotein cholesterol (LDL-C) levels throughout life, leading to premature atherosclerotic cardiovascular disease (ASCVD).[1,2] It is most commonly caused by autosomal dominant mutations in the LDL receptor gene, though other genes involved in LDL-C metabolism can also be implicated. FH is a relatively common condition, with modern registry studies estimating a prevalence of approximately 1 in 250 people, translating to around 14 to 34 million patients with FH worldwide.[1,2] However, fewer than 1% of patients with FH are diagnosed in most countries, a striking gap in diagnosis.[3]

Untreated heterozygous patients with FH face a 10- to 20-fold increased risk of premature coronary artery disease compared to the general population.[1,2] With the emergence of novel therapies, such as proprotein convertase subtilisin/kexin type 9 inhibitors and long-acting small interfering RNAs, which are highly effective in reducing serum LDL-C levels, the case for screening patients with FH early has never been stronger. Early identification and treatment offer a potential pathway to mitigating the lifelong ASCVD risk that patients with FH endure.

In recent years, there has been increasing interest in applying machine learning (ML) models to identify at-risk patients at a population level.[4-6] The wealth of data stored in electronic health records (EHRs), combined with advances in artificial intelligence (AI), has sparked the development of a wide array of new risk models and screening algorithms across various medical fields.[6] In cardiology, ML models have not only outperformed traditional risk algorithms, such as the pooled cohort equations for ASCVD risk prediction, but have also outpaced cardiology trainees in recognizing electrocardiogram abnormalities.[4,7]

In this issue of *JACC: Advances*, Osei et al[8] provided a thorough review of the current landscape of EHR-based screening for FH. Through a scoping review of 27 studies, which evaluated 14 different screening algorithms and tools, the authors aimed to assess the performance of existing EHR-based screening algorithms for FH, their utility in clinical practice, and the barriers associated with their adoption. This comprehensive overview offers a valuable synthesis of the varied screening tools and algorithms currently in use, summarizing their performance against reference standards at different prevalence assumptions.

This review subtly raises the question: Are EHR and ML-based algorithms for FH detection ready for routine clinical implementation? The authors identify several key challenges to their widespread adoption. One prominent issue is the lack of a definitive diagnostic test for FH. Diagnostic approaches vary, ranging from genetic testing to clinical assessment with LDL-C levels, and combined clinical-genetic approaches like the Simon Broome and Dutch Lipid Clinic Network criteria.[9] This lack of diagnostic uniformity complicates the reference standards used to evaluate EHR-based screening algorithms. Some studies in the review used genetic testing as the reference standard, while others relied on Simon Broome or Dutch Lipid Clinic Network criteria. Moreover, some studies that used different

From the Department of Cardiovascular Medicine, Mayo Clinic, Scottsdale, Arizona, USA.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

ISSN 2772-963X https://doi.org/10.1016/j.jacadv.2024.101296

algorithms for FH diagnosis performed well, suggesting that a one-size-fits-all approach may not be optimal when using EHR data to establish an FH diagnosis. This highlights the need for tailored methods that can account for diverse patient populations and clinical settings. With only one study providing a head-to-head comparison of algorithms, there remains insufficient data to recommend one EHR-based tool over another.[10]

Another significant challenge is defining the population that should be targeted for EHR-based screening. When applied to the general population through primary care data sets, the positive predictive value of the available algorithms can be quite low (as little as <1%).[11] However, when applied to high-risk cohorts, such as those in lipid clinics, the positive predictive value can be as high as 95.3%.[12] The implications of this are clear: applying these algorithms broadly would require substantial resources and infrastructure to manage the influx of patients flagged by positive screens. Furthermore, while the review highlights the utility of FH screening, there remains a paucity of health economic analyses. Additionally, the authors note the lack of data in minority populations and low-resource settings, which raises concerns about the generalizability of these tools across diverse populations. External validation of these algorithms remains underreported in this scoping review, limiting the general applicability. When there is discordance between the characteristics of the training data set, the validation data set, and the target population, the tool's effectiveness can be diminished.

Without careful consideration of these factors, the implementation of EHR-based screening may further widen the health disparity gap, particularly in identifying and treating high-risk individuals. Socioeconomic and racial disparities lead to reduced access to health care, including genetic testing and lipid screenings, which are integral components of FH diagnosis.[13] Furthermore, as EHR algorithms rely on patient data, these may not be as readily available or accurate for populations within underserved communities. These vulnerable populations face many gaps in health care coverage, incomplete data entries in EHR systems, and fragmented medical histories, many of which can impact the efficacy of these screening tools. Another consideration is the potential biases in the data sets that were used to train these algorithms, which may lead to misclassification in minority groups, further exacerbating existing health inequities.

Despite these hurdles, ML and AI are becoming indispensable tools in cardiovascular research, improving our ability to detect diseases early and enhance clinical decision-making.[6] With health care data becoming increasingly available, the potential for ML models to transform clinical practice is vast. Unstructured data, which comprises nearly 80% of all medical records, can now be processed more efficiently, attributable to advancements in natural language processing and automated data extraction.[14] For instance, natural language processing algorithms have been used to identify comorbidities, such as stroke, hypertension, coronary artery disease, diabetes, and dyslipidemia, emphasizing the broad applicability of AI technologies in cardiovascular risk stratification.[15]

Although the application of ML in FH detection is still in its early stages, substantial progress has already been made. As this field continues to develop, EHR-based screening may soon become a standard component of clinical practice. The increasing use of EHR modalities to extract and analyze medical data has significantly enhanced our ability to predict CVD risk factors and identify diseases more effectively. Advanced data-driven multiomics platforms that integrate clinical data, environmental factors, genomics, and epigenomics offer more precise and timely identification of the factors influencing CVD progression. This, in turn, will lead to the development of individualized profiles that allow for targeted public health interventions and more actionable strategies in high-risk areas. Similarly, the early detection and treatment of the underdiagnosed but prevalent condition of FH remain vital in the global effort to combat CVD.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Ramzi Ibrahim, Department of Cardiovascular Medicine, Mayo Clinic Arizona, 13400 East Shea Boulevard, Scottsdale, Arizona 85259, USA. E-mail: ramzi.ibrahim83@gmail.com.

REFERENCES

1. Zamora A, Masana L, Comas-Cufí M, et al. Familial hypercholesterolemia in a European Mediterranean population-Prevalence and clinical data from 2.5 million primary care patients. *J Clin Lipidol*. 2017;11(4):1013–1022.

2. Akioyamen LE, Genest J, Shan SD, Reel RL, Albaum JM, Chu A, Tu JV. Estimating the prevalence of heterozygous familial hypercholesterolaemia: a systematic review and meta-analysis. *BMJ Open*. 2017;7(9):e016461.

3. Nordestgaard BG, Chapman MJ, Humphries SE, et al. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: consensus statement of the European Atherosclerosis Society. *Eur Heart J*. 2013;34(45):3478–3490a.

4. Petrazzini BO, Chaudhary K, Márquez-Luna C, et al. Coronary risk estimation based on clinical data in electronic health records. *J Am Coll Cardiol*. 2022;79(12):1155–1166.

5. Chen Z, Dazard JE, Khalifa Y, Motairek I, Kreatsoulas C, Rajagopalan S, Al-Kindi S. Deep learning-based assessment of built environment from satellite images and cardiometabolic disease prevalence. *JAMA Cardiol*. 2024;9(6):556–564.

6. Ibrahim R, Pham HN, Ganatra S, Javed Z, Nasir K, Al-Kindi S. Social phenotyping for cardiovascular risk stratification in electronic health registries. *Curr Atheroscler Rep*. 2024;26(9):485–497.

7. Ribeiro AH, Ribeiro MH, Paixão GMM, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nat Commun*. 2020;11(1):1760.

8. Osei J, Razavi AC, Otchere B, et al. A scoping review of electronic health records-based screening algorithms for familial hypercholesterolemia. *JACC Adv*. 2024;3:101297.

9. McGowan MP, Hosseini Dehkordi SH, Moriarty PM, Duell PB. Diagnosis and treatment of heterozygous familial hypercholesterolemia. *J Am Heart Assoc*. 2019;8(24):e013225.

10. Gidding SS, Kirchner HL, Brangan A, et al. Yield of familial hypercholesterolemia genetic and phenotypic diagnoses after electronic health record and genomic data screening. *J Am Heart Assoc*. 2023;12(13):e030073.

11. Weng S, Kai J, Akyea R, Qureshi N. Detection of familial hypercholesterolaemia: external validation of the FAMCAT clinical case-finding algorithm to identify patients in primary care. *Lancet Public Health*. 2019;4(5):e256–e264.

12. Ruel I, Brisson D, Aljenedil S, et al. Simplified Canadian definition for familial hypercholesterolemia. *Can J Cardiol*. 2018;34(9):1210–1214.

13. Ibrahim R, Sainbayar E, Pham HN. Public health initiatives: addressing social vulnerability in research and practice. *J Invest Med*. 2024;72(1):159–161.

14. Kong HJ. Managing unstructured big data in healthcare system. *Healthc Inform Res*. 2019;25(1):1–2.

15. Berman AN, Biery DW, Ginder C, et al. Natural language processing for the assessment of cardiovascular disease comorbidities: the cardio-Canary comorbidity project. *Clin Cardiol*. 2021;44(9):1296–1304.

KEY WORDS artificial intelligence, hyperlipidemia, lipid disorders, machine learning