

EDITORIAL COMMENT

Familial Hypercholesterolemia Detection Through Machine Learning Algorithms



A Low-Hanging Fruit

Laurens F. Reeskamp, MD, PhD,^{a,b} Maxim E. Annink, MD, MSc,^a Willemijn A.M. Schonck, MD^a

Familial hypercholesterolemia (FH) is a genetic disorder characterized by markedly elevated low-density lipoprotein cholesterol (LDL-C) levels, which substantially increases the risk of premature cardiovascular disease.^{1,2} Recent studies have demonstrated that the prevalence of FH in the general population is approximately 1:300, with a 23-fold higher prevalence among persons with severe hypercholesterolemia.³ Effective lipid-lowering therapies, such as statins, ezetimibe, and proprotein convertase subtilisin/kexin 9 inhibitors, are readily available and initiation of these at a young age can greatly reduce the risk of atherosclerotic cardiovascular disease.⁴

Despite its high prevalence, associated serious cardiovascular sequelae, and available preventive therapies, FH remains underdiagnosed and undertreated. Current estimates suggest that as many as 90% of FH patients worldwide remain unidentified, and more than one-third may not be diagnosed until their first cardiovascular event has already occurred.^{5,6} Equally concerning are the recent findings from a large global FH registry, which revealed that no more than 3% of patients with FH reach the LDL-C targets recommended by guidelines.⁷ These concerning trends can be attributed to the fact that FH is an asymptomatic disorder until coronary artery disease events manifest, a lack of awareness of FH among physicians, combined with the labor-intensive

nature of screening families of FH patients and manually searching existing electronic health records for potential new cases.⁸

As such, novel approaches to address this problem are long overdue. One such approach is the use of machine learning algorithms tailored to the detection of FH. These algorithms employ automated data mining techniques to identify potential new cases of FH within health care systems. They do so by rapidly analyzing electronic health and laboratory records to identify features that are predictive of having yet-to-be-diagnosed FH. An example is the Family Heart Foundation's flag, identify, network, deliver FH (FIND FH) machine learning model, developed in 2019 and initially trained on electronic health records of clinically diagnosed FH patients from 4 large U.S. health care systems.⁹ Although only 50% of FH patients in this training data set had genetically confirmed FH, its potential value to discover new FH patients was recently demonstrated in a validation cohort, where 34 of the 573 flagged individuals had a (likely) pathogenic FH-causing variant after subsequent genetic validation.¹⁰

In this issue of *JACC: Advances*, Kim et al¹¹ provide further validation of the FIND FH model by applying it to a large U.S. health care system. In brief, the FIND FH algorithm was applied to the electronic health records of 167,955 patients with at least one known cardiovascular risk factor (eg, hypertension or hypercholesterolemia). This resulted in 471 potential FH patients who were automatically flagged by FIND FH and then manually reviewed. Of these, 32 patients had previously established heterozygous FH (10 known homozygous FH patients were excluded beforehand), and 121 patients were considered highly likely to have FH by the clinical reviewers based on elevated LDL-C values and family history. The

From the ^aDepartment of Vascular Medicine, Amsterdam Cardiovascular Sciences, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands; and the ^bDepartment of Internal Medicine, OLVG Oost, Amsterdam, the Netherlands.

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](#).

remaining 318 individuals selected by the algorithm were classified as “suspected FH.” One might question whether this designation is appropriate, as none of these individuals were clinically suspected of having FH by the reviewers.

A key finding of the study is the disparity in medical management between those with established FH diagnoses and those newly identified by the machine learning algorithm. Patients with established diagnoses were more likely to receive lipid panel monitoring, prescriptions for second-line lipid-lowering therapies (eg, ezetimibe and proprotein convertase subtilisin/kexin 9 inhibitors), and specialized cardiovascular evaluations, such as coronary artery calcium testing and lipoprotein(a) measurements. Despite these differences in management, there was no significant difference in the occurrence of major cardiovascular events between the diagnosed and undiagnosed groups.

Thus, integrating machine learning algorithms into routine clinical practice could facilitate the identification of FH patients, enabling timely and targeted interventions that can significantly improve patient outcomes. The algorithm’s ability to sift through vast quantities of electronic medical records and flag high-risk individuals represents a substantial advancement in automated technologies for personalized medicine.

However, it should be noted that this study is not without its limitations. First, it appears that validation against the gold standard of FH diagnostics, genetic testing, is lacking in the individuals that were flagged by the model. Instead, clinically unvalidated criteria based on LDL-C values and family history were used. Combined with the fact that there is only a limited overlap between patients identified with phenotypic FH (ie, severe hypercholesterolemia) and genetic FH,^{12,13} it is difficult to draw conclusions about the performance of this model to identify true, genetically confirmed, FH patients. This problem is likely compounded by the fact that less than half of individuals in the original FIND FH training cohort

had a genetic diagnosis to begin with.⁹ Secondly, we concur with the authors that it is unfortunate that the FIND FH model necessitates screened individuals exhibit at least one cardiovascular risk factor. This requirement constrains the model’s utility in identifying young, undiagnosed individuals who do not have other cardiovascular risk factors but would benefit the most from early initiation of lipid-lowering therapy.¹⁴ Finally, although the FIND FH approach is evidently much less time-consuming than manually screening entire health care systems’ populations, it seems inevitable that flagged records undergo rigorous manual review to verify the algorithm’s findings. This is a laborious process that might not be easily implemented in many health care systems.

Despite these limitations, the large number of previously undiagnosed potential FH patients identified by the algorithm in the study by Kim et al¹¹ highlights a critical gap in FH identification that might be effectively bridged using advanced computational tools. Moving forward, it is essential to validate and assess the algorithm’s sensitivity and specificity in other cohorts, as well as run practical implementation studies where such tools run prospectively and in parallel with clinical care. We look forward to follow-up research that builds on these results and foresee that this will improve clinical outcomes through early identification and treatment of FH in a real-world setting.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

Dr Reeskamp is cofounder of Lipid Tools; and has received speaker fees from Ultragenyx, Novartis, and Daiichi Sankyo.

ADDRESS FOR CORRESPONDENCE: Dr L.F. Reeskamp, Department of Vascular Medicine, Amsterdam Cardiovascular Sciences, Amsterdam UMC, University of Amsterdam, Meibergdreef 9, Amsterdam, 1105 AZ, the Netherlands. E-mail: l.f.reeskamp@amsterdamumc.nl.

REFERENCES

1. Tromp TR, Hartgers ML, Hovingh GK, et al. Worldwide experience of homozygous familial hypercholesterolaemia: retrospective cohort study. *Lancet*. 2022;399(10326):719-728.
2. Cuchel M, Bruckert E, Ginsberg HN, et al. Homozygous familial hypercholesterolaemia: new insights and guidance for clinicians to improve detection and clinical management. A position paper from the Consensus Panel on Familial Hypercholesterolaemia of the European Atherosclerosis Society. *Eur Heart J*. 2014;35(32):2146-2157.
3. Beheshti SO, Madsen CM, Varbo A, Nordestgaard BG. Worldwide prevalence of familial hypercholesterolemia: Meta-Analyses of 11 Million Subjects. *J Am Coll Cardiol*. 2020;75(20):2553-2566.
4. Luirink IK, Wiegman A, Kusters DM, et al. 20-Year follow-up of statins in Children with familial hypercholesterolemia. *N Engl J Med*. 2019;381(16):1547-1556.

5. Duell PB, Gidding SS, Andersen RL, et al. Longitudinal low density lipoprotein cholesterol goal achievement and cardiovascular outcomes among adult patients with familial hypercholesterolemia: the CASCADE FH registry. *Atherosclerosis*. 2019;289:85-93.
 6. 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):3478-3490.
 7. Vallejo-Vaz AJ, Stevens CAT, Lyons ARM, et al. Global perspective of familial hypercholesterolaemia: a cross-sectional study from the EAS Familial Hypercholesterolaemia Studies Collaboration (FHSC). *Lancet*. 2021;398(10312):1713-1725.
 8. Ibrahim S, Reeskamp LF, Stroes ESG, Watts GF. Advances, gaps and opportunities in the detection of familial hypercholesterolemia: overview of current and future screening and detection methods. *Curr Opin Lipidol*. 2020;31(6):347-355.
 9. Myers KD, Knowles JW, Staszak D, et al. Precision screening for familial hypercholesterolaemia: a machine learning study applied to electronic health encounter data. *Lancet Digital Health*. 2019;1(8):e393-e402.
 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. Kim K, Faruque SC, Lam S, et al. Implications of diagnosis through a machine learning algorithm on management of people with familial hypercholesterolemia. *JACC Adv*. 2024;3:101184.
 12. Saadatagah S, Jose M, Dikilitas O, et al. Genetic basis of hypercholesterolemia in adults. *NPJ Genom Med*. 2021;6(1):28.
 13. Björnsson E, Thorgeirsson G, Helgadóttir A, et al. Large-scale screening for Monogenic and clinically Defined familial hypercholesterolemia in Iceland. *Arterioscler Thromb Vasc Biol*. 2021;41(10):2616-2628.
 14. Ibrahim S, Reeskamp LF, de Goeij JN, et al. Beyond early LDL cholesterol lowering to prevent coronary atherosclerosis in familial hypercholesterolaemia. *Eur J Prev Cardiol*. 2024;31(7):892-900.
-
- KEY WORDS** cardiovascular risk, familial hypercholesterolemia, machine learning