American Heart Journal Plus: Cardiology Research and Practice 11 (2021) 100050

Contents lists available at ScienceDirect

American Heart Journal Plus: Cardiology Research and Practice

journal homepage: www.sciencedirect.com/journal/american-heart-journal-plus-cardiology-research-and-practice

Editorial

Emerging role of machine learning in cardiovascular disease investigation and translations

ARTICLE INFO

Keywords
Cardiovascular disease
Clinical diagnosis
Machine learning
Prediction models

ABSTRACT

Unexpected insights and practical advances in cardiovascular disease (CVD) are being discovered by rapidly advancing developments in supercomputers and machine learning (ML) software algorithms. These have been accelerated during the COVID-19 pandemic, and the resulting CVD translational implications of ML are steering new measures of prevention and treatment, new tools for objective clinical diagnosis, and even opportunities for rethinking basic foundations of CVD nosology. As the usual cardiovascular specialist may not be familiar with these tools, the editor has invited this brief overview.

Unexpected insights and practical advances in cardiovascular disease (CVD) are being discovered by rapidly advancing developments in supercomputers and machine learning (ML) software algorithms. These have been accelerated during the COVID-19 pandemic, and the resulting CVD translational implications of ML are steering new measures of prevention and treatment, new tools for objective clinical diagnosis, and even opportunities for rethinking basic foundations of CVD nosology [1–5]. As the usual cardiovascular specialist may not be familiar with these tools, the editor has invited this brief overview.

1. Types of machine learning

ML is a subcategory within the broad discipline of Artificial Intelligence (AI). AI is the general concept of employing machines and mathematical algorithms that aspire to ultimately make computers behave with the capabilities of the human mind. AI is a lofty moving-target goal spanning a wide umbrella of technological advances that includes various subgroups such as ML which are more practical and less existential. ML uses computer programs in a dynamic self-altering process to automatically improve outcomes through iterative self-training experiences of exploring data patterns [6]. There are four basic flavors of ML: unsupervised, supervised, reinforcement learning, and deep-learning.

- *Unsupervised ML* models learn from clustering and association patterns of unlabeled input data without human intervention. Here, the output delivers *de novo* meaningful insights to the table with minimal *a priori* assumptions. Such models include multivariate principal component analysis (PCA), k-means clustering, and hierarchical clustering. PCA is a method of reducing an unmanageably large number of complex variables in a dataset down to only two or three meaningful variables that still retain most of the original information. K-means clustering sorts through huge pools of raw data to form multiple groups each clustered around a common center value. Hierarchical clustering builds a branching tree of data organized by degree of similarity nested within a branch or comparisons between branches. One medical example of unsupervised ML is predicting or managing heart disease based on anomalies found within a battery of patient characteristics such as age, body mass index, sex, lab values, lifestyle, *etc.* Another example is the exploration of Fibonacci "Golden Ratio" relationships that occur in nature and throughout anatomy and physiology; for instance, invoking ML to look for fractal patterns has been proposed in interpreting transthoracic echocardiography data to unmask branching arrangements of coronary artery lesions responsible for ST-segment elevation myocardial infarctions, in order to assess abnormalities in cardiac dimensions and to examine blood pressure patterns that may predict decompensation in congestive heart failure patients [7,8].
- *Supervised ML* involves labeled data based on predictive models supplied by a human operator. Here, an algorithm arrives at useful insights based on self-learning from previous data sets. For example, support vector machine (SVM) models [9] can classify ischemic heart disease as simple binary states of "good or bad" based on template electrocardiograms (ECGs) or on ejection fraction data measured by single photon emission tomography and echocardiography images. Other supervised models incorporate Bayesian networks, random forests, or least absolute shrinkage and selection operator (LASSO) regressions.
- *Reinforcement ML* engages algorithms that either discover errors or bolster findings based on reward feedback from human intervention. This approach is being deployed for use in drug prescription recommendations and in imaging problems.
- *Deep-learning ML* can be either supervised or unsupervised in enabling a computer to teach itself by exposure to vast datasets using multiple layers of artificial neural network computational nodes. Such algorithms can stitch together small samples of data scalable into a big picture with a predictable pattern. For example, deep-learning algorithms that combine retinal blood vessel images with

https://doi.org/10.1016/j.ahjo.2021.100050
Received 10 June 2021; Received in revised form 2 August 2021; Accepted 1 September 2021
Available online 8 October 2021

clinical lab values can detect various CVD risk factors with accuracy comparable to human experts [10].

2. CVD machine learning on the street

Collectively, ML algorithms can use clinical and preclinical data of electrocardiography and echocardiography interpretations, MRI and computed tomography (CT) imaging diagnoses, scintigraphic myocardial perfusion imaging, and lab chemistry values to address predictions or diagnoses of stroke, heart failure, arrhythmias, coronary artery disease, acute myocardial infarction, hospital recidivism, transplant decisions, and other CVD issues requiring intervention decisions [11,12]. Many ML algorithms have been shown to validate such risk prediction models reliably, yielding receiver operating characteristic area under the curve (ROC AUC) values in the range of ~0.7 to 0.9 [11,12].

While numerous ML algorithms can explain the pathophysiological “why” underlying a given disorder, other approaches such as deep-learning engage “black-box” mathematical models that are of lesser value because they are devoid of extrapolation to actual medical insights. ML is gaining traction as an invaluable tool for advancing academic CVD basic science research discoveries, although literature meta-analyses reveal that considerably more work is required to translate ML algorithms into everyday clinical practices of reliable precision cardiology and patient management outside of major academic medical research institutions [2].

To improve accuracy and to validate results, large databases from high quality clinical trials are necessary to train the models. This is especially important for the elusive holy grail of assessing CVD risk in slowly emerging prodromic patients or mining risk assessment in electronic health records of asymptomatic patients. While the Framingham Risk Score was included in the American College of Cardiology/American Heart Association 2010 guideline [13], this traditional model is arguably not optimal because its incomplete variable set overestimates false positive CVD risk. To enhance diagnosis, the American Heart Association's Institute for Precision Cardiovascular Medicine [14] introduced patient genomic data into the mix as a laudable contributing factor.

Recent advances in microbiome metagenomics are contributing additional useful dimensions in applying ML to CVD. For example, assessing multifaceted gut microbiome–host interactions in mood disorders comorbid with CVD has surprisingly uncovered pathophysiological causal relationships that tie together heretofore strange bedfellow clinics: cardiology, psychiatry, and gastroenterology [15]. This is opening doors to integrate patient management across these traditionally siloed disciplines. One such illustration is the application of ML that unraveled a new putative phenotype of CVD denoted “depressive-hypertension” [15]. Here, hypertensive patients comorbid with depression harbor a unique signature of gastrointestinal microbial species' functional genomics in a model that disrupts central control of both blood pressure and mood. Further in these patients, ML analyses uninterestingly implied that oral ACE inhibitor drugs in the gut lumen exhibit antibiotic pharmacology that competitively inhibits key enzymes of proinflammatory gut bacterial cell wall biosynthesis associated with depression and hypertension [15]. Future expansion of these ML studies may lead to new or repurposed drugs to treat resistant hypertension.

3. Bioethical considerations

Unprecedented integration of facial expression, other “body language”, and volumes of lab data present unlimited opportunities for patient management. For example, imaging of facial expressions among patients in the critical care setting who may be too ill or sedated to express their discomfort has the potential to improve their management. The COVID-19 pandemic mandated virtual outpatient visit sessions revealing latent opportunities to apply ML to decipher facial expressions for the video images that would otherwise have not been captured during an in-person visit with a masked patient.

Nevertheless, making accurate predictions based on extensive use of personal biometric data raises the poignant issue that ML scientists and clinicians must rigorously address many ethical challenges. Attention must be paid to minimalizing biases and skewed representations relating to sex/gender, race, stereotypes, and physical limitations identities. Furthermore, in the light of the American Medical Association policy on AI in healthcare [16], ultimately clinicians must continue to treat patients humanistically regardless of the use of objective tools at their disposal. Emergence of journals such as *Intelligence-Based Medicine* (https://www.sciencedirect.com/journal/intelligence-based-medicine) shall facilitate this task.

4. Rolling out machine learning for CVD

ML holds much promise for the future of cardiology investigations and clinical translations. However, before ML becomes commonplace with FDA medical device approval, a considerable number of knowledge gaps exist. For example: i) numerous technological and societal/ethical challenges remain concerning standardization of gathering CVD data; and of ii) assessing model performances [12,17]. Once the models are established as robust and ethically unbiased, then iii) infrastructure matters must be addressed such as supercomputer processing requirements for ambulatory wearable sensors, clinic-based online cloud needs, massive data storage requirements, real-time needs for interventional procedures, and ethical/social implications of equitable data access. Nevertheless, the future is bright for AI/ML as applied to the cardiovascular system.

Funding

Dr. Pepine receives support from NIH/NHLBI grants R01 HL132448 and R01 HL146158.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R. Cuocolo, T. Perillo, E. De Rosa, L. Ugga, M. Petretta, Current applications of big data and machine learning in cardiology, J. Geriatr. Cardiol. 16 (2019) 601–607, https://doi.org/10.11909/j.issn.1671-5411.2019.08.002.

[2] C. Krittanawong, H.U.H. Virk, S. Bangalore, Z. Wang, K.W. Johnson, R. Pinotti, H. Zhang, S. Kaplin, B. Narasimhan, T. Kitai, U. Baber, J.L. Halperin, W.H.W. Tang, Machine learning prediction in cardiovascular diseases: a meta-analysis, Sci. Rep. 10 (2020) 16057, https://doi.org/10.1038/s41598-020-72685-1.

[3] K. Seetharam, S. Shrestha, P.P. Sengupta, Artificial intelligence in cardiovascular medicine, Curr. Treat. Options Cardiovasc. Med. 21 (2019) 25, https://doi.org/10.1007/s11936-019-0728-1.

[4] F. Lopez-Jimenez, Z. Attia, A.M. Arruda-Olson, R. Carter, P. Chareonthaitawee, H. Jouni, S. Kapa, A. Lerman, C. Luong, J.R. Medina-Inojosa, P.A. Noseworthy, P. A. Pellikka, M.M. Redfield, V.L. Roger, G.S. Sandhu, C. Senecal, P.A. Friedman, Artificial intelligence in cardiology: present and future, Mayo Clin. Proc. 95 (2020) 1015–1039, https://doi.org/10.1016/j.mayocp.2020.01.038.

[5] K.W. Johnson, J. Torres Soto, B.S. Glicksberg, K. Shameer, R. Miotto, M. Ali, E. Ashley, J.T. Dudley, Artificial intelligence in cardiology, J. Am. Coll. Cardiol. 71 (2018) 2668–2679, https://doi.org/10.1016/j.jacc.2018.03.521.

[6] M.I. Jordan, T.M. Mitchell, Machine learning: trends, perspectives, and prospects, Science 349 (2015) 255–260, https://doi.org/10.1126/science.aaa8415.

[7] C.M. Gibson, W.J. Gibson, S.A. Murphy, S.J. Marble, C.H. McCabe, M.P. Turakhia, A.J. Kirtane, J. Karha, J.M. Aroesty, R.P. Giugliano, E.M. Antman, And TIMI study group, Association of the Fibonacci Cascade with the distribution of coronary artery lesions responsible for ST-segment elevation myocardial infarction, Am. J. Cardiol. 92 (2003) 595–597, https://doi.org/10.1016/s0002-9149(03)00731-8.

[8] E. Yetkin, S. Ozturk, B. Cuglan, H. Turhan, Golden ratio in congestive heart failure: a promising proportion for prognosis and decompensation, Cardiol. J. 27 (2020) 904–905, https://doi.org/10.5603/CJ.2020.0177.

[9] S.S. Yadav, S.M. Jadhav, R.G. Bonde, S.T. Chaudhari, Automated cardiac disease diagnosis using support vector machine, in: 2020 3rd International Conference on

Communication System, Computing and IT Applications (CSCITA), 3-4 April 2020, 2020, pp. 56–61, https://doi.org/10.1109/CSCITA47329.2020.9137817.

[10] C.Y. Cheung, D. Xu, C.-Y. Cheng, C. Sabanayagam, Y.-C. Tham, M. Yu, T.H. Rim, C. Y. Chai, B. Gopinath, P. Mitchell, R. Poulton, T.E. Moffitt, A. Caspi, J.C. Yam, C. C. Tham, J.B. Jonas, Y.X. Wang, S.J. Song, L.M. Burrell, O. Farouque, L.J. Li, G. Tan, D.S.W. Ting, W. Hsu, M.L. Lee, T.Y. Wong, A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre, Nat. Biomed. Eng. (2020), https://doi.org/10.1038/s41551-020-00626-4.

[11] S.F. Weng, J. Reps, J. Kai, J.M. Garibaldi, N. Qureshi, Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS One. 12 (2017), e0174944 https://doi.org/10.1371/journal.pone.0174944.

[12] S. Ranka, M. Reddy, A. Noheria, Artificial intelligence in cardiovascular medicine, Curr. Opin. Cardiol. 36 (2021) 26–35, https://doi.org/10.1097/HCO.0000000000000812.

[13] P. Greenland, J.S. Alpert, G.A. Beller, E.J. Benjamin, M.J. Budoff, Z.A. Fayad, E. Foster, M.A. Hlatky, J.M. Hodgson, F.G. Kushner, M.S. Lauer, L.J. Shaw, S. C. Smith Jr., A.J. Taylor Jr., W.S. Weintraub Jr., N.K. Wenger Jr., A.K. Jacobs Jr., American College of Cardiology Foundation/American Heart Asscoiation Task Force on Practice Guidelines, 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, Circulation 122 (2010) 2748–2764. https://doi.org/10.1161/CIR.0b013e3182051bab.

[14] S.R. Houser, The American Heart Association's new institute for precision cardiovascular medicine, Circulation 134 (2016) 1913–1914, https://doi.org/10.1161/CIRCULATIONAHA.116.022138.

[15] B.R. Stevens, C.J. Pepine, E.M. Richards, S. Kim, M.K. Raizada, Depressive-hypertension: A proposed human endotype of brain/gut microbiome dysbiosis, Am. Heart J. (2021) 100540, https://doi.org/10.1016/j.ahj.2021.05.002.

[16] E. Crigger, C. Khoury, Making policy on augmented intelligence in health care, AMA J. Ethics 21 (2019) E188–E191, https://doi.org/10.1001/amajethics.2019.188.

[17] US_FDA, Artificial Intelligence/Machine Learning (AI/ML)-based_Software as a Medical Device (SaMD) action plan. https://www.fda.gov/media/145022/download, 2021.

Bruce R. Stevens[a], Carl J. Pepine[b,*]

[a] *Department of Physiology and Functional Genomics, University of Florida College of Medicine, Gainesville, FL, USA*

[b] *Division of Cardiovascular Medicine, University of Florida College of Medicine, Gainesville, FL, USA*

[*] Corresponding author at: Division of Cardiovascular Medicine, University of Florida College of Medicine, PO Box 100288, Gainesville, FL 23610, USA.

E-mail address: carl.pepine@medicine.ufl.edu (C.J. Pepine).