JACC: ADVANCES VOL. 3, NO. 9, 2024

EDITORIAL COMMENT

Can AI Find the Needle in a Haystack?

The Ongoing Search for Undiagnosed Cardiac Amyloidosis

Timothy J. Poterucha, MD,[a] Christopher M. Haggerty, PhD,[b,c] Pierre Elias, MD[a,b,c]

An accurate diagnosis is the *sine qua non* of medicine. With it, treatment and everything else follows. Without it, we are groping in the dark.

The story of cardiac amyloidosis has been one of relentless progress.[1] Two decades ago, the disease was thought to be both vanishingly rare and a death sentence. Increased awareness coupled with improved diagnostics has revealed the disease to be far more common than previously thought. Transthyretin cardiac amyloidosis (ATTR-CA), the most common subtype, is present in up to 11% of hospitalized patients with heart failure with preserved ejection fraction and may be present in upward of 1% to 2% of the general elderly population.[2] Improved understanding of the biology of ATTR-CA has unlocked new and successful therapeutics. These range from transthyretin stabilization using tafamidis or acoramidis to gene silencing using patisiran or inotersen to even gene therapy to delete the causative gene entirely.[3-5] However, all these therapies are inherently stabilizing in nature and no success has yet been had in treatments to remove amyloid that has already deposited. Thus, early recognition of the disease at its most treatable stage remains paramount. The field has increasingly asked the following question: could artificial intelligence (AI) play a role in identifying those patients with the early-stage disease most amenable to treatment?

In this issue of *JACC: Advances*, Vrudhula et al[6] attempt to answer this question and more fully understand how certain patient selection may impact the accuracy of AI models. In this single-center retrospective study, the authors studied the utility of 12-lead electrocardiogram (ECG) models to detect cardiac amyloidosis. These AI models analyze the ECG waveform to distinguish cases of cardiac amyloidosis from controls, a task they can accomplish relatively well with an area under the receiver operator characteristic curve (AU-ROC) as high as 0.898. Achieving such high performance is consistent with the general diagnostic capabilities that ECG-based AI models have demonstrated in recent years. Indeed, the top line results of this study are nearly identical to several studies classifying ATTR-CA from ECGs with deep learning with AU-ROCs of 0.85 to 0.91.[7-9]

The more important contributions from this analysis are the insights derived from the systematic evaluation of case and control selection on model output. As the authors convincingly demonstrate, this is not an idle question but instead one on which the success or failure of an AI program is critically dependent. Supervised learning, the most common type of AI, relies upon labeled data sets to train models to distinguish patterns. In the case of a model to detect cardiac amyloidosis, an AI model is developed by providing some type of data (in this case, an ECG) from patients with cardiac amyloidosis and the same type of data from controls without the disease. If provided enough data, the model can learn the ECG patterns associated with cardiac amyloidosis and may be able to identify cases of cardiac amyloidosis that it has not previously seen.

In this article, the authors tested multiple methods of identifying cases of cardiac amyloidosis such as curated cardiac amyloidosis clinic lists or amyloidosis

From the [a]Seymour, Paul, and Gloria Milstein Division of Cardiology, Department of Medicine, Columbia University Irving Medical Center, New York, New York, USA; [b]Department of Biomedical Informatics, Columbia University, New York, New York, USA; and the [c]Information Technology Data Science, NewYork-Presbyterian Hospital, New York, New York, 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.2024.100999

by diagnosis codes with or without supportive findings such as increased left ventricle thickness by echo or a positive troponin or brain natriuretic peptide.[6] These amyloidosis cases were combined with variously defined controls such as all patients; all patients matched by age, sex, wall thickness, or QRS amplitude; or patients with similar pathologic features, such as left ventricular hypertrophy or heart failure by various metrics. With these various combinations of cases and controls, the authors report an enormous variability in accuracy with AU-ROCs ranging from 0.467 to 0.898 depending on the selected model and test set.

Several trends emerged from this analysis. First, good performance in a test set mirroring the training/development set does not guarantee good performance in a broader population. While this is a known fundamental concept in data science and machine learning, empirical demonstration can still provide useful reinforcement. Second, incorporating a broad set of controls is critical—none of the models developed with narrow control definitions (ie, restricting to heart failure or left ventricular hypertrophy) were viable. Such a strategy could be reasonably motivated by forcing the model to learn to make difficult distinctions in a clinical diagnostic context, but such cases should still be part of a broader population if that is the intended use of the resulting model. Finally, while performance metrics varied for models across the spectrum of case definitions, all demonstrated a potential to generalize in this analysis. As the authors note, this suggests that specialized centers with highly curated registries are not a prerequisite for evaluating or developing models for rare diseases such as ATTR-CA.[6] While explicit generalization of these patterns for other disease states is needed, it is reasonable to presume that they may similarly apply to modeling studies of other rare cardiovascular conditions of interest, such as hypertrophic or arrhythmogenic cardiomyopathies.

What are the limitations of this analysis? First, this is a single-center study, and the external generalizability of these findings is uncertain. Second, the sensitivity and specificity of this approach may not support widespread deployment. At the Youden index, the model has a sensitivity of 61% and specificity of 72% for a positive predictive value of 0.018. Even when specificity is maximized at 97%, modeling demonstrates that out of 10,000 patients screened, a total of 275 would be flagged as positive and only 18 of those patients would be found to have amyloidosis. Maintenance of a funnel that starts with 10,000 patients and yields only 18 cases may prove untenable. Third and most importantly, this manuscript and other prior analyses are retrospective. We truly do not know if clinical pathways deploying these models will function effectively. It may be that the patients already diagnosed with cardiac amyloidosis are the "low hanging fruit" with the most phenotypically distinct features. The remaining cases may be those patients with subtle, difficult to detect findings. In addition, given the association between age and cardiac amyloidosis, it is probable that many undiagnosed patients will have concomitant diseases that preclude benefiting from the diagnosis, particularly given the exorbitant cost of cardiac amyloidosis therapies.

The authors are to be congratulated for carrying out this complex, rigorous analysis. The finding that using similar case and control definitions as the planned deployment is likely a generalizable one in AI research, and this manuscript continues to build the evidence basis for using AI to detect cardiac amyloidosis. The next frontier is clear: we must move beyond the retrospective and carry out the essential clinical trials to determine if we can diagnose cardiac amyloidosis with these technologies.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

Dr Poterucha owns stock in Abbott Laboratories and Baxter International with research support provided to his institution from the Amyloidosis Foundation, American Heart Association (Awards #933452 and #23SCISA1077494, https://doi.org/10.58275/AHA.23SCISA1077494.pc.gr.172160), Eidos Therapeutics, Janssen, Pfizer, and Edwards Lifesciences. Drs Poterucha, Haggerty, and Elias are inventors on a provisional patent to use deep learning analysis of the electrocardiogram to detect structural heart disease. This work was funded by the American Heart Association (Awards #933452 and #23SCISA1077494, https://doi.org/10.58275/AHA.23SCISA1077494.pc.gr.172160).

ADDRESS FOR CORRESPONDENCE: Dr Timothy J. Poterucha, Columbia University Medical Center, 177 Fort Washington Avenue, Milstein 2, New York, New York 10032, USA. E-mail: tp2558@cumc.columbia.edu.

REFERENCES

1. Ruberg FL, Grogan M, Hanna M, Kelly JW, Maurer MS. Transthyretin amyloid cardiomyopathy. *J Am Coll Cardiol*. 2019;73:2872-2891.

2. Gonzalez-Lopez E, Gallego-Delgado M, Guzzo-Merello G, et al. Wild-type transthyretin amyloidosis as a cause of heart failure with preserved ejection fraction. *Eur Heart J*. 2015;36:2585-2594.

3. Maurer MS, Schwartz JH, Gundapaneni B, et al. Tafamidis treatment for patients with transthyretin amyloid cardiomyopathy. *N Engl J Med*. 2018;379:1007-1016.

4. Gillmore JD, Judge DP, Cappelli F, et al. Efficacy and safety of acoramidis in transthyretin amyloid cardiomyopathy. *N Engl J Med*. 2024;390:132-142.

5. Ioannou A, Fontana M, Gillmore JD. RNA targeting and gene editing strategies for transthyretin amyloidosis. *BioDrugs*. 2023;37:127-142.

6. Vrudhula A, Stern L, Cheng PC, et al. Impact of case and control selection on training artificial intelligence screening of cardiac amyloidosis. *JACC Adv*. 2024;3:100998.

7. Goto S, Mahara K, Beussink-Nelson L, et al. Artificial intelligence-enabled fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms. *Nat Commun*. 2021;12:2726.

8. Grogan M, Lopez-Jimenez F, Cohen-Shelly M, et al. Artificial intelligence-enhanced electrocardiogram for the early detection of cardiac amyloidosis. *Mayo Clin Proc*. 2021;96:2768-2778.

9. Tison GH, Zhang J, Delling FN, Deo RC. Automated and interpretable patient ECG profiles for disease detection, tracking, and discovery. *Circ Cardiovasc Qual Outcomes*. 2019;12:e005289.

KEY WORDS amyloidosis, artificial intelligence, cardiac amyloidosis