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

VIEWPOINT

AI and Heart Failure

Present State and Future With Multimodal Large Language Models

Baljash Cheema, MD, MSCI, MSAI,[a,b] Jay Pandit, MD[c]

Machine learning and the larger field of artificial intelligence (AI) have the potential to revolutionize health care by overcoming the limitations of the human mind to ingest large quantities of clinical and scientific data and use it in real time for complex decision-making.[1] With advances in deep learning including a novel neural network architecture known as the transformer, AI can now take large sources of patient data, derived from both clinical care and a plethora of sensor-based devices from both the biome and exposome of each individual, and generate clinical insights.[2] This is built upon major advances over the last decade-plus across the field of AI, including breakthroughs in computer vision, reinforcement learning, natural language processing, and generative AI. This has led to the creation of large language models (LLMs), which have several surprising capabilities far beyond the interpretation and creation of language. These models now have the ability to incorporate multimodal data including text, video, sound, and sensor data, with many experts predicting a future with more timely diagnostics, optimized treatment protocols, reduced administrative load related to documentation and billing, less clinician burnout, and overall greater quality of care for patients with fewer medical mistakes.[2]

Yet, in 2024, claims that AI will change medicine for the better are far from new, and many in the field of medicine are wondering when they will see tools that impact the patients walking through the doors of their respective clinics each day. Indeed, while there is no shortage of new AI products ready for the marketplace, with over 650 Food and Drug Administration (FDA)-cleared AI algorithms at the time of writing this article, there is a gap in the integration and delivery of these technologies into clinical care.[3] This may be due in part to the breakneck speed at which AI technology is exploding forward, leaving clinicians without a clear understanding of what products exist and how they can be used right now, as well as several other challenges related to implementing AI in practice.

In this piece, we will share concrete examples of how AI can be used now to impact diagnostic and management decisions in patient care through the discussion of a real-life case of a patient with advanced heart failure. We will highlight both the strengths and challenges of the current state of AI in medicine by exploring a handful of representative technologies, before envisioning a future state with patient care augmented by multimodal LLMs.

Our case is that of a 65-year-old African American male, a veteran of the armed services, with recurrent hospitalizations for heart failure. Prior to his heart failure diagnosis, he had several years of multisystem symptoms including gastrointestinal distress, neuropathic pain, and fatigue, and he underwent bilateral carpal tunnel release surgery. Once diagnosed with heart failure, several years went by before the cause was found to be hereditary transthyretin amyloidosis. He had hospitalizations for heart failure triggered by an accumulation of lower extremity fluid and shortness of breath, although he often only recognized these changes after he was in a severely debilitated

From the [a]Bluhm Cardiovascular Institute, Center for Artificial Intelligence, Northwestern Medicine, Chicago, Illinois, USA; [b]Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA; and the [c]Scripps Translational Research Institute, Scripps Research, La Jolla, California, 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.2024.101029

state. As his cardiac amyloidosis was not discovered until late in the disease course, he did not benefit from amyloid-specific treatment, and he eventually required cardiac transplantation.

This case provides several examples of the challenges in managing patients with heart failure, where there remains substantial residual risk for morbidity and mortality even for those that tolerate all 4 pillars of guideline-directed medical therapy. In this specific case, a delay in identifying the cause of his cardiomyopathy played a major role in disease progression, as available therapies for cardiac amyloidosis can slow disease progression but cannot reverse the course, so they must be given in a timely fashion. Additionally, there were many periods of subclinical hemodynamic congestion, which may have provided windows of opportunity for intervention prior to hospitalizations for heart failure. Further, while he went on to successfully receive a heart transplant, others in similar circumstances are referred to specialty centers too late in the disease course of heart failure to merit consideration for advanced therapies such as ventricular assist devices and transplantation, thus requiring vigilance on the part of busy clinicians to identify late-stage disease within the “golden window” for referral.

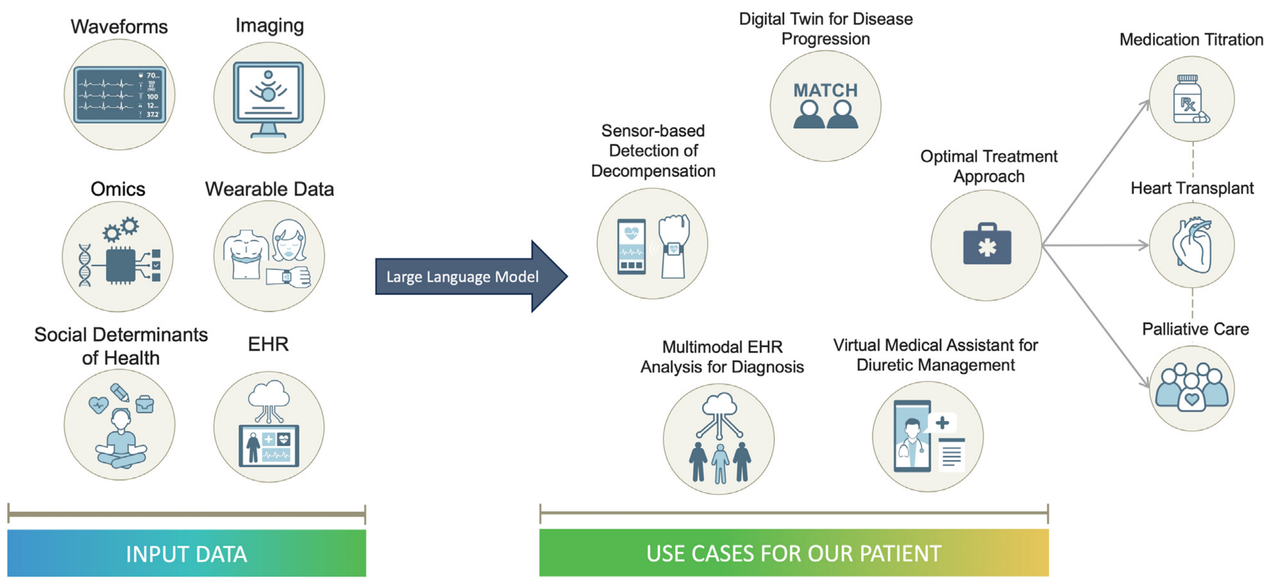
Each of these points could have been individually addressed with distinct AI tools in existence now, some available on the marketplace with FDA clearance and others created in academia and discussed in peer-reviewed journals but not yet available commercially. The patient’s electrocardiogram (ECG) was the first potential point of contact, as AI-driven ECG models have demonstrated an area under the receiver operating characteristic curves above 0.90 in diagnosing cardiac amyloidosis.[4] When echocardiograms are added, performance improves further, and models can distinguish between phenotypically similar disease states like hypertrophic cardiomyopathy, hypertensive heart disease, and left ventricular hypertrophy from end-stage renal disease.[5] Additionally, both invasive and noninvasive technologies that allow for reliable at-home monitoring for heart failure have the ability to clue patients and clinicians into impending heart failure events before they happen, allowing space for timely intervention.[6] While these technologies are currently only being used to detect heart failure exacerbations, their potential to identify novel biomarkers remains untapped. Lastly, AI models have been created that, when applied to the profile of a patient within an electronic health record, can predict whether or not they may have advanced heart failure and warrant consideration for advanced therapies in hopes of identifying patients at optimal times in their disease state.[7]

However, there are several challenges related to implementing these tools in practice, and all contribute to the poor uptake of AI-based technologies on the frontlines of clinical care. As these tools work in isolation for individual tasks, they require clinical champions to identify their relevance and push for adoption into existing workflows. At the same time, these tools are rarely investigated prospectively and in a randomized fashion, and even FDA-cleared devices often do not have available data for review on model performance in populations distinct from those in which the models were created. As such, there is warranted concern about model performance, bias, and drift, leaving clinicians to question whether their patients will truly benefit in the real world. Issues of model governance, reimbursement schemes, interoperability, privacy and protection of personal health information, trust in AI, and alert management for false positive results are just some of a number of other ongoing challenges.[8]

Importantly, none of these challenges are insurmountable, and with the potential impact AI may have on clinical medicine, it is imperative that we continue to work on them, especially when considering the added potential that multimodal LLMs bring.

Modern day LLMs are built upon the transformer model, a model that is particularly skilled at understanding sequential data by using the concept of attention and learning dependencies within the data, even if separated in time and space.[9] While unimodal data input into these models may allow for the detection of patterns the human eye may not readily see, multimodal data input has the potential to unearth patterns that were previously unidentified by applying attention across disparate but connected data sources.[2,9] In the current state of health care, the most advanced models integrated into clinical care are mostly unimodal or bimodal at best, using up to 2 data modalities such as the ECG and echocardiography. However, recent advances in LLMs allow for multimodality data input, creating the potential to combine text from clinical notes or patient-generated descriptions of their health state, sound from acoustic physiologic signals or voice recordings, images and videos captured by patients at home or in the clinic

FIGURE 1 Use Cases for Multimodal Large Language Models for Heart Failure



Depicted are several possible sources of input data, acknowledging that many more exist, and use cases for utilizing large language models capable of incorporating varied forms of data, each of which could have helped our patient. EHR = electronic health record.

with medical equipment, and multidimensional omics data all for input into a given model. This could have a significant impact on clinical decision-making, where clinicians have historically relied on much less complete knowledge representation before choosing an action.[2,10]

In the case of our patient, there may be a future in which data entry into a multimodal LLM improves his care in concrete ways (Figure 1). At the level of preventive screening, his demographics, basic blood test results, and text from clinical notes all within the electronic health record could have been used as input to models such as Generative Pre-trained Transformer 4 or Med-PaLM to predict high-risk features and clue clinicians into the fact that a unifying disease process was underlying his various symptoms.[2] Once a diagnosis of heart failure was established, data streams from digital biomarkers captured via wearable sensors such as heart rate variability, gait analysis, weight or congestion indices, and electromyography could have forecasted disease progression and periods of acute exacerbation and allowed for medication titration. He may have even utilized ECG capture via a smart device, a text or voice-based description of his symptoms analyzed by a publicly available LLM, a sensor-derived assessment of his hemodynamics, and a genetic screen from a direct-to-consumer commercial platform, fed into a multimodal AI model, to share with his doctor at his first appointment insights into diagnostic considerations and preferred treatment approaches they should discuss instead of hoping that a diagnosis would eventually be discovered for him.

While this future state is ideal and requires a substantial amount of work, the field is nascent, and there is real reason to believe the hype. It will be key to address consistent barriers around system-wide integration of these types of tools including but not limited to adequate governance structures, improvements in model explainability, frameworks designed to guide implementation and track post-implementation outcomes, and safety checks that ensure advanced models improve and not worsen inequity.[8] We believe working toward this end is in keeping with developing learning health systems, where patient-caregiver-provider relationships are evolving to share information rather than being one-dimensional, for the betterment of patient care.

There are great reasons to hope for a future state where clinical care continues to improve through the incorporation of machine learning and AI models into everyday practice. While certain AI-based tools could have improved the care of our patient when used by a skilled and informed clinician, the future may empower patients like him to make their own diagnostic and management decisions, guided by highly evolved AI systems. Addressing barriers to implementation is of critical concern, as more powerful AI models capable of handling multimodal data are available and markedly increase the capabilities of machine intelligence.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

Dr Cheema has received consulting fees from Caption Health, Inc and Viz.ai; served on an advisory board for Novo Nordisk; is an advisor with equity interest in Healthspan, Inc, and Zoe Biosciences; and has received speaking fees and honoraria for Bristol Myers Squibb. Dr Pandit has received consulting fees from QuidelOrtho; honoraria from Niterra; has participated on a data safety monitoring board or advisory board for Cardiosense, Preciseli, Bodyport, Angiotensin Tx, and Lytic; and has stock or stock options in Cardiosense, Preciseli, Bodyport, and Angiotensin Tx.

ADDRESS FOR CORRESPONDENCE: Dr Baljash Cheema, Northwestern University, Feinberg School of Medicine, 676 N. Saint Clair Street, Arkes Pavilion, Suite 600, Chicago, Illinois 60611, USA. E-mail: bcheema@nm.org. OR Dr Jay Pandit, Scripps Translational Research Institute, Scripps Research, 3334 North Torrey Pines Court, Suite 300, La Jolla, California 92037, USA. E-mail: jpandit@scripps.edu.

REFERENCES

1. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28:31-38.

2. Boonstra MJ, Weissenbacher D, Moore JH, Gonzalez-Hernandez G, Asselbergs FW. Artificial intelligence: revolutionizing cardiology with large language models. *Eur Heart J*. 2024;45:332-345.

3. Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *NPJ Digit Med*. 2023;6:120.

4. 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.

5. 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.

6. Boehmer JP, Hariharan R, Devecchi FG, et al. A multisensor algorithm predicts heart failure events in patients with implanted devices: results from the MultiSENSE study. *J Am Coll Cardiol HF*. 2017;5:216-225.

7. Cheema B, Mutharasan RK, Sharma A, et al. Augmented intelligence to identify patients with advanced heart failure in an integrated health system. *JACC: Adv*. 2022;1:100123.

8. Marwaha JS, Landman AB, Brat GA, Dunn T, Gordon WJ. Deploying digital health tools within large, complex health systems: key considerations for adoption and implementation. *NPJ Digit Med*. 2022;5:13.

9. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *Adv Neural Inf Process Syst*. 2017;30.

10. Topol EJ. As artificial intelligence goes multimodal, medical applications multiply. *Science*. 2023;381:eadk6139.

KEY WORDS artificial intelligence, heart failure, large language models, machine learning, multimodality data