Published in final edited form as:
*J Am Coll Cardiol*. 2024 June 18; 83(24): 2487–2496. doi:10.1016/j.jacc.2024.03.401.

# Advances in Artificial Intelligence for Cardiovascular Care—Part 2: Latest Developments & Deployment in Clinical Care

**Sneha Jain, MD, MBA**[*,a], **Pierre Elias, MD**[*,b,c], **Timothy Poterucha, MD**[b], **Michael Randazzo, MD**[d], **Francisco Lopez Jimenez, MD, MBA**[e], **Rohan Khera, MD, MS**[f], **Marco Perez, MD**[a], **David Ouyang, MD**[g], **James Pirruccello, MD**[h], **Michael Salerno, MD, PhD**[a], **Andrew Einstein, MD, PhD**[b], **Robert Avram, MD**[i], **Geoff Tison, MD, MPH**[h], **Girish Nadkarni, MD, MPH**[j], **Vivek Natarajan, MS**[k], **Emma Pierson, PhD**[l], **Ashley Beecy, MD**[m,n], **Deepa Kumaraiah, MD, MBA**[b,m], **Chris Haggerty, PhD**[c,m], **Jennifer N. Avari Silva, MD**[**,o], **Thomas M. Maddox, MD, MSc**[**,o]

[a]Division of Cardiology, Stanford University School of Medicine; Palo Alto, CA

[b]Seymour, Paul and Gloria Milstein Division of Cardiology, Columbia University Irving Medical Center; New York, NY

[c]Department of Biomedical Informatics Columbia University Irving Medical Center; New York, NY

[d]Division of Cardiology, University of Chicago Medical Center; Chicago, IL

[e]Department of Cardiology, Mayo Clinic College of Medicine; Rochester, MN

[f]Division of Cardiology, Yale School of Medicine; New Haven, CN

[g]Division of Cardiology, Cedars-Sinai Medical Center; Los Angeles, CA

[h]Division of Cardiology, University of California San Francisco, San Francisco, CA

[i]Division of Cardiology, Montreal Heart Institute, Montreal, CA

[j]Icahn School of Medicine at Mount Sinai; New York, NY

[k]Google Health; Mountain View, CA

[l]Department of Computer Science, Cornell Tech; New York, NY

[m]NewYork-Presbyterian Health System; New York, NY

[n]Division of Cardiology, Weill Cornell Medical College; New York, NY

[o]Division of Cardiology, Washington University School of Medicine; St. Louis, MO

## Abstract

Recent Artificial Intelligence (AI) advancements in cardiovascular care offer potential enhancements in effective diagnosis, treatment, and outcomes. Over 600 Food and Drug Administration (FDA)-approved clinical AI algorithms now exist, with 10% focusing on cardiovascular applications, highlighting the growing opportunities for AI to augment care.

**Address for Correspondence:** Dr. Thomas M. Maddox, BJC HealthCare/Washington University School of Medicine, Mailstop 90-29-933, 4590 Nash Way, St. Louis, MO, 63110. thomas.maddox@bjc.org.
[*]These co-first authors contributed equally to this manuscript
[**]These co-last authors contributed equally to this manuscript

This review discusses the latest advancements in the field of AI, with a particular focus on the utilization of multimodal inputs and the field of generative AI. Further discussions in this review involve an approach to understanding the larger context in which AI-augmented care may exist, and include a discussion of the need for rigorous evaluation, appropriate infrastructure for deployment, ethics and equity assessments, regulatory oversight, and viable business cases for deployment. Embracing this rapidly evolving technology while setting an appropriately high evaluation benchmark with careful and patient-centered implementation will be crucial for cardiology to leverage AI to enhance patient care and the provider experience.

## Condensed abstract

Recent AI advancements have the potential to alter and augment how we practice cardiovascular clinical care. In particular, multimodal inputs and generative AI are rapidly shifting the realm of what is possible. However, realizing AI's potential in cardiology requires rigorous evaluation, appropriate infrastructure, regulatory oversight, ethics and equity analyses, and viable business cases for deployment. This review provides an overview of existing work in these arenas, to better equip the cardiology community with the tools needed to effectively understand and evaluate the use of AI tools in clinical practice.

### Keywords

Artificial intelligence; machine learning; deep learning; implementation science; cardiac imaging; digital health; innovation; generative AI; large language models; health equity; clinical trials

## Introduction

Recent strides in artificial intelligence (AI) hold promise to augment the delivery of cardiovascular care. The U.S. Food and Drug administration has approved over 600 clinical AI algorithms since the first vetted algorithm in 1995, and more than half of these algorithms have been approved within the last five years. Cardiovascular applications account for 10% of Food and Drug Administration (FDA) approved AI algorithms, second only to radiology.[1] However, only a small portion of these tools have been successfully deployed to support clinical care. It is imperative for cardiologists to understand both the promise and limitations of AI in cardiovascular care, so that they can support its responsible deployment and use.

To address this need, part two of this review explores areas of current and future work at the intersection of AI and cardiology, with a focus on emerging areas for AI in cardiovascular care, including the latest advancements in multimodal inputs and generative AI, and a review of how to deploy AI successfully and responsibly. AI models will likely not meaningfully improve care without appropriate rigorous evaluation and effective implementation, and so we discuss the growing body of work in prospective evaluation, pragmatic effectiveness trials, implementation science, and health policy that support deployment and governance of AI in clinical medicine. We also discuss critical challenges, ethical considerations, and health equity implications of AI technologies that need to be navigated for successful and widespread adoption of these tools in clinical care. To aid the reader, we have included a glossary of terms relevant to AI (Table 1). Our goal is to equip the cardiology community with the latest knowledge and context around AI development, evaluation, regulation, and deployment.

## Emerging Areas for AI in Cardiovascular Care

### Multimodality: Combining Imaging, Genetics, Electronic Health Records, and More

Cardiovascular disease is heterogeneous and polygenic, requiring integration of multimodal data to understand and address complex underlying mechanisms of disease and subsequent personalized treatment. While there are still difficulties in integrating genomic data with other important clinical variables such as cardiac imaging, angiography, and biomarker data for both discovery and clinical care using multimodal AI, this integration is a focus of current research.[2] Combining data from various modalities such as MRI and ECG allows for, for example, identification of novel genetic associations with general cardiovascular phenotypes.[3]

Another benefit of multimodality is in building models that more clearly mimic the way data is currently interpreted in clinical care.[4-6] In determining if a patient with dyspnea has a pulmonary embolism, a number of distinct data modalities are important: presenting symptoms and their time of onset, risk factors in medical history, findings on ECG, lab work such as D-dimer, and imaging could all hold critical information that changes the likelihood they have an embolism. Today the majority of AI models are built to look at data from one modality. As such they may perform well in certain use cases but be prone to errors in real-world use. A similar limitation is the frequent inability to include data across time. A clinician may infer that a patient with persistent ST elevation on prior ECGs with normal troponin and no symptoms has a left ventricular aneurysm. Most models developed to date are unable to include all prior pertinent data to learn this inference. Building models that integrate multiple modalities across varying periods of time is a key area of research, with the goal of generating accurate predictions for a patient rather than for a study.[7]

### Generative AI

Generative AI, which is a subcategory of AI in which algorithms can generate new content, has become an area of immense interest in medical care and has the potential to revolutionize multiple facets of cardiovascular care, from streamlining administrative tasks, to helping make cardiovascular care more efficient and allowing doctors to work at the top of their license, to diagnostics and treatment planning, to drug discovery and education. There has been development in using large language models (LLMs) for medical imaging interpretation,[8, 9] text generation,[10] standardized medical test completion,[11] making informed consent more accessible to patients,[12] ambient scribing,[13] and multimodal data processing.[14] However, future development in clinical cardiovascular applications will require addressing several technical and evaluation challenges.

A critical bottleneck for the clinical application of generative AI models has been the lack of availability and access to large, high quality training datasets for machine learning.[15] Tools such as ChatGPT-4 are trained on heterogenous, disparate datasets, with uneven quality and comprehensiveness of medical knowledge. As such, general models may not generalize well to unseen medical tasks. Indeed, more specific clinical language models, despite being less than 1% the size of general models like ChatGPT-3, have so far had superior performance in a variety of clinical tasks.[15] Another important barrier to LLM development has been the relative lack of scientific evaluation and rigorous benchmarking of foundation models for medical use. This need has been complicated by the opaque nature of access to proprietary and commercial LLMs and foundation models via application programing interfaces. It is unclear when and how the models are being updated and their behavior has been found to change over time.[16]

As these challenges are addressed, there may be further development of finely tuned LLMs to improve performance for highly specialized clinical tasks. For example, vision-to-language models can automatically "read" and generate catheterization reports and coronary diagrams at the end of cases. Examples of this technology have shown early success in echocardiography.[17] AI-based telescribing can support clinical documentation and summarization. This ability may lead to reductions in documentation time and burnout.[18] Additionally, generative tools may able to automatically analyze a patient's medical needs and efficiently pair them with the appropriate clinicians or care teams for care. Furthermore, there may be advancements of existing general LLMs that may improve and expand their applications in cardiovascular care. Finally, there is growing exploration of the potential and limitations of synthetic data generation in clinical care and research.[19, 20]

The potential for generative AI in cardiovascular care is broad, and it is possible that these advancements in AI may fundamentally shift how we as clinicians interact with systems to understand patient data, generate new data during the course of their care, and communicate changes in care to patients and families.

## Deployment of AI in Cardiology

While the volume of clinical AI research has increased dramatically over the past 5 years, there is a large gap between this output and clinical integration. Reasons for this gap may include: 1) mismatch between the characteristics of training versus "real-world" data,[21] 2) lack of consideration for operational context and existing clinical workflow during the design phase of AI for healthcare,[22] 3) model outputs with uncertain implications for clinical action,[23] 4) regulatory or reimbursement barriers, and 5) lack of a business case for deployment.[24, 25] Cardiology also lags behind fields such as radiology in sufficient technical infrastructure to support AI model deployment, with only partial adoption of standards such as Digital Imaging and Communication in Medicine (DICOM), an international standard for medical images and the related information, and less integration of vendor technologies into centralized locations. Considering the necessary infrastructure paves the way for an ecosystem that can better adopt new AI technologies.[26] For example, a cardiology practice is much more likely to try new AI software that can integrate into their current echocardiogram machines, rather than one that requires acquisition of new echocardiogram machines.

Considering the cost and complexity of regulation, implementation, governance, and monitoring of AI, and the challenges of managing longer-term "technical debt" associated with such systems,[27] AI use should target those areas of clinical workflow where it can provide significant value for patients, providers, and/or the healthcare system.[28] For example, if a simple prediction model (e.g., regression analysis) can perform a given task to within a few performance-index points of a complex AI model (e.g., a neural network), then investment in the AI model is unlikely to make sense.[29] Intuitively, there are many tasks for which complex AI models create opportunities not available to simple models, but critical evaluation of its benefits and costs will be essential. A number of healthcare systems are developing assessments to guide the evaluation of value of AI deployment within their system, and using this evaluation as a way to prioritize which algorithms to deploy first.[30, 31]

The following subsections delve further into some examples and discussion of the key components of successful implementation, including prospective evaluation, understanding the governance and regulatory landscape, reimbursement considerations, ethical evaluation, and assessment of equity impact. Additionally, we provide a framework for model development, testing, and deployment to help guide institutions in operationalizing high-quality AI applications in Figure 1.

## Clinical Trials and Prospective Evaluation of AI in Cardiology

Randomized trials and other prospective evaluations of AI technology in clinical practice are starting to emerge. These include trials of smart device-embedded algorithms for atrial fibrillation detection,[32-34] a comparison of AI vs. sonographers in assessing cardiac function,[35] and an evaluation of diagnostic efficacy of ECG-AI to detect low left-ventricular ejection fraction.[36] Such efforts are vital to advancing the evidence for the clinical efficacy of AI or AI-assisted clinical workflows; that is, moving beyond diagnostic or prognostic accuracy to understanding clinical impact on diagnosis, treatment, and patient outcomes. Guidelines around best practices for these clinical trials, CONSORT-AI and SPIRIT-AI, have been helpful in standardizing required evidence generation.[37, 38] These along with standardized checklists for the development and evaluation of machine learning models, such as PRIME, provide a foundation for higher quality development and evaluation efforts for AI.[39]

There has also been an evolution of “pragmatic clinical trials” that focus on assessing effectiveness in real-world settings, and often use electronic health record data and randomization at a group unit to allow for more cost-efficient and expedite recruitment and enrollment. Compared to more traditional explanatory trials, pragmatic trials tend to have less strict exclusion criteria to allow for reflection of the general population. Many AI trials of today leverage these pragmatic trial approaches.[32-34, 36] While not all AI applications in healthcare will need randomized control trials to prove safety and value, we believe this kind of rigorous evaluation is warranted for applications that affect patient outcomes, and should be focused on evaluating patient safety, implementation success metrics, clinical efficacy, and clinical effectiveness.

## The Evolving Governance and Regulatory Landscape of AI

Commercial AI-enabled devices and technology platforms are proliferating rapidly. The increasing supply and demand for clinically relevant AI underscores the need for mature regulatory and governance structures to support the growth of accurate, safe, effective, and high-quality AI-based medical devices or algorithms.[40] In November 2023, Executive Order 14110 was published by President Biden,[16] which called for the development of AI assurance policy and infrastructure in the health and human services sector to evaluate both pre- and post-market performance of AI-enabled healthcare technology.[41]

AI presents unique challenges for governance compared with other medical devices due to the iterative and data-driven nature of their development and maintenance. For example, AI models can experience “model drift,” a decline in performance due to shifts in the external environment or changes in the underlying data that is used as inputs to the model. For example, an AI model could be designed to predict the risk of undiagnosed heart failure based on variables such as medical history, lab data, and real-time health metrics from wearable devices. Overtime, there may be changes in population health trends (such as obesity rates, or diet and exercise habits in the population) that could affect heart failure risk. New treatments for heart failure risk factors, such as hypertension, could alter patient outcomes and make the model’s predictions under an old treatment paradigm less accurate in the new age. Or, there could be upgrades or changes in the technology used to collect the data that serves as inputs to the model, such as a wearable that collects more precise data. This technologic advancement may result in data generation that differs significantly from the data the model was trained on. These changes could result in model drift, and poorer performance of the AI model. To mitigate this risk, periodic or continuous monitoring of a model’s performance is essential, and iterative updates to the model’s training will likely be required. The process of ensuring responsible upkeep of AI is one of a variety of responsibilities that falls under AI governance.

Governance exists at many levels from regulatory bodies, to national organizations and collaborations across healthcare systems, to hospitals and front-line clinicians. The clinician’s role as the human in the loop for the provision of clinical care requires knowledge of the AI system and ability to integrate model output effectively and safely into existing workflows; investing in training to explicitly foster these skills should be a strategic priority for medical education. For safe deployment, the healthcare system must review the product’s intended use and integration, model performance, data inputs, potential for technical errors after implementation, plan for newly augmented AI-assisted workflows, and assessment of possible bias, and long term monitoring plan. This assessment will require leveraging multi-disciplinary expertise throughout the product lifecycle, both during and after deployment into workflow. Previously, there have been calls for an AI “Nutrition Label”, which discloses the data used for an algorithm’s output and model evaluation metrics.[42] Recently, Shah et al. outlined the need for a network of AI assurance labs, which would be tasked with reporting on model performance in the population of intended use and across subgroups to assess for bias, and an achievable benefit report to assess the potential workflow, financial, and ethical implications of machine learning-assisted model deployment.[43]

At the federal level, the FDA has published and is iteratively building proposed standards for pre-market approval, reporting of anticipated modification over the lifetime of the AI algorithm, and algorithm change protocols.[44] The standards include, for example, proactive statements of anticipated modifications to the intended use, performance, or inputs of an AI software as medical device. These include changes that the developer/manufacturer of the AI model plan to pursue while the algorithm is operational, and depicts a "region of potential changes."[44] An accompanying algorithm change protocol would outline a risk mitigation strategy of anticipated modifications, with the goal of preserving the safety and efficacy of the AI algorithm.[44] The regulatory environment is likely to continue a rapid evolution as the demand for AI tools grows and the technological infrastructure for AI deployment within medical centers builds.

## Reimbursement Models for AI

The reimbursement landscape is also evolving, and there has been significant debate on the best ways to pay for the use of AI in clinical care.[45, 46] CMS has adopted per-use payments for AI as initial policy by two pathways, Current Procedural Terminology (CPT) codes created by the American Medical Association CPT editorial panel and by establishing new technology add-on payments (NTAPs) for inpatient use to provide additional reimbursement to a hospital above what is provided through a Medicare Severity Diagnosis-Related Group (DRG) payment.[45, 47] The AMA has published guidance for Current Procedural Terminology coding across three taxonomies: Assistive AI, Augmentative AI, and Autonomous AI.[48] Outside of CMS, a number of commercial insurance plans and other health service organizations are developing digital health formularies to curate lists of vetted digital health products, some of which have AI components.[49] Finally, some hospital systems are paying vendors directly for use of AI technology.[45]

Further clarity around compensation mechanisms will help support the possible establishment of AI use in daily clinical care. Until then, most focus is placed on deploying algorithms that serve a double bottom line that improves patient care and either improves revenue or reduces costs for the customer deploying the algorithm, to ensure financial sustainability.

## Challenges, Ethics, and Equity Considerations

In addition to its technical, regulatory, and implementation challenges, AI also introduces risks along the traditional four pillars of medical ethics: autonomy, beneficence, nonmaleficence, and justice. In particular, there are ongoing efforts and calls for improving patient data security and privacy.[50] There is a need to preserve patient agency, encourage advancements in methods of data anonymization and protection, and incentivize safe and equitable dissemination of these tools to prevent a digital divide While a complete exploration of the potential ethical risks of AI is beyond the scope of this review, there are broad principles that can guide AI deployment. First, effective patient and clinician autonomy mandate communication of provenance: AI results should be clearly labeled as "AI-generated" and explainable when possible, with the caveat that limitations in current explainability methods mean that this explanation is not enough to assure AI-assisted decisions are correct, and rigorous validation and workflow testing is required.[51] Second, in these early stages of AI deployment, particular attention should be paid towards avoiding errors. In the same way that computer-assisted ECG analysis has not replaced the need for cardiologist oversight, it is critical to maintain expert clinicians "in the loop" for now to avoid unintended AI-inflicted harm. Third, there is an important need for ethical focus on justice, equity, and fairness.

Six principles should guide the design of *equitable* AI algorithms, acknowledging that there may be a natural tension and trade-off between maximizing overall algorithmic performance and ensuring similar performance in certain groups.[52, 53] First, equity criteria should be chosen relative to the clinical decision being made. In general, different equity criteria cannot all be achieved at the same time,[54-56] so practitioners must choose criteria appropriate to their use case.[57, 58] Generic "debiasing" methods which seek to equalize metrics across groups should be used with caution, since they can worsen performance.[58] Second, algorithm development should occur on diverse datasets which represent the target patient population. Diverse datasets can improve model performance, and algorithms can underperform on groups underrepresented in the data.[59-61] Subsequent model validation should occur on patient subgroups of interest; evaluating only on the entire population can conceal disparities.[52, 62, 63] Third, inclusion of demographically sensitive features — in particular, patient race — as inputs in clinical algorithms warrants careful consideration.[64-66] Previous algorithms which have included race have often adequate supporting evidence, amplified disparities, or been based on racist misconceptions.[67] At the same time, removing sensitive features is not an automatic fix for racial or other biases, and can even worsen them.[68, 69] Fourth, choice of the prediction target for the algorithm also requires careful consideration, since imperfect proxies for the desired target — for example, health costs as a proxy for health needs,[70] or diagnosis as a proxy for true presence of the condition[52, 71, 72] — may introduce bias. Fifth, uninterpretable algorithms, whose predictions cannot be easily explained, should be used with caution and carefully audited to ensure their performance does not stem from confounding variables, so that they remain reliable and generalizable across data set shifts and subpopulations.[73, 74] Sixth, ensuring that clinical algorithms serve as a complement to, rather than a replacement for, clinician judgement is warranted.[75]

Regulatory bodies, including the World Health Organization, FDA, and American Medical Informatics Association have called for robust inclusion of ethical principles – such as those outlined above - with an emphasis on increasing inclusiveness and equity in deployment strategies.[40, 76, 77] In addition to an equity lens, one group built an 8-principle system of *responsible* AI deployment which not only includes alleviating health disparities, but also emphasizes concepts such as reporting clinically meaningful outcomes, reducing overdiagnosis and overtreatment, having high healthcare value, improving shared decision making, and easing algorithm tailoring to the local population through open access as core components.[78] There are currently a number of frameworks being developed to guide the development and deployment of reliable and trustworthy AI in healthcare.[79, 80]

The Central Illustration depicts modalities currently in use within the realm of AI in cardiology, along with core principles to guide successful use in clinical contexts discussed in the paragraphs above. Incorporation of these principles into clinical AI development can result in responsible and ethical deployment strategies which will maximize equitable and high-value impact.

## Conclusion

The work needed to harness the potential of AI in cardiovascular medicine is just beginning. As we see technologies able to classify, detect, diagnose, and augment the delivery of clinical care beyond what is currently possible, rigorous evaluation of AI models will be increasingly important, especially since the inputs and associations that AI models use to generate their insights are not always apparent. The capabilities of AI, underlying technology (such as generative AI), and the regulatory and reimbursement landscape for these technologies are rapidly evolving. These technologies have potential to augment our ability to practice medicine, and it is paramount that clinicians be actively involved in discussions on building relevant AI and testing real-world deployment.

The “black box” nature of this technology is a new challenge for our clinical community, since we’ve largely been able to ground technological insights into our basic understanding of cardiovascular anatomy and physiology. For example, the average clinician may not know the exact science behind how blood is analyzed to detect nanograms of troponin. However, they understand the physiologic basis of disease it represents and can integrate that understanding into their interpretation of troponin results along with the clinical trials that have been done to validate the results. In contrast, there are a growing number of AI models without a clear physiologic explanation, such as the ability to accurately predict reduced ejection fraction and valvular heart disease from electrocardiograms. In such cases, the clinical community will be even more reliant on robust prospective evaluation to understand the validity and safety of AI models in clinical care. Cardiologists will need to learn how to effectively evaluate AI studies, asking new questions such as “Were the models trained and validated appropriately?”, “Does the AI model provide sufficiently novel predictive insight to warrant the relative opacity of its methods?”, “How effective is the model when studied in real world populations?” and “Does the model maintain its accuracy over time?”. Fortunately, consensus guidelines on evaluation metrics and how to report AI evaluation results and other resources to guide rigorous evaluation are under development and will serve as important tools in assessing AI use.[39, 43, 81]

Realizing the potential for AI in cardiovascular medicine will require a concerted effort by our cardiology community to develop, test, and implement these technologies into our practice, continually evaluating where AI can and cannot provide significant value. As AI research progresses and the technology matures, it is vital for the cardiology community to engage actively with these developments and ensure AI tools and associated AI-assisted clinical workflows are validated, ethical, and effective for patient care.

## Disclosures:

SSJ reports consulting relationships with Bristol Myers Squibb, ARTIS Ventures, and Broadview Ventures. PE has research support provided to his institution from Eidos Therapeutics, Pfizer, Janssen, Edwards Lifesciences, New York Academy of Medicine, and Google. TJP owns stock in Abbott Laboratories and Baxter International with research support provided to his institution from the Amyloidosis Foundation, American Heart Association (Award #933452 and #23SCISA1077494, https://doi.org/10.58275/AHA.23SCISA1077494.pc.gr.172160), Eidos Therapeutics, Pfizer, Janssen, Edwards Lifesciences, and the Glorney-Raisbeck Fellowship Award from the New York Academy of Medicine. RA is a co-inventor in the patent 63/208,406 (Method and System for Automated Analysis of Coronary Angiograms). Dr. Avram reports receiving speaker fees from Abbott, Boston Scientific, Boehringer-Ingelheim and Novartis. JNAS in the co-founder and consultant to Sentiar, Inc and holds equity in

Sentiar. The technology has been licensed by Washington University to SentiAR.TMM discloses grant funding from the NIH (NHLBI UG3HL165065: The Rhythm Evaluation for Anticoagulation with Continuous Monitoring of Atrial Fibrillation Trial (REACT-AF)), honoraria and/or expense reimbursement in the past 3 years from the University of Chicago (Jan 2021), George Washington University (Jan 2021), Baylor College of Medicine (April 2021), the New York Cardiological Society (May 2021), and Medscape (Dec 2022). He has also received compensation and travel expense reimbursement for American College of Cardiology leadership roles and meetings. He is currently employed as a cardiologist and Vice President, Digital Products and Innovation at BJC HealthCare/Washington University School of Medicine. In this capacity, he is advising Myia Labs, for which his employer is receiving equity compensation in the company. He is receiving no individual compensation from the company. He is also a compensated director for a New Mexico-based foundation, the J.F Maddox Foundation. The remaining authors have nothing to disclose.

## Abbreviations

| | |
|---|---|
| **AI** | Artificial Intelligence |
| **CMS** | Center for Medicare & Medicaid Services |
| **CPT** | Current Proccedural Terminology |
| **CT** | Computed Tomography |
| **DICOM** | Digital Imaging and Communication in Medicine |
| **DRG** | Diagnosis-Related Group |
| **ECG** | Electrocardiogram |
| **FDA** | Food and Drug Administration |
| **LLM** | Large Language Model |
| **MRI** | Magnetic Resonance Imaging |
| **NTAP** | New Technology Add-on Payment |

## References

1. Anon. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. US Food and Drug Administration. 2022. Accessed September 27, 2023. https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices.
2. de Marvao A, Dawes TJW, O'Regan DP. Artificial intelligence for cardiac imaging-genetics research. Front Cardiovasc Med. 2019;6:195. [PubMed: 32039240]
3. Radhakrishnan A, Friedman SF, Khurshid S, et al. Cross-modal autoencoder framework learns holistic representations of cardiovascular state. Nat Commun. 2023;14:2436. [PubMed: 37105979]
4. Bannur S, Hyland S, Liu Q, et al. Learning to exploit temporal structure for biomedical vision-language processing. arXiv [csCV]. 2023.
5. Soin A, Merkow J, Long J, et al. CheXstray: Real-time multi-modal data concordance for drift detection in medical imaging AI. arXiv [eessIV]. 2022.
6. Soto JT, Hughes JW, Sanchez PA, Perez M, Ouyang D, Ashley E. Multimodal deep learning enhances diagnostic precision in left ventricular hypertrophy. bioRxiv. 2021.
7. Xie F, Yuan H, Ning Y, et al. Deep learning for temporal data representation in electronic health records: A systematic review of challenges and methodologies. J Biomed Inform. 2022;126:103980. [PubMed: 34974189]
8. Thawkar O, Shaker A, Mullappilly SS, et al. XrayGPT: Chest Radiographs Summarization using Medical Vision-Language Models. arXiv [csCV]. 2023.

9. Liu Z, Zhong A, Li Y, et al. Radiology-GPT: A Large Language Model for Radiology. arXiv [csCL]. 2023.
10. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature. 2023. Published onlineJuly 12, 2023. 10.1038/s41586-023-06291-2.
11. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digit Health. 2023;2:e0000198. [PubMed: 36812645]
12. Mirza Fatima N, Tang Oliver Y, Connolly Ian D, et al. Using ChatGPT to Facilitate Truly Informed Medical Consent. NEJM AI. 2024;1:AIcs2300145.
13. Anon. Ambient Artificial Intelligence Scribes to Alleviate the Burden. https://catalyst.nejm.org›doi›full›CAT230404https://catalyst.nejm.org›doi›full›CAT230404 10.1056/cat.23.0404.
14. Driess D, Xia F, Sajjadi MSM, et al. PaLM-E: An embodied multimodal language model. arXiv [csLG]. 2023.
15. Lehman E, Hernandez E, Mahajan D, et al. Do we still need clinical language models? arXiv [csCL]. 2023.
16. Chen L, Zaharia M, Zou J. How is ChatGPT's behavior changing over time? arXiv [csCL]. 2023.
17. Christensen M, Vukadinovic M, Yuan N, Ouyang D. Multimodal Foundation Models For Echocardiogram Interpretation. arXiv [csCV]. 2023.
18. Owens LM, Wilda JJ, Hahn PY, Koehler T, Fletcher JJ. The association between use of ambient voice technology documentation during primary care patient encounters, documentation burden, and provider burnout. Fam Pract. 2023. Published onlineSeptember 6, 2023. 10.1093/fampra/cmad092.
19. Gates D, Olajide O. Chapter 45 - Synthetic data in cardiovascular health research. In: Chang AC, Limon A, editors. Intelligence-Based Cardiology and Cardiac Surgery. Academic Press, 2024:407–413.
20. Loecher M, Perotti LE, Ennis DB. Using synthetic data generation to train a cardiac motion tag tracking neural network. Med Image Anal. 2021;74:102223. [PubMed: 34555661]
21. Beede E, Baylor E, Hersch F, et al. A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020:1–12.
22. van Smeden M, Heinze G, Van Calster B, et al. Critical appraisal of artificial intelligence-based prediction models for cardiovascular disease. Eur Heart J. 2022;43:2921–2930. [PubMed: 35639667]
23. Watson J, Hutyra CA, Clancy SM, et al. Overcoming barriers to the adoption and implementation of predictive modeling and machine learning in clinical care: what can we learn from US academic medical centers? JAMIA Open. 2020;3:167–172. [PubMed: 32734155]
24. Aristidou A, Jena R, Topol EJ. Bridging the chasm between AI and clinical implementation. Lancet. 2022;399:620. [PubMed: 35151388]
25. Schepart A, Burton A, Durkin L, et al. Artificial intelligence–enabled tools in cardiovascular medicine: A survey of current use, perceptions, and challenges. Cardiovascular Digital Health Journal. 2023;4:101–110. [PubMed: 37351333]
26. Kashyap S, Morse KE, Patel B, Shah NH. A survey of extant organizational and computational setups for deploying predictive models in health systems. J Am Med Inform Assoc. 2021;28:2445–2450. [PubMed: 34423364]
27. Sculley D, Holt G, Golovin D, et al. Hidden technical debt in Machine learning systems. In: Proceedings of the 28th International Conference on Neural Information Processing Systems.Vol 2. Cambridge, MA, USA: MIT Press, 2015:2503–2511.
28. Gonzalez-Smith J, Shen H, Singletary E, Silcox C. How health systems decide to use artificial intelligence for clinical decision support. NEJM Catal Innov Care Deliv. 2022;3.
29. Wessler BS, Lai Yh L, Kramer W, et al. Clinical prediction models for cardiovascular disease: Tufts predictive analytics and comparative effectiveness clinical prediction model database. Circ Cardiovasc Qual Outcomes. 2015;8:368–375. [PubMed: 26152680]

30. Mello MM, Shah NH, Char DS. President Biden's Executive Order on Artificial Intelligence-Implications for Health Care Organizations. JAMA. 2024;331:17–18. [PubMed: 38032634]
31. Bedoya AD, Economou-Zavlanos NJ, Goldstein BA, et al. A framework for the oversight and local deployment of safe and high-quality prediction models. J Am Med Inform Assoc. 2022;29:1631–1636. [PubMed: 35641123]
32. Perez MV, Mahaffey KW, Hedlin H, et al. Large-Scale Assessment of a Smartwatch to Identify Atrial Fibrillation. N Engl J Med. 2019;381:1909–1917. [PubMed: 31722151]
33. Lubitz SA, Faranesh AZ, Selvaggi C, et al. Detection of Atrial Fibrillation in a Large Population Using Wearable Devices: The Fitbit Heart Study. Circulation. 2022;146:1415–1424. [PubMed: 36148649]
34. Guo Y, Wang H, Zhang H, et al. Mobile Photoplethysmographic Technology to Detect Atrial Fibrillation. J Am Coll Cardiol. 2019;74:2365–2375. [PubMed: 31487545]
35. He B, Kwan AC, Cho JH, et al. Blinded, randomized trial of sonographer versus AI cardiac function assessment. Nature. 2023;616:520–524. [PubMed: 37020027]
36. Yao X, Rushlow DR, Inselman JW, et al. Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. Nat Med. 2021;27:815–819. [PubMed: 33958795]
37. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26:1364–1374. [PubMed: 32908283]
38. Cruz Rivera S, Liu X, Chan A-W, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Nat Med. 2020;26:1351–1363. [PubMed: 32908284]
39. Sengupta PP, Shrestha S, Berthon B, et al. Proposed requirements for cardiovascular imaging-related machine learning evaluation (PRIME): A checklist: Reviewed by the American college of cardiology healthcare innovation council. JACC Cardiovasc Imaging. 2020;13:2017–2035. [PubMed: 32912474]
40. Anon. Good Machine Learning Practice for Medical Device Development: Guiding Principles. US Food and Drug Administration Accessed September 27, 2023. https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles.
41. Executive Office of the President. Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Federal Register. 2023;88:75191–75226.
42. Sendak MP, Gao M, Brajer N, Balu S. Presenting machine learning model information to clinical end users with model facts labels. npj Digital Medicine. 2020;3:1–4. [PubMed: 31934645]
43. Shah NH, Halamka JD, Saria S, et al. A Nationwide Network of Health AI Assurance Laboratories. JAMA. 2024;331:245–249. [PubMed: 38117493]
44. Anon. Artificial Intelligence and Machine Learning in Software as a Medical Device. US Food and Drug Administration. 2019. Accessed September 27, 2023. www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device.
45. Parikh RB, Helmchen LA. Paying for artificial intelligence in medicine. npj Digital Medicine. 2022;5:1–5. [PubMed: 35013539]
46. Chen MM, Golding LP, Nicola GN. Who Will Pay for AI? Radiol Artif Intell. 2021;3:e210030. [PubMed: 34142090]
47. Kevin Wu, Eric Wu, Brandon Theodorou, et al. Characterizing the Clinical Adoption of Medical AI Devices through U.S. Insurance Claims. NEJM AI. 2023;1:AIoa2300030.
48. American Medical Association. CPT professional 2024. Chicago, IL: American Medical Association, 2023.
49. Salsabili M, Tesell M, Alcusky M, et al. Prescription digital therapeutics: Applying Medicaid experience to value assessment and formulary management. J Manag Care Spec Pharm. 2023;29:685–691. [PubMed: 37276040]
50. Khan B, Fatima H, Qureshi A, et al. Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. Biomed Mater Devices. 2023:1–8.

51. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. The Lancet Digital Health. 2021;3:e745–e750. [PubMed: 34711379]
52. Chen IY, Pierson E, Rose S, Joshi S, Ferryman K, Ghassemi M. Ethical machine learning in healthcare. Annu Rev Biomed Data Sci. 2021;4:123–144. [PubMed: 34396058]
53. Obermeyer Z, Nissan R, Stern M, Eaneff S, Bembeneck EJ, Mullainathan S. Algorithmic Bias Playbook. Chicago Booth Center for Applied Artificial Intelligence. 2021. Accessed September 27, 2023. https://www.chicagobooth.edu/research/center-for-applied-artificial-intelligence/research/algorithmic-bias/playbook.
54. Kleinberg J, Mullainathan S, Raghavan M. Inherent trade-offs in the fair determination of risk scores. arXiv [csLG]. 2016.
55. Corbett-Davies S, Pierson E, Feller A, Goel S, Huq A. Algorithmic decision making and the cost of fairness. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: ACM, 2017.
56. Chouldechova A. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. Big Data. 2017;5:153–163. [PubMed: 28632438]
57. Foryciarz A, Pfohl SR, Patel B, Shah N. Evaluating algorithmic fairness in the presence of clinical guidelines: the case of atherosclerotic cardiovascular disease risk estimation. BMJ Health Care Inform. 2022;29:e100460.
58. Pfohl SR, Foryciarz A, Shah NH. An empirical characterization of fair machine learning for clinical risk prediction. J Biomed Inform. 2021;113:103621. [PubMed: 33220494]
59. Pierson E, Cutler DM, Leskovec J, Mullainathan S, Obermeyer Z. An algorithmic approach to reducing unexplained pain disparities in underserved populations. Nat Med. 2021;27:136–140. [PubMed: 33442014]
60. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. Nat Genet. 2019;51:584–591. [PubMed: 30926966]
61. Daneshjou R, Vodrahalli K, Novoa RA, et al. Disparities in dermatology AI performance on a diverse, curated clinical image set. Sci Adv. 2022;8:eabq6147. [PubMed: 35960806]
62. Movva R, Shanmugam D, Hou K, et al. Coarse race data conceals disparities in clinical risk score performance. arXiv [csCY]. 2023.
63. Yadlowsky S, Hayward RA, Sussman JB, McClelland RL, Min Y-I, Basu S. Clinical implications of revised pooled cohort equations for estimating atherosclerotic cardiovascular disease risk. Ann Intern Med. 2018;169:20–29. [PubMed: 29868850]
64. Peterson PN, Rumsfeld JS, Liang L, et al. A validated risk score for in-hospital mortality in patients with heart failure from the American Heart Association get with the guidelines program. Circ Cardiovasc Qual Outcomes. 2010;3:25–32. [PubMed: 20123668]
65. Shahian DM, Jacobs JP, Badhwar V, et al. The Society of Thoracic Surgeons 2018 adult cardiac surgery risk models: Part 1-background, design considerations, and model development. Ann Thorac Surg. 2018;105:1411–1418. [PubMed: 29577925]
66. Tipton K, Leas BF, Flores E, et al. Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare. Agency for Healthcare Research and Quality (US), 2023.
67. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight - reconsidering the use of race correction in clinical algorithms. N Engl J Med. 2020;383:874–882. [PubMed: 32853499]
68. Khor S, Haupt EC, Hahn EE, Lyons LJL, Shankaran V, Bansal A. Racial and ethnic bias in risk prediction models for colorectal cancer recurrence when race and ethnicity are omitted as predictors. JAMA Netw Open. 2023;6:e2318495. [PubMed: 37318804]
69. Zink A, Obermeyer Z, Pierson E. Race corrections in clinical models: Examining family history and cancer risk. bioRxiv. 2023.
70. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366:447–453. [PubMed: 31649194]
71. Shanmugam D, Pierson E. Quantifying inequality in underreported medical conditions. arXiv [csCY]. 2021.
72. Nelson A. Unequal treatment: confronting racial and ethnic disparities in health care. J Natl Med Assoc. 2002;94:666–668. [PubMed: 12152921]

73. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Confounding variables can degrade generalization performance of radiological deep learning models. arXiv [csCV]. 2018.

74. Koh PW, Sagawa S, Marklund H, et al. WILDS: A benchmark of in-the-wild distribution shifts. arXiv [csLG]. 2020.

75. Mullainathan S, Obermeyer Z. Diagnosing physician error: A machine learning approach to low-value health care. Q J Econ. 2022;137:679–727.

76. Solomonides AE, Koski E, Atabaki SM, et al. Defining AMIA's artificial intelligence principles. J Am Med Inform Assoc. 2022;29:585–591. [PubMed: 35190824]

77. Anon. Ethics and governance of artificial intelligence for health: WHO guidance. Genève, Switzerland: World Health Organization, 2021.

78. Badal K, Lee CM, Esserman LJ. Guiding principles for the responsible development of artificial intelligence tools for healthcare. Commun Med (Lond). 2023;3:47. [PubMed: 37005467]

79. Anon. Blueprint for an AI bill of rights. The White House. 2022. Accessed February 16, 2024. https://www.whitehouse.gov/ostp/ai-bill-of-rights/.

80. Coalition for Health AI T. Blueprint for Trustworthy AI: Implementation Guidance and Assurance for Healthcare. 2022.

81. Ibrahim H, Liu X, Rivera SC, et al. Reporting guidelines for clinical trials of artificial intelligence interventions: the SPIRIT-AI and CONSORT-AI guidelines. Trials. 2021;22:11. [PubMed: 33407780]

## Highlights

- Areas of significant growth in AI and cardiology include multimodal inputs and generative AI
- Realizing AI's potential in cardiology requires rigorous evaluation, better infrastructure, regulatory oversight, and a viable business case
- Clinicians should have a basic understanding of AI and evaluation of AI to effectively assess the appropriateness of their use

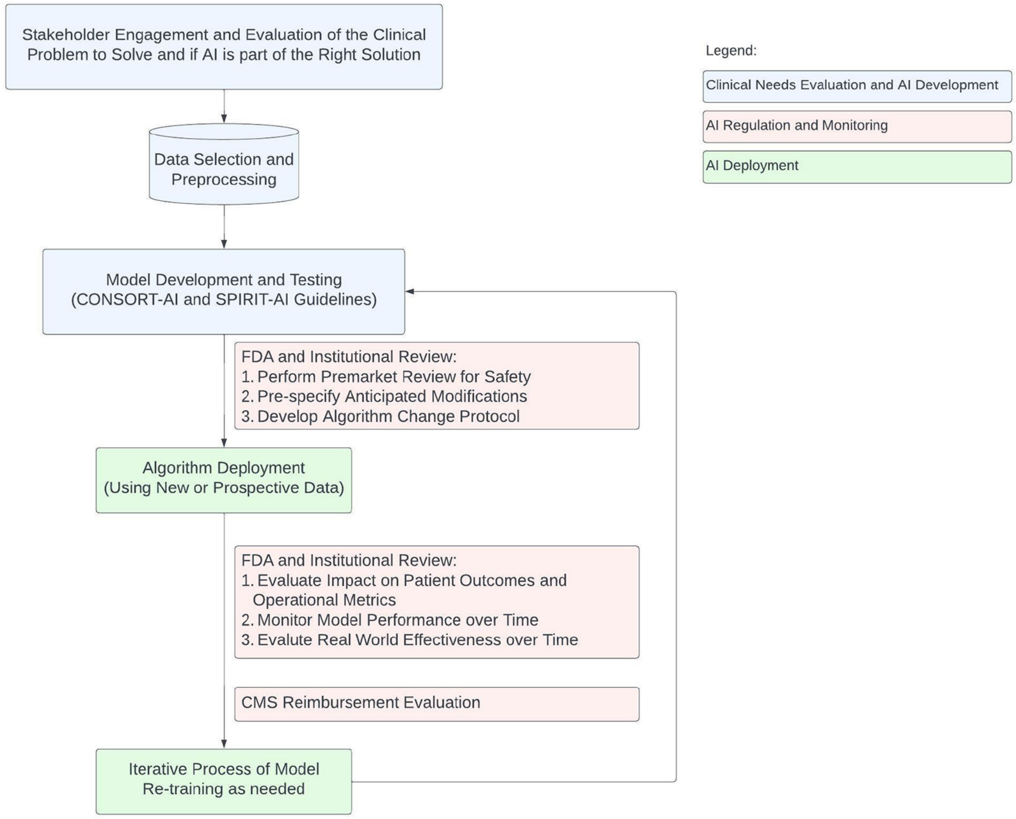
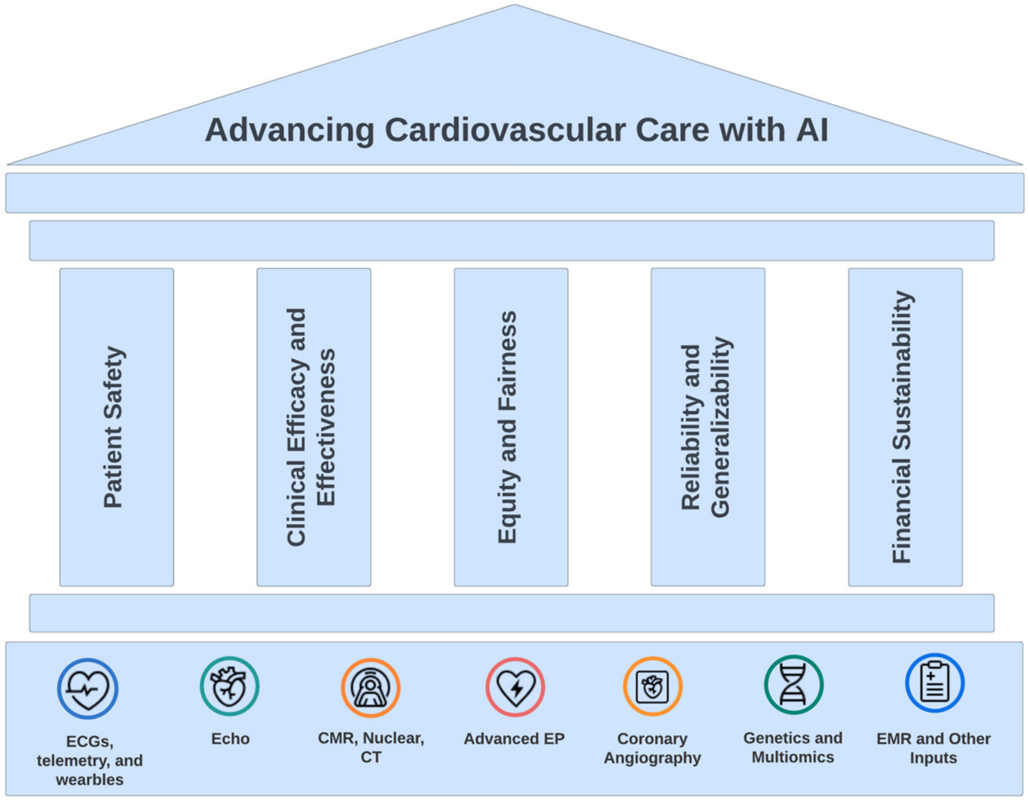



**Figure 1. AI Product Lifecycle.**

Framework for the development and deployment of clinically-relevant AI algorithms, including model generation, regulatory evaluation and quality control, and operationalization. CONSORT-AI and SPIRIT-AI are reporting guidelines for protocols of and completed randomized control trials involving AI.[37, 38] Per the Food and Drug Administration, anticipated modifications to the intended use, performance, or inputs of an AI software as medical device should be stated. [44] The algorithm change protocol is a plan for how to ensure success and appropriately control the risks of the anticipated modifications, with the goal of the AI algorithm remaining safe and effective after any modifications.[44] AI denotes Artificial Intelligence; CMS Centers for Medicare and Medicaid Services; FDA Food and Drug Administration.




**Central Illustration. Advancing Cardiovascular Care with AI.**
Depiction of current inputs for AI in cardiology, and pillars for success in advancing cardiovascular care with AI.

**Table 1.**

Glossary of Terms related to Artificial Intelligence.

| Term | Definition |
|---|---|
| Algorithm | A set of mathematical procedures used to learn patterns from data |
| Artificial Intelligence | The capability of a machine to imitate intelligent human behavior or perform tasks that typically require human intelligence |
| Assistive AI | The algorithm detects clinically relevant data without analysis or generated conclusions, and requires interpretation and report by a qualified healthcare professional |
| Augmentative AI | The algorithm analyzes and/or quantifies data to yield clinically meaningful output, and requires interpretation and report by a qualified healthcare professional |
| Autonomous AI | The algorithm automatically interprets data and independently generates clinically meaningful conclusions without concurrent qualified healthcare professional involvement |
| Deep Learning | A subset of machine learning that uses artificial neural networks |
| Fairness in Machine Learning | The study and practice of ensuring machine learning models are making decisions in an unbiased manner, avoiding perpetuating or introducing discriminatory outcomes |
| Features | Individual measurable properties of observed data that serve as input variables used by algorithms to learn patterns or make predictions |
| Foundation Models | Machine learning models trained on large amounts of unlabeled data that can be used for different tasks with very little fine-tuning |
| Large Language Models | A type of machine learning model that has been trained on large amounts of text to recognize, summarize, translate, predict, and/or generate content |
| Machine Learning | A subset of artificial intelligence in which computers learn from experience without explicit programming |
| Pragmatic Clinical Trials | Clinical trials designed to determine the effectiveness of an intervention using real-world conditions, as opposed to ideal or controlled conditions |