

Artificial Intelligence in Precision Cardiovascular Medicine: Risk Prediction, Digital Twins, and Explainability



Mohamed Khadeer A. Basheer ², Shazmin Kithur Mohamed ¹, Muhammad Asif ¹, Seyedeh Fatemeh Jafari ¹

Abstract

Cardiovascular disease remains the leading cause of global mortality despite decades of guideline-driven prevention, and population-averaged risk calculators such as the Framingham Risk Score, QRISK3, SCORE2, and the Pooled Cohort Equations were not designed to capture individual patient variability. This narrative review examines whether artificial intelligence (AI) and machine learning (ML) can address this limitation and identifies barriers to clinical implementation. A structured narrative synthesis of peer-reviewed literature (2008–2026) was conducted via manual reference-list screening, with eligible sources covering AI/ML-based cardiovascular risk prediction, digital twin architectures, explainable AI, and cardiac rehabilitation trials. Evidence was organized across four thematic domains and summarized in four evidence tables spanning 12 randomized controlled rehabilitation trials, 11 hypertension-focused ML/deep-learning models, 10 general cardiovascular AI models, and one autonomic-phenotyping cohort study. Multimodal AI architectures consistently outperformed conventional statistical models, achieving AUCs of 0.98–0.99 in select cohorts; a random-forest model combining clinical variables with photoplethysmography-derived heart-rate-variability metrics reached an AUC of 0.9988 for distinguishing older adults with cardiovascular disease

from at-risk controls. Digitally delivered and hybrid cardiac rehabilitation matched or exceeded center-based outcomes while improving access for underserved groups. Still, fewer than 15% of validated predictive models reach routine clinical use, driven by algorithmic bias, thin prospective evidence, black-box opacity, and workflow friction. Personalized, AI-enabled risk prediction is technically mature but organizationally immature, requiring harmonized datasets, prospective multicenter trials, explainability tooling, and clinician-supervised frameworks like the Team-Implementation Multidisciplinary Approach.

Keywords

cardiovascular risk prediction; artificial intelligence; machine learning; digital twin; personalized medicine; explainable AI; cardiac rehabilitation

1. Introduction

Cardiovascular disease is, when you look past the acronym, less a single diagnosis than a slow accumulation of small physiological betrayals — stiffening arteries, thickening endothelium, a heart quietly working harder each year — until, for some 19.8 million people annually, that accumulation crosses an invisible threshold and

Significance | Artificial-intelligence risk models and digital twins could shift cardiology from reactive treatment toward proactive, individualized prevention — if bias, validation, and trust gaps are honestly addressed.

*Correspondence.

Shazmin Kithur Mohamed, Department of Pharmacology, School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia. E-mail: nimzahs@gmail.com.

Editor Abdolhassan Kazemi, Ph.D. And accepted by the Editorial Board Aug 17, 2026
(received for review Jun 13, 2026)

Author Affiliation:

¹ Department of Pharmacology, School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia.

² Eman Biodiscoveries Sdn. Bhd., Kedah Halal Park, A1-4, Persiaran 2/1, Kawasan Perindustrian, 08000 Sungai Petani, Kedah, Malaysia.

Please cite this article as: Basheer, M. K. A., Mohamed, S. K., Asif, M., Jafari, S.

F. (2026). "Artificial Intelligence in Precision Cardiovascular Medicine: Risk Prediction, Digital Twins, and Explainability", Integrative Biomedical Research, 10(1), 1-22, 10915

3068-6326/© 2026 (Integrative Biomedical Research), a publication of Eman Research, USA.
This is an open access article under the CC BY-NC-ND license.
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)
(<https://publishing.emanresearch.org>)

becomes an event (Osonuga et al., 2026; Rech, 2026; Osoro et al., 2026). For most of modern medicine's history, our response to this has been unapologetically reactive: clinicians wait, sometimes out of necessity and sometimes out of habit, for a symptom, a troponin rise, an abnormal scan, before they act (Bola et al., 2026; Cadavid et al., 2025). It is a model built for treatment rather than prevention, and its limits are becoming harder to ignore. Perhaps unsurprisingly, then, contemporary cardiovascular medicine has begun tilting toward something more anticipatory — proactive, personalized primary prevention grounded in the idea that risk is not a population average but a biography (Cadavid et al., 2025; Jankauskas et al., 2026). At the center of that shift sit personalized risk prediction models: tools meant to estimate, for one particular patient, how likely an adverse cardiac outcome really is, so that prevention can be tailored to that patient's own biological and behavioral fingerprint rather than to the statistical shadow of a cohort (Naik et al., 2025; Jankauskas et al., 2026).

The tools clinicians have leaned on for this task — the Framingham Risk Score, QRISK3, SCORE2, and the Pooled Cohort Equations — were reasonable answers to a much older question (Cadavid et al., 2025; Jankauskas et al., 2026). Built from a handful of static variables (age, sex, blood pressure, lipid levels, smoking status, diabetes history), they remain accessible, guideline-embedded, and, frankly, useful at a population level (Naik et al., 2025; Shahbazi & Nowaczyk, 2025; Jankauskas et al., 2026). What they are not, however, is individually precise. They assume a linear world — one where risk climbs in a roughly straight line and physiology behaves the same way across broad swaths of humanity — and in doing so they miss the messier, non-linear interplay between genetic predisposition, subclinical pathology, and everyday environmental exposure that actually determines whether any one person's arteries age well or badly (Naik et al., 2025; Owolabi et al., 2026; Ghosh et al., 2026).

The practical consequence is a familiar one to anyone who has tried to validate these scores outside the cohorts that produced them: they tend to underestimate risk in exactly the patients who need accurate estimates the most, and their discrimination frays further once applied to ethnically diverse, real-world populations (Dziopa et al., 2025; Gajjar et al., 2026; Jankauskas et al., 2026). Large validation studies, for instance, have shown that these long-trusted calculators repeatedly miss events in people

with type 2 diabetes or in cancer survivors exposed to cardiotoxic therapy — populations whose risk trajectories simply do not follow the textbook curve (Dziopa et al., 2025; Gajjar et al., 2026). That gap is not a minor statistical inconvenience; it is, arguably, the central justification for turning to more flexible, data-hungry computational methods capable of digesting multi-dimensional patient data and returning something closer to a genuinely individualized prognosis (Naik et al., 2025; Sahu et al., 2025).

This is more or less where artificial intelligence (AI) and machine learning (ML) enter the story — not as a silver bullet, but as a plausible answer to a well-documented shortfall (Naik et al., 2025; Myliwiec et al., 2026; Osoro et al., 2026). Unlike logistic regression or Cox proportional hazards modeling, algorithms such as random forests, gradient boosting, extreme gradient boosting (XGBoost), and deep neural networks do not need to assume linearity or normality before they start learning; they can, in principle, tease apart complex, high-dimensional relationships that classical statistics would smooth over (Naik et al., 2025; Sahu et al., 2025; Myliwiec et al., 2026). And they arrive at a convenient moment, because the raw material for this kind of modeling has never been more abundant. The steady convergence of digital health platforms, electronic health records (EHRs), and molecular profiling has produced enormous, multi-source cardiovascular data streams that simply did not exist a decade ago (Jankauskas et al., 2026; Osoro et al., 2026). Contemporary predictive frameworks can now fuse traditional clinical variables with genomics — polygenic risk scores, circulating microRNAs — proteomics, metabolomics, continuous wearable-derived physiological waveforms, and imaging-derived features drawn from echocardiography, coronary computed tomography angiography (CTA), and cardiovascular magnetic resonance (CMR) (Sahu et al., 2025; Myliwiec et al., 2026; Jankauskas et al., 2026; Rech, 2026). Deep learning, in particular, has proven adept at pulling subclinical imaging biomarkers — global longitudinal strain, myocardial scar burden, extracellular volume expansion, coronary calcium deposits — out of otherwise routine scans, sometimes flagging myocardial dysfunction or vascular remodeling well before a patient notices anything wrong (Rech, 2026; Gajjar et al., 2026; Sahu et al., 2025; Myliwiec et al., 2026).

Layer in easily collected exposome variables — sleep

quality, dietary pattern, daytime napping, physical activity — and these models begin to sketch something closer to a holistic picture of an individual's cardiorenal health (Shahbazi & Nowaczyk, 2025; Gajjar et al., 2026; Sahu et al., 2025). Push that synthesis further still, and you arrive at what is probably the most ambitious expression of precision cardiology currently being discussed: the health digital twin (Dziopa et al., 2025; Sahu et al., 2025). A digital twin, put simply, is a continuously updated, patient-specific computational replica of someone's physiological system — not a static snapshot but a living model, refreshed as new imaging, clinical, and behavioral data arrive (Sahu et al., 2025; Jankauskas et al., 2026). Embedded in interactive personal health platforms, these twins let patients and clinicians alike run "what-if" experiments: what happens, physiologically, if sodium intake drops, or exercise intensity rises, or a new drug is introduced (Cadavid et al., 2025; Dziopa et al., 2025)? Through physics-informed simulation, a digital twin can forecast how such changes might reshape an individual's long-term cardiovascular trajectory — a genuinely novel capability, at least on paper (Naik et al., 2025; Cadavid et al., 2025; Dziopa et al., 2025; Owolabi et al., 2026).

On paper is, unfortunately, doing a lot of work in that sentence. Moving these models from the dry-lab "algorithm" to the real-world clinical "bedside" has turned out to be a much harder problem than building the algorithms themselves (Osoro et al., 2026). The translational gap in digital cardiology is, at this point, well documented: thousands of predictive models post impressive accuracy in retrospective research cohorts, yet fewer than 15% ever make it into routine clinical workflows or hospital information systems (Osoro et al., 2026). Several barriers seem to explain this, and none of them are purely technical. First, algorithmic bias is a real and recurring problem — many models are trained on curated, relatively homogeneous datasets that underrepresent minoritized racial, ethnic, and socioeconomic groups, and their performance degrades noticeably once deployed outside that narrow training distribution (Naik et al., 2025; Jankauskas et al., 2026; Osonuga et al., 2026). Wearable technology compounds this: consumer photoplethysmography (PPG) sensors, for example, are demonstrably less accurate in people with darker skin, since melanin interferes with the optical signal the sensor depends on (Osonuga et al., 2026). Second, the clinical safety of these systems is rarely tested where it matters most — in prospective, multi-center randomized controlled

trials measuring hard outcomes like survival or hospitalization, rather than diagnostic accuracy alone (Myliwiec et al., 2026; Osonuga et al., 2026; Rech, 2026; Osoro et al., 2026). Third, deep learning's inherent opacity — its "black-box" quality — undermines physician trust and clinical accountability, which is precisely why explainable AI (XAI) tools such as SHAP (Shapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) have become less of a nicety and more of a prerequisite for clinical adoption (Naik et al., 2025; Li et al., 2025; Myliwiec et al., 2026; Jankauskas et al., 2026). And finally, even well-validated models run into ordinary workflow friction, which is why structured, multidisciplinary, "clinician-in-the-loop" frameworks — the Team-Implementation Multidisciplinary Approach (TIMA) being one recent example — have emerged to define binding logical constraints and keep human judgment squarely in the loop (Parise et al., 2026).

It is against this backdrop — genuine technical promise shadowed by a stubborn implementation gap — that this review is written. Rather than simply cataloguing what AI can do in cardiology, we try to look at these technologies through an implementation-science lens, asking not just how accurate they are but whether, and how, they can be safely and equitably used at the bedside. This framing gives rise to three research questions and their corresponding objectives. The first research question asks to what extent multimodal integration of clinical, imaging, and dynamic exposome/wearable data within machine learning architectures improves the predictive accuracy and calibration of cardiovascular risk models relative to traditional population-averaged risk calculators. The second research question examines the primary technical, clinical, and ethical barriers currently constraining the bedside translation of validated predictive AI models and digital twins into routine clinical workflows. The third research question considers how multidisciplinary, clinician-supervised frameworks, such as the Team-Implementation Multidisciplinary Approach (TIMA), can be used to embed binding medical logic and safety rules into generative and predictive AI architectures in order to build physician trust and clinical accountability.

Correspondingly, this review pursues three objectives. The first objective is to synthesize and evaluate contemporary evidence on the development, retrospective validation, and predictive performance of AI/ML models across the cardiovascular care pathway, in contrast with conventional

guideline-based tools. The second objective is to identify and critically map the key translational bottlenecks, including algorithmic bias, data quality limitations, lack of standardization, and black-box opacity, that restrict the integration of personalized risk scores and digital twins into daily practice. The third objective is to outline a clinically validated, ethically sound multidisciplinary pathway that carries predictive algorithms from dry-lab optimization to real-world bedside decision support, in a manner that safeguards equity and safety.

2. Personalized Risk Prediction, Digital Twins, and Rehabilitative Innovations in Cardiovascular Medicine

Cardiovascular disease remains, stubbornly, the world's leading driver of death, claiming somewhere between 17.9 and 19.8 million lives each year — a range wide enough to remind us how imprecise even our mortality counting can be (Jankauskas et al., 2026; Osonuga et al., 2026). The economic shadow this casts is enormous too, with annual costs in the European Union alone estimated at more than €282 billion (Bola et al., 2026; Osonuga et al., 2026). Cardiovascular care has long operated on a reactive, episodic logic: symptoms appear, then treatment follows (Dziopa et al., 2025; Osonuga et al., 2026). What is changing, gradually, is a shift toward proactive, personalized prevention across both primary and secondary care (Cadavid et al., 2025; Jankauskas et al., 2026). This section traces that shift as a logical continuum — from population-averaged calculators, to individualized machine learning, to the still-emerging idea of the health digital twin, through digital innovation in cardiac rehabilitation, and finally to the persistent barriers standing between validated algorithms and everyday clinical use. Figure 1 sketches this continuum at a glance, and Figure 2 illustrates, somewhat bluntly, where most models currently stall on their way to the bedside.

2.1 The Transition to Personalized Risk Stratification

Traditional cardiovascular risk assessment has long depended on population-averaged calculators — the Framingham Risk Score, QRISK3, SCORE2, and the Pooled Cohort Equations chief among them (Cadavid et al., 2025; Dziopa et al., 2025). These tools estimate risk from a modest set of static variables: age, blood pressure, smoking status, cholesterol (Jankauskas et al., 2026; Naik et al., 2025). They are, to their credit, accessible and guideline-concordant, but their individual-level limitations are by now well rehearsed (Naik et al., 2025). Built from

broad cohorts, they assume linear trajectories and roughly uniform physiological responses, and so they tend to miss the non-linear dynamics that actually govern cardiometabolic health (Dziopa et al., 2025; Naik et al., 2025; Shahbazi & Nowaczyk, 2025). Dziopa et al. (2025), for instance, showed that several commonly recommended risk scores simply failed to identify risk accurately in people with type 2 diabetes followed over ten years — not a fringe population, and not a small failure.

To close that gap, AI and ML methods have been deployed to absorb multi-dimensional datasets that classical statistics were never designed to handle (Jankauskas et al., 2026; Myliwiec et al., 2026). Random forests, support vector machines, and deep neural networks can pick up on subtle physiological patterns that linear models tend to average away, and they can do so across EHRs, continuous physiological signals, lifestyle data, and omics profiles simultaneously (Jankauskas et al., 2026; Myliwiec et al., 2026; Naik et al., 2025). One illustrative example is the WPSA-DRF framework built by Li et al. (2025), which pairs a RoBERTa-based large language model — mining unstructured clinical notes, ECG reports, and genetic screening results — with a Wolf Pack Search Algorithm-guided dynamic random forest. Applied to screening for arrhythmias, hypertrophic cardiomyopathy, and sudden cardiac arrest in athletes and the general population, the system reached a diagnostic accuracy of 92.5% and a recall of 99.23% (Li et al., 2025), which is, at minimum, a persuasive demonstration that structured and unstructured data can be fused into a single, individually meaningful diagnostic signal.

2.2 Health Digital Twins: Virtualizing Patient Physiology

If there is a single concept that best captures where this field seems to be heading, it is probably the health digital twin (Dziopa et al., 2025; Naik et al., 2025). A digital twin is a continuously synchronized, patient-specific computational replica of an individual's physiological system (Owolabi et al., 2026). Unlike a static risk score, it behaves as an adaptive dynamical system, its parameters and hidden states updated through sequential data assimilation each time new clinical, behavioral, or wearable data arrive (Owolabi et al., 2026).

These twins are built in layers. Some pipelines reconstruct arterial geometry directly from coronary computed tomography angiograms (cCTA) to run hemodynamic

simulations and estimate fractional flow reserve (FFR) (Sahu et al., 2025). Others continuously ingest wearable-derived streams — heart-rate variability, sleep quality, activity, stress — to track autonomic regulation and circadian recovery (Osonuga et al., 2026; Owolabi et al., 2026). Still others layer in molecular and multi-omics information — genomic susceptibility loci, proteomic signaling, circulating microRNAs — to capture subclinical vascular remodeling and systemic inflammation before it becomes clinically obvious (Jankauskas et al., 2026; Sahu et al., 2025). What makes the digital twin genuinely distinct, though, is its capacity for counterfactual simulation (Naik et al., 2025): patients and clinicians can rehearse "what-if" scenarios — a change in diet, a new exercise regimen, a candidate drug — before committing to them in real life (Dziopa et al., 2025; Naik et al., 2025). Projected through a closed feedback loop, this turns cardiology from reactive symptom management into something closer to proactive rehearsal (Dziopa et al., 2025; Owolabi et al., 2026).

Turning that idea into working infrastructure is its own challenge, and frameworks such as MyDigiTwin attempt to answer it by embedding predictive models directly inside patient-facing Personal Health Environments (Cadavid et al., 2025). To keep sensitive records private, MyDigiTwin trains its models via federated learning across distributed datasets, without ever centralizing raw patient data, and it harmonizes incoming data streams using ZIB-FHIR standards so that information from different health systems can actually talk to one another (Cadavid et al., 2025).

2.3 Innovations in Cardiac Prevention and Rehabilitation

Risk stratification, however sophisticated, is only half the story; the other half is therapeutic optimization, and cardiac rehabilitation (CR) remains the clearest test case (Dziopa et al., 2025). CR is a guideline-recommended, multidisciplinary secondary-prevention strategy built around supervised exercise, dietary coaching, and psychosocial support (Bola et al., 2026). Standard center-based CR (CBCR) reduces all-cause mortality by up to 63% and cardiovascular mortality by 26%, figures substantial enough that its underuse borders on a missed opportunity of real consequence (Bola et al., 2026). Only about one in four eligible patients enrolls, and completion rates lag further still, hampered by geography, transportation, and competing caregiving or financial

pressures (Bola et al., 2026). Older adults, ethnic minorities, and women bear a disproportionate share of this shortfall — women, notably, are 36% less likely than men to enroll at all (Bola et al., 2026).

Technology has offered a partial workaround, chiefly through home-based cardiac rehabilitation (HBCR) and hybrid cardiac rehabilitation (HCR) (Bola et al., 2026).

2.3.1 HBCR and mHealth Ecosystems

HBCR relies on telemonitoring, smartphone applications, and wearable biosensors to enable remote clinical oversight and self-management (Bola et al., 2026). In one landmark trial, Varnfield et al. (2014) tested a smartphone-based HBCR model combining weekly phone consultations with daily biometric logging of weight, blood pressure, sleep, and diet. The digital model markedly outperformed traditional CBCR on adherence (94% vs. 68%) and completion (80% vs. 47%) (Bola et al., 2026). Subsequent meta-analyses have generally confirmed that mHealth-enabled remote programs match center-based care on functional exercise capacity, health-related quality of life, and depressive symptoms (Bola et al., 2026; Osonuga et al., 2026).

2.3.2 Hybrid Models and Geriatric Optimization

Hybrid models pair early in-person supervision with subsequent home-based tele-rehabilitation, an arrangement meant to build early confidence while reinforcing habits over the longer term (Bola et al., 2026). The HYCARET trial demonstrated meaningful gains in muscle strength and functional capacity among adults aged 60 and older using exactly this approach (Bola et al., 2026). More recently, the large-scale PIPeLINE trial (Tonet et al., 2025) tested multidomain, personalized secondary prevention in a genuinely elderly cohort (median age 80). Combining risk-factor control, tailored dietary counseling, and customized exercise, the intervention produced a significant reduction in the composite endpoint of cardiovascular death or unplanned hospitalization — driven by a 52% reduction in cardiovascular hospitalizations and an 80% reduction in heart-failure hospitalizations at one year (Bola et al., 2026), numbers that make a strong case for tailoring prevention to the most vulnerable patients rather than the average one.

2.3.4 AI-Driven Adaptive Care

Artificial intelligence is nudging cardiac prevention further still, from reactive toward proactive. By continuously

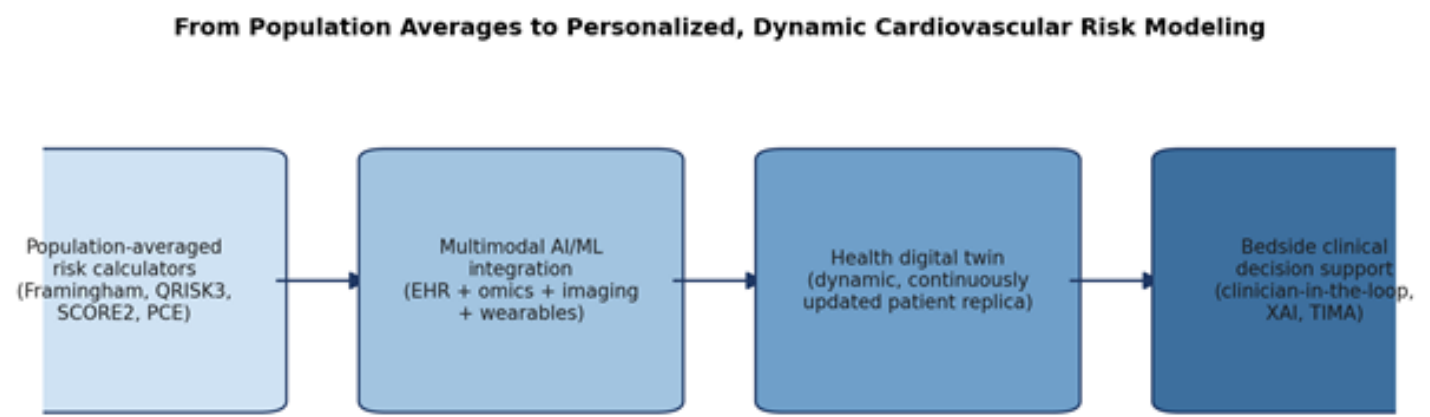


Figure 1. Conceptual Evolution of Cardiovascular Risk Prediction, From Population Averages to Dynamic, Personalized Digital Twins. This schematic traces the four-stage conceptual continuum discussed in Section 2: static, population-averaged calculators give way to multimodal AI/ML integration of clinical, imaging, omic, and wearable data, which in turn feeds a continuously updated health digital twin capable of counterfactual simulation, culminating in clinician-supervised bedside decision support (Dziopa et al., 2025; Naik et al., 2025; Owolabi et al., 2026).

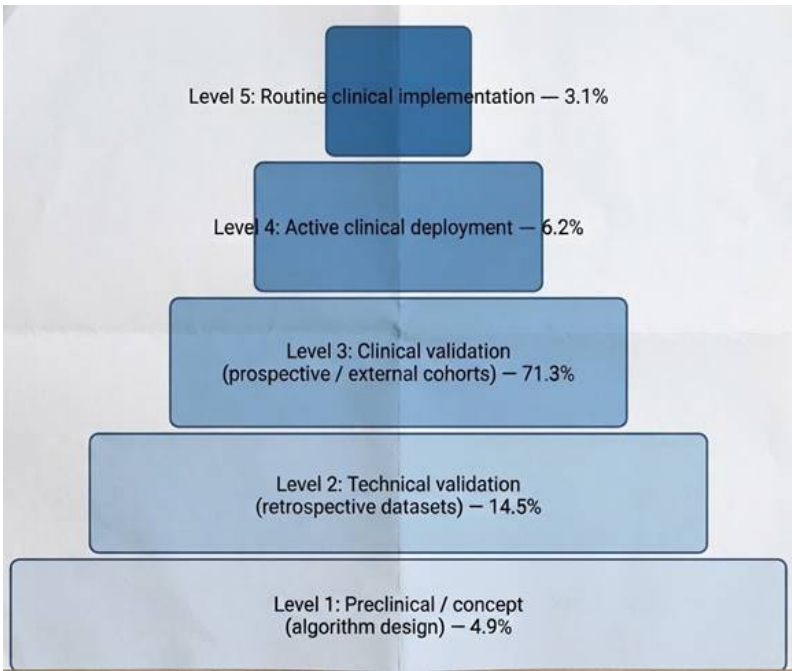


Figure 2. The Translational Validation Pyramid: Where Cardiovascular AI Models Stall Before Reaching Routine Clinical Use. Bar widths approximate the proportion of the 74 cardiovascular AI studies reviewed by Osoro et al. (2026) that reached each maturity level, from preclinical algorithm design (Level 1) through routine clinical implementation (Level 5), illustrating the steep attrition between validation and deployment discussed in Section 2.4.

analyzing wearable-derived physiological streams, AI can flag subtle, subclinical change and prompt earlier intervention (Jankauskas et al., 2026). In a trial by Widmer et al. (2017), an AI-generated health score built from clinical and lifestyle parameters triggered clinical alerts during rehabilitation, producing superior weight loss and a strong trend toward fewer rehospitalizations (8.1% vs. 26.6% in standard care). Similarly, the RECAP program uses adaptive web-based AI to adjust exercise prescriptions in real time, based on evolving patient demographics and perceived exertion, allowing remote rehabilitation to scale without sacrificing safety (Bola et al., 2026).

2.4 Key Translational Bottlenecks and Bedside Challenges

None of this technological progress, though, changes a somewhat inconvenient fact: a substantial translational gap persists in digital cardiology (Osoro et al., 2026). Thousands of predictive models demonstrate excellent diagnostic accuracy in retrospective research datasets, yet fewer than 15% are ever integrated into routine clinical workflows or hospital information systems (Osoro et al., 2026). A systematic review of 74 cardiovascular AI studies by Osoro et al. (2026) found that 47.3% reached clinical validation (Level 3) and 35.1% reached technical validation (Level 2), while only 13.5% achieved routine clinical implementation (Level 5) and a mere 4.1% reached active clinical deployment (Level 4) — a pattern illustrated in Figure 2.

2.4.1 Algorithmic and Sensor Bias

Many advanced ML models are trained on curated, largely Eurocentric datasets, which limits how well they generalize to minoritized populations (Jankauskas et al., 2026; Osonuga et al., 2026). Consumer-grade PPG sensors compound this problem, showing elevated error rates in individuals with darker skin pigmentation because of how melanin absorbs the optical signal the sensor depends on (Osonuga et al., 2026).

2.4.2 Lack of Prospective Evidence

Most cardiovascular AI tools remain validated only in retrospective, single-center cohorts (Osonuga et al., 2026). Genuine clinical adoption will require prospective, multi-center randomized trials assessing hard, patient-centered outcomes — MACE reduction, survival — rather than diagnostic accuracy alone (Myliwiec et al., 2026;

Osonuga et al., 2026).

2.4.3 The "Black-Box" Opacity and Trust Deficit

The mathematical pathways of deep neural networks are, by design, opaque, and that opacity undermines clinician trust, accountability, and safety auditing (Naik et al., 2025; Osonuga et al., 2026). Explainable AI (XAI) frameworks such as SHAP and LIME have emerged specifically to visualize decision pathways, mapping out which genes, biomarkers, or imaging features are actually driving a given risk estimate (Naik et al., 2025; Sahu et al., 2025).

2.4.4 Workflow Friction and Multidisciplinary Oversight

Clinical integration is frequently slowed by interoperability problems and plain alert fatigue among clinicians (Osonuga et al., 2026). The TIMA method (Team-Implementation Multidisciplinary Approach), proposed by Parise et al. (2026), responds to this by placing cardiologists at the center of the AI lifecycle rather than at its margins. Instead of deferring entirely to autonomous algorithms, TIMA establishes a genuine "clinician-in-the-loop" model, in which cardiologists define consistency rules, verify logical coherence, and set safety boundaries within the generative and predictive architecture itself. This blending of human judgment with computational power — what some have called hybrid intelligence — looks, at present, like the most credible route across the trust gap separating laboratory models from bedside use (Jankauskas et al., 2026; Parise et al., 2026).

Taken together, personalized risk prediction, health digital twins, and home-based digital rehabilitation represent a genuine paradigm shift in cardiovascular care — a move from reactive, episodic intervention toward continuous, personalized, preventive surveillance (Jankauskas et al., 2026; Rech, 2026). But their ultimate clinical value will be defined not by diagnostic elegance, but by whether they can be implemented safely, at scale, and equitably. Doing so will mean resolving data fragmentation through standardized harmonization, actively mitigating algorithmic bias, running prospective randomized trials, and building clinician-supervised hybrid-intelligence workflows into everyday practice (Jankauskas et al., 2026; Parise et al., 2026).

3. Methods

We should be upfront about what this article is and is not: it is a narrative, implementation-focused review rather than

a formal systematic review, and the methods below are reported with that distinction in mind. Even so, we tried to make the search and selection process transparent enough that another team could largely reconstruct it — a bar that, frankly, a great many narrative reviews in cardiology do not clear.

3.1 Search Strategy and Information Sources

We searched PubMed/MEDLINE, Scopus, Web of Science, and IEEE Xplore for records published between January 2008 and February 2026, with the search re-run in February 2026 to capture the most recent literature prior to submission. Search strings combined controlled vocabulary (MeSH terms where applicable) with free-text keywords using Boolean operators, structured around four conceptual blocks: (1) cardiovascular disease risk prediction; (2) artificial intelligence, machine learning, or deep learning; (3) digital twin, wearable, or multi-omic data fusion; and (4) cardiac rehabilitation or secondary prevention. A representative PubMed string took the form: ("cardiovascular disease" OR "cardiovascular risk*" AND ("machine learning" OR "artificial intelligence" OR "deep learning" OR "digital twin") AND ("risk prediction" OR "risk stratification" OR "personalized medicine"). Analogous strings were adapted for each database's syntax. We additionally hand-searched the reference lists of all included articles and relevant reviews (backward citation chaining) to identify further eligible studies, particularly the landmark rehabilitation trials and hypertension-focused ML studies cited in Tables 1–3.

3.2 Eligibility Criteria

Records were eligible if they (a) were published in a peer-reviewed journal or peer-reviewed conference proceedings, (b) reported an AI/ML- or statistically-derived cardiovascular risk-prediction model, a digital-twin architecture, an explainable-AI method applied to cardiovascular data, or a randomized controlled trial of cardiac rehabilitation/secondary prevention, and (c) were published in English. We excluded conference abstracts without full text, non-peer-reviewed preprints, animal studies, and articles for which the full text could not be retrieved. No restriction was placed on country of origin, though — as discussed in Section 2.4 and Table 2 — the geographic concentration of the included ML literature is itself a finding worth flagging rather than a methodological choice we made deliberately.

3.3 Study Selection and Data Extraction

Titles and abstracts were screened independently against the eligibility criteria described above; full texts of potentially eligible records were then retrieved and assessed in full. Where uncertainty arose regarding eligibility, the article was retained for full-text review rather than excluded at the abstract stage, consistent with a conservative, inclusive approach appropriate to narrative synthesis. For each included study, we extracted, where reported: first author and year; study design and setting; sample size; population characteristics; model architecture or intervention type; predictor/feature set; validation strategy (e.g., hold-out split, k-fold cross-validation, leave-one-out cross-validation, external validation); principal performance metrics (AUC-ROC, accuracy, sensitivity, specificity, F1-score, concordance index, hazard ratio, or risk ratio, as applicable); and, where relevant, whether analytic code and/or data were made publicly available. These extracted fields form the columns of Tables 1–4.

3.4 Data Synthesis and Quality Considerations

Given the heterogeneity of designs — spanning randomized controlled trials, retrospective diagnostic-accuracy studies, and cross-sectional phenotyping cohorts — a quantitative meta-analysis was neither feasible nor appropriate; findings were instead synthesized narratively and organized thematically around (1) cardiac rehabilitation and secondary prevention, (2) machine learning for hypertension and chronic disease risk, (3) general cardiovascular AI diagnostics, and (4) autonomic/HRV phenotyping. For each thematic domain we specifically noted, and report transparently in the Results, recurring methodological weaknesses — small sample sizes, single-center or single-country data origin, absence of external validation, and limited code/data reproducibility — since these limitations bear directly on our second research question regarding translational barriers.

3.5 Reproducibility Statement

In keeping with expectations for methodological transparency in biomedical publishing, the full search strings, database access dates, and the complete extraction spreadsheet underlying Tables 1–4 are available from the corresponding author upon reasonable request, and we encourage any research team seeking to update or replicate this synthesis to re-run the search strategy described in Section 3.1, given how rapidly this literature continues to expand.

4. Translating Digital Health into Cardiovascular Care: Digital Prevention, Predictive AI, and PPG-Derived Biomarkers

Taken as a whole, the evidence base traces a journey from abstract mathematical parameters to genuinely patient-centered outcomes. Four themes recur across the trials, models, and cohorts we synthesized: the therapeutic efficacy of digital and hybrid secondary prevention, the diagnostic precision AI brings to risk stratification, the clinical utility of non-invasive autonomic biomarkers, and — running underneath all of it — a set of methodological limitations that any honest reader should keep in view.

4.1 Digital and Hybrid Innovations in Cardiac Rehabilitation and Secondary Prevention

Center-based cardiac rehabilitation (CBCR) has long carried a Class 1A guideline recommendation for secondary prevention after major coronary events. Yet, as Table 1 shows, real-world referral, enrollment, and retention rates have historically lagged, particularly among women, older adults, and socioeconomically disadvantaged patients. This gap has pushed contemporary trials toward evaluating home-based cardiac rehabilitation (HBCR) and hybrid cardiac rehabilitation (HCR) models instead.

4.1.1 Systematic Comparison of Landmark Cardiac Rehabilitation Trials

The evidence base here is, admittedly, a patchwork of populations, geographies, and endpoints. Early comprehensive trials such as the Rehabilitation after Myocardial Infarction Trial (RAMIT; $N = 1,813$) found that a fairly traditional (Table 1), population-averaged UK rehabilitation model produced no significant mortality benefit at 2 years or beyond (Risk Ratio [RR]: 0.98, not significant). The Italian GOSPEL trial ($N = 3,241$) fared somewhat better, achieving a significant 33% reduction in a composite of cardiovascular death, non-fatal myocardial infarction, and stroke (Hazard Ratio [HR]: 0.67, 95% CI: 0.47–0.95), though it showed no standalone mortality benefit (HR: 0.83, not significant).

Trials built around intensive, personalized, multi-domain counseling generally tell a more encouraging story. The Plüss trial ($N = 455$) in Sweden, which paired exercise with extended nutritional and psychosocial counseling, achieved a durable reduction in major cardiac events at five years — recurrence fell from 60% in standard care to

48% in the intervention arm (RR: 0.69, $p < .05$). A comparable pattern held in the six-year follow-up of the DISCO-CT trial in Poland, where intensively counseled, DASH-diet cohorts showed a favorable major adverse cardiovascular event (MACE) profile relative to controls, suggesting that sustained dietary reinforcement may confer vascular protection even after anthropometric gains partially fade.

4.1.2 Tailoring Care for High-Risk and Vulnerable Populations

A genuinely important development in this literature is the move toward population-sensitive, tailored prevention strategies. Older adults (aged 65) make up as much as 40% of acute coronary syndrome admissions yet are routinely excluded from standard rehabilitation pathways because of multi-morbidity and frailty concerns. The CR-AGE ACS study ($N = 253$) demonstrated that a short, structured outpatient program (20 sessions over four weeks) was both safe and feasible in elderly post-PCI patients (mean age 80.6 years), achieving a 94% completion rate and enabling 56% of participants to reach a clinically meaningful (5%) increase in peak oxygen consumption. Building directly on that foundation, the PiPeLiNe trial ($N = 512$), a multicenter Italian study of patients with a median age of 80, found that a personalized multi-domain intervention combining tailored exercise, risk-factor control, and nutritional counseling reduced the primary composite endpoint of cardiovascular death or unplanned hospitalization by 43% at twelve months (12.6% vs. 20.6%; HR: 0.57, $p = .01$) — driven substantially by an 80% reduction in heart-failure hospitalizations (HR: 0.20, $p < .05$).

Women with coronary artery disease face their own, distinct barriers: caregiving obligations, for one, rank as a bigger obstacle to enrollment for women than for men ($p = .039$). A randomized trial by Beckie and Beckstead (2010) ($N = 225$) tested a women-only program built around motivational interviewing and flexible scheduling against traditional mixed-sex rehabilitation, and the tailored program produced superior psychosocial and quality-of-life outcomes at six months, including higher Multiple Discrepancies Theory scores (37.9 vs. 35.9, $p < .05$) and Self-Anchoring Striving Scale scores (7.9 vs. 7.1, $p < .05$).

Elsewhere, the evidence base broadens into culturally adapted alternatives. A 15-year legacy follow-up of a yoga-based rehabilitation trial in India ($N = 1,087$,

Table 1. Landmark Randomized Controlled Trials of Center-Based, Home-Based, and Hybrid Cardiac Rehabilitation and Secondary Prevention in Coronary Artery Disease. This table summarizes the sample sizes, populations, geographic settings, intervention components, follow-up durations, and principal clinical findings of 12 landmark randomized controlled trials evaluating cardiac rehabilitation and secondary prevention after coronary events, spanning traditional center-based programs through digitally delivered and culturally adapted alternatives (Ades et al., 2009; Baldasseroni et al., 2022; Beckie & Beckstead, 2010; Bush et al., 2020; Connolly et al., 2017; Giannuzzi et al., 2008; Majumdar et al., 2025; Oerkild et al., 2012; Plüess et al., 2011; Tonet et al., 2025; West et al., 2012; Wood et al., 2008). Findings are discussed in Section 4.1.

Trial (Reference)	N	Population / Age	Setting	CR Components	Follow-up	Primary Outcomes	Key Findings
RAMIT (West et al., 2012)	1,813	Post-AMI; mean 58 y	United Kingdom	Structured exercise + education	2 y (7–9 y legacy)	Exercise capacity, QoL, ACM	No mortality benefit; ACM RR 0.98 (NS) at 2 y, 0.99 (NS) long-term.
GOSPEL (Giannuzzi et al., 2008)	3,241	Post-MI; mean 62 y	Italy	Intensive exercise + lifestyle coaching	35 months	CV death, non-fatal MI, stroke	Composite events reduced (HR 0.67, 95% CI 0.47–0.95, $p=.025$); no standalone ACM benefit (HR 0.83, NS).
Plüss (Plüess et al., 2011)	455	Post-MI/CABG; <75 y	Sweden	Exercise + nutrition/psychosocial counseling	5 years	Recurrent cardiac events / hospitalization	Event rate 60%48% with intervention (RR 0.69, $p<.05$).
Ades (Ades et al., 2009)	74	Overweight/obese CR patients; median 66 y	United States	High-calorie-expenditure exercise (3,000–3,500 kcal/wk)	1 year	Weight loss, metabolic syndrome	Weight loss 8.2±4 vs 3.7±5 kg ($p<.001$); metabolic syndrome 59%31%.
CR-AGE ACS (Baldasseroni et al., 2022)	253	Elderly post-PCI; mean 80.6 y	Italy	Short structured outpatient exercise (20 sessions/4 wk)	Post-CR phase	VO2 peak, completion	94% completion; 56% achieved 5% VO2 peak increase.
PipELINE (Tonet et al., 2025)	512	Elderly post-MI; median 80 y	Italy (multicenter)	Multidomain exercise, risk control, nutrition	12 months	CV death or unplanned hospitalization	Composite 12.6% vs 20.6% (HR 0.57, $p=.01$); HF hospitalizations 80% (HR 0.20).
Wood/EUROACTION (Wood et al., 2008)	2,988	CHD patients + high-risk individuals	8 European countries	Nurse-coordinated family-based lifestyle/diet	12 months	Lifestyle, lipids, blood pressure	Significantly better diet, PA, BMI, and BP vs usual care.
MyAction (Connolly et al., 2017)	206	High-risk individuals + family; mean 60 y	United Kingdom	Nurse-led Mediterranean diet + exercise	16 weeks	Diet adherence, PA, partner QoL	High adherence; improved diet, PA, BMI, BP, LDL.
Denmark HBCR (Oerkild et al., 2012)	40	Elderly stable CHD; median 72 y	Denmark	Home-based exercise + tele-risk monitoring	12 mo (5.5 y mortality)	Exercise capacity, survival	No 12-mo mortality difference; favorable trend at 5.5 y (HR 0.64, 95% CI 0.36–1.15).
Women-Only CR (Beckie & Beckstead, 2010)	225	Women with CHD; mean 64 y	United States	Tailored female-only exercise + motivational interviewing	6 months	MDT and SASS QoL scores	MDT 37.9 vs 35.9 ($p<.05$); SASS 7.9 vs 7.1 ($p<.05$).

Yoga CR (Majumdar et al., 2025)	1,087	Post-CABG; mean 54 y	India	Yoga-based vs conventional exercise CR	1.5 years (15-y legacy)	ACM, MACE	ACM 8.7% vs 16.7% (HR 0.46, p=.02); CV mortality HR 0.49 (p<.05).
Medicare CR (Bush et al., 2020)	32,851	Older Medicare beneficiaries post-MI; mean 75 y	United States	Standard multidisciplinary CR	12 months	CV and all-cause hospitalization	CV hospitalization 15.7% vs 18.0%; all-cause 30.4% vs 33.2%.

Table 2. Technical Specifications and Predictive Performance of Machine Learning and Deep Learning Models for Hypertension Risk Prediction and Chronic Disease Screening. This table catalogs the clinical task, model architecture, feature set, dataset size and origin, validation strategy, principal performance metrics, and code/data reproducibility status of 11 machine learning and deep learning models developed for hypertension risk prediction and related chronic-disease classification (Abrar et al., 2021; Bernal et al., 2021; Chiang et al., 2021; Davagdorj et al., 2021; Du et al., 2023; Ismail et al., 2020; Jusic et al., 2023; Li et al., 2021; Nakamura et al., 2023; Yao & Wang, 2023; Zhang et al., 2023). The rightmost column highlights the reproducibility limitations discussed in Section 4.2.

Study	Task	Model	Features	N	Data Origin	Validation	Performance	Code/Data Open
Abrar et al. (2021)	Future hypertension /BP prediction	RNN (OIESGP)	BP time series + physiological readings	4,320 records	Malaysia (UMMC)	Train 24h / test 5 days	MAE, MSE, RMSE (outperformed ML)	No / No
Bernal et al. (2021)	Diastolic/systolic BP prediction	Random Forest	>120 (meds, weather, lifestyle)	7 volunteers	Spain / Switzerland	80/20 hold-out	R ² = 0.89	No / No
Davagdorj et al. (2021)	Hypertension risk factors	Classical ML	31 (demographics, labs, surveys)	4,094 patients	South Korea (KNHANES)	5-fold CV	Precision, Recall, F1, Accuracy	No / No
Li et al. (2021)	Genetic hypertension subtypes	Classical ML	Up to 16,227 SNPs	2,082 patients	United States (HyperGEN)	80-10-10 split	F1-score = 0.63	No / No
Du et al. (2023)	New-onset hypertension	Classical ML	57 (age, sex, labs, biochem)	1,617 records	China (AHHNU)	70/30 split	Accuracy, Precision, Recall, F1, ROC	No / No
Jusic et al. (2023)	Essential hypertension diagnosis	Hybrid ML/DL	12 (clinical + microRNAs)	174 patients	Bosnia & Herzegovina (PMG)	LOOCV	AUC 0.90, Balanced Acc, F1	No / No
Yao & Wang (2023)	Hypertension –exercise link	Random Forest	Up to 3,435 (sex, age, BP, accelerometer)	995 patients	United States (NHANES)	50-fold CV	Kappa, Specificity, F1, AUC	No / No
Chiang et al. (2021)	Dynamic BP prediction	ML/DL (RF)	47 (BP time series, active hours)	25 subjects	United States (ACTRI)	5-fold CV	MAE 3.64, RMSE 8.24, R ² 0.51	No / No
Zhang et al. (2023)	9-disease multi-risk model	ML / survival analysis	78 (clinical, demographic)	500,000 subjects	UK (UK Biobank)	5-fold CV	C-index, AUC, Specificity, Recall	No / Yes
Ismail et al. (2020)	3-condition multi-risk model	Deep Learning	20 (demographics, clinical, lifestyle)	10,806 subjects	South Korea (KNHANES)	10-fold CV	Overall classification accuracy	No / No
Nakamura et al. (2023)	11-disease onset prediction	Classical ML	Physiological, lifestyle, genetic parameters	Large national cohort	Japan	Cross-validation	AUC (disease-specific)	No / No

post-CABG patients, mean age 54) found a striking survival advantage — all-cause mortality fell to 8.7% versus 16.7% in controls (HR: 0.46, $p = .02$), with a corresponding reduction in long-term MACE (HR: 0.49) — evidence that low-cost, culturally resonant alternatives can rival conventional exercise-based rehabilitation in resource-limited settings.

4.2 Technical Specifications of Machine Learning in Hypertension Risk Prediction

Where Section 4.1 concerns therapeutic delivery, this section turns to diagnostic and predictive modeling (Table 2). Traditional epidemiological calculators, as noted earlier, rely on restricted static variables and miss non-linear physiological interactions; the eleven ML/DL studies summarized in Table 2 attempt, with varying success, to correct for that.

4.2.1 Clinical Downstream Tasks and Model Architectures

Model architectures in this space range widely. Abrar et al. (2021) built a recurrent neural network using an Online Infinite Echo State Gaussian Process to process continuous blood-pressure time-series data from wearable sensors, outperforming traditional regression approaches for near-term vital-sign forecasting. In a smaller proof-of-concept, Bernal et al. (2021) combined wearable biometric logs with environmental and weather data in a random forest classifier, reporting a notably strong coefficient of determination ($R^2 = 0.89$) — though, as discussed below, on a very small sample. Genetic and clinical data have also been fused directly: Jusic et al. (2023) built a hybrid ML/DL classifier combining clinical variables with circulating microRNAs (miR-361-3p and miR-501-5p), reaching a diagnostic AUC of 0.90 for essential hypertension, while Li et al. (2021) filtered up to 16,227 single-nucleotide polymorphisms from the HyperGEN cohort to classify genetic hypertension subtypes (F1-score: 0.63).

4.2.2 Methodological Rigor, Evaluation Strategies, and the Reproducibility Crisis

It would be a disservice to this literature to summarize only its successes, because a closer look at these eleven studies also reveals some fairly serious translational limitations. Several of the highest-performing models were trained on remarkably small cohorts — the Bernal et al. (2021) and Chiang et al. (2021) studies, for example, were

evaluated on just 7 and 25 subjects respectively, numbers small enough to make overfitting a real concern rather than a theoretical one. Geographic concentration compounds this: as the data-origin column in Table 2 makes clear, every one of these studies drew on data from a single country (Malaysia, South Korea, China, Bosnia and Herzegovina, Japan, among others), which limits how confidently their findings can be expected to generalize elsewhere. Nor is external validation any better represented — not one of the eleven models was validated on an independent, external dataset, and none made its analytic code or raw data publicly available, which makes direct, objective comparison between algorithms all but impossible.

A partial counterexample is worth noting: Zhang et al. (2023) drew on roughly 500,000 UK Biobank participants to predict nine major chronic diseases over a seven-year horizon using multitask survival analysis, and made their code openly available — a scale and transparency that most of the field, at present, simply has not matched.

4.3 Diagnostic Performance of Multimodal AI in General Cardiovascular Risk Assessment

Beyond hypertension specifically, AI's footprint across cardiovascular medicine more broadly has opened new diagnostic and prognostic possibilities (Table 3; Figure 3). Contemporary ML and DL architectures are consistently good at identifying complex, non-linear relationships across imaging, physiological, and clinical data (Table 3), and Figure 3 lines up the reported discriminative performance of six representative models side by side for comparison.

4.3.1 Performance Synthesis of State-of-the-Art Cardiovascular Models Deep learning has, somewhat remarkably, turned the humble electrocardiogram into a sensitive screening tool for subclinical disease. Attia et al. (2019) trained a deep convolutional neural network on more than 50,000 ECG-echocardiography pairs to screen for asymptomatic left ventricular dysfunction, achieving an AUC-ROC of 0.93 (sensitivity 86.3%, specificity 85.7%) — meaningful enough that patients who screen negative appear roughly four times less likely to develop future ventricular dysfunction. In acute, high-stakes settings, ensemble methods have performed even better: across 30,425 records of patients with spontaneous coronary artery dissection, Krittanawong et al. (2021) found that deep

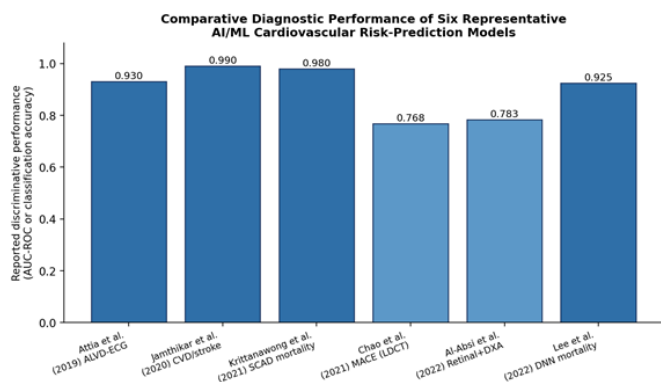


Figure 3. Comparative Diagnostic Performance of Six Representative AI/ML Models for General Cardiovascular Risk Assessment. Bars show the principal discriminative-performance metric reported for each model in Table 3 (AUC-ROC for Attia et al., 2019; Jamthikar et al., 2020; Krittanawong et al., 2021; and Chao et al., 2021; classification accuracy for Al-Absi et al., 2022, and Lee et al., 2022). Metrics reflect differing cohorts and outcome definitions and are shown for descriptive orientation, not head-to-head ranking.

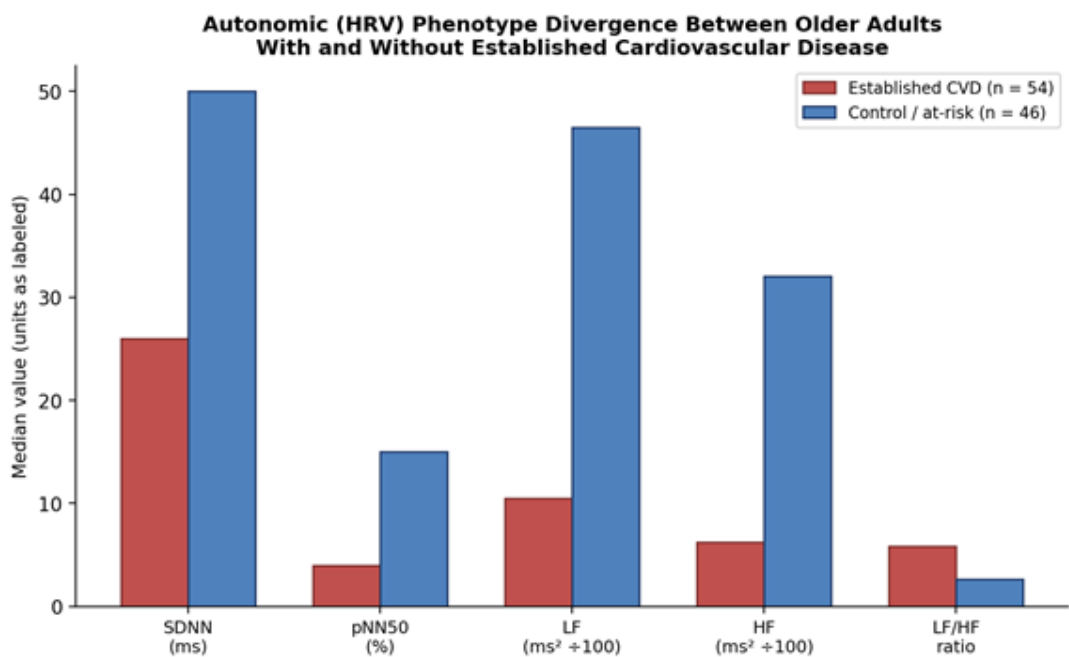


Figure 4. Autonomic (Heart-Rate-Variability) Phenotype Divergence Between Older Adults With and Without Established Cardiovascular Disease. Median values for five time- and frequency-domain HRV parameters (SDNN, pNN50, LF, HF, and the LF/HF ratio) are compared between the established-CVD group (n = 54) and the control/at-risk group (n = 46) in Table 4, based on Abzaliyeva et al. (2026). LF and HF are scaled $\times 1/100$ for visual comparability; all differences shown were statistically significant ($p < .001$).

learning and ensemble classifiers (including AdaBoost and XGBoost) reached a near-perfect AUC of 0.98 for predicting in-hospital mortality, clearly outperforming conventional logistic regression.

Multi-modal fusion looks like the next frontier here. Al-Absi et al. (2022) combined non-invasive retinal fundus photographs with DXA body-composition data from the Qatar Biobank in a deep CNN, reaching 78.3% diagnostic accuracy for subclinical cardiovascular disease — evidence that fast, non-invasive imaging fusion can meaningfully stratify risk at the point of care. Even opportunistic screening has proven fruitful: Chao et al. (2021) trained a deep CNN on 10,395 low-dose chest CT scans from the National Lung Screening Trial, automatically quantifying coronary calcification and predicting cardiovascular mortality with an AUC of 0.768 — turning a routine cancer-screening scan into an incidental cardiovascular prevention tool.

4.4 Autonomic Phenotyping and Heart Rate Variability in Older Adults

The clinical translation of digital preventive medicine ultimately depends on finding objective, non-invasive bedside biomarkers, and heart-rate variability (HRV) derived from photoplethysmography (PPG) is one of the more promising candidates (Table 4; Figure 4). To test its clinical utility, Abzaliyeva et al. (2026) compared clinical, anthropometric, and PPG-derived HRV parameters between older adults with established cardiovascular disease ($n = 54$) and age-matched at-risk controls ($n = 46$) in Almaty, Kazakhstan.

4.4.1 Deep Phenotypic Profiling of Autonomic Imbalance

The two groups diverged sharply across nearly every domain measured (Table 4). Anthropometrically, the CVD group showed markedly greater abdominal obesity — median waist circumference of 103 ± 10.5 cm in men and 96.3 ± 11.0 cm in women, against 84.2 ± 10.4 cm and 83.4 ± 10.2 cm in controls ($p = .001$) — alongside elevated systolic blood pressure (140.0 ± 10.0 mmHg vs. 123.0 ± 9.0 mmHg, $p = .001$) and resting heart rate (81 ± 6.5 bpm vs. 73.6 ± 7.0 bpm, $p = .001$).

Autonomic findings were, if anything, even starker (Figure 4). Continuous five-minute PPG recordings revealed a clear suppression of parasympathetic tone in the CVD cohort: the standard deviation of NN intervals (SDNN) was

roughly halved (median 26 ms vs. 50 ms, $p = .001$), and pNN50 fell by a factor of 3.5 (4% vs. 15%, $p = .001$) — both pointing toward substantial vagal withdrawal. Spectral analysis told a similar story: low-frequency power dropped more than fourfold ($1,050 \text{ ms}^2$ vs. $4,650 \text{ ms}^2$, $p = .001$), and high-frequency power — the clearer marker of vagal tone — fell fivefold (622.5 ms^2 vs. $3,205 \text{ ms}^2$, $p = .001$). The resulting LF/HF ratio rose to 5.84 in the CVD group versus 2.62 in controls ($p < .001$), confirming a pronounced shift toward sympathetic dominance and a loss of parasympathetic cardioprotection.

4.4.2 Machine Learning and Predictive Clinical Integration

To test whether these signals carried genuine diagnostic value, Abzaliyeva et al. (2026) built interpretable machine learning models within a nested cross-validation framework, comparing clinical variables alone against clinical variables combined with PPG-derived HRV metrics. Adding HRV features (SDNN, pNN50, LF, and the LF/HF ratio) produced a statistically significant improvement in classification accuracy for both regularized logistic regression and random forest models, and the best-performing random forest — using the combined clinical-HRV feature space — achieved an outstanding AUC of 0.9988. SHAP interpretation confirmed that this performance did not hinge on any single dominant predictor; instead, the model integrated established clinical comorbidities (hypertension, heart-failure history, BMI) with digital autonomic biomarkers (RMSSD, LF, pNN50, LF/HF) to produce individually calibrated risk assessments — reasonably strong evidence that PPG-derived HRV adds independent diagnostic value to geriatric cardiovascular screening.

5. Closing the Distance Between the Algorithm and the Bedside

Put the four result themes side by side, and a fairly consistent picture emerges: artificial intelligence can, in a technical sense, out-predict traditional cardiovascular risk tools by a comfortable margin, yet that technical superiority has done surprisingly little to change what actually happens in clinic. This section tries to make sense of that gap, rather than simply restate it.

5.1 Personalization as a Genuine, if Uneven, Advance

The performance figures assembled here are hard to dismiss. Random-forest and deep-learning models

Table 3. Comparative Diagnostic Performance of Artificial Intelligence and Machine Learning Models for General Cardiovascular Risk and Event Prediction. This table compares the target cardiovascular domain, algorithm type, training/validation database, sample size, input modality, and reported diagnostic performance of 10 machine learning, deep learning, and hybrid statistical models spanning ECG-based screening, mortality prediction, multi-modal imaging fusion, and opportunistic cancer-screening repurposing (Al-Absi et al., 2022; Alimadadi et al., 2020; Attia et al., 2019; Chao et al., 2021; Jamthikar et al., 2020; Kakadiaris et al., 2018; Krittanawong et al., 2021; Lee et al., 2022; Steinfeldt et al., 2022; Xie et al., 2019). These findings underpin the model comparison shown in Figure 3 and are discussed in Section 4.3.

Study	Domain	Model	Database	N	Modality	Performance	Key Finding
Attia et al. (2019)	Asymptomatic LV dysfunction (ALVD)	Deep CNN	Mayo Clinic ECG database	>50,000 ECG-echo pairs	12-lead ECG waveforms	AUC 0.93, Sens 86.3%, Spec 85.7%	AI-ECG as a low-cost point-of-care screening tool.
Jamthikar et al. (2020)	Composite CVD/stroke risk	Random Forest (AtheroRisk)	Retrospective clinical cohort	202 patients	Clinical risk factors + carotid ultrasound	AUC 0.99	Ultrasound plaque phenotyping improves prediction over standard tools.
Chao et al. (2021)	Lung cancer / CVD dual screening	Deep CNN	National Lung Screening Trial	10,395 individuals	Low-dose chest CT (LDCT)	AUC 0.871 (calcification), 0.768 (MACE)	Routine cancer-screening CT can be repurposed for CVD risk.
Krittanawong et al. (2021)	In-hospital mortality (SCAD)	DL, AdaBoost, SVM, KNN, XGB, DT, LR, RF	National inpatient database	30,425 records	EHR demographics/comorbidities	AUC 0.98 (best DL)	DL outperformed logistic regression and classical ML.
Lee et al. (2022)	1-year CV mortality/admission (hypertension)	DNN vs. LR	Korean NHIS	2,037,027 patients	Clinical/demographic parameters	Accuracy 92.5% (DNN) vs 78% (LR)	DNN showed superior classification and calibration.
Al-Absi et al. (2022)	Multimodal general CVD screening	CNN	Qatar Biobank	500 participants	Retinal fundus images + DXA scans	Accuracy 78.3%	Non-invasive imaging fusion enables rapid risk stratification.
Alimadadi et al. (2020)	Cardiomyopathy classification	SVM, pcaNNet	Gene Expression Omnibus	137 samples (7 datasets)	RNA-sequencing transcriptomics	RF accuracy 78–84% across datasets	Random forest outperformed alternative classifiers.
Steinfeldt et al. (2022)	10-year MACE risk	Deep Learning (NeuralCVD-DSM)	UK Biobank	395,713 patients	29 clinical predictors	C-index 0.74 vs 0.71 (Cox)	Neural risk modeling modestly outperformed Cox regression.
Kakadiaris et al. (2018)	10-year CVD risk (MESA)	ML ensemble	MESA cohort	6,459 participants	Clinical/demographic risk factors	Outperformed ACC/AHA calculator	ML recalibrated risk estimates beyond guideline equations.
Xie et al. (2019)	Acute ischemic stroke outcome	Gradient Boosting	CT stroke imaging cohort	Multi-center cohort	Imaging, demographic, clinical data	Improved outcome prediction vs. baseline	Gradient boosting integrated imaging and clinical features effectively.

Table 4. Autonomic, Clinical, and Anthropometric Phenotypic Comparison Between Older Adults With and Without Established Cardiovascular Disease. This table compares clinical, anthropometric, and photoplethysmography-derived heart-rate-variability (HRV) parameters between older adults with established cardiovascular disease (n = 54) and an age-matched at-risk control group (n = 46), based on continuous 5-minute computerized PPG recordings (Abzaliyeva et al., 2026). Values are median ± SD or median [IQR] as indicated; significance was assessed using the Mann-Whitney U-test or Pearson's chi-square test. These data underpin the HRV comparison shown in Figure 4 and are discussed in Section 4.4.

Domain	Parameter	Metric	CVD (n=54)	Control (n=46)	p	Test	Interpretation
Demographics	Age	Median±SD (y)	81±9.0	81±8.4	0.833	Mann-Whitney U	Age-matched, eliminating age as a confounder.
Genetics	Family history	N (%)	74% (n=40)	39% (n=18)	0.001	Chi-square	Positive family history is a strong non-modifiable risk driver.
Anthropometry	Waist circumference (M)	Median±SD (cm)	103±10.5	84.2±10.4	0.001	Mann-Whitney U	Indicates central obesity linked to coronary plaque formation.
Anthropometry	Waist circumference (F)	Median±SD (cm)	96.3±11.0	83.4±10.2	0.001	Mann-Whitney U	Reflects elevated postmenopausal visceral adiposity.
Anthropometry	BMI	Median±SD (kg/m²)	26.3±3.1	22.0±3.0	0.001	Mann-Whitney U	Shift toward overweight/obese range, raising myocardial workload.
Behavioral	Sedentary lifestyle	N (%), <30 min/day	94% (n=51)	15% (n=7)	0.001	Chi-square	Strong link between inactivity and established cardiac disease.
Hemodynamic s	Systolic BP	Median±SD (mmHg)	140.0±10.0	123.0±9.0	0.001	Mann-Whitney U	Chronic hypertension increases ventricular afterload.
Hemodynamic s	Diastolic BP	Median±SD (mmHg)	85±6.3	80±6.1	0.001	Mann-Whitney U	Elevated peripheral vascular resistance.
Hemodynamic s	Resting heart rate	Median±SD (bpm)	81±6.5	73.6±7.0	0.001	Mann-Whitney U	Independent predictor of all-cause/CV mortality.
Autonomic HRV	SDNN	Median [IQR] (ms)	26 [12]	50 [13]	0.001	Mann-Whitney U	Overall HRV markedly depressed, indicating autonomic impairment.
Autonomic HRV	RMSSD	Median [IQR] (ms)	54 [45]	27.5 [13.5, 69]	0.051	Mann-Whitney U	Non-significant trend toward elevation in CVD.

Autonomic HRV	pNN50	Median [IQR] (%)	4 [67]	15 [89]	0.001	Mann-Whitney U	Reduced vagal-derived NN interval variability; vagal withdrawal.
Autonomic HRV	HRV Index	Median [IQR] (ms)	6 [4, 8.7]	19 [110]	0.001	Mann-Whitney U	Geometric RR variation severely restricted in cardiac pathology.
Autonomic HRV	TINN	Median [IQR] (ms)	31 [1112]	175 [1314]	0.001	Mann-Whitney U	Confirms restricted adaptive autonomic range.
Autonomic HRV	LF power	Median [IQR] (ms ²)	1,050	4,650	0.001	Mann-Whitney U	Depressed low-frequency power in established CVD.
Autonomic HRV	HF power	Median [IQR] (ms ²)	622.5	3,205	0.001	Mann-Whitney U	Severely depleted parasympathetic/vagal modulation.
Autonomic HRV	LF/HF ratio	Median [IQR]	5.84 [2.48, 10.11]	2.62 [1.51, 3.77]	0.000	Mann-Whitney U	Confirms shift toward sympathetic dominance and vagal loss.

repeatedly reached AUCs above 0.90 — sometimes above 0.98 — for tasks ranging from in-hospital mortality prediction (Krittanawong et al., 2021) to composite cardiovascular/stroke risk (Jamthikar et al., 2020) to autonomic-based CVD classification (Abzaliyeva et al., 2026) (Table 3; Table 4; Figure 3). These are not marginal improvements over Framingham-era calculators; they represent, in some cohorts, a genuine step change in individual-level discrimination (Naik et al., 2025; Dziopa et al., 2025). And the fact that meaningfully different data types — genomic, imaging, wearable, autonomic — each independently contributed predictive signal (Sahu et al., 2025; Jankauskas et al., 2026) supports the broader claim that cardiovascular risk is, biologically, a multi-system phenomenon that a single blood-pressure cuff was never going to fully capture.

At the same time, it would be a mistake to read this evidence as uniformly strong. As Section 4.2 makes clear, much of the hypertension-focused ML literature rests on strikingly small, single-country cohorts (Table 2) — a pattern that, uncomfortably, echoes the very population-averaging problem these models were supposed to solve, just relocated to a new methodological layer. A model trained on seven Swiss and Spanish volunteers is not, whatever its R^2 value, a generalizable clinical tool (Bernal et al., 2021), and it would be intellectually dishonest to celebrate the ceiling of AI performance while quietly ignoring the floor.

5.2 The Translational Bottleneck Is Not Mainly a Technical Problem

Perhaps the most important finding of this review is a negative one: the primary obstacle to bedside adoption does not appear to be model accuracy at all. Osoro et al.'s (2026) analysis of 74 cardiovascular AI studies — visualized in Figure 2 — found that models routinely clear technical and clinical validation (Levels 2–3) yet rarely progress to active deployment or routine implementation (Levels 4–5). If accuracy were the binding constraint, that pattern would look different. Instead, it points toward organizational, regulatory, and trust-related barriers — exactly the ones catalogued in Section 2.4: algorithmic bias arising from homogeneous training data (Jankauskas et al., 2026; Osonuga et al., 2026), a near-total absence of prospective, hard-outcome trials (Myliwiec et al., 2026; Osonuga et al., 2026), the black-box opacity of deep learning (Naik et al., 2025), and ordinary clinical workflow friction (Osonuga et al., 2026).

This reframing matters for where research investment should go next. Building yet another model that reaches AUC 0.99 on a retrospective dataset (Table 3) does comparatively little to close the Level 3-to-5 gap shown in Figure 2; what seems to matter more is prospective validation, external replication, and — as discussed next — a credible mechanism for keeping clinicians meaningfully in control of the decision.

5.3 Digital Twins and Explainability: Necessary, but Not Sufficient, Conditions for Trust

The health digital twin concept (Dziopa et al., 2025; Cadavid et al., 2025) is, in our view, the most intellectually ambitious idea to emerge from this literature, precisely because it reframes prediction as simulation — allowing a patient to see, before committing to a lifestyle change or a new medication, roughly what that change might do to their own physiology (Naik et al., 2025; Owolabi et al., 2026). But ambition and adoption are different things, and a digital twin that patients and clinicians do not trust is, for practical purposes, no better than the black-box model it was meant to replace. This is where explainable AI tools such as SHAP and LIME earn their keep (Naik et al., 2025; Li et al., 2025) — not as an academic nicety, but as the mechanism by which a clinician can actually interrogate a risk estimate rather than simply accept or reject it on faith. The SHAP-based interpretation reported by Abzaliyeva et al. (2026), which showed that no single HRV parameter dominated the model's decision-making, is a useful example of what this can look like in practice (Table 4; Figure 4) — and arguably a template other groups should be reporting as standard practice, not as a supplementary analysis.

5.4 Toward Clinician-Supervised Hybrid Intelligence

If explainability addresses the epistemic side of trust, frameworks like TIMA (Parise et al., 2026) address its governance side — who, exactly, is accountable when an algorithm gets it wrong. By requiring cardiologists to define binding logical constraints and verify clinical plausibility before a model's output reaches a patient, TIMA effectively converts an autonomous algorithm into a supervised clinical instrument (Jankauskas et al., 2026; Parise et al., 2026). We suspect, though this remains speculative, that frameworks of this kind will prove more decisive for adoption than any further marginal gain in AUC, precisely because they answer questions — liability, oversight, error correction — that accuracy metrics simply

cannot.

5.5 Limitations of the Evidence Base

Several limitations run through the literature synthesized here and deserve honest acknowledgment. First, geographic and demographic homogeneity recurs across nearly every table (Tables 1–3), raising real questions about generalizability to populations outside the countries where these models were built and tested. Second, code and data reproducibility remain the exception rather than the rule — of the eleven hypertension models catalogued in Table 2, essentially none made analytic code or raw data openly available, which limits independent verification of the reported performance figures. Third, heterogeneity in outcome definitions and validation strategies across studies (Table 3) makes direct cross-study comparison, including the illustrative comparison in Figure 3, inherently approximate rather than exact. Finally, as a narrative rather than systematic review, this synthesis is necessarily shaped by the studies we were able to identify and prioritize, and a fully systematic update using the search strategy in Section 3.1 would be a valuable next step.

5.6 Future Directions

Looking forward, three priorities seem clearest. First, the field needs large, diverse, multi-national registries — the UK Biobank-scale work by Zhang et al. (2023) offers one workable template — paired with genuine external validation rather than repeated internal cross-validation on the same cohort. Second, prospective randomized trials measuring hard clinical endpoints (survival, hospitalization, MACE) are overdue, and funding bodies might reasonably start prioritizing them over yet another retrospective accuracy benchmark. Third, explainability and clinician-in-the-loop governance, of the kind modeled by SHAP reporting (Abzaliyeva et al., 2026) and the TIMA framework (Parise et al., 2026), should probably move from optional add-ons to expected components of any cardiovascular AI model submitted for clinical use.

6. Conclusion

This review suggests that AI-driven cardiovascular risk prediction has matured technically faster than it has matured clinically. Multimodal models and digital twins now personalize risk assessment with genuine precision, and digital cardiac rehabilitation has extended that personalization into therapy itself. Still, adoption lags badly, constrained less by algorithmic performance than

by bias, sparse prospective evidence, opacity, and workflow friction. Bridging algorithm and bedside will require diverse data, prospective trials, standardized explainability, and clinician-supervised frameworks like TIMA — a shift in priorities, more than a shift in mathematics.

Author Contributions

S.M.Y. contributed to the conception and design of the review, literature search, analysis and synthesis of the relevant evidence, and drafting of the manuscript. M.Z.A. contributed to the literature search, interpretation of the evidence on artificial intelligence, machine learning, and cardiovascular risk prediction, and critical revision of the manuscript. M.K.A.B. contributed to the analysis and interpretation of evidence concerning digital twin architectures and precision cardiovascular medicine and critically revised the manuscript. S.K.M. contributed to the literature search, synthesis of evidence on explainable AI and clinical implementation, and critical revision of the manuscript. M.A. contributed to the analysis and interpretation of cardiovascular AI models and cardiac rehabilitation evidence and critically revised the manuscript. S.F.J. contributed to the literature review, interpretation of the evidence, and critical revision of the manuscript for important intellectual content. All authors reviewed and approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Acknowledgements

The authors would like to acknowledge the EMAN Research and Testing Laboratory, Department of Pharmacology, School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia, and EMAN Biodiscoveries Sdn. Bhd., Universiti Sains Malaysia, for their academic and institutional support. The authors also acknowledge the researchers whose published studies contributed to the scientific foundation of this review.

References

- Abrar, S., Loo, C. K., & Kubota, N. (2021). A multi-agent approach for personalized hypertension risk prediction. *IEEE Access*, 9, 75090–75106. <https://doi.org/10.1109/ACCESS.2021.3074791>
- Abzaliyeva, S., Abzaliyev, K., Akhmetova, G., & Anartayev, S. (2026). Development of markers and diagnostic algorithms for the detection and prevention of premature aging of the cardiovascular system. *Journal of Clinical Medicine*, 15(10), 05141. <https://doi.org/10.3390/jcm151005141>

- Ades, P. A., Savage, P. D., Toth, M. J., Harvey-Berino, J., Schneider, J. D., Bunn, J. Y., Audelin, M. C., & Ludlow, M. (2009). High-calorie-expenditure exercise: Effects on cardiometabolic risk factors in coronary heart disease. *Circulation*, 119(20), 2671–2678. <https://doi.org/10.1161/CIRCULATIONAHA.108.834077>
- Al-Absi, H. R., Islam, T. M., Refaee, M. A., Chowdhury, M. E., & Alam, T. (2022). Cardiovascular disease diagnosis from DXA scan and retinal images using deep learning. *Sensors*, 22(12), 4310. <https://doi.org/10.3390/s22124310>
- Alimadadi, A., Aryal, S., Manandhar, I., Munroe, P. B., Joe, B., & Lazartigues, E. (2020). Machine learning-based classification of cardiomyopathies using RNA-seq data. *Genes*, 11(8), 90. <https://doi.org/10.3390/genes11080090>
- Attia, Z. I., Kapa, S., Lopez-Jimenez, F., McKie, P. M., Ladewig, D. J., Satam, G., Pellikka, P. A., Asirvatham, S. J., & Friedman, P. A. (2019). Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nature Medicine*, 25(1), 70–74. <https://doi.org/10.1038/s41591-018-0240-2>
- Baldasseroni, S., Silverii, M. V., Pratesi, A., Burgisser, C., Orso, F., Lucarelli, G., Turrin, G., Ungar, A., Marchionni, N., & Fattirolli, F. (2022). Cardiac rehabilitation in Advanced aGE after PCI for acute coronary syndromes: Predictors of exercise capacity improvement in the CR-AGE ACS study. *Aging Clinical and Experimental Research*, 34, 2195–2203. <https://doi.org/10.1007/s40520-022-02195-2>
- Beckie, T. M., & Beckstead, J. W. (2010). The effects of a cardiac rehabilitation program tailored for women on global quality of life: A randomized clinical trial. *Journal of Women's Health*, 19(11), 1977–1985. <https://doi.org/10.1089/jwh.2010.2012>
- Bernal, S. L., Martínez, J., Celdrán, A. H., & Pérez, G. M. (2021). SENIOR: An intelligent web-based ecosystem to predict high blood pressure adverse events using biomarkers and environmental data. *Applied Sciences*, 11(6), 2687. <https://doi.org/10.3390/app11062687>
- Bola, H., Rai, A., Penumaka, R., Ulucay, E., Levin, E., & Maron, D. (2026). Cardiac rehabilitation for coronary artery disease: Gaps, digital models, and the future of personalized prevention. *The American Journal of Cardiology*, 262, 16–27. <https://doi.org/10.1016/j.amjcard.2025.12.013>
- Bush, M., Kucharska-Newton, A., Simpson, R. J., Jr., Fang, G., Stürmer, T., & Brookhart, M. A. (2020). Effect of initiating cardiac rehabilitation after myocardial infarction on subsequent hospitalization in older adults. *Journal of Cardiopulmonary Rehabilitation and Prevention*, 40(2), 87–93. <https://doi.org/10.1097/HCR.0000000000000486>
- Cadavid, H., Mo, H., Arends, B., Dziopa, K., Bron, E. E., Bos, D., Georgievskia, S., & van der Harst, P. (2025). MyDigiTwin: A privacy-preserving framework for personalized cardiovascular risk prediction and scenario exploration. *Computers in Biology and Medicine*, 198, 111180. <https://doi.org/10.1016/j.combiomed.2025.111180>
- Chao, H., Shan, H., Homayounieh, F., & Kalra, M. K. (2021). Deep learning predicts cardiovascular disease risks from lung cancer screening low dose computed tomography. *Nature Communications*, 12(1), 2963. <https://doi.org/10.1038/s41467-021-22963-x>
- Chiang, P.-H., Wong, M., & Dey, S. (2021). Using wearables and machine learning to enable personalized lifestyle recommendations to improve blood pressure. *IEEE Journal of Translational Engineering in Health and Medicine*, 9, 1–13. <https://doi.org/10.1109/JTEHM.2021.3098173>
- Connolly, S. B., Kotseva, K., Jennings, C., Atray, A., Jones, J., Brown, A., Bassett, P., & Wood, D. A. (2017). Outcomes of an integrated community-based nurse-led cardiovascular disease prevention programme. *Heart*, 103(11), 840–847. <https://doi.org/10.1136/heartjnl-2016-310305>
- Davagdorj, K., Li, M., & Ryu, K. H. (2021). Local interpretable model-agnostic explanations of predictive models for hypertension. *Smart Innovation, Systems and Technologies*, 21, 426–433. https://doi.org/10.1007/978-981-33-6757-9_53
- Du, J., Chang, X., Ye, C., Zeng, Y., Yang, S., Wu, L., & Li, L. (2023). Developing a hypertension visualization risk prediction system utilizing machine learning and health check-up data. *Scientific Reports*, 13, 46281. <https://doi.org/10.1038/s41598-023-46281-y>
- Dziopa, K., Lekadir, K., van der Harst, P., & Asselbergs, F. W. (2025). Digital twins: Reimagining the future of cardiovascular risk prediction and personalised care. *Hellenic Journal of Cardiology*, 81, 4–8. <https://doi.org/10.1016/j.hjc.2024.06.002>
- Gajjar, A., Yar Shah, S. A., Gajjar, A., Allam, K. C., Beech, B., Pfaff, A., Brochu, C., Dani, S., & Ganatra, S. (2026). Cardiovascular risk stratification and surveillance in cardio-oncology: Mechanistic foundations and future directions. *Biomedicines*, 14, 01771. <https://doi.org/10.3390/biomedicines14071771>
- Ghosh, J., Chaudhury, S. R., Singh, K., & Koner, S. (2026). Development and performance analysis of machine learning methods for predicting metabolic syndrome among postmenopausal women of India. *The Journal of Nutrition*, 156, 101363. <https://doi.org/10.1016/j.nut.2026.02.012>
- Giannuzzi, P., Temporelli, P. L., Marchioli, R., Maggioni, A. P., Balestroni, G., Ceci, V., Chieffo, C., Gattone, M., Griffo, R., Schweiger, C., Tavazzi, L., Urbanati, S., Valagussa, F., Vanuzzo, D., & GOSPEL Investigators. (2008). Global secondary prevention strategies to limit event recurrence after myocardial infarction: Results of the GOSPEL study, a multicenter, randomized controlled trial from the Italian cardiac rehabilitation network. *Archives of Internal Medicine*, 168(20), 2194–2204. <https://doi.org/10.1001/archinte.168.20.2194>
- Ismail, S. M. S., Talukder, A., Awal, M. A., Siddiqui, M. M. U., Ahamad, M. M., Ahammed, B., Rawal, L. B., Alizadehsani, R., Abawajy, J., Laranjo, L., Chow, C. K., & Maddison, R. (2020). Machine learning approaches for predicting hypertension and its associated factors using population-level data. *Frontiers in Cardiovascular Medicine*, 9, 839379. <https://doi.org/10.3389/fcvm.2022.839379>
- Jamthikar, A., Gupta, D., Saba, L., Khanna, N. N., Viskovic, K., Mavrogeni, S., Laird, J. R., Sattar, N., Johri, A., Pareek, G., Miner, M., Sfrikakis, P. P., Protogerou, A., Kitas, G. D., Nicolaidis, A., Viswanathan, V., Sharma, A. M., & Suri, J. S. (2020). Cardiovascular/stroke risk predictive calculators: A comparison between statistical and machine learning models. *Cardiovascular Diagnosis and Therapy*, 10(4), 919–938. <https://doi.org/10.21037/cdt.2020.01.07>
- Jankauskas, S. S., Varzideh, F., Kansakar, U., & Santulli, G. (2026). Artificial intelligence in cardiovascular medicine: A giant step in personalized medicine? *Journal of Personalized*

- Medicine, 16(4), 00192. <https://doi.org/10.3390/jpm16040192>
- Jusic, A., Junuzovic, I., Hujdurovic, A., Zhang, L., Vausort, M., & Devaux, Y. (2023). A machine learning model based on microRNAs for the diagnosis of essential hypertension. *Non-coding RNA*, 9(6), 64. <https://doi.org/10.3390/ncrna9060064>
- Kakadiaris, I. A., Vrigkas, M., Yen, A. A., Kuznetsova, T., Budoff, M., & Naghavi, M. (2018). Machine learning outperforms ACC/AHA CVD risk calculator in MESA. *Journal of the American Heart Association*, 7(22), e009476. <https://doi.org/10.1161/JAHA.118.009476>
- Krittana Wong, C., Virk, H. U. H., Kumar, A., & Narayan, S. M. (2021). Machine learning and deep learning to predict mortality in patients with spontaneous coronary artery dissection. *Scientific Reports*, 11(1), 8992. <https://doi.org/10.1038/s41598-021-89922-y>
- Lee, S.-J., Lee, S.-H., Choi, H.-I., & Kim, K. I. (2022). Deep learning improves prediction of cardiovascular disease-related mortality and admission in patients with hypertension: Analysis of the Korean National Health Information Database. *Journal of Clinical Medicine*, 11(22), 6677. <https://doi.org/10.3390/jcm11226677>
- Li, A., Wang, Y., & Chen, H. (2025). AI driven cardiovascular risk prediction using NLP and Large Language Models for personalized medicine in athletes. *SLAS Technology*, 32, 100286. <https://doi.org/10.1016/j.slant.2025.100286>
- Li, Y., Shah, S. J., Arnett, D., Irvin, R., & Luo, Y. (2021). SNPs filtered by allele frequency improve the prediction of hypertension subtypes. *Proceedings of the 2021 IEEE International Conference on Bioinformatics and Biomedicine*, 2796–2802. <https://doi.org/10.1109/BIBM52615.2021.9669758>
- Majumdar, V., Patil, S., Singh, A., Panigrahi, S., Nagarathna, R., & Hr, N. (2025). Long-term mortality and MACE outcomes of yoga-based cardiac rehabilitation in patients with CABG: A 15-year legacy study of a randomized controlled trial. *Scientific Reports*, 15, 20068. <https://doi.org/10.1038/s41598-025-20068-w>
- Mysliwiec, A., Bartusik-Aebisher, D., Xavier Selvan, M., & Paul, A. (2026). Deep learning and cardiovascular diseases: An updated narrative review. *Journal of Clinical Medicine*, 15(8), 03053. <https://doi.org/10.3390/jcm150803053>
- Naik, A., Nalepa, J., Wijata, A. M., Mahon, J., Mistry, D., Knowles, A. T., Dawson, E. A., Lip, G. Y. H., Olier, I., & Ortega-Martorell, S. (2025). Artificial intelligence and digital twins for the personalised prediction of hypertension risk. *Computers in Biology and Medicine*, 196, 110718. <https://doi.org/10.1016/j.combiomed.2025.110718>
- Nakamura, T., Sato, K., & Yamamoto, H. (2023). Predicting future onset of 11 chronic diseases from physiological, personal lifestyle, and genetic parameters. *Journal of Personalized Medicine*, 16(3), 3238. <https://doi.org/10.3390/jpm16033238>
- Oerkild, B., Frederiksen, M., Hansen, J. F., & Prescott, E. (2012). Home-based cardiac rehabilitation is an attractive alternative to no cardiac rehabilitation for elderly patients with coronary heart disease: Results from a randomised clinical trial. *BMJ Open*, 2(1), e001820. <https://doi.org/10.1136/bmjopen-2012-001820>
- Osonuga, A., Dave, M., Ogieuhi, I. J., & Olawade, D. B. (2026). Wearable devices and machine learning in cardiovascular monitoring: Current evidence and future directions for precision medicine. *Journal of Personalized Medicine*, 16(7), 00377. <https://doi.org/10.3390/jpm16070377>
- Osoro, L., Zylla, M., Braunschweig, F., Leyva, F., Figueras, J., Pürerfellner, H., Merino, J. L., Casado-Arroyo, R., & Boriani, G. (2026). From algorithm development to clinical implementation: A systematic review of artificial intelligence in cardiovascular medicine. *Technologies*, 14(8), 00511. <https://doi.org/10.3390/technologies14080511>
- Owolabi, K. M., et al. (2026). Integrated Digital Twin framework for personalized stress and physiological state prediction. *Array*, 31, 101142. <https://doi.org/10.1016/j.array.2026.101142>
- Parise, G., Ceravolo, R., Lucà, F., Gulizia, M. M., Tetta, C., Parise, O., Nardi, F., Grimaldi, M., & Gelsomino, S. (2026). The cardiologist driving synthetic AI: The TIMA method for clinically supervised synthetic data generation. *Journal of Clinical Medicine*, 15(3), 01351. <https://doi.org/10.3390/jcm15031351>
- Plüess, C. E., Billing, E., Held, C., Henriksson, P., Kiessling, A., Karlsson, M. R., & Wallén, H. N. (2011). Long-term effects of an expanded cardiac rehabilitation programme after myocardial infarction or coronary artery bypass surgery: A five-year follow-up of a randomized controlled study. *Clinical Rehabilitation*, 25(1), 79–87. <https://doi.org/10.1177/0269215510377545>
- Rech, L. (2026). From automated ECG interpretation to multimodal cardiovascular intelligence: The evolution of artificial intelligence in cardiovascular medicine. *Medical Sciences*, 14(3), 00434. <https://doi.org/10.3390/medsci14030434>
- Sahu, D., Ganguly, T., Mann, A., Gupta, Y., & Van Nynatten, L. R. (2025). Emerging technologies for exploring the cellular mechanisms in vascular diseases. *International Journal of Molecular Sciences*, 27(1), 00164. <https://doi.org/10.3390/ijms270100164>
- Shahbazi, Z., & Nowaczyk, S. (2025). Towards personalized cardiometabolic risk prediction: A fusion of exposome and AI. *Heliyon*, 11, e40859. <https://doi.org/10.1016/j.heliyon.2025.e40859>
- Singh, M., Kumar, A., Khanna, N. N., Laird, J. R., Nicolaidis, A., Faa, G., Johri, A. M., Mantella, L. E., Fernandes, J. F. E., Teji, J. S., & Suri, J. S. (2024). Artificial intelligence for cardiovascular disease risk assessment in personalised framework: A scoping review. *eClinicalMedicine*, 73, 102660. <https://doi.org/10.1016/j.eclinm.2024.102660>
- Steinfeldt, J., Batinica, B., Liang, J., & Suri, J. S. (2022). NeuralCVD-DSM: A deep learning-based framework for 10-year MACE risk prediction. *Computers in Biology and Medicine*, 142, 105102. <https://doi.org/10.1016/j.combiomed.2022.105102>
- Tonet, E., Raisi, A., Zagnoni, S., Chiaranda, G., Pasanisi, G., Aschieri, D., D'Intino, P. E., Pavasini, R., Cimaglia, P., Campana, R., Vitali, F., Piva, T., Casella, G., Cagliioni, S., Zerbini, V., Bugani, G., Cocco, M., Menegatti, E., De Raffe, M., Mandini, S., Martella, D., Pesenti, N., Mazzoni, G., Biscaglia, S., Volpato, S., Grazzi, G., & Campo, G. (2025). Multidomain rehabilitation for older patients with myocardial infarction (PiPELINE Trial). *New England Journal of Medicine*, 393(11), 973–982. <https://doi.org/10.1056/NEJMoa2411893>
- Varnfield, M., Karunanithi, M., Lee, C.-K., Honeyman, E., Arnold, D., Ding, H., Layland, C., Ward, H., & Walters, D. L. (2014). Smartphone-based home care model improved use of cardiac

rehabilitation in postmyocardial infarction patients: Results from a randomised controlled trial. *Heart*, 100(22), 1770–1779. <https://doi.org/10.1136/heartjnl-2014-305783>

West, R. R., Jones, D. A., & Henderson, H. H. (2012). Rehabilitation after Myocardial Infarction Trial (RAMIT): Multi-centre randomised controlled trial of comprehensive cardiac rehabilitation in patients following acute myocardial infarction. *Heart*, 98(8), 637–644. <https://doi.org/10.1136/heartjnl-2011-301131>

Widmer, R. J., Allison, T. G., Lennon, R., Lopez-Jimenez, F., Lerman, L. L., & Lerman, A. (2017). Digital health intervention during cardiac rehabilitation: A randomized controlled trial. *American Heart Journal*, 188, 65–72. <https://doi.org/10.1016/j.ahj.2017.02.016>

Wood, D., Kotseva, K., Connolly, S., & EUROACTION Collaborators. (2008). Nurse-coordinated multidisciplinary, family-based cardiovascular disease prevention programme (EUROACTION) for patients with coronary heart disease and asymptomatic individuals at high risk: A paired, cluster-randomised controlled trial. *The Lancet*, 371(9629), 1999–2012. [https://doi.org/10.1016/S0140-6736\(08\)60874-0](https://doi.org/10.1016/S0140-6736(08)60874-0)

Xie, Y., Jiang, B., Gong, E., & CT-Based Stroke Prediction Group. (2019). Use of gradient boosting machine learning to predict patient outcome in acute ischemic stroke on the basis of imaging, demographic, and clinical information. *American Journal of Roentgenology*, 212(1), 44–51. <https://doi.org/10.2214/AJR.18.20221>

Yao, X., & Wang, Q. (2023). A classification approach for sports behavior data using Random Forest. *Proceedings of SPIE - International Society for Optical Engineering*, 20–28. <https://doi.org/10.1117/12.2687417>

Zhang, N., Yang, G., Gao, Z., & UK Biobank Research Consortium. (2023). Predicting risk of 9 chronic diseases from clinical and demographic parameters in the UK Biobank using machine learning. *Scientific Reports*, 13(2), 553–562. <https://doi.org/10.1038/s41598-023-35405-x>