

Viewpoint

Operationalizing AI-Enabled Cardiovascular Biomarkers: A Clinician-Centered Framework for Validation, Governance, and Workflow Integration

Dabeluchi Chiedozi Ngwu, MD

Department of Surgery, King Abdullah Hospital, Bisha, Asir, Saudi Arabia

Corresponding Author:

Dabeluchi Chiedozi Ngwu, MD

Department of Surgery

King Abdullah Hospital

Al Nakhil

Bisha, Asir 67714

Saudi Arabia

Phone: 966 543423453

Email: doctordabs@gmail.com

Abstract

Cardiovascular biomarkers are increasingly extracted or interpreted using artificial intelligence (AI) applied to electrocardiograms, imaging, laboratory measurements, electronic health records, wearable devices, and longitudinal data. Predictive performance alone, however, does not establish that a measurement is valid, that a model is transportable and calibrated, or that acting on its output improves care. Existing resources provide essential but complementary foundations: the US Food and Drug Administration–National Institutes of Health (FDA-NIH) BEST (Biomarkers, EndpointS, and other Tools) resource standardizes biomarker terminology; the V3 and V3+ frameworks address verification, analytical validation, clinical validation, and usability of digitally measured signals; prediction-model and trustworthy-AI guidance addresses reporting, risk of bias, early clinical evaluation, and deployability; and regulatory qualification pathways evaluate evidence within a defined context of use. A remaining practical challenge is translating these complementary requirements into accountable clinical action within cardiovascular workflows. This viewpoint proposes a clinician-centered operational framework organized around validation, governance, and workflow integration. For each pillar, it identifies accountable actors, required steps, documented outputs, and escalation or stop rules. Validation establishes whether the input measurement and model are fit for the intended population and decision. Governance assigns institutional authorization, clinical ownership, version control, monitoring, and authority to restrict, pause, or withdraw the intervention. Workflow integration specifies who receives the output, what confirmatory action follows, how disagreement is handled, and how decisions are documented. A worked example of an AI-enabled electrocardiographic screening output for possible left ventricular systolic dysfunction illustrates the pathway from local evaluation to echocardiographic confirmation and lifecycle monitoring. The framework does not replace established validation or regulatory standards; it operationalizes them as a clinician-centered and institutionally accountable pathway from validated signal to governed, patient-facing action.

JMIR Cardio 2026;10:e103401; doi: [10.2196/103401](https://doi.org/10.2196/103401)

Keywords: artificial intelligence; cardiovascular biomarkers; digital biomarker; clinical validation; clinical decision support systems; workflow integration; governance

Introduction

Cardiovascular biomarkers influence consequential decisions, including diagnosis, triage, preventive treatment, referral, device therapy, and longitudinal monitoring. Artificial intelligence (AI) is expanding this landscape by extracting or interpreting signals from electrocardiograms (ECGs),

imaging, laboratory measurements, electronic health records (EHRs), wearable devices, omics platforms, and longitudinal clinical data. These methods may reveal patterns that conventional thresholds or models do not capture and support earlier recognition or more individualized assessment [1–12]. Their clinical value, however, depends on more than performance in a development dataset.

Terminology is important because several related objects are often conflated. Following the US Food and Drug Administration–National Institutes of Health (FDA-NIH) BEST (Biomarkers, EndpointS, and other Tools) resource, a biomarker is a defined characteristic measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention [13]. A digital biomarker is a biomarker measured or derived using digital technology [14,15]. In this viewpoint, an AI-enabled cardiovascular biomarker is a cardiovascular biomarker whose extraction, quantification, combination, or interpretation materially depends on an AI method. Digital biomarkers and AI-enabled biomarkers overlap, but neither wholly contains the other: a digitally captured heart-rate signal may require no AI, whereas AI may interpret conventional laboratory or imaging measurements. A prediction or risk score is a model-generated estimate and is not automatically a biomarker. A complete clinical decision-support intervention comprises the data input, model or interpretive process, output, interface, intended user, response pathway, and oversight arrangements.

These distinctions matter because an AI-enabled biomarker output may be associated with disease without being suitable for clinical action. A model may discriminate well yet be poorly calibrated, lose performance across settings or patient groups, or trigger testing and treatment whose benefits and harms have not been established. Reliability concerns the consistency or reproducibility of a measurement, whereas validity concerns whether it accurately captures the intended characteristic. External validation tests performance in data separated from development by place, time, population, or setting. Calibration concerns agreement between predicted probabilities and observed event frequencies in the intended population, particularly around thresholds used for action. Clinical utility asks whether using the output improves decisions or patient-relevant outcomes enough to justify its consequences [2-4,16-23]. Predictive performance in a development dataset is therefore only one component of validation and does not establish measurement validity, transportability, calibration, or clinical utility in the intended setting.

Existing frameworks provide complementary foundations. BEST standardizes biomarker terminology and categories [13]. V3 addresses verification, analytical validation, and clinical validation for biometric monitoring technologies, while V3+ extends this foundation to usability validation for sensor-based digital health technologies [14,15]. Cardiovascular guidance addresses the predictive and clinical utility of novel biomarkers and models [2-4]. TRIPOD+AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis + Artificial Intelligence) and PROBAST+AI (Prediction Model Risk of Bias Assessment Tool + Artificial Intelligence) address reporting, risk of bias, and applicability; DECIDE-AI (Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence) addresses early-stage clinical evaluation; and FUTURE-AI (Fairness, Universality, Traceability, Usability, Robustness, and

Explainability in Artificial Intelligence) addresses trustworthy and deployable AI [23-26]. FDA and European Medicines Agency qualification pathways also anchor evidence to a defined context of use in medical product development [27, 28]. V3+ focuses on usability, whereas the workflow-integration pillar considered here addresses the broader clinical pathway through which an output is received, confirmed, acted upon, documented, and monitored.

This viewpoint does not propose a new biomarker taxonomy or replace established validation standards. It integrates these complementary requirements into a clinician-centered implementation pathway that links fit-for-purpose evidence to named clinical and institutional responsibilities, predefined clinical responses, and lifecycle monitoring. The framework is organized around 3 interconnected pillars: validation, governance, and workflow integration. A visual abstract of the proposed framework is presented in [Multimedia Appendix 1](#).

From Biomarker Signal to Clinical Intervention

Biomarker translation can fail before AI is introduced. The underlying assay, sensor, image acquisition process, or data stream may be unstable; the development cohort may be narrow; the outcome may be poorly defined; or the proposed use may not correspond to a meaningful clinical decision. Sample handling, device characteristics, preprocessing, timing, biological variability, kidney function, comorbidity, medication use, and missing data can all alter the behavior of a cardiovascular signal. If the input is unreliable or invalid, a more sophisticated model cannot repair the evidentiary foundation [2,3,16-19].

Transportability presents a second challenge. Cardiovascular populations and workflows vary across hospitals, laboratories, ECG systems, imaging platforms, EHRs, referral practices, and treatment pathways. Average performance may conceal important failures across age, sex, comorbidity, ancestry or ethnicity (when these data are appropriately collected), kidney function, device access, and care setting. Because local prevalence also changes positive and negative predictive values, a threshold developed in a tertiary referral cohort may generate an unacceptable false-positive burden in primary care. External and local validation must therefore examine calibration, subgroup performance, and the consequences of the proposed threshold, not discrimination alone [23,24,29-31].

Clinical utility remains distinct from statistical performance. A reproducible and externally validated output may add little beyond existing assessment, identify patients for whom no beneficial response is available, or generate unnecessary testing, anxiety, treatment, and cost. Incremental performance, decision-curve analysis, pragmatic evaluation, and patient-relevant outcomes can help determine whether acting on the output is preferable to current care [4,17,20,22, 32]. The relevant unit of evaluation is therefore the com-

plete clinical decision-support intervention: who receives the output, what action follows, and what happens to the patient.

AI can nevertheless add meaningful capabilities by combining multimodal signals, identifying nonlinear relationships, supporting phenotype discovery, and enabling longitudinal interpretation. It may also support lifecycle monitoring by detecting calibration drift, changes in data quality, and subgroup deterioration. These capabilities become clinically meaningful only when the output is tied to a defined intended use, accountable ownership, and an evaluable response pathway. Otherwise, AI may scale weak evidence or ambiguous action more rapidly than conventional biomarker development [5,8-12,21,25,26,32-34].

A Clinician-Centered Operational Framework

Overview

The 3 pillars are connected, rather than parallel, checklists. Validation determines what the output can legitimately support. Governance determines who may authorize, monitor, modify, pause, or withdraw it. Workflow integration determines what happens when the output enters clinical care. Each pillar is therefore defined by accountable actors, required steps, documented outputs, and escalation or stop rules.

Validation

Accountable actors include the developer or vendor, a local clinical lead, an expert in the relevant measurement domain, and a biostatistics or data-science team. Their first task is to define the intended use: the target population, setting, decision, timing, comparator, and actions that may follow. Verification and analytical validation should establish that the acquisition system, assay, device, preprocessing pipeline, and data transfer perform as intended. Clinical validation should establish the relationship between the measured characteristic or model output and the relevant clinical state or outcome [13-15].

Where an AI model produces a probability, classification, or risk score, evaluation should include separation from development data, discrimination, calibration, clinically relevant subgroup performance, robustness to local devices and data pipelines, and positive or negative predictive value in the intended population. Thresholds should be evaluated in relation to downstream action rather than selected for statistical convenience. Clinical utility should determine whether the output adds value beyond current assessment and whether expected benefits justify confirmatory testing, referral, treatment, or monitoring [4,20,22-24,26,27,32]. When important subgroup differences are identified, evaluation should move beyond reporting them to determining whether data revision, threshold adjustment, workflow safeguards, restricted use, or nondeployment is required.

The resulting local evidence report should specify the approved population, exclusions, model version, data

requirements, threshold, uncertainty, expected downstream burden, and residual limitations. Deployment should not proceed when the input is insufficiently reliable or valid, performance is not transportable or calibrated for the intended setting, important subgroup risks remain unresolved, or no justifiable clinical response follows.

Governance

Governance requires both a named clinical service owner and a multidisciplinary institutional oversight function. The exact structure may vary, but it should include clinical, informatics, data-science, quality and safety, privacy, and legal or risk-management expertise, with patient or public perspectives incorporated when warranted by the intended use and patient-facing consequences. Vendors may provide technical documentation and model updates, but the institution retains responsibility for local authorization and oversight and should maintain an independent record of the deployed version.

Before deployment, the oversight body should approve the intended use, eligible population, prohibited uses, data dependencies, clinical owner, monitoring metrics, review cadence, incident-reporting route, and procedure for material updates. Version changes, altered input pipelines, new devices, population shifts, or workflow redesign may invalidate prior assumptions and should trigger proportionate reassessment. Prespecified monitoring criteria may include material calibration deterioration, unacceptable subgroup disparities, increased false-positive burden, clinically important missed cases, harmful downstream action, or substantial changes in data quality or software behavior [25,26,32-37].

Governance should produce formal authorization, a named accountable owner, a version and change log, a monitoring protocol, and procedures for recalibration, restricted use, temporary suspension, and withdrawal. It is effective only when someone has both the information and authority to act when performance or clinical consequences become unacceptable.

Workflow Integration

Workflow integration is owned by the clinical pathway rather than the algorithm. The clinical service lead, intended users, informatics team, and relevant nursing or allied professionals should determine who receives the output, when and where it appears, what confirmatory action is appropriate, how urgency is communicated, and how uncertainty or disagreement is documented. The output should be designed around a decision point rather than added as an unprioritized dashboard item or interruptive alert.

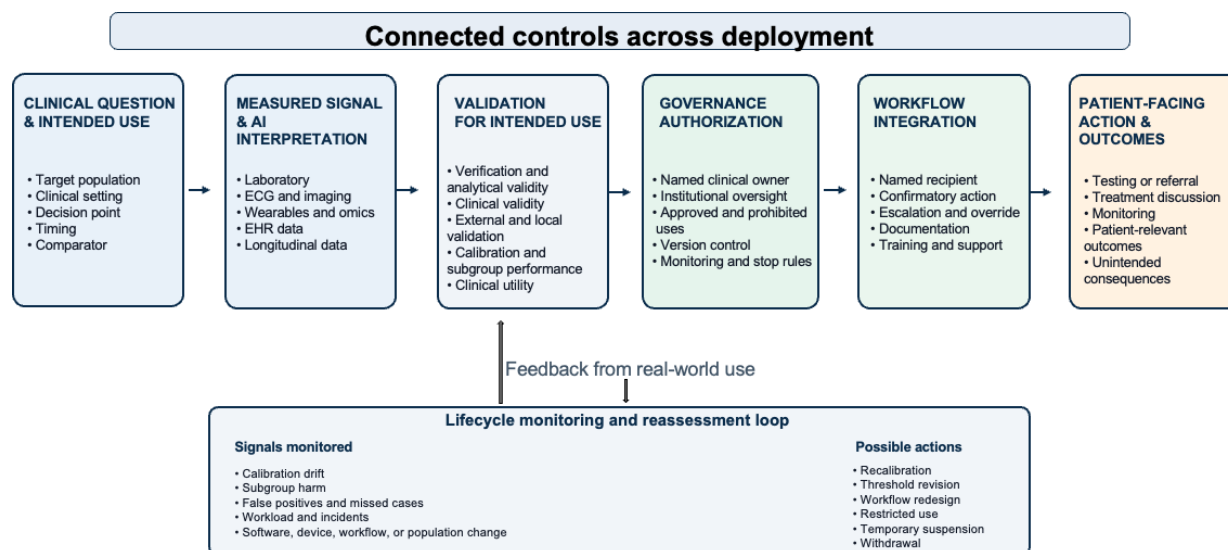
A defined response pathway specifies the recipient; the threshold or conditions that trigger review; permissible actions; any required confirmation before treatment; escalation and de-escalation rules; when an override rationale is required; and how responsibility transfers across teams. Documentation should be proportionate to clinical risk. When an AI output conflicts with clinical judgment, the pathway should preserve clinician authority while making high-risk disagreements visible for review. Patient communication

should be specified when the result leads to testing, referral, or treatment discussion.

The pathway should be written, tested, and embedded in the relevant clinical system. Implementation readiness also requires role-specific training so that intended users understand the output, its limitations, the required confirmatory actions, and the circumstances in which escalation or override is appropriate. Deployment should be delayed when the recipient is unclear, confirmatory capacity is unavailable,

the alert cannot be acted on within the required time, the output encourages unsupported treatment, or responsibility is fragmented across teams. Workflow simulation, silent or shadow-mode testing, and early prospective evaluation can identify these failures before broad implementation [9-12,26,32]. Figure 1 summarizes the clinician-centered operational pathway from intended use and signal measurement through validation, governance authorization, workflow integration, patient-facing action, and lifecycle reassessment.

Figure 1. Clinician-centered operational pathway for AI-enabled cardiovascular biomarkers. AI: artificial intelligence; ECG: electrocardiogram; EHR: electronic health record.



The pathway begins with a defined clinical question and intended use, followed by measurement and AI-enabled interpretation, validation for the intended population and decision, governance authorization, workflow integration, and patient-facing action. Validation, governance, and workflow integration operate as connected controls across deployment. Postdeployment monitoring evaluates calibration drift, subgroup harm, false-positive and missed-case patterns, workload, incidents, and changes in software, devices, workflows, or populations. Findings may trigger recalibration, threshold revision, workflow redesign, restricted use, temporary suspension, or withdrawal. The pathway, therefore, does not end at model validation; authorization, clinical response, and lifecycle reassessment determine whether an output becomes a safe clinical decision-support intervention.

Applying the Framework: AI-ECG Screening for Left Ventricular Systolic Dysfunction

Consider an AI-enabled ECG model intended to identify patients with possible left ventricular systolic dysfunction who may benefit from echocardiography. Development studies have shown that AI can detect a signal associated with low ejection fraction from a standard ECG, and pragmatic evaluation has shown that presenting this output in routine primary care can increase detection of low ejection fraction [38,39]. Its intended use must remain narrow: the output

is a screening prompt for confirmatory assessment, not a diagnosis of heart failure or an independent indication for treatment.

Validation would require confirmation that the locked model is compatible with local ECG acquisition, preprocessing, and data-transfer systems, followed by testing in a local or otherwise suitably representative population independent of model development. Evaluation should include discrimination, calibration, positive predictive value, subgroup performance, and the expected number of echocardiograms generated per confirmed case. A threshold acceptable in a referral population may produce excessive false positives or overwhelm echocardiography capacity in lower-prevalence primary care.

Governance would require institutional authorization, a named clinical owner, and documented responsibility for version control and monitoring. Relevant measures would include false-positive and false-negative patterns, subgroup performance, echocardiography yield, time to confirmation, alert burden, and inappropriate downstream action. Material performance deterioration or a change in the software or data pipeline should trigger reassessment and, when necessary, restricted use or suspension.

Within the workflow, the output would reach a named ordering clinician or designated results-management team with a clear recommendation for clinical review and, where appropriate, confirmatory echocardiography. The pathway

would specify urgency, referral options, documentation, and how to respond when recent imaging or an alternative explanation is already available. No diagnosis of heart failure or disease-specific treatment should follow solely from the AI-ECG output. The example illustrates how a technically valid prediction becomes part of a safer clinical decision-support intervention only when evidence, ownership, confirmation, and monitoring are connected.

The requirements are role-based rather than tied to a single committee structure because institutions differ in size and resources. Across settings, explicit ownership, documented evidence, a tested response pathway, and authority to restrict or stop use should remain nonnegotiable.

Operational Requirements

Table 1 translates the 3 pillars and the cross-cutting lifecycle-monitoring function into minimum operational requirements.

Table 1. Minimum operational requirements for validation, governance, workflow integration, and lifecycle monitoring of AI-enabled cardiovascular biomarkers.

Domain	Accountable actors	Minimum operational requirements	Required outputs and action threshold
Validation	Developer or vendor; local clinical lead; measurement-domain expert; biostatistics or data-science team	Define intended use and population; verify acquisition and preprocessing; conduct analytical and clinical validation; assess transportability, calibration, subgroup performance, and clinical utility	Local evidence report specifying population, exclusions, model version, decision threshold, uncertainty, and limitations. Do not deploy when validity, calibration, transportability, subgroup safety, or actionable utility is inadequate.
Governance	Named clinical service owner; multidisciplinary institutional oversight body; quality, safety, and informatics representatives	Authorize intended and prohibited uses; document data dependencies and version; assign monitoring; review incidents and material updates; define restriction, pause, and withdrawal procedures	Authorization record, accountable owner, change log, and monitoring plan. Reassess, restrict, pause, or withdraw use after material performance deterioration, inequity, harmful consequences, or major system change.
Workflow integration	Clinical pathway owner; intended clinicians; informatics; nursing and allied professionals as relevant	Identify recipient, timing, interface, confirmatory action, urgency, escalation, override, documentation, patient communication, and training requirements	Written and tested response pathway, workflow map, and training materials. Delay deployment when responsibility is unclear, action capacity is unavailable, alert burden is unsafe, or unsupported treatment may result.
Lifecycle monitoring (cross-cutting) ^a	Clinical owner; governance body; analytics or quality team; vendor as applicable	Track calibration, yield, data quality, subgroup performance, missed cases, false positives, downstream actions, workload, incidents, and version changes	Monitoring report and corrective-action record. Trigger recalibration, threshold revision, workflow redesign, restricted use, suspension, or withdrawal according to prespecified risk-based criteria.

^aLifecycle monitoring is cross-cutting and supports all 3 pillars.

Implications for Evaluation and Reporting

Investigators, peer reviewers, health systems, and regulators should evaluate AI-enabled cardiovascular biomarker studies beyond discrimination alone. Reports should clarify whether the object under evaluation is a biomarker, digital biomarker, prediction model, or complete clinical decision-support intervention; define the intended use and target population; describe measurement verification and validity; report external validation, calibration, and clinically relevant subgroup performance; and explain the comparator and downstream consequences. When deployment is proposed, reports should also identify the accountable owner, response pathway, prohibited uses, monitoring plan, and criteria for recalibration, suspension, or withdrawal.

The evidentiary question, therefore, shifts from “Does the model predict?” to “Are the measured signal, model, institution, and response pathway jointly fit for the intended use?” A valid measurement and a well-performing model

do not by themselves establish the clinical value or safety of deployment, and health systems should not infer either from regulatory status, technical performance, or publication. Prospective evaluation should examine patient-relevant outcomes, unintended consequences, clinician workload, equity, and the behavior of the complete clinical decision-support intervention over time [25,26,32].

Conclusion

AI can expand the information obtainable from cardiovascular biomarkers and related clinical data, but technical performance alone does not establish that an input is valid, an output is transportable and calibrated, or that acting on it improves care.

The proposed framework links established evidentiary standards to clinical implementation by connecting validation, accountable governance, defined response pathways, and lifecycle monitoring. The worked AI-ECG example shows why these functions must be designed together. Clinical trust

should therefore be earned at the level of the complete ownership, a defined response pathway, and continuing clinical decision-support intervention: a validated signal, a evidence that the intervention remains safe and useful in fit-for-purpose model, accountable clinical and institutional practice.

Acknowledgments

The author thanks colleagues and external reviewers who provided informal feedback during manuscript development. The author also acknowledges the faculty and course contributors of the MIT Artificial Intelligence in Health Care program and the Harvard programs in health care transformation, whose teaching and discussions informed the author's thinking about clinical artificial intelligence, implementation, and organizational change.

Funding

No funding was received for this work.

Data Availability

No new data were generated or analyzed in this article. Data sharing is not applicable because this is a viewpoint based on the interpretation of published literature and the development of a conceptual framework.

Authors' Contributions

DCN conceived the manuscript, developed the clinician-centered framework, drafted the manuscript, revised the intellectual content, and approved the final version for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Visual abstract.

[[PNG File \(Portable Network Graphics File\), 119 KB-Multimedia Appendix 1](#)]

References

1. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. Mar 2001;69(3):89-95. [doi: [10.1067/mcp.2001.113989](#)] [Medline: [11240971](#)]
2. Morrow DA, de Lemos JA. Benchmarks for the assessment of novel cardiovascular biomarkers. *Circulation*. Feb 27, 2007;115(8):949-952. [doi: [10.1161/CIRCULATIONAHA.106.683110](#)] [Medline: [17325253](#)]
3. Hlatky MA, Greenland P, Arnett DK, et al. Criteria for evaluation of novel markers of cardiovascular risk: a scientific statement from the American Heart Association. *Circulation*. May 5, 2009;119(17):2408-2416. [doi: [10.1161/CIRCULATIONAHA.109.192278](#)] [Medline: [19364974](#)]
4. Khan SS, Greenland P, Hayman LL, et al. Criteria to assess the predictive and clinical utility of novel models, biomarkers, and tools for risk of cardiovascular disease: a scientific statement from the American Heart Association. *Circulation*. Mar 17, 2026;153(11):e953-e970. [doi: [10.1161/CIR.0000000000001401](#)] [Medline: [41669831](#)]
5. Krittanawong C, Rogers AJ, Johnson KW, et al. Integration of novel monitoring devices with machine learning technology for scalable cardiovascular management. *Nat Rev Cardiol*. Feb 2021;18(2):75-91. [doi: [10.1038/s41569-020-00445-9](#)] [Medline: [33037325](#)]
6. Attia ZI, Noseworthy PA, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet*. Sep 7, 2019;394(10201):861-867. [doi: [10.1016/S0140-6736\(19\)31721-0](#)] [Medline: [31378392](#)]
7. Perez MV, Mahaffey KW, Hedlin H, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med*. Nov 14, 2019;381(20):1909-1917. [doi: [10.1056/NEJMoa1901183](#)] [Medline: [31722151](#)]
8. Chaturvedi A, Prabhakaran D. Transforming cardiovascular care with digital health: the past, progress, and promise. *JACC Adv*. Sep 2024;3(9):101183. [doi: [10.1016/j.jacadv.2024.101183](#)] [Medline: [39220713](#)]
9. Stremmel C, Breitschwerdt R. Digital transformation in the diagnostics and therapy of cardiovascular diseases: comprehensive literature review. *JMIR Cardio*. Aug 30, 2023;7:e44983. [doi: [10.2196/44983](#)] [Medline: [37647103](#)]
10. Abedi A, Verma A, Jain D, et al. AI-driven real-time monitoring of cardiovascular conditions with wearable devices: scoping review. *JMIR Mhealth Uhealth*. Nov 11, 2025;13:e73846. [doi: [10.2196/73846](#)] [Medline: [41218203](#)]
11. Bozyel S, Şimşek E, Koçyiğit Burunkaya D, et al. Artificial intelligence-based clinical decision support systems in cardiovascular diseases. *Anatol J Cardiol*. Jan 7, 2024;28(2):74-86. [doi: [10.14744/AnatolJCardiol.2023.3685](#)] [Medline: [38168009](#)]

12. Cho Y, Yoon M, Kim J, et al. Artificial intelligence-based electrocardiographic biomarker for outcome prediction in patients with acute heart failure: prospective cohort study. *J Med Internet Res*. Jul 3, 2024;26:e52139. [doi: [10.2196/52139](https://doi.org/10.2196/52139)] [Medline: [38959500](https://pubmed.ncbi.nlm.nih.gov/38959500/)]
13. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource. Food and Drug Administration (US); National Institutes of Health (US); 2016. URL: <https://www.ncbi.nlm.nih.gov/books/NBK326791> [Accessed 2026-07-14]
14. Goldsack JC, Coravos A, Bakker JP, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *NPJ Digit Med*. 2020;3:55. [doi: [10.1038/s41746-020-0260-4](https://doi.org/10.1038/s41746-020-0260-4)] [Medline: [32337371](https://pubmed.ncbi.nlm.nih.gov/32337371/)]
15. Bakker JP, Barge R, Centra J, et al. V3+ extends the V3 framework to ensure user-centricity and scalability of sensor-based digital health technologies. *NPJ Digit Med*. Jan 24, 2025;8(1):51. [doi: [10.1038/s41746-024-01322-2](https://doi.org/10.1038/s41746-024-01322-2)] [Medline: [39856145](https://pubmed.ncbi.nlm.nih.gov/39856145/)]
16. Institute of Medicine (US) Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease. Micheel CM, Ball JR, editors. Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease. National Academies Press (US); 2010. [doi: [10.17226/12869](https://doi.org/10.17226/12869)] [Medline: [25032382](https://pubmed.ncbi.nlm.nih.gov/25032382/)]
17. Ioannidis JPA, Bossuyt PMM. Waste, leaks, and failures in the biomarker pipeline. *Clin Chem*. May 2017;63(5):963-972. [doi: [10.1373/clinchem.2016.254649](https://doi.org/10.1373/clinchem.2016.254649)] [Medline: [28270433](https://pubmed.ncbi.nlm.nih.gov/28270433/)]
18. Saleh S, George J, Kott KA, Meikle PJ, Figtree GA. The translation and commercialisation of biomarkers for cardiovascular disease-a review. *Front Cardiovasc Med*. 2022;9:897106. [doi: [10.3389/fcvm.2022.897106](https://doi.org/10.3389/fcvm.2022.897106)] [Medline: [35722087](https://pubmed.ncbi.nlm.nih.gov/35722087/)]
19. Cunningham JW, Pfeiffer MA. Clinical translation of cardiovascular biomarkers - a long and uncertain road. *NEJM Evid*. Oct 2022;1(10):EVIDe2200222. [doi: [10.1056/EVIDe2200222](https://doi.org/10.1056/EVIDe2200222)] [Medline: [38319843](https://pubmed.ncbi.nlm.nih.gov/38319843/)]
20. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res*. 2019;3(1):18. [doi: [10.1186/s41512-019-0064-7](https://doi.org/10.1186/s41512-019-0064-7)] [Medline: [31592444](https://pubmed.ncbi.nlm.nih.gov/31592444/)]
21. Zehra B, Vinod N, BinEshaq S, et al. Multiomics approaches to cardiovascular disease: technological innovations and clinical translation. *Am J Physiol Heart Circ Physiol*. Jun 1, 2026;330(6):H1752-H1784. [doi: [10.1152/ajpheart.00024.2026](https://doi.org/10.1152/ajpheart.00024.2026)] [Medline: [42013055](https://pubmed.ncbi.nlm.nih.gov/42013055/)]
22. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. Jan 30, 2008;27(2):157-172. [doi: [10.1002/sim.2929](https://doi.org/10.1002/sim.2929)] [Medline: [17569110](https://pubmed.ncbi.nlm.nih.gov/17569110/)]
23. Moons KGM, Damen JAA, Kaul T, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*. Mar 24, 2025;388:e082505. [doi: [10.1136/bmj-2024-082505](https://doi.org/10.1136/bmj-2024-082505)] [Medline: [40127903](https://pubmed.ncbi.nlm.nih.gov/40127903/)]
24. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. Apr 16, 2024;385:e078378. [doi: [10.1136/bmj-2023-078378](https://doi.org/10.1136/bmj-2023-078378)] [Medline: [38626948](https://pubmed.ncbi.nlm.nih.gov/38626948/)]
25. Lekadir K, Frangi AF, Porras AR, et al. FUTURE-AI: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. *BMJ*. Feb 5, 2025;388:e081554. [doi: [10.1136/bmj-2024-081554](https://doi.org/10.1136/bmj-2024-081554)] [Medline: [39909534](https://pubmed.ncbi.nlm.nih.gov/39909534/)]
26. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med*. May 2022;28(5):924-933. [doi: [10.1038/s41591-022-01772-9](https://doi.org/10.1038/s41591-022-01772-9)] [Medline: [35585198](https://pubmed.ncbi.nlm.nih.gov/35585198/)]
27. Context of use. US Food and Drug Administration. Jan 21, 2026. URL: <https://www.fda.gov/drugs/biomarker-qualification-program/context-use> [Accessed 2026-07-14]
28. Qualification of novel methodologies for medicine development. European Medicines Agency. URL: <https://www.ema.europa.eu/en/human-regulatory-overview/research-development/scientific-advice-protocol-assistance/qualification-novel-methodologies-medicine-development> [Accessed 2026-07-14]
29. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. Oct 25, 2019;366(6464):447-453. [doi: [10.1126/science.aax2342](https://doi.org/10.1126/science.aax2342)] [Medline: [31649194](https://pubmed.ncbi.nlm.nih.gov/31649194/)]
30. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. *Ann Intern Med*. Dec 18, 2018;169(12):866-872. [doi: [10.7326/M18-1990](https://doi.org/10.7326/M18-1990)] [Medline: [30508424](https://pubmed.ncbi.nlm.nih.gov/30508424/)]
31. Yagi R, Goto S, Katsumata Y, MacRae CA, Deo RC. Importance of external validation and subgroup analysis of artificial intelligence in the detection of low ejection fraction from electrocardiograms. *Eur Heart J Digit Health*. Dec 2022;3(4):654-657. [doi: [10.1093/ehjdh/ztac065](https://doi.org/10.1093/ehjdh/ztac065)] [Medline: [36710903](https://pubmed.ncbi.nlm.nih.gov/36710903/)]

32. Jain SS, Goto S, Hall JL, et al. Pragmatic approaches to the evaluation and monitoring of artificial intelligence in health care: a science advisory from the American Heart Association. *Circulation*. Dec 9, 2025;152(23):e433-e442. [doi: [10.1161/CIR.0000000000001400](https://doi.org/10.1161/CIR.0000000000001400)] [Medline: [41208719](#)]
33. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med*. Sep 2019;25(9):1337-1340. [doi: [10.1038/s41591-019-0548-6](https://doi.org/10.1038/s41591-019-0548-6)] [Medline: [31427808](#)]
34. Davis SE, Lasko TA, Chen G, Matheny ME. Calibration drift among regression and machine learning models for hospital mortality. *AMIA Annu Symp Proc*. 2017;2017:625-634. [Medline: [29854127](#)]
35. Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847. *EUR-Lex*. Feb 11, 2025. URL: <http://data.europa.eu/eli/reg/2025/327/oj> [Accessed 2026-05-31]
36. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). *EUR-Lex*. Jun 13, 2024. URL: <http://data.europa.eu/eli/reg/2024/1689/oj> [Accessed 2026-05-31]
37. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. *EUR-Lex*. Apr 5, 2017. URL: <http://data.europa.eu/eli/reg/2017/745/oj> [Accessed 2026-05-31]
38. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nat Med*. Jan 2019;25(1):70-74. [doi: [10.1038/s41591-018-0240-2](https://doi.org/10.1038/s41591-018-0240-2)] [Medline: [30617318](#)]
39. 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*. May 2021;27(5):815-819. [doi: [10.1038/s41591-021-01335-4](https://doi.org/10.1038/s41591-021-01335-4)] [Medline: [33958795](#)]

Abbreviations

AI: artificial intelligence

BEST: Biomarkers, EndpointS, and other Tools

DECIDE-AI: Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence

ECG: electrocardiogram

EHR: electronic health record

FDA: US Food and Drug Administration

FUTURE-AI: Fairness, Universality, Traceability, Usability, Robustness, and Explainability in Artificial Intelligence

NIH: National Institutes of Health

PROBAST+AI: Prediction Model Risk of Bias Assessment Tool + Artificial Intelligence

TRIPOD+AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis + Artificial Intelligence

Edited by Felix Muehlensiepen; peer-reviewed by Amr A Arafat, Rory Lambe; submitted 02.Jun.2026; accepted 11.Aug.2026; published 28.Aug.2026

Please cite as:

Ngwu DC

Operationalizing AI-Enabled Cardiovascular Biomarkers: A Clinician-Centered Framework for Validation, Governance, and Workflow Integration

JMIR Cardio 2026;10:e103401

URL: <https://cardio.jmir.org/2026/1/e103401>

doi: [10.2196/103401](https://doi.org/10.2196/103401)

© Dabeluchi Chiedozi Ngwu. Originally published in *JMIR Cardio* (<https://cardio.jmir.org>), 28.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Cardio*, is properly cited. The complete bibliographic information, a link to the original publication on <https://cardio.jmir.org>, as well as this copyright and license information must be included.