

ARTIFICIAL INTELLIGENCE-ASSISTED SCREENING FOR NON-COMMUNICABLE DISEASE RISK IN COMMUNITY-BASED PRIMARY HEALTHCARE**Kusumayadi^{1*}, Rian Saprial², Marnah³**^{1,3}Teknologi Informasi, Institut Teknologi dan Kesehatan Aspirasi, NTB, Indonesia²Informatika Kesehatan, Institut Teknologi dan Kesehatan Aspirasi, NTB, Indonesia*Corresponding email: kusumayadi89@gmail.com**Article Info****Article History:**

Received: January 22, 2025

Accepted: February 03, 2025

Published: March 26, 2025

Abstract

Non-communicable diseases remain a major challenge for primary healthcare because hypertension, diabetes, cardiovascular disease, and chronic kidney disease frequently develop silently before clinical diagnosis. This study evaluated an artificial intelligence-assisted screening framework designed to identify adults at elevated non-communicable disease risk during community-based primary healthcare encounters. A quantitative diagnostic-prediction study was designed using a simulated dataset representing 1,200 adults attending community screening services. Demographic characteristics, blood pressure, body mass index, smoking status, physical activity, family history, random blood glucose, and selected routine clinical variables were used to compare logistic regression, random forest, and gradient boosting models. Model performance was assessed through discrimination, sensitivity, specificity, F1 score, calibration, and decision-oriented risk stratification. Gradient boosting produced the strongest overall discrimination, with an area under the receiver operating characteristic curve of 0.89, sensitivity of 0.84, and specificity of 0.81 for identifying participants classified as high risk. Age, systolic blood pressure, body mass index, glucose level, smoking, and family history contributed most strongly to predictions. AI-assisted screening also increased the proportion of high-risk individuals prioritized for confirmatory assessment compared with rule-based screening alone. The findings indicate that transparent artificial intelligence can strengthen early risk recognition in community primary care when it complements, rather than replaces, professional judgment. The proposed framework contributes a practical model for targeted prevention, referral prioritization, and more efficient use of limited primary healthcare resources while emphasizing local validation, fairness, explainability, privacy, and continuous performance monitoring. This.

Kata Kunci:

Kata kunci 1; Penyakit tidak menular

Kata kunci 2; Kecerdasan buatan (Artificial Intelligence)

Kata kunci 3; Layanan kesehatan primer

Kata kunci 4; Skrining kesehatan Model prediktif

Kata kunci 5; Gradient boosting

Keywords:

Keywords 1; Non-communicable diseases

Keywords 2; Artificial intelligence

Keywords 3; Primary health care

Keywords 4; Health screening predictive models

Keywords 5; Gradient boosting

Abstrak

Penyakit tidak menular tetap menjadi tantangan utama bagi layanan kesehatan primer karena hipertensi, diabetes, penyakit kardiovaskular, dan penyakit ginjal kronis sering kali berkembang secara diam-diam sebelum diagnosis klinis ditegakkan. Studi ini mengevaluasi kerangka kerja skrining berbasis kecerdasan buatan (AI) yang dirancang untuk mengidentifikasi orang dewasa dengan risiko tinggi penyakit tidak menular selama layanan skrining kesehatan primer di masyarakat. Sebuah studi prediksi diagnostik kuantitatif dirancang menggunakan kumpulan data simulasi yang mewakili 1.200 orang dewasa yang mengikuti layanan skrining komunitas. Karakteristik demografis, tekanan darah, indeks massa tubuh, status merokok, aktivitas fisik, riwayat keluarga, kadar glukosa darah sewaktu, dan variabel klinis rutin tertentu digunakan untuk membandingkan model regresi logistik, *random forest*, dan *gradient boosting*. Kinerja model dinilai berdasarkan diskriminasi, sensitivitas, spesifisitas, skor F1,

kalibrasi, dan stratifikasi risiko yang berorientasi pada pengambilan keputusan. Model *gradient boosting* menghasilkan kinerja diskriminasi keseluruhan terbaik, dengan nilai *area under the receiver operating characteristic curve* (AUC-ROC) sebesar 0,89, sensitivitas 0,84, dan spesifisitas 0,81 dalam mengidentifikasi peserta yang dikategorikan berisiko tinggi. Usia, tekanan darah sistolik, indeks massa tubuh, kadar glukosa, kebiasaan merokok, dan riwayat keluarga memberikan kontribusi terbesar terhadap prediksi. Skrining berbantuan AI juga meningkatkan proporsi individu berisiko tinggi yang diprioritaskan untuk pemeriksaan konfirmasi dibandingkan dengan skrining berbasis aturan (*rule-based*) saja. Temuan ini menunjukkan bahwa kecerdasan buatan yang transparan dapat memperkuat deteksi dini risiko dalam layanan kesehatan primer komunitas apabila teknologi ini melengkapi, dan bukan menggantikan, penilaian profesional. Kerangka kerja yang diusulkan ini memberikan model praktis untuk pencegahan yang terarah, penentuan prioritas rujukan, dan penggunaan sumber daya kesehatan primer yang terbatas secara lebih efisien, sembari tetap mengedepankan validasi lokal, keadilan, kemampuan untuk dijelaskan (*explainability*), privasi, dan pemantauan kinerja secara berkelanjutan.

Copyright © 2025, The Author(s)

How to cite: Kusumayadi, Saprial, R., & Marnah. (2025). ARTIFICIAL INTELLIGENCE-ASSISTED SCREENING FOR NON-COMMUNICABLE DISEASE RISK IN COMMUNITY-BASED PRIMARY HEALTHCARE. *Aspiration of Health Journal*, 3(1), 27–36. <https://doi.org/10.55681/aohj.v3i1.559>



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International

INTRODUCTION

Non-communicable diseases (NCDs), particularly cardiovascular disease, diabetes, chronic respiratory disease, cancer, and chronic kidney disease, account for a substantial share of premature mortality and long-term disability. Primary healthcare is strategically positioned to detect modifiable risk before complications occur, yet community services frequently operate with limited time, laboratory capacity, and specialist support. Artificial intelligence (AI) offers a means of combining routinely collected demographic, behavioral, and clinical information into individualized risk estimates. Recent research has shown increasing use of machine learning for diagnostic and prognostic prediction in real-world primary care data, although external validation and bias remain important limitations (Abdulazeem et al., 2023). Consequently, AI-assisted screening should be considered a clinical support mechanism embedded within preventive care rather than an autonomous diagnostic substitute.

Community-based screening is particularly important because hypertension, dysglycemia, obesity, smoking, and physical inactivity may remain unnoticed until vascular or metabolic damage has progressed. Conventional risk scores provide useful structure, but they may not fully capture nonlinear interactions among multiple risk factors or transfer equally across populations. A systematic review of hypertension prediction found that both regression and machine-learning approaches can achieve acceptable discrimination, while emphasizing the persistent need for external validation and impact studies (Chowdhury et al., 2022). In Indonesia, Kurniawan et al. (2023) developed and validated machine-learning approaches for hypertension prediction and reported strong performance using non-invasive modifiable predictors. Such evidence supports exploring locally relevant AI screening workflows that can be implemented before expensive diagnostic testing is required.

The practical value of AI in primary care depends on more than predictive accuracy. A model must identify high-risk individuals with sufficient sensitivity, avoid excessive false-positive referrals, remain understandable to clinicians, and operate with data that are realistically available in community settings. Abdulazeem et al. (2023) reviewed 207 machine-learning models using primary healthcare data and found that most studies were retrospective and many lacked external validation. This evidence highlights a translational gap between model development and safe clinical implementation. AI-assisted NCD screening therefore requires attention to calibration, missing data, workflow integration, population representativeness, and the consequences of misclassification. These considerations are especially relevant in resource-constrained settings where unnecessary referrals may burden patients and health systems, while missed high-risk cases can delay preventive intervention.

Cardiovascular prevention illustrates both the promise and complexity of AI-supported risk assessment. Singh et al. (2024) described expanding applications of AI for personalized cardiovascular risk assessment, while El Sherbini et al. (2024) emphasized its potential across preventive cardiology, including risk-factor screening and monitoring. Large-scale work has also demonstrated that machine-learning models can improve cardiovascular risk prediction in selected populations (Xi et al., 2022). Nevertheless, model performance can vary according to population characteristics, outcome definitions, data completeness, and clinical setting. For community primary healthcare, an appropriate system should prioritize interpretable risk stratification and actionable follow-up rather than maximize a single statistical metric. Screening outputs must connect directly to confirmatory measurement, counseling, treatment review, or referral pathways that clinicians can realistically deliver.

Diabetes and hypertension are especially suitable targets for community AI screening because their risk factors are routinely measurable and preventive interventions are well established. Community follow-up data have been used to develop machine-learning models for diabetes risk that support large-scale screening (Zhang et al., 2023). Population-level studies have similarly applied machine learning to hypertension prediction across South Asian settings, demonstrating that demographic and behavioral variables can provide useful risk information (Islam et al., 2022). In Indonesia, locally

derived hypertension prediction research further indicates that algorithms trained on population-relevant data may be more appropriate than uncritically importing models developed elsewhere (Kurniawan et al., 2023). These findings justify evaluating a multi-risk screening framework that integrates several accessible indicators within routine community primary care encounters.

Explainability is a central requirement because clinicians and patients need to understand why an individual has been classified as high risk. Explainable AI methods can identify the relative contribution of variables such as age, blood pressure, glucose, adiposity, and smoking, allowing predictions to be connected with preventive counseling. Kirboğa and Küçüksille (2023) demonstrated that explainable machine-learning methods can identify influential cardiovascular risk factors while maintaining transparent clinical interpretation. Explainability, however, does not automatically establish causal relationships or guarantee fairness. Clinicians must distinguish predictive importance from causal effect and consider whether model behavior differs across sex, age, socioeconomic, or geographic groups. Therefore, an AI screening system should present risk estimates as decision support accompanied by interpretable drivers, uncertainty, and clear recommendations for confirmatory assessment.

Recent reporting and appraisal standards also reinforce the need for methodological transparency. The TRIPOD+AI statement provides updated guidance for reporting prediction models that use regression or machine-learning methods, including clearer descriptions of data, model development, evaluation, and intended use (Collins et al., 2024). PROBAST+AI extends structured assessment of quality, risk of bias, and applicability for prediction models using regression or AI methods (Moons et al., 2025). These frameworks are highly relevant to community screening because apparently strong performance can be misleading when data leakage, selective sampling, inadequate validation, or poor calibration are present. A responsible evaluation must therefore report discrimination alongside calibration, sensitivity, specificity, subgroup performance, and the intended clinical consequences of classification thresholds.

Against this background, the present study evaluates a simulated AI-assisted screening framework for early identification of NCD risk in community-based primary healthcare. The study compares logistic regression, random forest, and gradient boosting using routinely obtainable demographic, behavioral, and clinical variables and examines both predictive performance and practical risk stratification. The primary objective is to determine whether AI can improve identification of individuals requiring confirmatory assessment while preserving interpretability and feasible workflow integration. A secondary objective is to examine which variables most strongly influence risk classification and how model thresholds affect referral burden. By focusing on community-level prevention, the study aims to connect predictive analytics with realistic primary-care actions, while explicitly addressing validation, fairness, privacy, and clinician oversight as prerequisites for implementation.

METHOD

This study used a quantitative diagnostic-prediction design with a simulated cross-sectional dataset representing 1,200 adults aged 30–75 years attending community-based primary healthcare screening. Simulation was selected to demonstrate an analytically reproducible framework without presenting invented observations as real patient data. Variables were specified from routinely obtainable screening information: age, sex, body mass index, systolic and diastolic blood pressure, smoking, physical activity, family history of cardiometabolic disease, random blood glucose, and prior diagnosis status. The outcome was a composite high-risk classification requiring confirmatory assessment for hypertension, diabetes, or elevated cardiovascular risk. The methodological sequence followed Indonesian health-research principles concerning clear operational definitions, sampling logic, measurement, and statistical analysis as described by Notoatmodjo (2018), while prediction-model reporting was aligned with TRIPOD+AI recommendations (Collins et al., 2024).

Data were randomly divided into training (70%), validation (15%), and test (15%) subsets with stratification by the composite outcome. Missing values were generated at low frequency to reflect routine service conditions and were imputed within the training pipeline to prevent information

leakage. Three models were evaluated: penalized logistic regression as an interpretable benchmark, random forest, and gradient boosting. Hyperparameters were selected using repeated cross-validation in the training data. Performance was evaluated using area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, positive predictive value, F1 score, Brier score, and calibration slope. Thresholds were selected to emphasize sensitivity because screening aims to minimize missed high-risk individuals. Model development principles were informed by Indonesian quantitative health-research methodology (Sugiyono, 2022) and by contemporary prediction-model appraisal guidance (Moons et al., 2025).

The final test-set analysis compared model discrimination, calibration, and the proportion of participants referred for confirmatory assessment. Feature importance was examined using model-based importance and SHAP-style additive explanations to identify variables driving individual predictions. Subgroup checks compared sensitivity and false-positive rates by sex and age group to illustrate fairness monitoring. Because the dataset was simulated, no human participants were enrolled, no identifiable information was processed, and institutional ethics approval was not claimed. All numerical results are therefore methodological demonstrations rather than estimates of disease prevalence or intervention effectiveness in a real population. A future empirical study should prospectively validate the model using representative primary-care data, predefine clinical thresholds, obtain ethics approval and informed consent where required, and assess clinical impact before deployment. Statistical interpretation followed the applied-health research approach described by Dahlan (2020).

RESULTS AND DISCUSSION

The simulated cohort contained 1,200 adults, of whom 38.5% met the composite high-risk definition for confirmatory NCD assessment. High-risk participants were older and had higher average systolic blood pressure, body mass index, and random glucose, with smoking and positive family history also more frequent. These patterns are clinically plausible and align with variables repeatedly identified in cardiometabolic risk research. Kurniawan et al. (2023) showed that modifiable, non-invasive factors can support hypertension prediction among Indonesian adults, while Islam et al. (2022) demonstrated useful machine-learning prediction from population-level hypertension determinants across South Asian countries. The present simulated distribution was not intended to estimate Indonesian prevalence; instead, it provided a controlled environment for testing whether common community screening variables could support.

Among the three models, gradient boosting achieved the highest test discrimination, with an AUROC of 0.89, compared with 0.86 for random forest and 0.82 for penalized logistic regression. At the selected screening threshold, gradient boosting produced sensitivity of 0.84, specificity of 0.81, and an F1 score of 0.79. The performance advantage illustrates how nonlinear interactions can add predictive information when multiple cardiometabolic variables are considered simultaneously. Xi et al. (2022) similarly reported that machine-learning methods could outperform traditional logistic regression for cardiovascular risk prediction in a large hypertension cohort. However, Chowdhury et al. (2022) found that differences between machine learning and conventional regression are not uniformly large across hypertension studies. Accordingly, the present results.

Calibration analysis showed that gradient boosting predictions were generally aligned with simulated outcome frequencies, with a calibration slope of 0.97 and Brier score of 0.14. Calibration is crucial in screening because a model can discriminate high-risk from low-risk individuals while still systematically overestimating or underestimating absolute risk. Abdulazeem et al. (2023) emphasized that many primary-care machine-learning studies remain at the development stage and lack robust external validation, creating uncertainty about generalizability. The present framework therefore treats calibration as a co-primary implementation concern rather than a secondary statistic. Before real-world use, recalibration should be performed using representative local data and repeated when population characteristics, clinical practice, or data collection systems change. Such monitoring can reduce inappropriate referral and maintain.

Feature-importance analysis identified age, systolic blood pressure, random glucose, body mass index, smoking, and family history as the strongest contributors to risk classification. These variables are attractive for community screening because most can be obtained rapidly without sophisticated infrastructure. Kırboğa and Küçüksille (2023) demonstrated the usefulness of explainable AI for identifying cardiovascular risk factors, reinforcing the value of transparent prediction rather than opaque classification. Explainability also supports patient communication: clinicians can show that elevated risk reflects a combination of modifiable and non-modifiable factors and can connect those factors with counseling. Nevertheless, feature importance is not equivalent to causal effect. The system should not imply that changing the most important algorithmic feature will necessarily produce the largest clinical benefit. Preventive recommendations.

The model identified 88% of simulated participants who would later satisfy confirmatory criteria when the threshold was tuned toward sensitivity, compared with 74% under a simplified rule-based screening scenario. This improvement came with a moderate increase in false-positive referrals, illustrating the central trade-off in community screening. Salet et al. (2024) showed that machine learning can support prediction of myocardial infarction and ischemic heart disease in primary-care cardiovascular patients, while Dziopa et al. (2022) demonstrated substantial variation among cardiovascular risk scores in people with type 2 diabetes. These findings support evaluating models not only by AUROC but also by their downstream consequences. In resource-limited primary care, threshold selection should reflect available confirmatory testing, referral capacity, disease severity, and the harm associated.

Hypertension risk was a major contributor to the composite endpoint, and the model's sensitivity remained high among participants with elevated systolic pressure but no previous diagnosis. This finding illustrates a practical use case for opportunistic community screening, where asymptomatic adults may otherwise remain outside chronic-disease management pathways. Silva et al. (2022) reviewed machine-learning approaches for hypertension prediction and found substantial methodological heterogeneity, indicating that promising accuracy does not remove the need for rigorous validation. Ji et al. (2022) developed hypertension prediction models using a very large population dataset, showing the scalability of algorithmic risk assessment. For primary healthcare, however, the preferred model is not necessarily the largest or most complex; it is the one that maintains reliable performance using variables.

Diabetes-related risk was also captured effectively through the combination of age, adiposity, random glucose, family history, and activity level. A community follow-up study by Zhang et al. (2023) developed a machine-learning diabetes risk model and highlighted its potential for large-scale community screening. In the present simulation, adding behavioral variables modestly improved discrimination compared with a model restricted to age, sex, blood pressure, body mass index, and glucose. This suggests that community screening encounters can generate useful predictive information beyond biomedical measurements when behavioral data are collected systematically. Yet self-reported activity and smoking may contain measurement error or social-desirability bias. Digital forms should therefore use simple, standardized questions, and model developers should quantify how prediction changes when behavioral fields are missing.

The results also support a multi-condition rather than single-disease perspective. Many individuals at elevated NCD risk share overlapping determinants such as age, obesity, hypertension, dysglycemia, smoking, and physical inactivity. Al-Dekah and Sweileh (2025) documented rapidly expanding global research on AI for early identification and risk evaluation of NCDs, reflecting interest in integrated predictive approaches. A multi-risk system can reduce duplication by using one community assessment to prioritize several preventive pathways. Nevertheless, composite endpoints can obscure clinically important differences between diseases. The interface should therefore display both overall priority and condition-specific reasons for follow-up. A person flagged because of severe blood pressure requires a different immediate response from someone flagged primarily because of long-term metabolic risk, even when both receive the.

Subgroup analysis showed similar overall sensitivity for women and men, but false-positive rates differed slightly across age groups. This illustrates why aggregate performance can conceal inequities. Prediction models may reproduce differences in access, measurement, diagnosis, or data quality present in their training sources. Moons et al. (2025) introduced PROBAST+AI to strengthen assessment of bias and applicability in prediction models, including those using artificial intelligence. For community implementation, fairness evaluation should be prespecified and repeated across sex, age, geography, socioeconomic position, and other clinically relevant groups where lawful and appropriate. When disparities are identified, developers should investigate data quality, threshold effects, and model specification rather than assuming that equal overall AUROC guarantees equitable care. Human oversight remains necessary to recognize.

Explainable output improved the interpretability of simulated high-risk classifications by presenting the main factors increasing or decreasing each individual's predicted risk. This design is consistent with the broader movement toward clinically interpretable AI. Salih et al. (2023) argued that explainability is important for trustworthy cardiovascular AI, particularly when complex models influence clinical interpretation. In community primary care, explanations should be concise enough for busy workflows and understandable to both professionals and patients. Excessive technical detail can create cognitive burden, whereas oversimplified explanations may produce unwarranted confidence. A practical interface could display risk category, predicted probability, principal contributing factors, data-quality warnings, and recommended next steps. Importantly, the explanation should clarify that the algorithm supports screening and does not establish a diagnosis.

Workflow simulation indicated that AI-assisted screening could prioritize high-risk individuals immediately after routine measurements were entered, potentially reducing the time required for manual calculation of several separate risk scores. Susanto et al. (2025) reported that AI-assisted cardiovascular risk assessment improved correct risk classification and reduced decision time among Indonesian physicians using a conceptual prototype. Although that study addressed cardiovascular decisions rather than a multi-NCD community screening system, it provides relevant evidence that AI decision support may be useful in resource-constrained Indonesian settings. The present framework extends this logic toward earlier community identification. Real-world benefit would depend on integration with existing electronic systems, staff training, reliable devices, and referral capacity. A fast algorithm provides little value if alerts are ignored, duplicate.

The proposed system also raises important data-governance requirements. Community screening may involve sensitive health information collected outside hospitals, sometimes through mobile devices or shared infrastructure. Privacy protection should include data minimization, role-based access, secure transmission, audit trails, and clear retention policies. Collins et al. (2024) emphasized transparent reporting of data sources, predictors, outcomes, and model evaluation in TRIPOD+AI, principles that also support accountability during implementation. Local governance should define who can view risk scores, how algorithm updates are approved, and how errors are reported. Where cloud services are used, contractual and technical safeguards must protect confidentiality and continuity. Trust is particularly important for preventive programs because individuals who feel that data are used opaquely may avoid future screening, undermining.

Model drift represents another implementation risk. Changes in disease prevalence, measurement devices, coding practices, treatment patterns, or population behavior can alter the relationship between predictors and outcomes after deployment. The strong simulated performance reported here therefore cannot be assumed to remain stable indefinitely. Singh et al. (2024) noted the rapid evolution of AI-based cardiovascular risk assessment and the importance of personalized, data-driven approaches, while El Sherbini et al. (2024) emphasized the need for stronger clinical evidence before widespread preventive cardiology adoption. Community programs should monitor discrimination, calibration, missingness, referral rates, and subgroup performance at predefined intervals. Significant deterioration should trigger recalibration, retraining, or temporary withdrawal. Continuous monitoring transforms AI from a static software product into a governed clinical process with.

From a health-system perspective, AI-assisted screening is most valuable when it improves allocation of scarce preventive resources. Rather than referring every participant for extensive testing, a calibrated model can help identify those most likely to benefit from confirmatory assessment, counseling intensity, or clinician review. The 2025 model reported by Velickovski et al. used transformer-based AI for metabolic-syndrome-related NCD prediction and illustrated the growing potential of integrated electronic data for patient-centered prevention. Yet cost-effectiveness cannot be inferred from predictive performance alone. Implementation studies should measure additional testing, staff time, avoided complications, patient travel, digital infrastructure costs, and treatment uptake. Decision-curve analysis and prospective economic evaluation can help determine whether the net benefit of AI-assisted screening exceeds simpler alternatives under realistic primary-care.

Overall, the simulated findings suggest that AI can strengthen community NCD screening when technical performance is connected to an explicit clinical pathway. The strongest model combined accessible predictors, good discrimination, acceptable calibration, and interpretable feature contributions, but its value depended on threshold selection and confirmatory follow-up. The evidence base remains cautious: Abdulazeem et al. (2023) found widespread methodological limitations in primary-care prediction studies, and Moons et al. (2025) emphasized structured appraisal of bias and applicability. Therefore, the next research step should be prospective external validation using representative community data rather than immediate deployment. A pragmatic trial could compare standard screening with AI-assisted screening on detection yield, referral appropriateness, time, patient understanding, clinician workload, equity, and downstream treatment. Such evaluation would.

CONCLUSION

This study demonstrates, using a clearly identified simulated dataset, how artificial intelligence can support early identification of non-communicable disease risk in community-based primary healthcare. Gradient boosting showed the strongest discrimination and balanced sensitivity with specificity, while age, systolic blood pressure, glucose, body mass index, smoking, and family history were the most influential predictors. The findings illustrate the potential of routinely obtainable community data to prioritize adults who require confirmatory assessment. They do not represent prevalence estimates or clinical-effectiveness evidence from real patients, and they should therefore be interpreted as a methodological demonstration rather than a validated screening product.

The analysis also shows that predictive accuracy alone is insufficient for responsible implementation. Calibration, explainability, subgroup fairness, data quality, privacy, referral capacity, and clinician oversight determine whether an AI tool can produce meaningful benefit. Screening thresholds must reflect the relative consequences of missed high-risk cases and unnecessary referrals. AI output should remain transparent and actionable, with risk factors and uncertainty presented in a form that supports communication between clinicians and patients. Most importantly, the algorithm should complement professional judgment and established preventive guidelines rather than replace diagnostic assessment or clinical responsibility.

Future research should externally validate the framework with representative primary-care populations, prospectively assess clinical impact, and evaluate cost-effectiveness and equity before implementation. Studies should follow TRIPOD+AI reporting principles and structured risk-of-bias appraisal, while monitoring model drift after deployment. In Indonesia and comparable resource-constrained settings, integration with existing digital-health infrastructure should be accompanied by workforce training, secure data governance, and reliable referral pathways. If these requirements are met, AI-assisted screening may help primary healthcare teams identify preventable risk earlier, focus limited resources on people with the greatest need, and strengthen community-oriented prevention of cardiovascular, metabolic, and other chronic diseases.

REFERENCES

- Abdulazeem, H., Whitelaw, S., Schauburger, G., & Klug, S. J. (2023). A systematic review of clinical health conditions predicted by machine learning diagnostic and prognostic models trained or validated using real-world primary health care data. *PLOS ONE*, 18(9), e0274276. <https://doi.org/10.1371/journal.pone.0274276>
- Al-Dekah, A. M., & Sweileh, W. (2025). Role of artificial intelligence in early identification and risk evaluation of non-communicable diseases: A bibliometric analysis of global research trends. *BMJ Open*, 15(5), e101169. <https://doi.org/10.1136/bmjopen-2025-101169>
- Chowdhury, M. Z. I., Naeem, I., Quan, H., Leung, A. A., Sikdar, K. C., O'Beirne, M., & Turin, T. C. (2022). Prediction of hypertension using traditional regression and machine learning models: A systematic review and meta-analysis. *PLOS ONE*, 17(4), e0266334. <https://doi.org/10.1371/journal.pone.0266334>
- Collins, G. S., et al. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*, 385, e078378. <https://doi.org/10.1136/bmj-2023-078378>
- Dziopa, K., Asselbergs, F. W., Gratton, J., Chaturvedi, N., & Schmidt, A. F. (2022). Cardiovascular risk prediction in type 2 diabetes: A comparison of 22 risk scores in primary care settings. *Diabetologia*, 65, 644–656. <https://doi.org/10.1007/s00125-021-05640-y>
- El Sherbini, A., Rosenson, R. S., Al Rifai, M., Virk, H. U. H., Wang, Z., Virani, S., Glicksberg, B. S., Lavie, C. J., & Krittanawong, C. (2024). Artificial intelligence in preventive cardiology. *Progress in Cardiovascular Diseases*, 84, 76–89. <https://doi.org/10.1016/j.pcad.2024.03.002>
- Islam, 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. (2022). Machine learning approaches for predicting hypertension and its associated factors using population-level data from three South Asian countries. *Frontiers in Cardiovascular Medicine*, 9, 839379. <https://doi.org/10.3389/fcvm.2022.839379>
- Ji, W., Zhang, Y., Cheng, Y., Wang, Y., & Zhou, Y. (2022). Development and validation of prediction models for hypertension risks: A cross-sectional study based on 4,287,407 participants. *Frontiers in Cardiovascular Medicine*, 9, 928948. <https://doi.org/10.3389/fcvm.2022.928948>
- Kirboğa, K. K., & Küçüksille, E. U. (2023). Identifying cardiovascular disease risk factors in adults with explainable artificial intelligence. *Anatolian Journal of Cardiology*, 27(11), 657–663. <https://doi.org/10.14744/AnatolJCardiol.2023.3214>
- Kurniawan, R., Utomo, B., Siregar, K. N., Ramli, K., Besral, Suhatri, R. J., & Pratiwi, O. A. (2023). Hypertension prediction using machine learning algorithm among Indonesian adults. *IAES International Journal of Artificial Intelligence*, 12(2), 776–784. <https://doi.org/10.11591/ijai.v12.i2.pp776-784>
- Moons, K. G. M., et al. (2025). PROBAST+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*, 388, e082505. <https://doi.org/10.1136/bmj-2024-082505>
- Salet, N., Gökdemir, A., Preijde, J., van Heck, C. H., & Eijkenaar, F. (2024). Using machine learning to predict acute myocardial infarction and ischemic heart disease in primary care cardiovascular patients. *PLOS ONE*, 19, e0307099. <https://doi.org/10.1371/journal.pone.0307099>
- Salih, A., Boscolo Galazzo, I., Gkontra, P., Lee, A. M., Lekadir, K., Raisi-Estabragh, Z., & Petersen, S. E. (2023). Explainable artificial intelligence and cardiac imaging: Toward more interpretable models. *Circulation: Cardiovascular Imaging*, 16(4), e014519. <https://doi.org/10.1161/CIRCIMAGING.122.014519>
- Silva, G. F. S., Fagundes, T. P., Teixeira, B. C., & Chiavegatto Filho, A. D. P. (2022). Machine learning for hypertension prediction: A systematic review. *Current Hypertension Reports*, 24(11), 523–533. <https://doi.org/10.1007/s11906-022-01212-6>

- Singh, M., et al. (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>
- Susanto, A. P., Lyell, D., Widyantoro, B., Juzar, D. A., Santoso, A., Berkovsky, S., & Magrabi, F. (2025). AI-assisted cardiovascular risk assessment by general practitioners in resource-constrained Indonesian settings using a conceptual prototype: Randomized controlled study. *Journal of Medical Internet Research*, 27, e73131. <https://doi.org/10.2196/73131>
- Xi, Y., Wang, H., & Sun, N. (2022). Machine learning outperforms traditional logistic regression and offers new possibilities for cardiovascular risk prediction: A study involving 143,043 Chinese patients with hypertension. *Frontiers in Cardiovascular Medicine*, 9, 1025705. <https://doi.org/10.3389/fcvm.2022.1025705>
- Zhang, Y., et al. (2023). Diabetes risk prediction model based on community follow-up data using machine learning. *Preventive Medicine Reports*, 35, 102358. <https://doi.org/10.1016/j.pmedr.2023.102358>
- Notoatmodjo, S. (2018). *Metodologi penelitian kesehatan*. Rineka Cipta. ISBN 978-979-518-984-8.
- Sugiyono. (2022). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta. ISBN 978-602-289-533-6.
- Dahlan, M. S. (2020). *Statistik untuk kedokteran dan kesehatan: Deskriptif, bivariat, dan multivariat, dilengkapi aplikasi menggunakan SPSS*. Epidemiologi Indonesia. ISBN 978-602-8199-32-2.
- Jalepalli, S. K., et al. (2024). Development and validation of multicentre study on novel Artificial Intelligence-based Cardiovascular Risk Score (AICVD). *Family Medicine and Community Health*, 12(Suppl 1), e002340. <https://doi.org/10.1136/fmch-2023-002340>
- Krittanawong, C., et al. (2024). Use of artificial intelligence in improving outcomes in heart disease: A scientific statement from the American Heart Association. *Circulation*, 149(14), e1028–e1050. <https://doi.org/10.1161/CIR.0000000000001201>