ORIDSAN

Nigerian Journal of Operations Research (NIJOR)

https://nijor.org.ng/
pISSN: 2141-7938
eISSN: 3156-5956

Fuzzy and Neuro-Fuzzy Models for Coronary Artery Disease Detection: A Systematic Review of Development, Validation and Risk of Bias in Open-Access Literature

Shehu Mohammed Ahmad[1], Timothy Moses[2], Mukthar Muhammad Liman[3], Mahmud Ahmad Bamanga[1]*

[1,2,3,4] Department of Computer Science, Faculty of Computing, Federal University of Lafia, Nigeria

ARTICLE INFO

Article history:
Received 03 October 2026
Revised form 12 November 2026
Accepted 15 November 2026
Available online 21 November 2026

Keywords: *Coronary artery disease; Fuzzy logic; Neuro-fuzzy systems; ANFIS; Clinical prediction models; Risk of bias; Systematic review; Decision support; Multi-criteria decision-making*

ABSTRACT

Fuzzy and neuro-fuzzy models are often proposed for the detection of coronary artery disease (CAD), because their reasoning is expressed as rules that clinicians can read. How these models are developed and validated, and how much risk of bias they carry, has not been appraised systematically. We searched PubMed and OpenAlex from inception to 3 October 2026. Screening, data extraction and risk of bias assessment with PROBAST applied predefined rules and were checked by the research team, with every decision and its reason recorded; no second reviewer repeated these steps independently. Full texts were limited to those available in open access. Of 1,925 unique records, 794 were sought in full text, and 132 could be assessed. In total, 62 reports were included: 26 used ANFIS or another neuro-fuzzy network and 24 a fuzzy rule-based system. Of these, 28 used the UCI heart disease data, and 12 did not clearly report their data. Only 1 report validated its model in a separate cohort, 3 reported an area under the ROC curve, and 1 reported calibration. In all, 61 of the 62 models were at high risk of bias, in every case including the analysis domain, and one was unclear. Reported accuracies of up to 99.66% came from internal validation, and no meta-analysis was possible. In the open-access literature, fuzzy and neuro-fuzzy models for CAD are not ready for clinical use. Because 61 of the 62 models (98.4%) did not report calibration, their predicted probabilities cannot yet support threshold-based triage or the allocation of scarce diagnostic resources. Future studies should use clinical cohorts of adequate size, report the full rule base, assess calibration, and test their models in new populations.

* *Corresponding author.*
E-mail address: mabamanga@gmail.com; https://doi.org/10.31181/nijor/0304.031

1. Introduction

Coronary artery disease (CAD) is the largest single contributor to the global burden of cardiovascular disease (Roth *et al.*, 2020). Its diagnosis depends on the clinical history, risk factors and a sequence of non-invasive and invasive tests, and models that estimate the probability of CAD from routine findings could help clinicians decide whom to refer for further testing. For such models to be used, clinicians need to trust them, and trust is easier to earn when the reasoning of a model can be read and checked rather than explained after the fact (Rudin, 2019).

Fuzzy models were designed for this kind of reasoning. Fuzzy sets allow a finding to belong partly to a category such as "high cholesterol" or "typical angina" (Zadeh, 1965), and fuzzy rule bases combine such categories in linguistic IF-THEN rules. Early systems encoded expert rules in the Mamdani form (Mamdani & Assilian, 1975); the Takagi-Sugeno-Kang form made rule outputs functions of the inputs (Takagi & Sugeno, 1985); and the adaptive neuro-fuzzy inference system (ANFIS) learned the rules and membership functions from data (Jang, 1993). Later extensions, including type-2 fuzzy sets (Mendel & John, 2002) and intuitionistic fuzzy sets (Atanassov, 1986), were introduced to represent uncertainty about the membership functions themselves. Heart disease has been a frequent application of all of these families.

From the standpoint of operations research and decision analysis, a fuzzy or neuro-fuzzy diagnostic model is a multi-criteria decision support system that operates under uncertainty. Each clinical finding acts as a criterion, the membership functions express the degree to which a patient satisfies each criterion, the rule base aggregates the criteria, and the output is turned into an action, such as reassurance, further testing or referral, by comparison with a decision threshold. Fuzzy multi-criteria methods of this kind are well established in operations research, for example fuzzy TOPSIS for pharmaceutical supplier selection (Modibbo *et al.*, 2022), multi-criteria prioritization of border-control strategies (Barma *et al.*, 2026) and Mamdani inference for adaptive agricultural decisions (Madu *et al.*, 2026), and ANFIS has been used to model the transmission dynamics of infectious disease (Raina *et al.*, 2026). Recent work on fuzzy theory has examined how these models represent uncertainty and how their rule bases can be kept interpretable (Varshney & Torra, 2023; Wu *et al.*, 2023). In a clinical decision, however, the value of such a model depends on the quality of the probability it produces: a model whose probabilities are poorly calibrated, or whose rule base is too large to be checked, cannot support sound triage or the efficient allocation of scarce diagnostic resources such as angiography.

Many of these studies report very high accuracy, often on a small number of public datasets such as the Cleveland heart disease data (Detrano *et al.*, 1989) and the Z-Alizadeh Sani data (Alizadehsani *et al.*, 2013). High accuracy on a small public dataset does not show that a model will work in a new population. Reviews of prediction models built with machine learning have found that most were at high risk of bias, mainly because of the analysis: small samples, internal validation only, and calibration that was not assessed (Andaur Navarro *et al.*, 2021, 2023). Data leakage, in which information from the test data enters model development, can inflate performance further (Kapoor & Narayanan, 2023). Reporting standards and appraisal tools now exist for such models (Collins *et al.*, 2024; Moons *et al.*, 2019, 2025), and calibration is recognized as essential for clinical use (Van Calster *et al.*, 2019).

Existing reviews do not tell us whether fuzzy models of CAD meet these standards. Reviews of machine learning for cardiovascular disease and CAD pool many model types and say little about fuzzy models (Krittanawong *et al.*, 2020; Garavand *et al.*, 2023). Reviews of fuzzy methods describe the techniques and their interpretability but do not assess risk of bias or validation: Cao *et al.* (2024) reviewed fuzzy rule-based systems

for disease diagnosis, Talpur *et al.* (2023) surveyed deep neuro-fuzzy systems, and Menaceur *et al.* (2025) reviewed fuzzy logic for arrhythmia rather than CAD. Three systematic reviews have included fuzzy models of heart disease among wider topics: Dudchenko and Kopanitsa (2017) reviewed decision support systems in cardiology, and Thomas *et al.* (2019) and Thukral and Gupta (2026) reviewed fuzzy logic for the diagnosis of many diseases. None was limited to CAD, and none assessed risk of bias or the type of validation. A search of PubMed, OpenAlex and the Open Science Framework registry on 3 October 2026 found no completed or registered review of fuzzy models for CAD that did so.

This review therefore asked, among studies that developed or validated fuzzy or neuro-fuzzy models for the diagnosis of CAD or heart disease in adults: (i) how the models were built; (ii) how they were validated and how completely they were reported; (iii) what their risk of bias and applicability were, assessed with PROBAST; (iv) what performance they reported; and (v) what gaps remain in the reporting of rule bases, calibration, uncertainty and external validation.

1.1 Formal background

This subsection gives the formal definitions on which the appraisal rests, so that the links between model construction, calibration and decision-making are explicit.

A fuzzy set A on a universe X is defined by a membership function $\mu_A: X \to [0, 1]$. A Gaussian membership function with centre c and width σ, the form most often learned by neuro-fuzzy models, is

$$\mu A(x) = exp\left[-(x - c)\frac{2}{(2\sigma 2)}\right] \quad (1)$$

A Mamdani rule base (Mamdani & Assilian, 1975) consists of J rules of the form Rj: IF x_1 is A_{1j} AND … AND x_n is A_{nj} THEN y is B_j. The firing strength of rule j and the aggregated output set are

$$wj = mini\,\mu Aij(xi), \quad \mu B'(y) = maxj\,min[wj, \mu Bj(y)])$$

and the crisp output is obtained by centroid defuzzification,

$$y* = \int y\,\mu B'(y)\,dy \,/\, \int \mu B'(y)\,dy \quad (3)$$

In the Takagi-Sugeno-Kang form (Takagi & Sugeno, 1985), the consequent of rule j is a function of the inputs, usually linear, $yj = pj0 + \Sigma i\, pjixi$, and the output is the weighted average

$$\hat{Y} = \frac{\Sigma j\, wjyj}{\Sigma j\, wj}, \quad wj = \Pi i\, \mu Aij(xi) \quad (4)$$

ANFIS (Jang, 1993) implements Eq. (4) as a five-layer network whose premise parameters α (the centres and widths in Eq. (1)) and consequent parameters $\theta = \{p_{ji}\}$ are learned from N training pairs (x_k, y_k) by minimizing the squared error

$$E(\alpha, \theta) = ½\,\Sigma k\,(yk - \hat{y}k)2 \quad (5)$$

Hybrid learning alternates two steps. In the forward pass, with α fixed, $\hat{y}$ is linear in θ, which is estimated by least squares,

$$\hat{\theta} = (ATA) - 1ATy \quad (6)$$

where A is the matrix of normalized firing strengths multiplied by the inputs. In the backward pass, with θ fixed, the premise parameters are updated by gradient descent with learning rate η,

$$\alpha \leftarrow \alpha - \eta\frac{\partial E}{\partial \alpha} \quad (7)$$

For a binary diagnosis the output is passed through a logistic function or a threshold to give a predicted probability $\hat{p}$ or a class. Many of the reviewed models replaced Eq. (7) with a genetic or swarm algorithm, but the objective was the same: an error measured on the development data.

In an interval type-2 fuzzy set (Mendel & John, 2002) the membership grade of x is an interval $[\underline{\mu}(x), \bar{\mu}(x)]$, and the region between the lower and upper membership functions is the footprint of uncertainty. Each rule then fires with an interval $[\underline{w}_j, \bar{w}_j]$, and type reduction by the Karnik-Mendel algorithm (Karnik & Mendel, 2001) gives an output interval whose midpoint is the crisp output,

$$[yl, yr] = [min, max] of \frac{\Sigma j\ wjyj}{\Sigma j\ wj\ over\ wj} \in [\underline{wj}, \bar{wj}],\quad y* = \frac{yl + yr}{2}$$

diagnostic model supports decisions only if its predicted probabilities agree with the risk observed. Calibration is assessed by logistic recalibration in the validation data (Van Calster *et al.*, 2019),

$$logit\ P(Y = 1 \mid \hat{p}) = \alpha c + \beta c\ logit(\hat{p}) \tag{9}$$

where the calibration intercept α_c (ideally 0) measures systematic over- or underestimation of risk and the calibration slope β_c (ideally 1) shows whether predictions are too extreme ($\beta c < 1$, typical of overfitting) or too moderate ($\beta c > 1$). The overall accuracy of the probabilities is summarized by the Brier score,

$$BS = \left(\frac{1}{N}\right) \Sigma k\ (\hat{p}k - yk)2 \tag{10}$$

Because Eqs. (5) to (7) are minimized on the development data only, a model can fit those data closely and still be miscalibrated in a new population, where overfitting appears as $\beta_c < 1$ and a different prevalence as $\alpha_c \neq 0$.

In decision analysis, a patient is referred or treated when $\hat{p} \geq p_t$, where the threshold probability p_t reflects the relative harm of an unnecessary and the benefit of a necessary intervention, through the odds $p_t/(1 - p_t)$. The net benefit of a model at that threshold (Vickers & Elkin, 2006) is

$$NB(pt) = \frac{TP}{N} - \left(\frac{FP}{N}\right) \times \frac{pt}{1 - pt} \tag{11}$$

where TP and FP are the numbers of true and false positives among N patients. Decision curve analysis plots NB(p_t) over a clinically sensible range of thresholds and compares it with referring all patients or none. A model affects Eq. (11) only through the decision $\hat{p} \geq p_t$, so a miscalibrated model applies a different threshold from the one intended, and its net benefit cannot be judged from accuracy alone.

PROBAST (Wolff *et al.*, 2019) is a structured judgement tool rather than a loss function. Each of its four domains d (participants, predictors, outcome and analysis) is rated at low, high or unclear risk of bias from signalling questions, and the overall rating follows a fixed aggregation rule,

$$RoB = High\ if\ \exists d: RoBd = High;\ Low\ if\ \forall d: RoBd = Low;\ Unclear\ otherwise \tag{12}$$

Which was applied in this review (Section 2.7).

2. Methods

This section describes how the review was planned and carried out. It presents the protocol (Section 2.1), the eligibility criteria (Section 2.2), the information sources and search strategy (Section 2.3), the management

of records and removal of duplicates (Section 2.4), study selection (Section 2.5), data extraction (Section 2.6), the assessment of risk of bias and applicability (Section 2.7), the synthesis (Section 2.8) and the changes made from the protocol (Section 2.9).

2.1 Protocol

The review followed a protocol prepared according to PRISMA-P (Moher *et al.*, 2015) before the searches were run. The protocol was not registered on PROSPERO, because screening had started before registration was completed. To make it publicly available, the protocol (version 0.1, 2 October 2026) is provided unchanged as supplementary material, so that the planned methods can be compared directly with those reported. The review is reported according to PRISMA 2020 (Page *et al.*, 2021). Changes from the protocol are listed in Section 2.9.

2.2 Eligibility criteria

Eligibility was defined in terms of population, index model, outcome, timing and study type (Table 1). Reports were eligible if they developed or validated at least one fuzzy or neuro-fuzzy model whose output was a diagnosis or a probability of CAD or heart disease, and reported a quantitative performance result on patient data.

Table 1: Eligibility criteria.

Element	Included	Excluded
Population	Adults evaluated for CAD or heart disease, in clinical cohorts or public datasets	Children; animals; simulated patients only
Index model	Any fuzzy or neuro-fuzzy model whose output is a diagnosis or probability of disease: Mamdani, TSK, ANFIS and variants, type-2 and interval-valued, intuitionistic, neutrosophic, evolving, deep neuro-fuzzy, and hybrids with optimization	Fuzzy methods used only for preprocessing, clustering or feature selection, with a non-fuzzy classifier
Outcome	CAD, ischaemic heart disease, myocardial infarction, acute coronary syndrome, or "heart disease" defined by angiography or a clinical diagnosis of CAD (as in the UCI datasets)	Arrhythmia, heart failure, valvular or congenital disease; ECG beat or heart sound classification without a CAD outcome
Timing	Diagnostic models (disease present at assessment)	Prognostic models of future events: included but reported separately
Study type	Development or validation studies with a quantitative performance result	Reviews, editorials, abstracts without results, studies without any evaluation on patient data
Language and date	English; any date up to the search date	Full texts not in English

Source: Authors' review protocol (version 0.1, 2 October 2026), structured by population, index model, outcome, timing and study type.

2.3 Information sources and search strategy

We searched PubMed (MEDLINE) and OpenAlex (Priem *et al.*, 2022) from inception to 3 October 2026. PubMed was chosen for its controlled vocabulary and clinical coverage, and OpenAlex because it indexes the engineering and computing journals and conference proceedings in which most fuzzy models are published. Both were searched through their public programming interfaces (NCBI E-utilities and the OpenAlex API), so that the searches can be repeated exactly. The choice of sources was tested beforehand: of 29 eligible studies

known to the authors in advance, only 4 were indexed in PubMed, and on 2 October 2026 the two search strings together retrieved 27 of the 29 (93%), including all 4 in PubMed.

The strategy combined two concepts, fuzzy methods and the target condition, searched in titles and abstracts, with Medical Subject Headings added in PubMed. No filters for language, date or publication type were applied. A third concept for the diagnostic task was tested and dropped, because it removed eligible studies without reducing the screening load much. The term "type-2" was also dropped, because in OpenAlex it matched tens of thousands of records on type 2 diabetes; type-2 fuzzy studies are retrieved through the term "fuzzy".

2.4 Record management and removal of duplicates

The raw exports from both sources were saved unchanged. Duplicates were removed in two steps with scripts written for the review. First, records were matched automatically on the digital object identifier (DOI), then on the PubMed identifier, then on the normalized title together with the year of publication; where a record was found in both sources, the PubMed record was kept and its missing fields were filled from the OpenAlex record. Second, records with the same normalized title but a different year were flagged and checked one by one. Only records of the same publication were removed at this step: repository copies, mirrors, records without bibliographic data and early-online versions of the same article under a different DOI. Republished versions of the same work, such as reprints, preprints and conference papers later published in a journal, were kept and linked to each other.

2.5 Study selection

All 1,925 records were screened, first on their titles and then, for the 988 records that were not clearly irrelevant on the title, on their abstracts. First-pass decisions were drafted by the first author with the assistance of the team, using the predefined eligibility criteria, and a reason was recorded for every decision. Each decision was then checked and corrected where necessary. Each record was assessed against the predefined eligibility criteria and marked Include, Maybe or Exclude with a reason. Records without an abstract, records probably not in English and records in which the role of the fuzzy method was unclear were marked maybe, and all records marked Include or Maybe were sought in full text. Full texts were obtained from the open-access copies listed by OpenAlex and from copies already held by the authors. Each full text was assessed against the criteria and excluded with a single reason, taken in a fixed order: not in English; not a primary study; not a fuzzy classifier; outcome not CAD or heart disease; no evaluation on patient data; duplicate publication. Reports in which eligibility could not be determined from the text, for example because the part played by the fuzzy component in the final classification was not described, were not included, and the reason was recorded. Selection was not repeated independently by a second reviewer, so inter-reviewer agreement could not be measured (Section 4.5).

2.6 Data extraction

Data were extracted into a form based on the CHARMS checklist (Moons *et al.*, 2014), with items added for fuzzy models (Table 2). For each report, the researchers recorded the data source, the fuzzy family, whether type-2 fuzzy sets were used, whether the model was diagnostic or prognostic, the type of validation, the performance measures reported, and whether an area under the ROC curve (AUC) and calibration were reported, with a short note of the main findings taken from the text. The data source was classed as the public UCI heart disease data (Cleveland, Statlog or combined), UCI data together with the authors' own clinical data, another public dataset, the authors' own clinical data only, or not clearly reported. The unit of analysis was the report. Where several reports described the same model or used the same patient data, each report was extracted and appraised separately, because each made its own claims, and the overlap was recorded and reported (Section 3.1). Overlapping use of the public UCI data was counted at the level of reports only: no patient-level data or performance measures were pooled, so the same patients could not be counted twice in any synthesis.

Table 2: Data items extracted from each report.

Group	Items
Study	Authors, year, journal or venue
Data	Source (clinical cohort or named public dataset), sample size where reported, outcome definition, use of several datasets
Model	Fuzzy family; use of type-2 fuzzy sets; rule generation and number of rules where reported; learning or optimization method
Timing	Diagnostic or prognostic
Validation	Apparent, split-sample, cross-validation, or external (different population or time)
Performance	Accuracy, sensitivity, specificity, AUC, calibration measures; comparator models

Source: Adapted from the CHARMS checklist (Moons *et al.*, 2014), with items for fuzzy models added by the authors.

2.7 Risk of bias and applicability

Each model was assessed with PROBAST (Wolff *et al.*, 2019; Moons *et al.*, 2019) in the domains of participants, predictors, outcome and analysis, giving an overall judgement of low, high or unclear risk of bias and of concern about applicability. The judgements were made with explicit rules applied to the extracted items, so that the same evidence always led to the same judgement. In the participants domain, a data source that was not reported was judged high risk, and secondary public benchmark data or patient selection that was not described were judged unclear. In the outcome domain, outcome labels derived from the same signal used as a predictor were judged high risk, and a reference standard based on expert judgement or an abnormal test rather than angiography was judged unclear. In the analysis domain, internal validation only, with or without calibration, was judged high risk, following PROBAST and PROBAST+AI (Moons *et al.*, 2025). A model was at high overall risk of bias if any domain was at high risk as shown in Eq. (12).

2.8 Synthesis

Reports were described in tables and summarized by fuzzy family, data source, type of validation and the reporting of key items. The protocol planned to pool AUCs with a random-effects model where at least five

reports gave an AUC with a confidence interval. Too few reports gave an AUC for this (Section 3.6), so no meta-analysis was done, and accuracies were not pooled because they depend on disease prevalence and on the threshold used.

2.9 Changes from the protocol

The review differs from its protocol in six ways. First, screening, data extraction and risk of bias assessment were carried out as described in Section 2.5, with decisions and reasons recorded throughout the process, instead of independently by two reviewers. Second, full texts were limited to those available in open access or already held by the authors. Third, the planned backward and forward citation searches were not done. Fourth, authors of the included reports were not contacted for missing data. Fifth, records were deduplicated with scripts rather than in a reference manager, and screened in a spreadsheet form rather than in Rayyan; scripts made every step reproducible. Sixth, the planned meta-analysis was not possible. The eligibility criteria were not changed.

3. Results

This section reports the findings of the review in the order of the review questions set out in Section 1. It describes the selection of reports (Section 3.1), the characteristics of the included reports (Section 3.2), the models they developed (Section 3.3), their validation and completeness of reporting (Section 3.4), their risk of bias and applicability (Section 3.5), the performance they reported (Section 3.6), and their interpretability, handling of uncertainty and coverage of premature disease (Section 3.7).

3.1 Study selection

The database searches on 3 October 2026 returned 2,215 records: 281 from PubMed and 1,934 from OpenAlex. Automatic matching removed 283 duplicate records (207 by DOI, 26 by PubMed identifier and 50 by title and year). Of 51 records flagged as possible duplicates, in 25 groups, 7 were removed after the manual check, leaving 1,925 unique records, of which 397 had no abstract in either source (Fig. 1).

Screening excluded 1,131 records, 937 of them on the title alone. The exclusions made on the abstract were for: not a fuzzy classifier (88), outcome not CAD or heart disease (45), not a diagnostic or prognostic model (40), not a primary study (17), no evaluation on patient data (3) and population not adults (1). Of the 794 reports sought in full text, 384 had been marked Include and 410 Maybe, including 127 without an abstract and 34 probably not in English.

An open-access full text or a copy held by the authors was available for 135 of the 794 reports. Of these, 3 could not be used (a wrong file under the record, a scanned thesis without text, and a thesis of which only the front matter was readable, whose journal version was assessed instead), leaving 132 assessed. Of these, 44 were excluded: no evaluation on patient data (15), not a fuzzy classifier (11), not in English (11) and outcome not CAD or heart disease (7). In 26 further reports eligibility could not be determined from the text, most often because the fuzzy component was a combiner, a severity grader or a step whose contribution to the classification was not reported. In total, 62 reports were included (Table 3). Two of them, by the same author in the same year, report the same accuracy and may describe one model (Singh, 2024a; Singh, 2024b), and two others used

the same patient data (Apostolopoulos & Groumpos, 2020; Apostolopoulos *et al.*, 2020); all four were kept as separate reports.

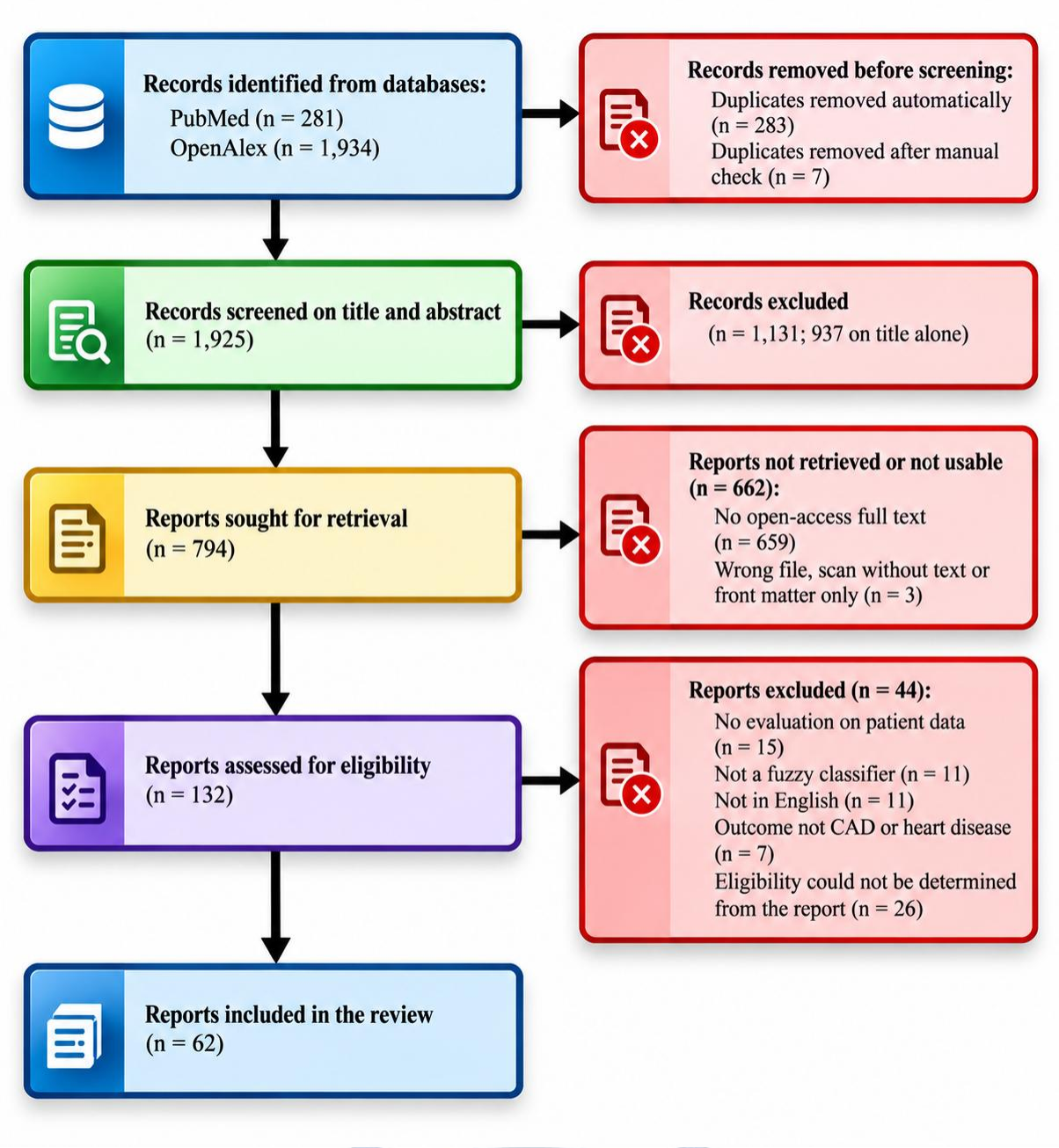

Fig. 1. PRISMA 2020 flow diagram of the search and selection of reports.
Source: Authors' search and screening results (3 October 2026), presented in the PRISMA 2020 flow-diagram format (Page *et al.*, 2021).

3.2 Characteristics of included reports

The included reports were published between 2000 and 2026: 40 from 2020 onwards, 20 in 2010 to 2019 and 2 before 2010. Most appeared in computing and engineering journals or conference proceedings rather than in clinical journals. In all, 25 reports used only the public UCI heart disease data (Cleveland, Statlog or the combined UCI data), and 3 combined them with their own clinical data, so that 28 of 62 (45%) relied at least in

part on the same public data. Another 6 used another public dataset (Kaggle, IEEE Dataport, Framingham or PTB), 16 used only their own clinical data, collected in Brazil, China, the Czech Republic, Greece, Indonesia, Iran, Nigeria and Ukraine among other countries, and 12 did not clearly report where their data came from. The largest clinical samples came from Brazil, where a Sugeno model for selecting patients for myocardial perfusion scintigraphy was developed on 1,053 patients and tested on 1,045 more (Duarte *et al.*, 2006), and from the Czech Republic, where fuzzy and weighted risk models predicted death within 24 months of myocardial infarction in 308 patients (Honzik *et al.*, 2010). Coronary angiography was the reference standard in only a few hospital studies, such as an ANFIS model for 152 patients in Shiraz (Ayat & Khosravanian, 2014) and a fuzzy cluster analysis of 166 positive exercise tests (Peters *et al.*, 2000). At the other extreme, three systems were evaluated on 10 to 12 patients (Arslankaya & Çelik, 2021; Al Hakim *et al.*, 2023; Rizkyta *et al.*, 2026). Of the included reports, 58 described diagnostic models and 4 prognostic models, of later unstable angina, death after myocardial infarction or ten-year cardiovascular risk.

Table 3: Characteristics of the included reports.

Report	Data source	Fuzzy family	Timing
Abdeldjouad *et al.*, 2020	UCI heart disease	Genetic fuzzy	Diagnostic
Abushariah *et al.*, 2014	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Aggarwal & Kumar, 2022	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Al Hakim *et al.*, 2023	Not clearly reported	Rule-based	Diagnostic
Alanazi & Khamis, 2024	Other public dataset	Genetic fuzzy	Diagnostic
Ali *et al.*, 2024	UCI and own clinical data	Rule-based	Diagnostic
Apostolopoulos & Groumpos, 2020	Own clinical data	Fuzzy cognitive map	Diagnostic
Apostolopoulos *et al.*, 2020	Own clinical data	Fuzzy cognitive map	Diagnostic
Ardan *et al.*, 2019	Other public dataset	Rule-based	Diagnostic
Arslankaya & Çelik, 2021	Own clinical data	Rule-based	Diagnostic
Ayat & Khosravanian, 2014	Own clinical data	ANFIS or neuro-fuzzy	Diagnostic
Barman & Chaudhury, 2013	UCI heart disease	Rule-based	Diagnostic
Damodharan *et al.*, 2023	UCI heart disease	Fuzzy clustering	Diagnostic
Diwan *et al.*, 2026	UCI heart disease	Rule-based	Diagnostic
Dong *et al.*, 2014	Own clinical data	Genetic fuzzy	Prognostic
Duarte *et al.*, 2006	Own clinical data	Rule-based	Diagnostic
Ephzibah, 2011	UCI heart disease	Rule-based	Diagnostic
Ephzibah, 2012	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Febriani *et al.*, 2023	Own clinical data	Other fuzzy	Diagnostic
Feng *et al.*, 2021	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Honzik *et al.*, 2010	Own clinical data	Other fuzzy	Prognostic
Huynh, 2020	Other public dataset	Other fuzzy	Prognostic
Janani *et al.*, 2023	Not clearly reported	Rule-based	Diagnostic
Jha *et al.*, 2023	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Jian *et al.*, 2014	UCI heart disease	Rule-based	Diagnostic
Kasbe, 2022	UCI heart disease	Rule-based	Diagnostic
Khan & Algarni, 2020	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Khormehr & Maihami, 2016	Own clinical data	Rule-based	Diagnostic
Kordnoori *et al.*, 2021	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Kresnawati *et al.*, 2024	Not clearly reported	Fuzzy decision tree	Diagnostic
Krishnam & Shafiulilah, 2015	Not clearly reported	Rule-based	Diagnostic
Manohar *et al.*, 2021	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Meda & Bhogapathi, 2021	Not clearly reported	ANFIS or neuro-fuzzy	Diagnostic
Mohamed *et al.*, 2025	Other public dataset	ANFIS or neuro-fuzzy	Diagnostic
Muflikhah *et al.*, 2013	Own clinical data	Rule-based	Diagnostic
Muhammad & Algehyne, 2021	Own clinical data	Rule-based	Diagnostic
Muwal & Kumar, 2016	UCI heart disease	Rule-based, type-2	Diagnostic
Narigina *et al.*, 2026	Other public dataset	ANFIS or neuro-fuzzy, type-2	Diagnostic
Naseer *et al.*, 2018	UCI heart disease	Rule-based	Diagnostic
Obanijesu & Emuoyibofarhe, 2012	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Palanisamy *et al.*, 2025	Not clearly reported	ANFIS or neuro-fuzzy	Diagnostic
Pasam *et al.*, 2025	Other public dataset	ANFIS or neuro-fuzzy	Prognostic
Peters *et al.*, 2000	Own clinical data	Fuzzy clustering	Diagnostic

Report	Data source	Fuzzy family	Timing
Pham *et al.*, 2022	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Prasad *et al.*, 2019	Not clearly reported	ANFIS or neuro-fuzzy	Diagnostic
Priya, 2024	Not clearly reported	Fuzzy decision tree	Diagnostic
Ragab *et al.*, 2024	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Rajab, 2018	UCI and own clinical data	ANFIS or neuro-fuzzy	Diagnostic
Rizkyta *et al.*, 2026	Own clinical data	Rule-based	Diagnostic
Sadiyamole & Manju Priya, 2023	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Sajiah *et al.*, 2016	UCI heart disease	Rule-based, type-2	Diagnostic
Saranya & Manavalan, 2019	Not clearly reported	ANFIS or neuro-fuzzy	Diagnostic
Senthil Kumar, 2013	UCI heart disease	ANFIS or neuro-fuzzy	Diagnostic
Setiawan *et al.*, 2020	UCI and own clinical data	Rule-based	Diagnostic
Singh *et al.*, 2021	Not clearly reported	ANFIS or neuro-fuzzy	Diagnostic
Singh, 2024a	Not clearly reported	ANFIS or neuro-fuzzy	Diagnostic
Singh, 2024b	Not clearly reported	ANFIS or neuro-fuzzy	Diagnostic
Sumarlinda *et al.*, 2023	Own clinical data	Rule-based	Diagnostic
Sumarlinda *et al.*, 2024	Own clinical data	ANFIS or neuro-fuzzy	Diagnostic
Tiwari & Garg, 2021	UCI heart disease	Rule-based	Diagnostic
Wojcik *et al.*, 2023	Own clinical data	Rule-based	Diagnostic
Zhang *et al.*, 2023	UCI heart disease	Rule-based	Diagnostic

Source: Authors' data extraction from the 62 included reports (search date 3 October 2026).

3.3 Model characteristics

Of the 62 reports, 26 (42%) used ANFIS or another neuro-fuzzy network, often tuned with a genetic or swarm algorithm (Feng *et al.*, 2021; Khan & Algarni, 2020) or built into deep convolutional, recurrent and wavelet fuzzy neural networks (Mohamed *et al.*, 2025; Kordnoori *et al.*, 2021; Ragab *et al.*, 2024), and 24 (39%) a fuzzy expert or rule-based system of the Mamdani, Sugeno, Takagi-Sugeno-Kang or Tsukamoto type. The rest used genetic or evolutionary fuzzy rule learning (3), fuzzy cognitive maps (2), fuzzy clustering as the classifier (2), fuzzy decision trees (2) or another approach (3: fuzzy logistic regression, a fuzzy-weighted risk score and a fuzzy ontology with case-based reasoning). Only 3 used type-2 fuzzy sets: Sajiah *et al.* (2016) compared interval type-2 and type-1 systems on the same data, Muwal and Kumar (2016) combined an interval type-2 classifier with nature-inspired attribute reduction, and Narigina *et al.* (2026) attached an interval type-2 ANFIS to a convolutional network. Rules were set by experts in many of the rule-based systems and learned from data in the neuro-fuzzy and genetic systems. Where the number of rules was reported it varied widely, from 18 in one Mamdani system (Rizkyta *et al.*, 2026) to more than 13,000 in a hybrid model (Jha *et al.*, 2023), and several expert systems had several hundred rules (Jian *et al.*, 2014; Kasbe, 2022; Arslankaya & Çelik, 2021), far more than a clinician could read. Many neuro-fuzzy models were tuned with genetic or other metaheuristic algorithms, or combined with deep networks in which the fuzzy layer was one part of a larger model.

3.4 Validation and completeness of reporting

Only 1 report (Huynh, 2020) tested its model in a cohort separate from the one used to develop it: a fuzzy ontology with case-based reasoning for ten-year cardiovascular risk, developed in the Framingham Offspring cohort and tested in the Framingham Original cohort. Its sensitivity held up (0.87 internally, 0.82 externally), but its specificity fell from 0.83 to 0.50, the kind of loss that internal validation cannot reveal. All the others reported performance from a single train-test split, from cross-validation or on the development data, and several did not say which. Reporting of the measures needed to judge a clinical model was poor (Table 4). Only 3 reports gave an AUC and 1 reported calibration (Narigina *et al.*, 2026). Most reports gave accuracy alone or

with sensitivity and specificity, without confidence intervals. The handling of missing data and of class imbalance, and whether feature selection or resampling was done before or after the data were split, were often not described, so data leakage could not be ruled out in most reports. Where fuzzy classifiers were compared with conventional ones on the same data, they did not consistently do better: on the IEEE Dataport data a multi-objective evolutionary fuzzy classifier reached about 81.6% accuracy against 85.6% for a decision table with naive Bayes (Alanazi & Khamis, 2024), and on the Cleveland data none of four genetic fuzzy classifiers outperformed AdaBoost (Abdeldjouad *et al.*, 2020).

Table 4: Reporting of key items across included reports.

Item	Number of reports	Percentage
Data source clearly reported	50	81%
Own clinical data used (alone or with public data)	19	31%
Validation in a separate cohort	1	2%
Area under the ROC curve reported	3	5%
Calibration reported	1	2%
Type-2 fuzzy sets used	3	5%

Source: Authors' analysis of the 62 included reports.

3.5 Risk of bias and applicability

In all, 61 of the 62 models (98%) were at high overall risk of bias and 1 was unclear; none was at low risk. Every high judgement included a high risk in the analysis domain, because performance was assessed only on the development data or by internal validation; the one model validated externally did not report calibration and was judged unclear. In the participants domain, 12 models were at high risk because the data source was not reported, and the rest were unclear, because public benchmark data and most hospital samples did not describe how patients were recruited. In the outcome domain, 2 models were at high risk because the outcome labels were derived from the same electrocardiogram used as the predictor, and the rest were unclear, mainly because the definition of "heart disease" in the public data and the reference standard in several hospital studies (expert judgement or an abnormal perfusion scan rather than angiography) were not fully described. The predictors domain was unclear for all models, because predictor definitions and their timing were rarely given. Concern about applicability was high for 3 models and unclear for the rest.

3.6 Reported performance

Reported accuracies ranged from 58% (Al Hakim *et al.*, 2023, on 12 test cases) to 99.66% (Sadiyamole & Manju Priya, 2023), and many reports on the UCI and Kaggle data gave accuracies above 95%. These figures came from internal validation on public data, often after feature selection and tuning, and are not comparable across reports. Where both training and test results were given, the gap could be large: one ANFIS model reached 100% accuracy on its training data and 75.93% on its test data (Abushariah *et al.*, 2014). The three AUCs reported were for a myocardial infarction class on an electrocardiogram dataset (0.933), a hybrid rule and fuzzy model on the Cleveland data, and a multi-disease IoT classifier. Because only three reports gave an AUC, fewer than the five needed, no meta-analysis was done.

3.7 Interpretability, uncertainty and premature disease

Although interpretability was the reason most reports gave for choosing a fuzzy model, few showed the rule base in a form a clinician could check, and none used a formal measure of interpretability, although a few kept the rule base small for that reason (for example Dong *et al.*, 2014). In the neuro-fuzzy and deep fuzzy models the learned rules were generally not reported. One model gave a per-patient measure of confidence, through an interval type-2 fuzzy component that decided when to trust a neural network (Narigina *et al.*, 2026). No report analysed premature (young-onset) CAD separately.

4. Discussion

This section interprets the findings of the review. It summarizes the principal findings (Section 4.1), compares them with previous reviews (Section 4.2), sets out their implications for research and practice (Section 4.3) and for decision-making under uncertainty (Section 4.4), and discusses the strengths and limitations of the review (Section 4.5).

4.1 Principal findings

This review found 62 open-access reports of fuzzy and neuro-fuzzy models for CAD or heart disease. Most were recent, most used ANFIS-type networks or expert rule bases, and almost half relied on the same public UCI heart disease data, which contain a few hundred patients collected decades ago. Only one report tested a model in a separate cohort, almost none reported an AUC or calibration, and every model but one was at high risk of bias. The high accuracies that are often quoted for these models therefore show how well they fit a familiar dataset, not how well they would detect CAD in a new clinic.

The single external validation illustrates the problem: a model that looked sound on its development data lost about 40% of its specificity in another cohort from the same study (Huynh, 2020). There is no reason to expect the far smaller models built on the UCI data to generalize better. The main argument for fuzzy models, that clinicians can read their rules, was rarely put into practice. Expert systems with hundreds of rules, and neuro-fuzzy and deep fuzzy networks whose learned rules were never shown, are no easier to check than other machine learning models. The few reports that collected their own clinical data were often very small, and some used reference standards other than angiography, which limits what their results can show.

Several features of the UCI Cleveland data explain why performance on this benchmark does not transfer. The dataset contains 303 patients examined in the 1980s (Detrano *et al.*, 1989), of whom 139 had disease, so a model with dozens of parameters or hundreds of rules has very few outcome events per parameter, far fewer than a stable model requires (Riley *et al.*, 2020). Two of its 13 attributes, the thallium scintigraphy result and the number of vessels coloured by fluoroscopy, come from tests performed late in the diagnostic pathway, and are not available when a first referral decision is made. The data have also been reused so often that design choices are, in effect, tuned to the same few hundred patients: each report that selects features, membership functions and hyperparameters by their accuracy on this dataset adds to a collective search whose best results are likely to be overfitted to it. Combined files add a further hazard: a widely used file that merges the UCI

cohorts contains 1,190 records, of which 272 are duplicates (Kumar *et al.*, 2025), so the same patient can appear in both training and test data.

Data leakage compounds these problems. When feature selection or tuning uses all the data before cross-validation, the selected features and parameters have already seen the test folds, so every fold gives an optimistic estimate; resampling such as SMOTE before the split places synthetic copies of test patients in the training data and has the same effect (Kapoor & Narayanan, 2023). In this review, the timing of these steps was usually not reported (Section 3.4). Accuracies of 98% or more on the Cleveland data are far above the discrimination achieved when a model built on 95,935 electronic health records was tested in an external cohort (AUC 0.91; Forrest *et al.*, 2023), and are more likely to reflect leakage and repeated tuning than diagnostic skill.

4.2 Comparison with previous reviews

These findings match those of reviews of machine learning prediction models in general, which found that most models were at high risk of bias because of small samples, internal validation only and calibration that was not assessed (Andaur Navarro *et al.*, 2021, 2023). Fuzzy models do not escape these weaknesses, and their reliance on a single public dataset is, if anything, greater than in the cardiovascular machine learning literature described by Krittanawong *et al.* (2020) and Garavand *et al.* (2023). Reviews of fuzzy methods (Cao *et al.*, 2024; Talpur *et al.*, 2023; Thukral & Gupta, 2026) described the techniques and their promise of interpretability but did not appraise validation or bias; this review shows that the promise has not yet been tested in the way clinical use would require.

4.3 Implications for research and practice

Fuzzy and neuro-fuzzy models for CAD should not yet be used to guide clinical decisions. Future studies should be planned with an adequate sample size (Riley *et al.*, 2020) and use clinical cohorts with a clear reference standard, preferably angiography; keep all feature selection, resampling and tuning inside the training data to avoid leakage; report discrimination with confidence intervals and assess calibration (Van Calster *et al.*, 2019); test the model in a new population; publish the full rule base and the number of rules, and keep it small enough to be read; and report according to TRIPOD+AI (Collins *et al.*, 2024). Further reuse of the UCI data is unlikely to add new evidence; new, shared clinical datasets, including patients with premature disease, would do more to show whether fuzzy models have a place in practice.

4.4 Implications for decision-making under uncertainty

Viewed as decision support, these findings have direct operational consequences. Only 1 of the 62 models reported calibration, so for 61 (98.4%) the agreement between predicted and observed risk is unknown. A model used for triage acts through a threshold (Eq. 11): if its probabilities are systematically too high, more patients cross the threshold than intended, and scarce resources such as stress testing or angiography are spent on patients at low risk; if they are too low, patients with disease are reassured. Without calibration, the net benefit of these models at a clinically relevant threshold cannot be estimated, and accuracy alone cannot show whether using a model would lead to better decisions than referring all patients or none. Where angiography is scarce, as in much of sub-Saharan Africa, these are exactly the trade-offs that a decision model is meant to manage.

Rule-base inflation has a similar cost. In multi-criteria decision analysis, a decision rule is useful partly because decision makers can see how each criterion contributes to the result, as in fuzzy TOPSIS (Modibbo *et al.*, 2022). Rule bases of several hundred to more than 13,000 rules (Section 3.3) remove this property: the decision can no longer be traced to the criteria, errors cannot be found by inspection, and accountability for a referral decision is lost. For fuzzy models to serve as decision support in health operations, they should be reported and evaluated as decision rules, with a calibrated probability, a stated threshold, a decision curve, and a rule base small enough to be checked.

4.5 Strengths and limitations

The searches can be rerun exactly through public interfaces, their recall was tested beforehand, and the raw exports, scripts and every screening, extraction and risk of bias decision was recorded with its reason, so each step can be checked and repeated. This is, to our knowledge, the first review to apply PROBAST to fuzzy models of CAD.

The review has important limitations. First, and most important, although screening, data extraction and risk of bias assessment were checked by the research team, the process did not involve a second independent reviewer or independent duplicate assessment of all decisions. Therefore, some errors of selection or judgement may remain, and formal inter-reviewer agreement could not be measured. Every decision was recorded with its reason so that it can be checked, and the findings should be confirmed by independent reviewers. The judgements that drive the main conclusion, internal validation only and calibration not reported, are matters of fact that can be checked directly in each report, which limits the effect that errors of judgement could have on that conclusion. Second, only 132 of 794 reports sought could be assessed, because full texts were limited to open-access copies; reports in subscription journals, which may differ in quality, are under-represented, and the counts given here describe the open-access literature rather than all published models. Third, 26 reports whose eligibility could not be determined were not included. Fourth, only two bibliographic sources were searched and citation searching was not done, although the recall test retrieved 27 of 29 known eligible studies. Fifth, 397 records had no abstract and were judged on their titles, and reports not in English were excluded. Sixth, the PROBAST judgements were made by explicit rules from the extracted items, which makes them consistent but may miss detail that a reviewer reading the whole paper would weigh. Finally, "heart disease" in the public datasets does not always mean angiographic CAD, so outcome definitions differ between reports. None of these limitations is likely to change the main conclusion, since the problems found were common and consistent across the reports assessed.

5. Conclusions

In the open-access literature, fuzzy and neuro-fuzzy models for the detection of CAD are almost all at high risk of bias, mostly rely on one small public dataset, and have rarely been tested in a new population or checked for calibration. Their readable rules, the main reason for choosing them, are seldom reported in a usable form. Before such models can be used in practice, they must be developed on adequate clinical cohorts, reported in full and validated externally.

Author Contributions

Conceptualization, methodology, data curation, formal analysis, writing and original draft, S.M.A; T.M., and M.M.L: validation, review, editing S.M.A; T.M.; M.M.L and M.A.B.; supervision, T.M., and M.M.L.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement

In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. You might choose to exclude this statement if the study did not report any data.

Conflicts of Interest

The authors declare no conflict of interest.

Data availability

The raw search exports, the deduplication scripts, the duplicate-check decisions, the screening decisions with reasons, the extracted data, and the PROBAST judgements are available from the corresponding author upon request.

Acknowledgement

We want to appreciate all the researchers and academics who wrote papers that helped us create this big review.

References

Abdeldjouad, F. Z., Menaouer, B., & Matta, N. (2020). A hybrid approach for heart disease diagnosis and prediction using machine learning techniques. In M. Jmaiel, M. Mokhtari, B. Abdulrazak, H. Aloulou, & S. Kallel (Eds.), The impact of digital technologies on public health in developed and developing countries (Lecture Notes in Computer Science, Vol. 12157, pp. 299-306). Springer. https://doi.org/10.1007/978-3-030-51517-1_26

Abushariah, M. A. M., Alqudah, A. A. M., Adwan, O., & Yousef, R. (2014). Automatic heart disease diagnosis system based on artificial neural network (ANN) and adaptive neuro-fuzzy inference systems (ANFIS) approaches. *Journal of Software Engineering and Applications, 7*(12), 1055-1064. https://doi.org/10.4236/jsea.2014.712093

Aggarwal, R., & Kumar, S. (2022). Classification model for meticulous presaging of heart disease detection through SDA and NCA using machine learning: CMSDANCA. *International Journal on Recent and Innovation Trends in Computing and Communication, 10*(1s), 217-224. https://doi.org/10.17762/ijritcc.v10i1s.5827

Al Hakim, M. F., Fajriati, N., & Pratama, R. N. (2023). Heart disease diagnosis using Tsukamoto fuzzy method. *Journal of Advances in Information Systems and Technology, 5*(1), 12-22. https://doi.org/10.15294/jaist.v5i1.67565

Alanazi, S. M., & Khamis, G. S. M. (2024). Optimizing machine learning classifiers for enhanced cardiovascular disease prediction. *Engineering, Technology & Applied Science Research, 14*(1), 12911-12917. https://doi.org/10.48084/etasr.6684

Ali, M. L., Sadi, M. S., & Goni, M. O. (2024). Diagnosis of heart diseases: A fuzzy-logic-based approach. *PLOS ONE, 19*(2), e0293112. https://doi.org/10.1371/journal.pone.0293112

Alizadehsani, R., Habibi, J. A., Hosseini, M. J., Mashayekhi, H., Boghrati, R., Ghandeharioun, A., Bahadorian, B., & Sani, Z. A. (2013). A data mining approach for diagnosis of coronary artery disease. *Computer Methods and Programs in Biomedicine, 111*(1), 52-61. https://doi.org/10.1016/j.cmpb.2013.03.004

Andaur Navarro, C. L., Damen, J. A., Takada, T., Nijman, S. W. J., Dhiman, P., Ma, J., Collins, G. S., Bajpai, R. C., Riley, R. D., Moons, K. G. M., & Hooft, L. (2021). Risk of bias in studies on prediction models developed using supervised machine learning techniques: Systematic review. *BMJ, 375*, n2281. https://doi.org/10.1136/bmj.n2281

Andaur Navarro, C. L., Damen, J. A., van Smeden, M., Takada, T., Nijman, S. W. J., Dhiman, P., Ma, J., Collins, G. S., Bajpai, R. C., Riley, R. D., Moons, K. G. M., & Hooft, L. (2023). Systematic review identifies the design and

methodological conduct of studies on machine learning-based prediction models. *Journal of Clinical Epidemiology, 154*, 8-22. https://doi.org/10.1016/j.jclinepi.2022.11.015

Apostolopoulos, I. D., & Groumpos, P. P. (2020). Non-invasive modelling methodology for the diagnosis of coronary artery disease using fuzzy cognitive maps. *Computer Methods in Biomechanics and Biomedical Engineering, 23*(12), 879-887. https://doi.org/10.1080/10255842.2020.1768534

Apostolopoulos, I. D., Groumpos, P. P., & Apostolopoulos, D. (2020). State Space Advanced Fuzzy Cognitive Map approach for automatic and non Invasive diagnosis of Coronary Artery Disease. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2004.03372

Ardan, A. N., Ma'arif, M., Aisyah, Z. H., Olivia, M., & Titin, S. M. (2019). Myocardial infarction detection system from PTB diagnostic ECG database using fuzzy inference system for S-T waves. *Journal of Physics: Conference Series, 1204*, 012071. https://doi.org/10.1088/1742-6596/1204/1/012071

Arslankaya, S., & Çelik, M. T. (2021). Prediction of heart attack using fuzzy logic method and determination of factors affecting heart attacks. *International Journal of Computational and Experimental Science and Engineering, 7*(1), 1-8. https://doi.org/10.22399/ijcesen.837731

Atanassov, K. T. (1986). Intuitionistic fuzzy sets. *Fuzzy Sets and Systems, 20*(1), 87-96. https://doi.org/10.1016/S0165-0114(86)80034-3

Ayat, S., & Khosravanian, A. (2014). Identification and classification of coronary artery disease patients using neuro-fuzzy inference systems. *Journal of Mathematics and Computer Science, 13*(2), 136-141. https://doi.org/10.22436/jmcs.013.02.05

Barma, M., Babando, A. H., & Emmanuel, J. J. (2026). Multi-criterion decision-making approach to prioritize strategies for combating smuggling and border control along the north-east borders of Nigeria. *Nigerian Journal of Operations Research, 3*(2), 56-72. https://doi.org/10.67868/nijor/0302.006

Barman, M., & Chaudhury, J. (2013). A framework for selection of membership function using fuzzy rule base system for the diagnosis of heart disease. *International Journal of Information Technology and Computer Science, 5*(11), 62-70. https://doi.org/10.5815/ijitcs.2013.11.07

Cao, J. L., Zhou, T., Zhi, S., Lam, S., Ren, G., Zhang, Y., Wang, Y., Dong, Y., & Cai, J. Y. (2024). Fuzzy inference system with interpretable fuzzy rules: Advancing explainable artificial intelligence for disease diagnosis-A comprehensive review. *Information Sciences, 662*, 120212. https://doi.org/10.1016/j.ins.2024.120212

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A., Camaradou, J., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R., Harvey, H., Heinze, G., . . . Logullo, P. (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

Damodharan, D., Goel, A. K., Agrawal, K. K., Johri, S., & Kumar, A. (2023). CFLCA: High performance based heart disease prediction system using fuzzy learning with neural networks. *International Journal on Recent and Innovation Trends in Computing and Communication, 11*(4), 98-112. https://doi.org/10.17762/ijritcc.v11i4.6392

Detrano, R., Janosi, A., Steinbrunn, W., Pfisterer, M. E., Schmid, J.-J., Sandhu, S. S., Guppy, K. H., Lee, S. J., & Froelicher, V. F. (1989). International application of a new probability algorithm for the diagnosis of coronary artery disease. *The American Journal of Cardiology, 64*(5), 304-310. https://doi.org/10.1016/0002-9149(89)90524-9

Diwan, A., Karmakar, S., & Soni, S. (2026). Design and development of a hybrid model for prediction of heart disease using back propagation neural network (BPNN), weighted classification based association (WCBA) and fuzzy logic (FL). *International Journal of Drug Delivery Technology, 16*(35s). https://doi.org/10.25258/ijddt.16.35s.102

Dong, W., Huang, Z., Ji, L., & Duan, H. (2014). A genetic fuzzy system for unstable angina risk assessment. *BMC Medical Informatics and Decision Making, 14*(1), 12. https://doi.org/10.1186/1472-6947-14-12

Duarte, P. S., Mastrocolla, L. E., Farsky, P. S., Sampaio, C. R., Tonelli, P. A., Barros, L. C., Ortega, N. R., & Pereira, J. C. (2006). Selection of patients for myocardial perfusion scintigraphy based on fuzzy sets theory applied to clinical-epidemiological data and treadmill test results. *Brazilian Journal of Medical and Biological Research, 39*(1), 9-18. https://doi.org/10.1590/s0100-879x2006000100002

Dudchenko, A., & Kopanitsa, G. (2017). Decision support systems in cardiology: A systematic review. *Studies in Health Technology and Informatics, 237*, 209-214. https://doi.org/10.3233/978-1-61499-761-0-209

Ephzibah, E. P. (2011). Time complexity analysis of genetic- fuzzy system for disease diagnosis. *Advanced Computing: An International Journal, 2*(4), 23-31. https://doi.org/10.5121/acij.2011.2403

Ephzibah, E. P. (2012). A neuro fuzzy expert system for heart disease diagnosis. *Computer Science & Engineering: An International Journal, 2*(1), 17-23. https://doi.org/10.5121/cseij.2012.2103

Febriani, V., Lestari, D. I., Mardiyati, S., & Lilyasari, O. (2023). Fuzzy logistic regression application on predictions coronary heart disease. *BAREKENG: Jurnal Ilmu Matematika dan Terapan, 17*(1), 0571-0580. https://doi.org/10.30598/barekengvol17iss1pp0571-0580

Feng, J., Wang, Q., & Li, N. (2021). An intelligent system for heart disease prediction using adaptive neuro-fuzzy inference systems and genetic algorithm. *Journal of Physics: Conference Series, 2010*(1), 012172. https://doi.org/10.1088/1742-6596/2010/1/012172

Forrest, I. S., Petrazzini, B. O., Duffy, Á., Park, J. K., Márquez-Luna, C., Jordan, D. M., Rocheleau, G., Cho, J. H., Rosenson, R. S., Narula, J., Nadkarni, G. N., & Do, R. (2023). Machine learning-based marker for coronary artery disease: Derivation and validation in two longitudinal cohorts. *The Lancet, 401*(10372), 215-225. https://doi.org/10.1016/S0140-6736(22)02079-7

Garavand, A., Behmanesh, A., Aslani, N., Sadeghsalehi, H., & Ghaderzadeh, M. (2023). Towards diagnostic aided systems in coronary artery disease detection: A comprehensive multiview survey of the state of the art. *International Journal of Intelligent Systems, 2023*(1), 6442756. https://doi.org/10.1155/2023/6442756

Honzik, P., Krivan, L., Lokaj, P., Labrova, R., Novakova, Z., Fiser, B., & Honzikova, N. (2010). Logit and fuzzy models in data analysis: Estimation of risk in cardiac patients. *Physiological Research, 59*(Suppl. 1), S89-S96. https://doi.org/10.33549/physiolres.932008

Huynh, S. M. (2020). *Fuzzy ontology case-based reasoning approaches to prediction of cardiovascular disease* [Doctoral thesis, Auckland University of Technology]. Tuwhera. https://hdl.handle.net/10292/13416

Janani, S. R., Subramanian, R., Karthik, S., & Vimalarani, C. (2023). Healthcare monitoring using machine learning based data analytics. *International Journal of Computers Communications & Control, 18*(1). https://doi.org/10.15837/ijccc.2023.1.4973

Jang, J.-S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics, 23*(3), 665-685. https://doi.org/10.1109/21.256541

Jha, R. K., Henge, S. K., Mandal, S. K., Menaka, C., Mehta, D., Upadhyay, A., Saini, A. K., & Mishra, N. (2023). Personating GA neural fuzzy hybrid system for computing HD probability. *International Journal of Advanced Computer Science and Applications, 14*(7). https://doi.org/10.14569/ijacsa.2023.0140771

Jian, S. J., Liou, K., & Chiu, R. K. (2014). A hybrid intelligent system for the disease risk evaluation. In *Proceedings of the Management, Knowledge and Learning International Conference 2014* (pp. 837-844). ToKnowPress. http://www.toknowpress.net/ISBN/978-961-6914-09-3/papers/ML14-644.pdf

Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns, 4*(9), 100804. https://doi.org/10.1016/j.patter.2023.100804

Karnik, N. N., & Mendel, J. M. (2001). Centroid of a type-2 fuzzy set. *Information Sciences, 132*(1-4), 195-220. https://doi.org/10.1016/S0020-0255(01)00069-X

Kasbe, T. (2022). FRBF: A fuzzy rule based framework for heart disease diagnosis. *Inteligencia Artificial, 25*(69), 122-138. https://doi.org/10.4114/intartif.vol25iss69pp122-138

Khan, M. A., & Algarni, F. (2020). A healthcare monitoring system for the diagnosis of heart disease in the IoMT cloud environment using MSSO-ANFIS. *IEEE Access, 8*, 122259-122269. https://doi.org/10.1109/access.2020.3006424

Khormehr, A., & Maihami, V. (2016). A novel fuzzy expert system design for predicting heart diseases. *International Journal of Computer Applications, 138*(4), 33-38. https://doi.org/10.5120/ijca2016908789

Kordnoori, S., Mostafaei, H., Rostamy-Malkhalifeh, M., Ostadrahimi, M., & Banihashemi, S. A. (2021). Diagnosis of heart disease using feature selection methods based on recurrent fuzzy neural networks. *IPTEK The Journal for Technology and Science, 32*(2), 64-73. https://doi.org/10.12962/j20882033.v32i2.3176

Kresnawati, E. S., Zayanti, D. A., Amran, A., Resti, Y., Aisyah, N., & Salsabila, A. (2024). Effect of fuzzy criteria on the performance of decision tree models for heart disease classification. In Proceedings of the 3rd Sriwijaya International Conference on Basic and Applied Sciences (SICBAS 2023). EAI. https://doi.org/10.4108/eai.3-11-2023.2347907

Krishnam, M., & Shafiulilah, S. K. (2015). An enhanced risk prediction system for cardiovascular disease in India using fuzzy classification. *International Journal of Computer Applications, 126*(7), 11-16. https://doi.org/10.5120/ijca2015906087

Krittanawong, C., Virk, H. U. H., Bangalore, S., Wang, Z., Johnson, K. W., Pinotti, R., Zhang, H., Kaplin, S. L., Narasimhan, B., Kitai, T., Baber, U., Halperin, J. L., & Tang, W. H. W. (2020). Machine learning prediction in cardiovascular diseases: A meta-analysis. *Scientific Reports, 10*(1), 16057. https://doi.org/10.1038/s41598-020-72685-1

Kumar, A., Dhanka, S., Sharma, A., Bansal, R., Fahlevi, M., Rabby, F., & Aljuaid, M. M. (2025). A hybrid framework for heart disease prediction using classical and quantum-inspired machine learning techniques. *Scientific Reports, 15*(1), 25040. https://doi.org/10.1038/s41598-025-09957-1

Madu, H. B., Yakubu, J., & Baba, U. A. (2026). A Mamdani fuzzy inference system for intelligent greenhouse climate control: Integrating temperature, humidity, and soil moisture for adaptive precision agricultural decision-making. *Nigerian Journal of Operations Research, 3*(3), 41-62. https://doi.org/10.67868/nijor/0303.021

Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies, 7*(1), 1-13. https://doi.org/10.1016/S0020-7373(75)80002-2

Manohar, M., Pani, A. K., & Kumar, P. (2021). A big data analysis using fuzzy deep convolution network based model for heart disease classification. *International Journal of Intelligent Engineering and Systems, 14*(2), 147-156. https://doi.org/10.22266/ijies2021.0430.13

Meda, S., & Bhogapathi, R. B. (2021). An efficient and scalable heart disease diagnosis system with attribute impact based weights and genetic correlation analysis. *Revue d'Intelligence Artificielle, 35*(1), 47-53. https://doi.org/10.18280/ria.350105

Menaceur, N. E., Kouah, S., Derdour, M., Ouanes, K., & Ammi, M. (2025). Fuzzy logic in arrhythmia detection: A systematic review of techniques, applications, and clinical interpretability. *Applied Computer Science, 21*(3), 162-181. https://doi.org/10.35784/acs_7657

Mendel, J. M., & John, R. I. (2002). Type-2 fuzzy sets made simple. *IEEE Transactions on Fuzzy Systems, 10*(2), 117-127. https://doi.org/10.1109/91.995115

Modibbo, U. M., Hassan, M., Ahmed, A., & Ali, I. (2022). Multi-criteria decision analysis for pharmaceutical supplier selection problem using fuzzy TOPSIS. *Management Decision, 60*(3), 806-836. https://doi.org/10.1108/MD-10-2020-1335

Mohamed, N., Lafi Almutairi, R., Abdelrahim, S., Alharbi, R., Alhomayani, F. M., Alsulami, A., & Alkhalaf, S. (2025). Deep convolutional fuzzy neural networks with stork optimization on chronic cardiovascular disease monitoring for pervasive healthcare services. *Scientific Reports, 15*(1), 19008. https://doi.org/10.1038/s41598-025-02924-w

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*(1), 1. https://doi.org/10.1186/2046-4053-4-1

Moons, K. G. M., Damen, J. A. A., Kaul, T., Hooft, L., Andaur Navarro, C. L., Dhiman, P., Beam, A. L., Van Calster, B., Celi, L. A., Denaxas, S., Denniston, A. K., Ghassemi, M., Heinze, G., Kengne, A. P., Maier-Hein, L., Liu, X., Logullo, P., McCradden, M. D., Liu, N., . . . van Smeden, M. (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

Moons, K. G. M., de Groot, J. A. H., Bouwmeester, W., Vergouwe, Y., Mallett, S., Altman, D. G., Reitsma, J. B., & Collins, G. S. (2014). Critical appraisal and data extraction for systematic reviews of prediction modelling studies: The CHARMS checklist. *PLoS Medicine, 11*(10), e1001744. https://doi.org/10.1371/journal.pmed.1001744

Moons, K. G. M., Wolff, R. F., Riley, R. D., Whiting, P. F., Westwood, M. E., Collins, G. S., Reitsma, J. B., Kleijnen, J., & Mallett, S. V. (2019). PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. *Annals of Internal Medicine, 170*(1), W1-W33. https://doi.org/10.7326/m18-1377

Muflikhah, L., Wahyuningsih, Y., & Marji. (2013). Fuzzy rule generation for diagnosis of coronary heart disease risk using substractive clustering method. *Journal of Software Engineering and Applications, 6*(7), 372-378. https://doi.org/10.4236/jsea.2013.67046

Muhammad, L. J., & Algehyne, E. A. (2021). Fuzzy based expert system for diagnosis of coronary artery disease in Nigeria. *Health and Technology, 11*(2), 319-329. https://doi.org/10.1007/s12553-021-00531-z

Muwal, S., & Kumar, N. (2016). A hybrid nature-inspired classification technique for medical diagnosis. *International Journal of Computer Applications, 153*(4), 32-38. https://doi.org/10.5120/ijca2016912003

Narigina, M., Romānovs, A., & Merkurjevs, J. (2026). A neuro-fuzzy digital twin for interpretable cardiac disease recognition. *Applied Sciences, 16*(15), 7371. https://doi.org/10.3390/app16157371

Naseer, I., Khan, B., Saqib, S., Tahir, S., Tariq, S., & Akhter, M. (2018). Diagnosis heart disease using Mamdani fuzzy inference expert system. *ICST Transactions on Scalable Information Systems, 162736*. https://doi.org/10.4108/eai.15-1-2020.162736

Obanijesu, O., & Emuoyibofarhe, J. O. (2012). Development of neuro-fuzzy system for early prediction of heart attack. *International Journal of Information Technology and Computer Science, 4*(9), 22-28. https://doi.org/10.5815/ijitcs.2012.09.03

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Palanisamy, S., Thangaraju, V., Kandasamy, J., & Salau, A. O. (2025). Towards precision in IoT-based healthcare systems: A hybrid optimized framework for big data classification. *Journal of Big Data, 12*(1), 190. https://doi.org/10.1186/s40537-025-01243-1

Pasam, G., Gupta, S., Srinivas, G., & Shaik, H. (2025). AIoMT-based multi-model soft computing techniques for the prediction of heart disease. *Iraqi Journal of Science, 66*(6), 2593-2602. https://doi.org/10.24996/ijs.2025.66.6.32

Peters, R. M., Peters, J. C., & Shanies, S. A. (2000). Predicting extent of coronary disease: Fuzzy cluster analysis vs. Duke treadmill score. *Journal für Kardiologie, 3*(1), 39-41. http://www.kup.at/kup/pdf/83.pdf

Pham, D., Nguyen, V., & Le, T. M. (2022). Fuzzy brain emotional controller for heart disease diagnosis. *Annals of Computer Science and Information Systems, 33*, 33-36. https://doi.org/10.15439/2022r50

Prasad, P. P., Francis, F. S., & Huq, S. (2019). An effective fuzzy neural network with reinforcement learning approach for medical data classification. *International Journal of Intelligent Engineering and Systems, 13*(1), 344-353. https://doi.org/10.22266/ijies2020.0229.32

Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv*, 2205.01833. https://doi.org/10.48550/arXiv.2205.01833

Priya, M. (2024). A knowledge based cloud-expert system for disease diagnosis. *Journal of Electrical Systems, 20*(10s), 8086-8092. https://doi.org/10.52783/jes.7042

Ragab, M., Binyamin, S. S., Alghamdi, W., Althaqafi, T., Assiri, F. Y., Al-Hanawi, M. K., & Al-Ghamdi, A. A. (2024). Leveraging fuzzy embedded wavelet neural network with multi-criteria decision-making approach for coronary artery disease prediction using biomedical data. *Scientific Reports, 14*(1), 31087. https://doi.org/10.1038/s41598-024-82019-0

Raina, A. A., Ali, S., Kalra, P., & Modibbo, U. M. (2026). Adaptive neuro-fuzzy inference system for transmission dynamics of HIV with an SEIRS epidemic model. *New Mathematics and Natural Computation, 22*(4), 1303-1325. https://doi.org/10.1142/S1793005726500651

Rajab, W. (2018). A hybrid AI approach based on genetic algorithm and ANFIS for heart disease diagnosis. *International Journal for Research in Applied Science and Engineering Technology, 6*(7), 789-798. https://doi.org/10.22214/ijraset.2018.7133

Riley, R. D., Ensor, J., Snell, K. I. E., Harrell, F. E., Martin, G. P., Reitsma, J. B., Moons, K. G. M., Collins, G. S., & van Smeden, M. (2020). Calculating the sample size required for developing a clinical prediction model. *BMJ, 368*, m441. https://doi.org/10.1136/bmj.m441

Rizkyta, A. D., Ulum, M. B., & Tobing, F. A. T. (2026). Fuzzy expert system for early heart disease diagnosis using Mamdani method in web-based system. *Ultimatics: Jurnal Teknik Informatika, 18*(1), 30-36. https://doi.org/10.31937/ti.v18i1.4367

Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., Barengo, N. C., Beaton, A., Benjamin, E. J., Benziger, C. P., Bonny, A., Brauer, M. D., Brodmann, M., Cahill, T. J., Carapetis, J. R., Catapano, A. L., Chugh, S. S., Cooper, L. T., Coresh, J., Fuster, V. (2020). Global burden of cardiovascular diseases and risk factors, 1990–2019. *Journal of the American College of Cardiology, 76*(25), 2982-3021. https://doi.org/10.1016/j.jacc.2020.11.010

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*(5), 206-215. https://doi.org/10.1038/s42256-019-0048-x

Sadiyamole, P. A., & Manju Priya, S. (2023). A robust cardiovascular disease predictor based on genetic feature selection and ensemble learning classification. *International Journal of Electrical and Computer Engineering Systems, 14*(7), 799-809. https://doi.org/10.32985/ijeces.14.7.7

Sajiah, A. M., Setiawan, N. A., & Wahyunggoro, O. (2016). Interval type-2 fuzzy logic system for diagnosis coronary artery disease. *Communications in Science and Technology, 1*(2), 51-55. https://doi.org/10.21924/cst.1.2.2016.26

Saranya, S., & Manavalan, R. (2019). Computational framework for heart disease prediction using deep belief neural network with fuzzy logic. *International Journal of Computer Applications, 181*(49), 26-34. https://doi.org/10.5120/ijca2019918675

Senthil Kumar, A. V. (2013). Generating rules for advanced fuzzy resolution mechanism to diagnosis heart disease. *International Journal of Computer Applications, 77*(11), 6-12. https://doi.org/10.5120/13436-0477

Setiawan, N. A., Venkatachalam, P. A., & Hani, A. F. M. (2020). Diagnosis of Coronary Artery Disease Using Artificial Intelligence Based Decision Support System. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2007.02854

Singh, J. (2024a). Comparison of proposed intelligent systems with existing models for monitoring heart diseases. *Journal of Information Systems Engineering and Management, 10*(1s), 34-46. https://doi.org/10.52783/jisem.v10i1s.99

Singh, J. (2024b). Intelligent hybrid model for predicting and monitoring heart disease. *Communications on Applied Nonlinear Analysis, 32*(1s), 335-345. https://doi.org/10.52783/cana.v32.2189

Singh, P., Kaur, A., Batth, R. S., Kaur, S., & Gianini, G. (2021). Multi-disease big data analysis using beetle swarm optimization and an adaptive neuro-fuzzy inference system. *Neural Computing and Applications, 33*(16), 10403-10414. https://doi.org/10.1007/s00521-021-05798-x

Sumarlinda, S., Rahmat, A., Awang Long, Z., & Lestari, W. I. (2023). Implementation of fuzzy inference system (FIS) for cardiovascular diseases prediction. *Proceeding of International Conference on Science, Health, and Technology*, 389-397. https://doi.org/10.47701/icohetech.v4i1.3418

Sumarlinda, S., Rahmat, A., Awang Long, Z., & Lestari, W. I. (2024). The comparative early prediction model for cardiovascular disease using machine learning. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 10*(1), 24-33. https://doi.org/10.32628/cseit241011

Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-15(1), 116-132. https://doi.org/10.1109/TSMC.1985.6313399

Talpur, N., Abdulkadir, S. J., Alhussian, H., Hasan, M. H., Aziz, N., & Bamhdi, A. M. (2023). Deep neuro-fuzzy system application trends, challenges, and future perspectives: A systematic survey. *Artificial Intelligence Review, 56*(2), 865-913. https://doi.org/10.1007/s10462-022-10188-3

Thomas, L. L., Goni, I., & Emeje, G. D. (2019). Fuzzy models applied to medical diagnosis: A systematic review. *Advances in Networks, 7*(2), 45. https://doi.org/10.11648/j.net.20190702.15

Thukral, S., & Gupta, A. (2026). Fuzzy logic applications in chronic disease in decision-making: A comprehensive healthcare review. *Informatics for Health and Social Care, 51*(1), 122-138. https://doi.org/10.1080/17538157.2025.2596948

Tiwari, S. K., & Garg, S. K. (2021). Comparison of coronary heart disease prediction models using various machine learning algorithms. *Journal of Engineering Research*. https://doi.org/10.36909/jer.icari.15323

Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L., & Steyerberg, E. W. (2019). Calibration: The Achilles heel of predictive analytics. *BMC Medicine, 17*(1), 230. https://doi.org/10.1186/s12916-019-1466-7

Varshney, A. K., & Torra, V. (2023). Literature review of the recent trends and applications in various fuzzy rule-based systems. *International Journal of Fuzzy Systems, 25*(6), 2163-2186. https://doi.org/10.1007/s40815-023-01534-w

Vickers, A. J., & Elkin, E. B. (2006). Decision curve analysis: A novel method for evaluating prediction models. *Medical Decision Making, 26*(6), 565-574. https://doi.org/10.1177/0272989X06295361

Wojcik, W., Mezhiievska, I., Pavlov, S. V., Lewandowski, T., Vlasenko, O. V., Maslovskyi, V., Volosovych, O., Kobylianska, I., Moskovchuk, O., Ovcharuk, V., & Lewandowska, A. (2023). Medical fuzzy-expert system for assessment of the degree of anatomical lesion of coronary arteries. *International Journal of Environmental Research and Public Health, 20*(2), 979. https://doi.org/10.3390/ijerph20020979

Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., Reitsma, J. B., Kleijnen, J., & Mallett, S. (2019). PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine, 170*(1), 51-58. https://doi.org/10.7326/M18-1376

Wu, D., Peng, R., & Mendel, J. M. (2023). Type-1 and interval type-2 fuzzy systems [AI-eXplained]. *IEEE Computational Intelligence Magazine, 18*(1), 81-83. https://doi.org/10.1109/MCI.2022.3223496

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zhang, X., Yin, Z., Jiang, Y., Jiang, Y., Sun, D., & Liu, Y. (2023). Fuzzy Knowledge Distillation from High-Order TSK to Low-Order TSK. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2302.08038