IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol.20, No.1, January 2026, pp. 71~82
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: https://doi.org/10.22146/ijccs.114611

Hybrid Manhattan Distance-Certainty Factor for Early Cardiovascular Diagnosis in a Hospital

Michel Farrel Tomatala*[1], Grace Sahoa[2], Mihuandayani[3], Dewi Christa Kobis[4]
[1,2]Department of Informatics Engineering, STMIK Multicom, Kotamobagu, Indonesia
[3]Department of Information System, STMIK Multicom, Kotamobagu, Indonesia [4]Department of English Literature, FIB Unsrat, Manado, Indonesia
e-mail: ***[1]michel@stmikmulticom.ac.id**, [2]gracesahoa@stmikmulticom.ac.id, [3]mihuandayani@stmikmulticom.ac.id, [4]kobisdewichrista@unsrat.ac.id

Abstrak

Penyakit kardiovaskular berkontribusi besar terhadap tingginya angka kematian dan beban operasional di rumah sakit publik, sehingga dibutuhkan alat bantu yang mampu menstandarkan pengambilan keputusan dini secara cepat. Penelitian ini mengevaluasi sistem pakar hibrida yang mengintegrasikan Manhattan Distance (MD) untuk mengukur kemiripan berbasis kasus dengan kerangka Certainty Factor (CF) untuk menggabungkan bukti berbasis aturan. Dengan memanfaatkan basis pengetahuan lokal yang dikurasi (110 kasus, 69 gejala, 13 kondisi) serta pengujian hold-out sebanyak 26 kasus yang dibandingkan dengan rujukan dokter spesialis, sistem mengambil kasus terdekat menggunakan MD pada vektor gejala, kemudian menghitung tingkat keyakinan tiap diagnosis menggunakan CF. Hasil evaluasi menunjukkan 23 dari 26 prediksi sesuai secara tepat (akurasi 88,46%), dengan rentang nilai keyakinan 71,90–99,99%. Nilai keyakinan cenderung lebih tinggi ketika pola kasus terdekat dan aturan saling menguatkan, serta lebih moderat pada presentasi klinis yang ambigu (misalnya dugaan aneurisma). Keluaran sistem bersifat interpretatif dan sesuai untuk konsultasi 15–30 menit, sehingga dapat mendukung triase yang lebih konsisten ketika kapasitas spesialis terbatas. Temuan ini menunjukkan jalur praktis untuk meningkatkan ketepatan waktu layanan dan mengurangi variasi keputusan. Penelitian lanjutan disarankan mencakup validasi multisitus, perluasan basis pengetahuan untuk fenotipe atipikal, serta tata kelola implementasi yang aman dan berkeadilan

Kata kunci— *Manhattan Distance, Certainty Factor, Sistem Pakar, Jantung dan Pembuluh Darah, Sistem Pendukung Keputusan.*

Abstract

Cardiovascular disease drives high mortality and operational strain in public hospitals, underscoring the need for tools that standardize rapid early decisions. We evaluated a hybrid expert system that integrates Manhattan Distance (MD) for case-based similarity with a Certainty Factor (CF) framework for rule-based evidence aggregation. Using a locally curated knowledge base (110 cases, 69 symptoms, 13 conditions) and a 26-case hold-out against specialist references, the system retrieves nearest cases via MD on symptom vectors and then computes per-diagnosis confidence with CF. The system achieved 23/26 exact matches (accuracy 88.46%), with confidence values spanning 71.90–99.99% higher when nearest-case patterns and rules converged and moderated in ambiguous presentations (e.g., suspected aneurysm). Outputs were interpretable and suitable for 15–30-minute consultations, supporting consistent triage where specialist capacity is limited. These findings suggest a practical pathway to improve timeliness and reduce variability. Future work should pursue multi-site validation, knowledge-based expansion for atypical phenotypes, and governance for safe, equitable deployment.

Keywords— *Manhattan Distance, Certainty Factor, Expert System, Cardiovascular, Decision Support*

Received December 19th,2025; Revised January 21th, 2026; Accepted January 28th, 2026

1. INTRODUCTION

Cardiovascular diseases (CVDs) remain the leading cause of death globally, accounting for approximately 17.9 million deaths annually, and they impose substantial clinical and economic burdens on health systems with constrained resources [1]. Early recognition and timely initiation of appropriate management are pivotal to reducing mortality and disability; yet in many low- and middle-income settings, the speed and consistency of diagnostic decisions are impeded by workforce limitations, variable expertise, and operational bottlenecks within public hospitals [2, 3]. In Indonesia and the broader Southeast Asian region, recent reports highlight persistent inequities in specialist availability between urban and rural areas, prolonged waiting times, and fragmented triage and referral pathways, all of which collectively delay definitive care for time-sensitive cardiovascular conditions [4].

The local context of RSUD Kota Kotamobagu exemplifies these challenges. Routine clinical encounters often require 15–30 minutes per case, and patient surges frequently overwhelm limited cardiology resources—at times staffed by only two specialists—leading to delays in initial decision-making that can adversely affect outcomes when minutes matter [5, 6]. Beyond capacity constraints, the overlapping presentation of cardiovascular syndromes (e.g., ischemic versus hemorrhagic stroke, hypertensive urgency versus pain-mediated tachycardia) complicates rapid differentiation at first contact. Within this environment, a standardized decision support approach that rapidly proposes likely diagnoses from presenting symptoms and articulates an interpretable degree of confidence could mitigate delays and reduce unwarranted variability in early care [7], [8]. Accordingly, this study addresses the following problem: how to design an expert system that accelerates early CVD diagnosis while explicitly accommodating clinical uncertainty so that outputs are both fast and informative for frontline clinicians. The system must operate effectively under constraints of incomplete information, mixed symptom quality, and fluctuating workload, and it should integrate seamlessly into existing triage and consultation pathways within a public hospital setting [9, 10]. From a methodological standpoint, the tool should balance computational efficiency, data parsimony, and interpretability to be sustainable in routine use.

General solutions proposed in the literature emphasize two complementary pillars. First, similarity-based reasoning maps a patient's symptom profile to archived cases to surface plausible diagnoses swiftly, a strategy well suited to resource-constrained contexts where curated case knowledge can be leveraged for rapid guidance. Second, rule-based certainty aggregation quantifies diagnostic confidence using clinician-elicited weights, thereby translating expert knowledge into transparent, auditable outputs that can be communicated at the bedside. Together, these paradigms promise speed and interpretability—key for clinical adoption [11, 12, 13]. Within the similarity family, Manhattan Distance (L1 norm) is frequently favored for high-dimensional clinical symptom vectors due to its computational simplicity and relative robustness to outliers compared with Euclidean distance; studies in diagnostic and classification tasks have reported competitive or superior performance for Manhattan Distance in settings with overlapping features [14, 15, 16]. Its focus on coordinate-wise deviations aligns with the way clinicians often reason about incremental symptom differences across cases. In complement, cosine similarity has utility in sparse, direction-sensitive spaces (e.g., text or some genomic profiles), yet its angle-based nature may be less informative for continuous symptom scales dominant in early cardiovascular assessment. Selecting Manhattan Distance therefore reflects a pragmatic alignment between mathematical properties and the problem structure of triage-oriented, symptom-driven diagnosis [17, 18].

For uncertainty handling, the Certainty Factor (CF) framework—originating from rule-based expert systems—offers an interpretable mechanism to encode expert belief about how specific symptoms alter confidence in candidate diagnoses. CF's additive and multiplicative update rules are easy to implement and explain, allowing clinicians to audit how evidence accumulates toward or against hypotheses. Although probabilistic Bayesian networks, fuzzy logic, or Dempster–Shafer theory provides richer uncertainty semantics, they typically require heavier model construction, parameterization, and computational resources that may impede realtime deployment in high-throughput public hospitals. In contexts where transparency and speed are paramount, CF achieves a practical balance between fidelity to clinical reasoning and operational feasibility [19, 20, 21]. A converging body of literature indicates that expert systems can effectively encode specialist knowledge to standardize consultations and expedite decisions when human experts are scarce, if inputs and outputs remain intelligible to end users. Empirical reports on Manhattan Distance in classification and triage-like tasks demonstrate strong performance under feature overlap and limited data regimes, while studies deploying CF in cardiovascular signal or symptom domains report high diagnostic agreement, underscoring CF's suitability for early, symptom-centric decisions.

Despite these advances, most implementations are evaluated on generic datasets or outside the realities of Indonesian public hospitals, where specialist scarcity, uneven referral pathways, and variable documentation quality complicate translation. This gap motivates contextualized solutions and localized knowledge bases that reflect the epidemiology and workflow of specific hospitals. Building on this foundation, the present study develops and evaluates a hybrid expert system that integrates Manhattan Distance for case–symptom similarity with CF-based rule aggregation orchestrated via forward chaining. The system is grounded in a localized knowledge base comprising 110 hospital cases across 13 cardiovascular conditions (7 cardiac, 6 vascular) and 69 symptoms curated with specialist input, reflecting routine presentations at RSUD Kota Kotamobagu.

The working hypothesis is that the combined approach will yield rapid, interpretable early diagnoses with sufficient accuracy to support triage and initial clinical decisions in a resource-constrained environment. The scope is deliberately focused on early diagnostic guidance rather than definitive diagnosis or treatment recommendation, acknowledging both the intended decision point and safety considerations. In summary, this introduction advances four contributions. First, it articulates a pressing clinical problem—delays and variability in early cardiovascular diagnosis within an overburdened Indonesian public hospital—framed by contemporary evidence on global and regional CVD burden and system constraints. Second, it motivates a hybrid, interpretable methodology Manhattan Distance for efficient similarity search and CF for transparent uncertainty expression selected for congruence with symptom-driven triage and operational realities.

Third, it identifies a concrete research gap: the paucity of integrated Manhattan Distance+ CF systems evaluated with localized knowledge bases and workflows in Indonesian public hospitals. Fourth, it states the study's objectives, novelty, and scope: to design, implement, and validate such a system for early CVD diagnosis at RSUD Kota Kotamobagu, hypothesizing that the hybrid approach will deliver timely, accurate, and interpretable outputs aligned with triage needs. By pursuing a solution that privileges transparency, computational parsimony, and contextual fit, this work aims to contribute a deployable decision aid and a replicable design pattern for similar hospitals confronting comparable constraints.

2. METHODS

2.1 Research design

This study employed a non-experimental, cross-sectional systems engineering design to develop and validate a rule- and case-based clinical decision support system (CDSS) for early cardiovascular diagnosis in a public hospital context. The methodological choice reflects the need for speed, interpretability, and operational fit in resource-constrained environments where specialist availability and consultation time are limited. The CDSS integrates a similarity engine based on Manhattan Distance to retrieve historical case analogues and a rule-based Certainty Factor (CF) mechanism to aggregate evidential support, consistent with prior literature advocating interpretable, triage-oriented inference under uncertainty. The general model of this research can be seen in Figure 1.

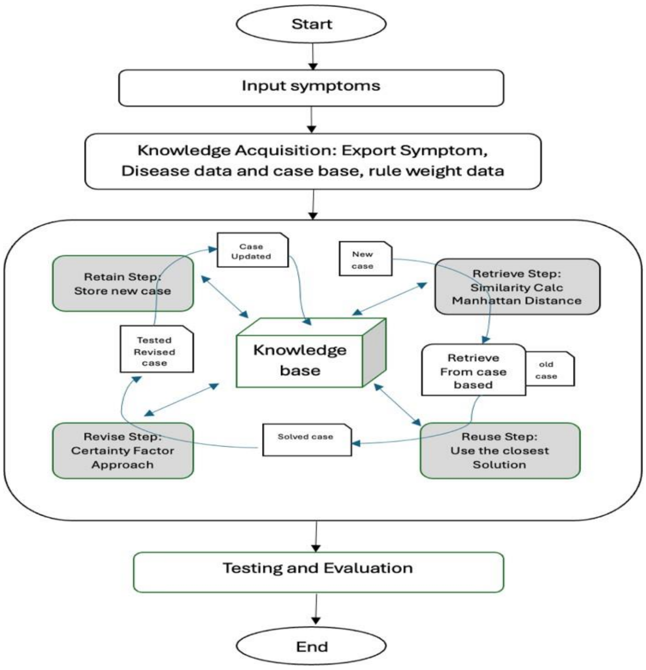

Figure 1. General Model Manhattan Distance-Certainty Factor

2.2 Setting, population, and sample.

The development setting was RSUD Kota Kotamobagu, a public hospital where cardiovascular care is challenged by high patient volumes and limited specialist staff. The population of interest comprises adult patients presenting with signs and symptoms suggestive of cardiac or vascular conditions in emergency and outpatient services. A convenience sample of 110 de-identified medical records was assembled to build the knowledge base, spanning 1 January–31 December 2023, and covering 13 target conditions seven cardiac and six vascular encoded across 69 symptoms curated with specialist input. For validation, 26 additional cases were selected as a test subset for expert comparison, representing a pragmatic sample aligned with daily case mix and workload constraints at the site.

Table 1. Dataset Summary

Component	Value
Diseases (cardiac / vascular)	13 (7 / 6)
Symptoms	69
Case base (period)	110 cases (1 Jan–31 Dec 2023)
Validation set	26 cases

2.3 Similarity-based method using Manhattan Distance.

The observed accuracy and confidence behaviors are congruent with the literature motivating the design. Similarity-based methods using Manhattan Distance have repeatedly demonstrated competitive accuracy in classification tasks that resemble triage, especially when features overlap and outliers can distort Euclidean metrics. CF-based rule aggregation has achieved high agreement in cardiovascular signal contexts, supporting its use for synthesizing symptom evidence into interpretable, graded recommendations. Against the backdrop of global CVD burden and local workflow constraints, the combination provides a reasonable, evidence- aligned mechanism for early decision support. Validation set concordance and confidence range for the MD+CF expert system on 26 cases from RSUD Kota Kotamobagu. Accuracy is defined as the proportion of exact matches between the system's top recommendation and the specialist reference; confidence is the per-diagnosis Certainty Factor percentage. The Historical cases can be seen in Table 2.

Table 2. The Case Base Table (Historical Cases)

Case	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	Disease
1	1	0	0	0	0	1	0	1	1	0	P1
2	1	0	0	0	1	0	1	0	1	0	P2

This section illustrates how a new patient case can be represented and matched against a stored case base using the Manhattan Distance (L1) metric. In this representation, each symptom feature is encoded in a binary form where a value of 1 indicates the symptom is present ("yes") and a value of 0 indicates the symptom is absent ("no"). The objective is to compute the distance between a new case and each historical case; the smallest distance indicates the most similar case, which is then used as the primary retrieval candidate in the case-based reasoning stage.

The stored case base excerpt is shown in Table 2. Each historical case is described by a symptom vector across features G1–G10 and is associated with a disease label (P1 or P2).

A new incoming case (Case A / KA) is encoded using the same G1–G10 schema, as presented in Table 3. This ensures that the new case and stored cases share a comparable feature space, enabling direct distance computation.

Table 3. Example new case table (Case / KA)

Case	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10
KA	1	0	1	1	0	1	0	1	1	0

Manhattan Distance is computed as the sum of absolute differences across all symptom features. For a new case vector xand a stored case vector c, the distance is:

$$D(x, c) = \sum_{i=1}^{m} |x_i - c_i| \tag{1}$$

where mis the number of symptom attributes (in this example, m = 10 for G1–G10). Because the attributes are binary, each term $| x_i - c_i |$ is either 0 (same value) or 1 (different value), making the distance directly interpretable as the number of symptom mismatches.

The knowledge based for the expert systems used in this method consist of Diseases Table can be seen in Table 4. The Symptons data consist of codes and names of this system is 69 symtoms and can be seen in Table 5.

Table 4. Diseases Table (Codes and Names).

No	Disease (English/Indonesian)	Code
1	Atherosclerotic heart disease (Penyakit Jantung Koroner)	P01
2	Congestive heart failure (Gagal Jantung)	P02
3	Angina Pectoris (Angin duduk)	P03
4	Atrial fibrillation and flutter (Gangguan Irama Jantung)	P04
5	Coronary (Serangan Jantung)	P05
6	Stroke ischemic	P06
7	Valvular heart disease (Penyakit Katup Jantung)	P07
8	Vena varicose (Varases)	P08
9	Thrombophlebitis	P09
10	Stroke hemorrhagic	P10
11	Hypertension	P11
12	Hypotensive	P12
13	Aneurism aorta	P13

Table 5. Symptom Codes and Names.

Symptom Code	Symptom Name
G1	Chest pain (pressure-like)
G2	Unusual fatigue without exertion
G3	Chest pain radiating to back and neck
G4	Epigastric pain (heartburn-like)
G5	Pain in right arm
G6	Pain in left arm
G7	Leg pain
G8	Heel pain
G9	Abdominal pain
G10	Back pain
G11	Flank/groin pain
G12	Shortness of breath
G13	Dyspnea on exertion
G14	Orthopnea
G15	Paroxysmal nocturnal dyspnea
G16	Palpitations
G17	Low blood pressure (hypotension)
G18	High blood pressure (hypertension)
G19	Cold sweats/diaphoresis

2.4 Certainty Factor: Discussion and Worked Example

The Certainty Factor (CF) approach is a classic mechanism in rule-based expert systems for expressing and combining uncertainty when evidence is incomplete or partially reliable. In clinical domains, uncertainty is common because many diseases share overlapping symptoms and because some findings are subjective or depend on patient-reported intensity. The CF model encodes the strength of evidence as a numerical value and then combines multiple pieces of evidence to obtain an overall confidence for a diagnostic hypothesis. In this study context, CF is used after candidate diagnoses are nominated (e.g., from similarity retrieval), so that the system can present an interpretable degree of belief for each diagnosis. To make CF entry consistent and reproducible, the original material defines a mapping from linguistic certainty terms to numeric values for expert weights, and a mapping for patient symptom intensity. These scales can be used to standardize elicitation during interviews with clinicians and to reduce variability in data entry.

2.5 Core CF definitions

In the CF framework, the impact of a symptom (evidence) E on a diagnostic hypothesis H can be represented in terms of belief and disbelief.

$$CF(H, E) = MB(H, E) - MD(H, E) \qquad (2)$$

where MB(H, E)denotes a *measure of belief* and MD(H, E)denotes a *measure of disbelief*. The resulting CF is often interpreted on a scale from 0 to 1 (supporting evidence) in simplified implementations, although the general CF framework can also represent negative values when evidence refutes a hypothesis. In the hospital implementation described by the attached material, uncertainty is operationalized in an intuitive way: the patient (or clinician enering the case) provides a *user confidence* (how strongly the symptom is present), and the expert provides an *expert CF weight* (how strongly that symptom supports a disease). The CF contribution of a symptom is then computed as Singe Symptom rule:

$$CF_{Symtom} = CF_{user} \; X \; CF_{Expert} \qquad (3)$$

This multiplication gives a transparent interpretation: evidence contributes strongly only when the symptom is both strongly present (high CF_{user}) and diagnostically meaningful according to the expert (high CF_{expert}). When multiple symptoms support the same diagnosis, their CF values are combined iteratively. For supportive (positive) CF values, the combination rule is:

$$CF_{Combine} = CF_{old} + CF_{new} \text{ X } (1 - CF_{old}) \tag{4}$$

Finally, the combined CF can be displayed as a percentage:

$$CF_{\%} = CF_{combine} \text{ X } 100 \tag{5}$$

To make CF entry consistent and reproducible, the original material defines a mapping from linguistic certainty terms to numeric values for expert weights, and a mapping for patient symptom intensity. These scales can be used to standardize elicitation during interviews with clinicians and to reduce variability in data entry. Then the expert Certainty Factor values can be seen in Table 6.

Table 6. Expert Certainty Factor Values

Expert Certainty Level	CF Value
Uncertainty	0.0
Almost uncertainty	0.2
Possibly not	0.4
Possible	0.6
Almost certain	0.8
Certain	1.0

defines how patient-reported symptom intensity is converted into a Certainty Factor (CF) value. In this study, the CF value represents the degree of confidence (from 0 to 1) that a symptom is truly present or clinically meaningful, based on how strongly the patient experiences it.

Table 7. Patient symptom intensity parameter values

Patient Symptom Intensity	CF Value
Mild	0.4
Moderate	0.6
Severe	0.8
Very severe	1.0

2.6 Worked example: applying CF to a sample patient

The attached material provides an example patient (Patient A) presenting with three symptoms: chest pain described as pressure (G1 with user weight 0.8), unusual fatigue without activity (G2 with user weight 0.6), and left arm pain (G6 with user weight 0.8). The following table summarizes how these selected symptoms map to candidate diagnoses and expert CF weights in the knowledge base.

Table 8. Selected Symptom and corresponding diseases with Expert CF Weight (Patient A)

Selected Symptom(user	Candidate Disease (rule	Expert CF
G1 (0,8)	P01 (rule 1)	1,0
	P03 (rule 3)	0,6
	P04 (rule 4)	0,4
	P05 (rule 5)	0,8
G2 (0,6)	P01 (rule 1)	0,6
	P02 (rule 2)	0,8
	P03 (rule 3)	0,6
	P04 (rule 4)	0,8
	P05 (rule 5)	0,8
G6 (0,8)	P01 (rule 1)	0,8
	P03 (rule 3)	0,6
	P05 (rule 5)	0,8

The CF contribution for each (symptom, disease) pair is obtained using Equation (3). For example, for diagnosis P01, symptom G1 contributes 0.8 × 1.0 = 0.8, symptom G2 contributes 0.6 × 0.6 = 0.36, and symptom G6 contributes 0.8 × 0.8 = 0.64. For diagnosis P03, the corresponding contributions are 0.8 × 0.6 = 0.48, 0.6 × 0.6 = 0.36, and 0.8 × 0.6 = 0.48. For diagnosis P05, contributions are 0.8 × 0.8 = 0.64, 0.6 × 0.8 = 0.48, and 0.8 × 0.8 = 0.64. After computing symptom-level CF values, each diagnosis aggregates evidence using Equation. (4) in sequence.

In the attached calculation, the final combined value is reported as $CF_2 = 0.808192$, corresponding to $CF_{\%} = 80.81\%$. In implementation, it is important to compute $CF_1$and CF_2directly from Equation (4) (and to handle rounding carefully), because small arithmetic or rounding inconsistencies in intermediate steps can propagate to the final percentage.

3. RESULTS AND DISCUSSION

The expert system was evaluated on a de-identified local corpus drawn from RSUD Kota Kotamobagu. The knowledge base comprised 110 historical cases across 13 cardiovascular conditions seven cardiac and six vascular encoded by 69 symptoms curated with specialist input. A hold-out validation subset of 26 cases was assembled to benchmark system recommendations against specialist conclusions. This configuration reflects the clinical reality of constrained specialist availability and high patient volumes in the setting, where rapid, standardized early decisions are needed to mitigate delay-related harms. The evaluation focused on early diagnostic concordance, acknowledging that definitive diagnosis and management remain clinician responsibilities.

3.1 Manual evaluation protocol and comparison criteria

The diagnostic performance of the proposed hybrid expert system was verified through a manual evaluation procedure that explicitly compared the system's output against specialist diagnoses recorded at RSUD Kota Kotamobagu. The assessment began by matching the system's predicted condition for each test case with the corresponding expert diagnosis, while retaining the system-generated confidence value expressed as a percentage. This manual workflow follows the same computational stages described earlier in the thesis results section, where relevant rules are executed according to the patient's input symptoms and the resulting values are then checked against the expert reference.

3.2 Accuracy calculation

Overall accuracy was computed using the same definition provided in the thesis, where accuracy ais the proportion of test cases whose system diagnosis matches the specialist diagnosis. Let n_{match} denote the number of consistent cases and nthe total number of evaluated cases. The accuracy is calculated as:

$$a = \frac{n_{\text{match}}}{n} \times 100$$

With $n_{\text{match}} = 23$ and $n = 26$, the computed accuracy is:

$$a = \frac{23}{26} \times 100 = 88.46\%.$$

The performance table reports disease-level precision, recall, and F1-score computed from the 26- case manual hold-out evaluation by treating the specialist diagnosis as the reference label. Across the evaluated conditions, precision ranges from 50.0% to 100.0%, recall ranges from 50.0% to 100.0%, and F1-scores range from 66.7% to 100.0%. Several categories show perfect scores because every test instance for those diseases was correctly classified and no other disease was mistakenly predicted as them within this test set. Overall, this pattern indicates that, for many conditions, the symptom configurations encoded in the knowledge base and retrieved through Manhattan Distance are sufficiently distinctive to support consistent early diagnostic suggestions. The diasnotic performance per disease can be seen in Table 10.

Table 10. Diagnostic performance per disease (precision, recall, F1-score)

Disease Code	Disease	Precision (%)	Recall (%)	F1-score (%)
P01	Atherosclerotic heart disease (Coronary Heart Disease)	100.00	100.00	100.00
P02	Congestive heart failure	100.00	100.00	100.00
P03	Angina pectoris	100.00	100.00	100.00
P04	Atrial fibrillation/flutter	100.00	100.00	100.00
P05	Myocardial infarction (Heart attack)	100.00	50.00	66.67
P06	Heart valve disease	50.00	100.00	66.67
P07	Congenital heart defects	100.00	50.00	66.67
P08	Peripheral artery disease	100.00	66.67	80.00
P09	Cardiomyopathy	50.00	100.00	66.67
P10	Aortic aneurysm	100.00	100.00	100.00
P11	Ischemic stroke	100.00	100.00	100.00
P12	Hemorrhagic stroke	100.00	100.00	100.00
P13	Hypertensive disorder (High blood pressure)	100.00	100.00	100.00

Lower scores are concentrated in clinically overlapping presentations. Heart valve disease and cardiomyopathy exhibit reduced precision (50.0%) because the system produced false-positive predictions for these labels, implying that certain symptom combinations (and their certainty weights) are not yet specific enough to separate them from close differential diagnoses. Myocardial infarction and congenital heart defects show reduced recall (50.0%) because one case in each class was missed (both shifted toward heart valve disease), while peripheral artery disease shows recall of 66.7% due to one misclassification toward cardiomyopathy. The accompanying bar chart visualizes the same phenomenon: diseases with shorter precision bars are impacted by false positives, whereas those with shorter recall bars are impacted by false negatives, making it easier to see whether refinement should prioritize reducing over-diagnosis or preventing missed cases. Because several diseases are represented by only 1–3 cases, these per-disease metrics should be interpreted cautiously and used primarily to guide knowledge-base and rule-weight refinement; a larger multi-site evaluation would better stabilize and generalize the estimates. The performance metrics per Cardiovascular Disease can be seen in Figure 2.

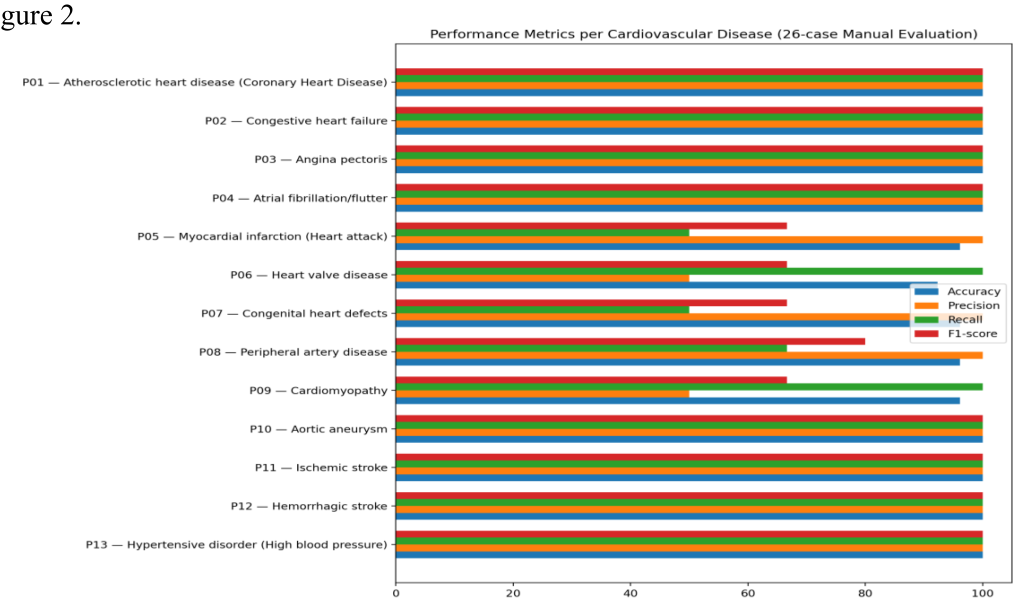

Figure 2. Performance Metric per Cardiovascular Disease(26 Cases Manual Evaluation)

Despite its limitations, the manual testing results support the system's intended role as a decision-support tool for early triage and preliminary assessment, particularly in environments where specialist capacity is constrained and timely early decisions matter.

The system's ability to provide a ranked diagnostic output with an interpretable confidence percentage can help clinicians or trained staff standardize initial reasoning, document the rationale for preliminary suspicion, and decide when to escalate to specialist review. In this sense, the system's primary value lies not in replacing clinical judgment, but in supporting consistency, transparency, and speed in the early diagnostic workflow, while maintaining a clear boundary that final decisions remain with qualified medical professionals. The observed inconsistencies point to refinement priorities that are typical for expert systems and case-based reasoning in medical domains. First, discriminative symptom definitions and completeness of symptom coverage are critical, because overlapping symptoms can cause multiple hypotheses to receive similar support. Second, the certainty weights attached to symptom–disease relationships may need recalibration through iterative expert review and additional data, particularly for conditions that share clinical manifestations. Third, the evaluation suggests that expanding the knowledge base to include more diverse cases and atypical presentations could reduce false matches by improving the granularity of similarity and rule evidence. Finally, while the 88.46% accuracy indicates promising internal performance, broader external validation across different time periods or clinical units would be necessary to confirm generalizability and to ensure that the system remains robust when encountering patient profiles that were underrepresented in the initial data source.

4. CONCLUSIONS

This study demonstrates that a hybrid expert system combining Manhattan Distance for case- based similarity and Certainty Factor for rule-based uncertainty aggregation can deliver accurate and interpretable support for early cardiovascular diagnosis in a resource-constrained public hospital. Using a locally curated knowledge base (110 cases, 69 symptoms, 13 conditions) and a 26-case validation, the system achieved 88.46% diagnostic concordance with specialists, with calibrated confidence outputs spanning 71.90–99.99%. These results indicate that similarity retrieval anchored in L1 distance can reliably surface clinically coherent differentials, while CF provides transparent degrees of belief that align with clinician expectations. Together, they offer a practical pathway to standardize early decisions within typical 15–30-minute consultations without supplanting clinical judgment. The findings contribute to the evidence base on explainable clinical decision support by showing that low-complexity, interpretable hybrids can meet operational requirements where specialist capacity and data resources are limited. Future research should evaluate multi-site generalizability, quantify time-to-decision and downstream outcomes, expand knowledge bases to better cover atypical phenotypes, and examine governance, safety, and equity impacts during real- world deployment. Methodological refinements distance weighted k-nearest neighbors, adaptive symptom weighting, and uncertainty calibration may further enhance performance while preserving interpretability.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the clinicians and staff of RSUD Kota Kotamobagu for their support during data curation and validation. We thank the cardiology and vascular specialists for expert knowledge elicitation and case-review sessions that informed the rule base and calibration of certainty weights. We are indebted to the medical records team and research assistants for meticulous abstraction, de-identification, and quality checks of the 2023 case base. We also appreciate the information technology unit for assistance with data security, system prototyping, and integration testing within routine clinical workflows.

We extend our thanks to colleagues who provided language help, writing assistance, and manuscript proofreading, which improved clarity and readability of the paper. We further acknowledge ChatGPT (OpenAI) as an AI assistant used for language polishing, drafting support, and formatting under the authors' supervision; responsibility for the scientific content remains solely with the authors. Finally, we acknowledge the hospital administration for facilitating access to non-identifiable records and for administrative coordination throughout the study. No individuals listed here bear responsibility for any remaining errors or interpretations.

REFERENCES

[1] G. A. Roth *et al.*, "Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study," Dec. 22, 2020, *Elsevier Inc.* doi: 10.1016/j.jacc.2020.11.010.

[2] A. Wahid, H. Prichoiruna, Ainun, and M. Muzakir, "Assessment of Risk and EmergencyKnowledge Level - Cardiovascular Emergencies in Rural Communities Sungai Nagara, MargasariVillage, Tapin, South Kalimantan," *Journal of Nursing Practice*, vol. 4, no. 2, pp. 196–203, 2021, doi: 10.30994/jnp.v4i2.121.

[3] Y. Koto, S. Solehudin, and M. O. Kurniawan, "Self Efficacy on Self Care in Coronary Heart Patients," *Int J Health Sci (Qassim)*, vol. 2, no. 1, pp. 1–11, 2024, doi: 10.59585/ijhs.v2i1.222.

[4] S. Wenang *et al.*, "Availability and Accessibility of Primary Care for the Remote, Rural, and Poor Population of Indonesia," *Front Public Health*, vol. 9, 2021, doi: 10.3389/fpubh.2021.721886.

[5] A. A. G. T. G. Pratama and A. Iskandar, "The Function of Endothelin-1 (ET-1) as a Biomarker for the Early Identification of Endothelial Dysfunction in National Health Security (NHS) and Sustainable Development Goals (SDGs): A Shift in Paradigm," 2025, doi: 10.21203/rs.3.rs6639485/v1.

[6] A. Menshawi and M. M. Hassan, "A Novel Triage Framework for Emergency Department Based on Machine Learning Paradigm," *Expert Syst*, vol. 42, no. 2, 2024, doi: 10.1111/exsy.13735.

[7] D. K. Kargirwar, D. A. Dange, and D. Pandit, "The Role of Artificial Intelligence and Machine Learning in Decision-Making in the ICU," *International Journal of Medical Science and Clinical Research Studies*, vol. 04, no. 12, 2024, doi: 10.47191/ijmscrs/v4-i12-31.

[8] D. Choi, J. J. Park, T. Ali, and S. Lee, "Artificial Intelligence for the Diagnosis of Heart Failure," *NPJ Digit Med*, vol. 3, no. 1, 2020, doi: 10.1038/s41746-020-0261-3.

[9] S. N. Zisad, E. Chowdhury, M. S. Hossain, R. U. Islam, and K. Andersson, "An Integrated Deep Learning and Belief Rule-Based Expert System for Visual Sentiment Analysis Under Uncertainty," *Algorithms*, vol. 14, no. 7, p. 213, 2021, doi: 10.3390/a14070213.

[10] N. Zoanetti and J. Pearce, "The Potential Use of Bayesian Networks to Support Committee Decisions in Programmatic Assessment," *Med Educ*, vol. 55, no. 7, pp. 808–817, 2020, doi: 10.1111/medu.14407.

[11] F. Xie *et al.*, "Development and Assessment of an Interpretable Machine Learning Triage Tool for Estimating Mortality After Emergency Admissions," *JAMA Netw Open*, vol. 4, no. 8, p. e2118467, 2021, doi: 10.1001/jamanetworkopen.2021.18467.

[12] Y. Miyachi, O. Ishii, and K. Torigoe, “Design, Implementation, and Evaluation of the ComputerAided Clinical Decision Support System Based on Learning-to-Rank: Collaboration Between Physicians and Machine Learning in the Differential Diagnosis Process,” 2022, doi: 10.21203/rs.3.rs-2068201/v2.

[13] S. Amin *et al.*, “Developing and Demonstrating the Viability and Availability of the Multilevel Implementation Strategy for Syncope Optimal Care Through Engagement (MISSION) Syncope App: Evidence-Based Clinical Decision Support Tool,” *J Med Internet Res*, vol. 23, no. 11, p. e25192, 2021, doi: 10.2196/25192.

[14] Z. Baratz and Y. Assaf, “Optimizing FreeSurfer’s Surface Reconstruction Parameters for Anatomical Feature Estimation,” 2023, doi: 10.1101/2023.01.02.522457.

[15] E. R. Batubara, M. Zarlis, and R. Rosnelly, “Hyperteroid Disease Analysis With Backpropagation Artificial Neural Network,” *International Conference on Information Science and Technology Innovation (Icostec)*, vol. 2, no. 1, pp. 59–64, 2023, doi: 10.35842/icostec.v2i1.38.

[16] S. Xie, K. Li, M. Xiao, L. Zhang, and W. Li, “Key Quality Indicators Prediction for Web Browsing With Embedded Filter Feature Selection,” *Applied Sciences*, vol. 10, no. 6, p. 2141, 2020, doi: 10.3390/app10062141.

[17] D. Peng *et al.*, “Pose Determination of the Disc Cutter Holder of Shield Machine Based on Monocular Vision,” *Sensors*, vol. 22, no. 2, p. 467, 2022, doi: 10.3390/s22020467.

[18] H. Tu *et al.*, “Improving Lung Cancer Risk Prediction Using Machine Learning: A Comparative Analysis of Stacking Models and Traditional Approaches,” *Cancers (Basel)*, vol. 17, no. 10, p. 1651, 2025, doi: 10.3390/cancers17101651.

[19] A. M. Madni and A. T. Bahill, “Handling Uncertainty in Engineered Systems,” pp. 1–34, 2020, doi: 10.1007/978-981-13-0370-8_22-1.

[20] S. Supardi, C. Kirana, and F. Ferdian, “Expert System for Early Childhood Talent Detection Using Certainty Factor and Dempster Shafer Algorithms,” *Journal of Information Systems and Informatics*, vol. 7, no. 3, pp. 3025–3045, 2025, doi: 10.51519/journalisi.v7i3.1283.

[21] M. Aldjawad, S. Andryana, and A. Andrianingsih, “Penerapan Metode Perbandingan DempsterShafer Dengan Certainty Factor Pada Aplikasi Sistem Pakar Deteksi Dini Penyakit Alzheimer Pada Lansia Berbasis Web,” *Jurnal Jtik (Jurnal Teknologi Informasi Dan Komunikasi)*, vol. 5, no. 2, pp. 144–152, 2021, doi: 10.35870/jtik.v5i2.206.