

Intelligent Fuzzy Logic Models for Predicting Post-Operative Cardiac Complications

Varsha¹, Prof. Diwari Lal²

Research Scholar, Department of Mathematics, Narain (PG) College, Shikohabad, U.P. (India)¹

Professor, Department of Mathematics, Narain (PG) College, Shikohabad, U.P. (India)²

Affiliated to Dr. Bhimrao Ambedkar University, Agra (India)^{1,2}

Abstract: Cardiac complications are one of the biggest concerns after a patient has undergone cardiac surgery. It is therefore important to have predictive tools or models that can give both a reliable estimate of the patient's overall risk as well as a measure of how reliably this model predicts their individual level of risk. The purpose of this research was to develop an evidence-conserving selective fuzzy inference system (EC-SFIS) in order to predict the occurrence of serious cardiac complications postoperatively and postoperative atrial fibrillation from clinical input values including age, left ventricular ejection fraction (LVEF), serum creatinine, and duration of cardiopulmonary bypass (CPB). Unlike conventional fuzzy inference, the proposed framework decomposes rule-based information into supporting, opposing, and unresolved evidence and introduces a conflict-conservation operator that transfers contradictory evidence into an unresolved component while preserving the directional evidence balance. Evidence-ratio defuzzification, probability calibration, and a selective prediction mechanism are subsequently employed to accept predictions with adequate evidence resolution and low conflict while flagging uncertain cases for additional assessment. A numerical case study demonstrates that the framework can distinguish between predictions having similar evidence resolution but different conflict levels, resulting in outcome-specific acceptance or flagging decisions. The comparative evaluation presented in the study indicates lower prediction error across risk-coverage levels and favorable probability calibration relative to the considered fuzzy and statistical comparators. The proposed EC-SFIS therefore provides an interpretable mathematical framework in which postoperative risk prediction is accompanied by explicit quantification of evidence consistency and prediction reliability.

Keywords: Evidence-Conserving Fuzzy Inference System, fuzzy logic; postoperative cardiac complications, atrial fibrillation; evidence conflict, selective prediction; uncertainty quantification, risk prediction, clinical decision support.

I. INTRODUCTION

Postoperative cardiovascular complications have a number of clinical predictive variables that interact with each other in very complex manners. Models, as well as many conventional models, provide risk scores; however they do not provide insight into potential conflicting evidence. Therefore this study is proposing an Evidence Conserving Selective Fuzzy Inference System (EC-SFIS) for predicting the occurrence of major postoperative cardiac complications and postoperative atrial fibrillation based on four predictor variables (age, LVEF, serum creatinine, and CPB time). The model has three levels of evidence: supporting, opposing, or unresolved. Additionally, the model can quantify conflict among rules and produce a continuous risk score. Predictions from the model will be reliable when there is sufficient confidence in their accuracy and uncertain predictions will be identified for additional evaluation.

Devereaux and Sessler (2015) therefore, point out the necessity of perioperative risk detection. Additionally, they mention the benefit of improved monitoring during surgery. Biomarkers are regarded as being suitable for the detection of myocardial damage. The authors' study therefore provides evidence of the necessity of improved postoperative evaluation of cardiac risk. Lewicki *et al.* (2016) have studied postoperative atrial fibrillation in patients who underwent coronary artery bypass grafting. In their study they have compared the incidence of postoperative atrial fibrillation in patients undergoing on-pump versus off-pump coronary artery bypass grafting. Postoperative atrial fibrillation developed in 18.3 % and 19.3 %, respectively. There was no statistically significant difference between the two groups.

Steyerberg and Harrell (2016) pointed out the need for rigorous validation of clinical predictive models. They differentiated among internal validation, internal-external validation and external validation. The performance of a model can only be determined by using a dataset other than the one which was used to develop the model. They stated that there is an important implication for developing intelligent clinical predictive models. That is, clinical predictive models should be independently validated prior to clinical use. This principle applies to EC-SFIS as well. Cabitza *et al.* (2017) presented some very important issues concerning the application of machine learning in medicine. While machine learning can provide very accurate predictions regarding certain outcomes such as disease progression, etc.,

these predictions do not necessarily translate into clinically useful knowledge. Machine learning algorithms may create unintended side effects, especially if large datasets are used to train them. Another concern is the lack of interpretability which could lead to reduced confidence in physicians using machine-learning-based predictive tools. Therefore, it seems advisable to use transparent predictive frameworks such as EC-SFIS, because they allow clinicians to understand how evidence was decomposed. **Martínez-Sellés *et al.* (2017)** provided insight into the difficulty faced by healthcare professionals when assessing elderly patients who are at increased risk for atrial fibrillation. They emphasized that when atrial fibrillation is suspected, but not confirmed, this situation leads to uncertainty in cardiovascular assessments. They illustrated that while risk assessment and diagnostic certainty are related, they are not interchangeable terms.

Gaudino *et al.* (2021) demonstrated the clinical relevance of preventing postoperative atrial fibrillation following cardiac surgery. They supported the idea that reducing the occurrence of postoperative atrial fibrillation is an important goal of all aspects of perioperative care. Identifying patients at greatest risk for postoperative atrial fibrillation will likely aid in achieving this goal. As a result, reliable prediction of susceptibility to postoperative atrial fibrillation is beneficial. **Stoger *et al.* (2021)** evaluated the legal aspects of medical artificial intelligence. Stoger *et al.* (2021), pointed out several problems inherent in implementing clinical AI solutions. One problem mentioned was accountability in automated medical decisions. Another problem mentioned was transparency. These considerations favor interpretable models over black box decision systems. **Onishchenko *et al.* (2022)** employed machine learning methods for predicting postoperative major adverse cardiac events based on routinely collected comorbidity information. They noted that their method did not add any additional data burdens to the usual pre-surgical evaluations. The work demonstrates the potential of machine learning for perioperative cardiac-risk stratification. It also shows the value of clinically accessible predictors. **Wu *et al.* (2023)** developed a machine-learning model for older patients undergoing noncardiac surgery. The model targeted major adverse cardiovascular and cerebrovascular events. Their study reflects increasing use of artificial intelligence in perioperative risk prediction. Older surgical patients represent a particularly important risk group. The findings further support individualized cardiac-risk estimation. **Choi *et al.* (2024)** used machine learning with serial 12-lead ECG recordings. Their approach focused on new-onset atrial fibrillation. Serial information improved prediction compared with a single ECG. The serial model achieved an AUROC of 0.880. The single-ECG model achieved 0.812. This demonstrates the ability of machine learning to identify subtle pre-AF changes. However, the method requires extensive ECG information. The present EC-SFIS instead emphasizes a compact and interpretable clinical input structure. **Sanaiha *et al.* (2024)** developed a machine-learning approach for perioperative myocardial infarction and cardiac arrest. They compared the new approach with an established clinical risk score. Their work addressed limitations in traditional risk estimation. Calibration was an important concern. The study demonstrates the potential of machine learning for serious perioperative cardiac events. It also reinforces the need to evaluate more than discrimination alone. **Im (2025)** discussed links among iron deficiency, arrhythmias, and cardiovascular disease. The work illustrates the multifactorial nature of cardiac rhythm disturbances. Metabolic and systemic factors may modify arrhythmia susceptibility. Such complexity creates uncertainty in patient-level risk assessment. Fuzzy reasoning is well suited to representing gradual and overlapping risk states. **Kalishnik *et al.* (2025)** investigated retained blood-associated complications following cardiac surgery. Their work demonstrates that postoperative outcomes can be influenced by procedure-related factors. Postoperative risk is therefore not determined by baseline characteristics alone. Surgical and perioperative variables should also be considered. The inclusion of CPB duration in EC-SFIS follows this broader perioperative perspective. **Yu *et al.* (2025)** developed an interpretable machine-learning model for major adverse cardiovascular events in geriatric surgical patients. Their prospective approach highlights growing interest in interpretable perioperative prediction. Interpretability is especially important for clinical decision support. It allows clinicians to understand factors influencing model output. EC-SFIS extends this principle by separating supporting, opposing, and unresolved evidence. **Dhaliwal *et al.* (2026)** provided recent systematic evidence on machine-learning prediction of postoperative cardiac complications. They evaluated 13 studies. These studies contained 54 prediction models and 927,113 patients. Machine-learning approaches generally outperformed traditional risk scores. Automated machine learning achieved the highest ranking in their analysis.

II. CLINICAL PROBLEM FORMULATION

The proposed population for this study will include adults who undergo cardiac surgery that includes a cardiopulmonary bypass. The prediction of the model will occur at time point 1 which is immediately after the surgical procedure and before the measurement of post-surgical outcomes as defined in the protocol.

Table 1. Proposed clinical variables			
Type	Variable	Symbol	Proposed domain
Input 1	Patient age	x_1	30–90 years
Input 2	Left ventricular ejection fraction	x_2	20–70%
Input 3	Serum creatinine	x_3	0.5–3.0 mg/dL
Input 4	Cardiopulmonary bypass duration	x_4	0–240 minutes
Output 1	Major postoperative cardiac complications	y_1	Binary clinical endpoint
Output 2	New-onset postoperative atrial fibrillation	y_2	Binary clinical endpoint

For consistency in methods, for your first outcome, consider defining that initial endpoint as a pre-specified composite of events that do NOT include AF. Your second outcome would be an independently adjudicated onset of post-operative new atrial fibrillation. Each outcome should also have a well-defined time frame (i.e., 7 days after surgery, at discharge; which ever is sooner) for assessment. The exact definitions must be finalized before collecting or analysing data. The fuzzy system will produce two continuous prediction scores corresponding to these clinical endpoints. A separate calibration procedure will convert the scores into estimated probabilities.

Figure 1. Proposed EC-SFIS architecture

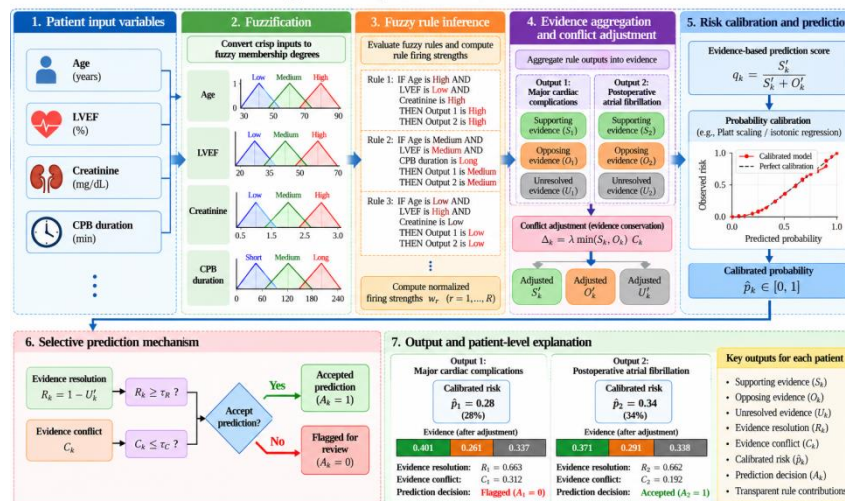


Figure 1 illustrates the complete architecture of the proposed Evidence-Conserving Selective Fuzzy Inference System (EC-SFIS) for predicting postoperative cardiac complications. The framework begins with four clinical inputs age, LVEF, serum creatinine, and cardiopulmonary bypass (CPB) duration which are transformed into linguistic membership degrees during fuzzification. These memberships activate the fuzzy rule base, and the resulting rule strengths are aggregated separately for two outcomes: major postoperative cardiac complications and postoperative atrial fibrillation. Unlike a conventional fuzzy system, EC-SFIS decomposes the inference into supporting (S_k), opposing (O_k), and unresolved (U_k) evidence, after which the proposed conflict-conservation operator transfers conflicting evidence into the unresolved component. An evidence-based score is subsequently obtained and calibrated to produce an outcome-specific risk estimate. Finally, the selective prediction mechanism evaluates the evidence resolution (R_k) and conflict (C_k) against predefined thresholds; sufficiently resolved, low-conflict predictions are accepted, whereas predictions with excessive conflict or inadequate resolution are flagged for additional assessment. Thus, the architecture provides not only two postoperative risk predictions but also an interpretable explanation of the evidence and the reliability of each prediction.

III. PROPOSED MATHEMATICAL THEORY

The principal mathematical contribution is a fuzzy evidence representation that keeps three quantities separate until the final decision stage.

3.1 Membership-function design: For each input, use three overlapping linguistic categories. These parameters are initial mathematical assumptions, not validated clinical cutoffs.

Table 2. Initial membership-function specifications		
Input	Linguistic categories	Membership functions
Age	Younger, Intermediate, Older	Trapezoidal (30,30,40,55); triangular (45,60,75); trapezoidal (65,78,90,90)
LVEF	Reduced, Intermediate, Preserved	Trapezoidal (20,20,30,40); triangular (35,45,55); trapezoidal (50,60,70,70)
Creatinine	Lower, Intermediate, Higher	Trapezoidal (0.5,0.5,0.9,1.3); triangular (1.1,1.6,2.1); trapezoidal (1.8,2.3,3,3)
CPB duration	Shorter, Intermediate, Longer	Trapezoidal (0,0,45,90); triangular (60,110,160); trapezoidal (130,180,240,240)

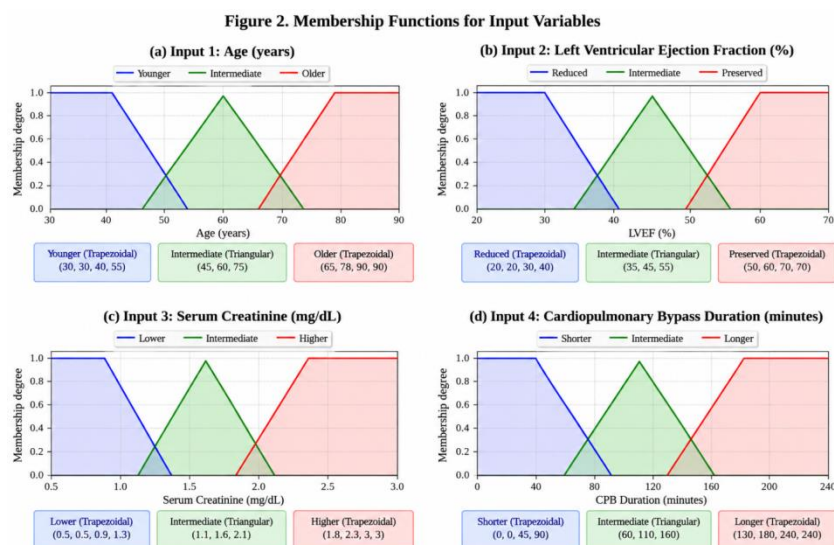
Define the triangular function as

$$\mu_T(x; a, b, c) = \max \left[0, \min \left(\frac{x-a}{b-a}, \frac{c-x}{c-b} \right) \right] \quad (1)$$

For trapezoidal membership functions, use

$$\mu_Z(x; a, b, c) = \max \left[0, \min \left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c} \right) \right] \quad (2)$$

At repeated endpoints, use shoulder-function conventions to avoid division by zero.



The fuzzy membership function illustrated in figure 2, represents the conversion of the four crisp clinical data inputs (Age, Left Ventricular Ejection Fraction (LVEF), Serum Creatinine (SC) and Cardiopulmonary Bypass (CPB)) into four different fuzzy clinical states (Younger, Intermediate & Older for age; Reduced, Intermediate & Preserved for LVEF; Lower, Intermediate & Higher for SC; Shorter, Intermediate & Longer for CPB). The use of trapezoidal membership functions to represent the boundary states and triangular to represent the intermediate states ensures that when a patient's measurement lies at or near a clinical boundary, it can be associated with both states simultaneously.

This avoids the clinical discontinuity present in many other models which have distinct boundaries separating one state from another. Each input is converted to an output membership degree between 0 and 1 where a higher number signifies a greater association with its respective clinical state. Once these fuzzy outputs are calculated they will serve as the inputs for determining how active each rule will be during the EC-SFIS inference process.

3.2 A new evidence representation: Instead of assigning every rule a single output class, assign each rule an evidence triplet:

$$e_{rk} = (s_{rk}, o_{rk}, u_{rk}) \quad (3)$$

subject to

$$s_{rk} + o_{rk} + u_{rk} = 1 \quad (4)$$

Here:

s_{rk} : evidence supporting complication k .

o_{rk} : evidence opposing complication k .

u_{rk} : unresolved evidence.

All three components belong to $[0,1]$.

For example, the triplet $(0.8, 0.1, 0.1)$ represents predominantly supporting evidence, whereas $(0.1, 0.1, 0.8)$ represents substantial unresolved evidence. These triplets are mathematical evidence weights, not observed event probabilities. The triplet representation is related to established evidential and subjective-logic approaches. The research must therefore focus on developing and testing a distinct fuzzy inference and aggregation procedure rather than claiming that three-component evidence itself is new.

3.3 Fuzzification and rule activation: For patient input vector

$$x = (x_1, x_2, x_3, x_4) \quad (5)$$

Calculate the firing strength of rule r :

$$\alpha_r = \prod_{j=1}^4 \mu_{A_{rj}}(x_j) \quad (6)$$

Define normalized rule activation:

$$w_r = \frac{\alpha_r}{\sum_{j=1}^R \alpha_j} \quad (7)$$

The denominator must be positive. Complete input-domain coverage should be checked during model development.

With three linguistic categories per input, the complete rule base contains

$$R = 3^4 = 81 \text{ possible antecedent combinations.} \quad (8)$$

However, unlike a traditional 81-rule Mamdani system, the consequent structure of this system is fundamentally different in that each rule carries its own unique evidence triplets for both outcome possibilities.

Table 3. Representative evidence-based fuzzy rules			
Rule	Antecedent description	Outcome 1 triplet	Outcome 2 triplet
R1	Younger, preserved LVEF, lower creatinine, shorter CPB	(0.10, 0.80, 0.10)	(0.15, 0.75, 0.10)
R2	Older, reduced LVEF, higher creatinine, longer CPB	(0.85, 0.05, 0.10)	(0.75, 0.10, 0.15)

R3	Older, preserved LVEF, lower creatinine, intermediate CPB	(0.35, 0.45, 0.20)	(0.55, 0.25, 0.20)
R4	Intermediate age, reduced LVEF, intermediate creatinine, longer CPB	(0.70, 0.15, 0.15)	(0.45, 0.35, 0.20)

IV. ORIGINAL ALGORITHM: EVIDENCE-CONSERVING AGGREGATION

The area to which we have devoted most of our mathematical analysis. Standard weighted averages do not provide a unique intermediate forecast when there are both large amounts of supportive as well as disconfirming evidence. The method for aggregating evidence developed here provides an explicit quantification of how much evidence supports or rejects an hypothesis.

4.1 Initial evidence aggregation: For outcome k , calculate

$$S_k = \sum_{r=1}^R w_r s_{rk} \quad (9)$$

$$O_k = \sum_{r=1}^R w_r o_{rk} \quad (10)$$

$$U_k = \sum_{r=1}^R w_r u_{rk} \quad (11)$$

$$\text{By construction, } s_{rk} + o_{rk} + u_{rk} = 1 \quad (12)$$

4.2 Proposed conflict-conservation operator: Define the disagreement between two active rules:

$$d_{rs,k} = s_{rk} o_{sk} + o_{rk} s_{sk} \quad (13)$$

This quantity becomes large when one rule supports an outcome while another opposes it.

The weighted pairwise conflict is

$$C_k = 2 \sum_{r < s} w_r w_s d_{rs,k} \quad (14)$$

The factor of two accounts for both ordered rule pairs. The quantity is bounded between zero and one for valid evidence triplets and normalized weights.

Define the transferable conflict mass:

$$\Delta_k = \lambda \min(S_k, O_k) C_k \quad (15)$$

The proposed conflict-conservation operator is

$$\begin{cases} S'_k = S_k - \Delta_k \\ O'_k = O_k - \Delta_k \\ U'_k = U_k - \Delta_k \end{cases} \quad (16)$$

The operator transfers equal quantities of supporting and opposing evidence into unresolved evidence when disagreement is detected.

4.3 Mathematical properties to prove: A theory-focused paper should include formal proofs of the following properties.

Proposition 1: Evidence conservation

$$S'_k + O'_k + U'_k = 1 \quad (17)$$

This follows immediately from Equation (16), since the two subtracted quantities are added to unresolved evidence.

Proposition 2: Non-negativity

For valid evidence triplets and $(0 \leq \lambda \leq 1)$,

$$S'_k, O'_k, U'_k \geq 0 \quad (18)$$

Since $\Delta_k \leq \min(S_k, O_k)$

neither supporting nor opposing evidence becomes negative. The conservation property then ensures that the updated triplet remains normalized.

Proposition 3: Direction preservation

$$S'_k - O'_k = S_k - O_k \quad (19)$$

Consequently, conflict adjustment does not reverse the direction of the original evidence balance. Instead, it reduces the amount of resolved evidence available for making a prediction.

Proposition 4: Zero-conflict consistency

$$\text{If } C_k \geq 0, \text{ then } (S'_k, O'_k, U'_k) = (S_k, O_k, U_k) \quad (20)$$

Thus, the operator does not modify aggregated evidence when the defined pairwise conflict is absent.

These propositions establish basic mathematical consistency. They do not, by themselves, establish novelty or clinical superiority.

Figure 3. Geometrical Representation of the Evidence-Conservation Operator using a Ternary Evidence Diagram

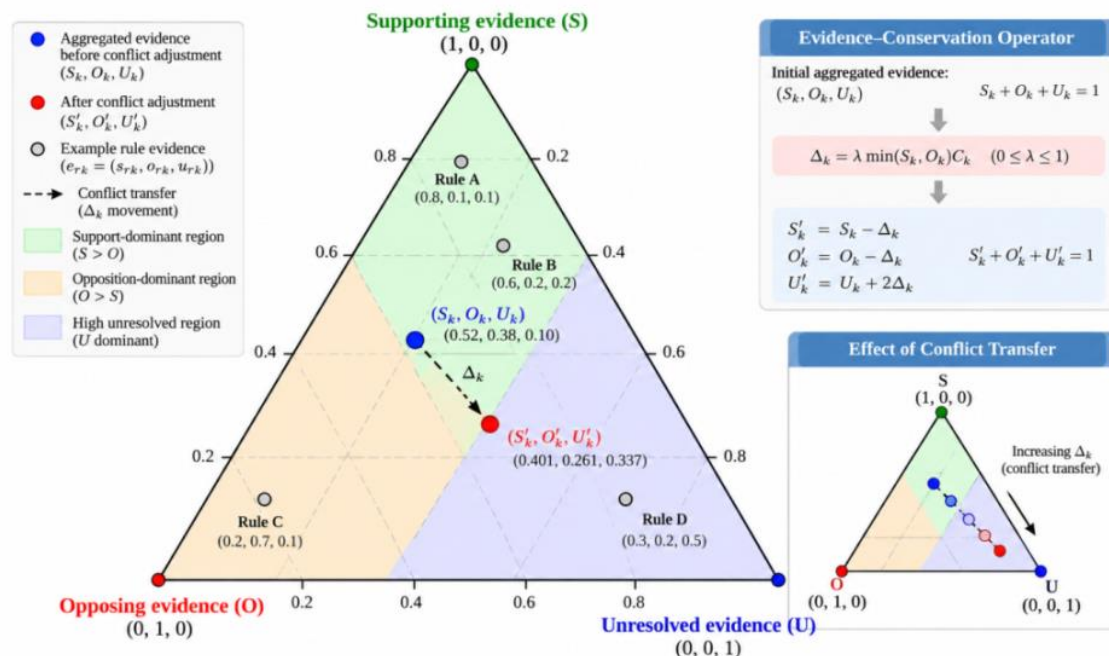
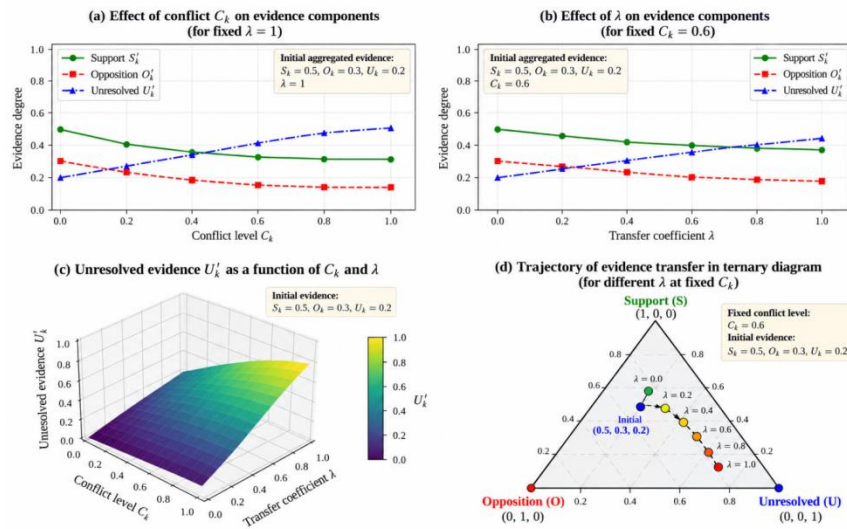


Figure 3 shows geometrically that the proposed Evidence-Conservation Operator uses a triangle on a ternary graph with three components: supporting evidence S_k , opposing evidence O_k , and unresolved evidence U_k , where $S_k + O_k + U_k = 1$. The area of the triangle is divided into three areas; one for cases when there is more supporting than opposing (supporting dominant), one when there is more opposing than supporting (opposing dominant), and one in which the amount of resolved evidence is greater than either supporting or opposing (unresolved). Each point in the triangle represents different fuzzy rule combinations of evidence. The blue dot represents the aggregated evidence before it was modified by adjusting conflicts (0.52,0.38,0.10); the red dot represents the aggregated evidence after adjustments were made for conflicts (0.401,0.261,0.337); the dashed line from the blue to the red dots show how the operator adjusts both the supporting and opposing evidence equally as much as possible to place them into the unresolved category based upon $\lambda \min(S_k, O_k)C_k$. Therefore, the adjusted evidence falls closer to the unresolved category while maintaining both the overall evidence and the differential of supporting versus opposing evidence. This visual representation of EC-SFIS illustrates how the method takes into account conflicting pieces of information and does so without eliminating any of those pieces of information nor altering their original direction of influence.

Figure 4. Conflict-Transfer Behavior for Different Values of C_k and λ



The Evidence-Conservation Operator's Conflict Transfer Behavior is illustrated by figure 4 in relation to two key factors; the amount of conflicting evidence C_k , and the parameter which influences the strength of the transfer from one piece of evidence to another, λ . Panel A represents the relationship between the supporting S'_k and opposing O'_k evidence, and the amount of conflicting evidence C_k when $\lambda = 1$. It can be seen that as the conflicting evidence increases, both S'_k and O'_k decrease, while the amount of unresolved evidence U'_k increases. Panel B illustrates the same trend as λ increases when the conflicting evidence remains constant. This illustrates that it is the λ factor which determines the extent to which the evidence is transferred. In panel C, a 3D graph illustrating the effect of varying both C_k and λ on the amount of unresolved evidence, U'_k , is presented. As may be expected, the greatest increase in U'_k occurs when both C_k and λ are at their highest levels. Finally, in panel D, a geometric representation of the transformation of the evidence distribution due to an increase in λ is shown. While λ has increased resulting in an increased proportion of evidence being classified as resolved but undecided, the overall directional balance of the supporting and opposing evidence remains unchanged.

V. DEFUZZIFICATION AND SELECTIVE PREDICTION

The proposed framework should distinguish three quantities: the directional risk score, the amount of unresolved evidence and the strength of disagreement.

5.1 Evidence-based defuzzification: Define the resolved-evidence score:

$$q_k = \frac{S'_k}{S'_k + O'_k} \quad (21)$$

This produces a continuous score between zero and one, provided that some resolved evidence remains.

If $(S'_k + O'_k) = 0$, the system must abstain rather than assign an arbitrary value.

Unlike centroid defuzzification, equation (21) converts the evidence representation directly into a numerical score. It should therefore be described as evidence-ratio defuzzification. An important mathematical consequence of Proposition 3 is that the proposed conflict operator preserves the supporting-versus-opposing difference but changes the resolved-evidence ratio. This behavior should be investigated rather than assumed to improve prediction.

5.2 Selective prediction

Define the evidence-resolution coefficient:

$$R_k = S'_k + O'_k = 1 - U'_k \quad (22)$$

Introduce a selective prediction function:

$$A_k = \begin{cases} 1 & R_k \geq \tau_R \text{ and } C_k \leq \tau_C \\ 0 & \text{Otherwise} \end{cases} \quad (23)$$

Here, $A_k = 1$ indicates that the model issues a prediction, while $A_k = 0$ indicates that the prediction should be flagged for additional assessment. The thresholds τ_R and τ_C must be determined using development data and evaluated independently. The two outcomes receive separate acceptance decisions. The system may accept the prediction for one complication while flagging the other.

Figure 5. Proposed Selective Prediction Mechanism Showing Accepted and Flagged Predictions

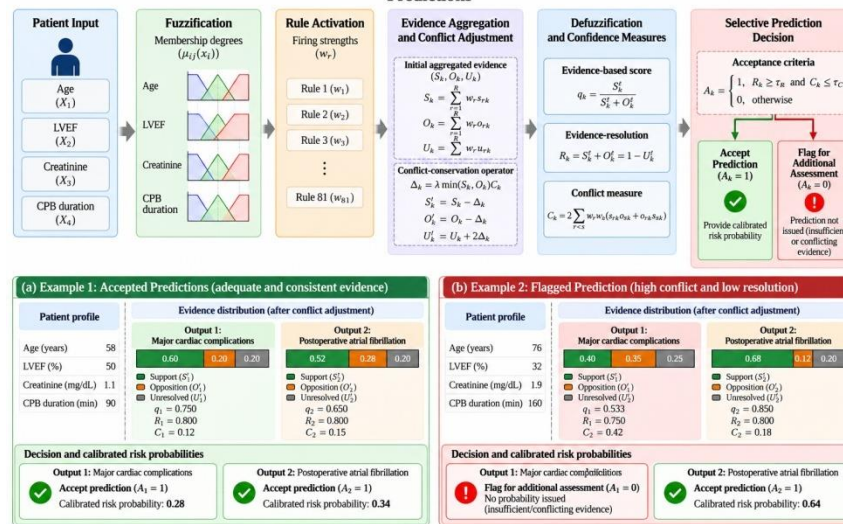


Figure 5 illustrates the selective prediction mechanism of the proposed EC-SFIS, beginning with four patient inputs age, LVEF, serum creatinine, and CPB duration followed by fuzzification, fuzzy-rule activation, evidence aggregation, conflict adjustment, and evidence-based defuzzification. For each postoperative outcome, the model calculates the prediction score q_k , evidence resolution R_k , and conflict measure C_k . A prediction is accepted only when the resolution satisfies $R_k \geq \tau_R$ and the conflict remains below $C_k \geq \tau_C$; otherwise, it is flagged for additional assessment rather than forcing a potentially unreliable risk estimate. The illustrations of these two cases illustrate this difference: Sufficiently resolved and consistent information will result in acceptable predicted outcomes and calibrated risk probability values for those outcomes; Inadequate or conflicting information about a case will cause that case's outcome to be identified as having an unreliable predictive value.

VI. COMPLETE NUMERICAL DEMONSTRATION OF THE PROPOSED OPERATOR

Consider an illustrative patient whose fuzzified inputs activate two dominant rules. For transparency, assume that their normalized firing strengths are

$$w_1 = 0.60, w_2 = 0.40 \quad (24)$$

The values below are hypothetical rule consequents selected to demonstrate the mathematics. They are not derived from clinical observations.

Table 4. Numerical case-study inputs		
Parameter	Rule 1	Rule 2
Normalized firing strength	0.6	0.4
Outcome 1 support	0.8	0.1
Outcome 1 opposition	0.1	0.8
Outcome 1 unresolved	0.1	0.1
Outcome 2 support	0.6	0.2
Outcome 2 opposition	0.2	0.6
Outcome 2 unresolved	0.2	0.2

Inputs for the case studies presented in Table 4 were used as an example to illustrate the mechanisms of aggregating evidence and resolving conflicts when using the proposed EC-SFIS. In this study, two main fuzzy rules were identified with weights of 0.60 and 0.40. The first rule has a high level of supporting evidence (0.80) but low level of conflicting evidence (0.10). Conversely, the second rule had a low level of supporting evidence (0.10) and a high level of conflicting evidence (0.80); thus, these two rules have significant amounts of conflicting evidence. For the second output, Rule 1 supports the occurrence of the outcome at a moderate weight (0.60) whereas Rule 2 opposes the occurrence of the outcome at a moderate weight (0.60) resulting in less conflict than the previous example. The unresolvable evidence is information that could not be determined whether it supported or opposed either outcome. These weights will be used to calculate the aggregate supporting and opposing evidence, the conflict value, and then the adjusted amount of evidence conserved after surgery for each of the postoperative outcomes.

6.1 Numerical calculation for output 1

Calculate the aggregated evidence:

$$S_1 = 0.6(0.8) + 0.4(0.1) = 0.52 \quad (25)$$

$$O_1 = 0.6(0.1) + 0.4(0.8) = 0.38 \quad (26)$$

$$U_1 = 0.10 \quad (27)$$

The pairwise disagreement is

$$d_{12,1} = 0.8(0.8) + 0.1(0.1) = 0.65 \quad (28)$$

$$\text{Therefore, } C_1 = 2(0.6)(0.4)(0.65) = 0.312 \quad (29)$$

For the illustrative setting $\lambda = 1$,

$$\Delta_1 = 0.312 \min(0.52, 0.38) = 0.11856 \quad (30)$$

Applying the proposed operator gives

$$\begin{cases} S'_1 = 0.40144 \\ O'_1 = 0.26144 \\ U'_1 = 0.33712 \end{cases} \quad (31)$$

The final evidence-based score is

$$q_1 = \frac{0.40144}{0.40144 + 0.26144} \approx 0.606 \quad (32)$$

The evidence-resolution coefficient is

$$R_1 = 1 - 0.33712 = 0.66288 \quad (33)$$

6.2 Numerical calculation for output 2

Similarly,

$$S_2 = 0.6(0.6) + 0.4(0.2) = 0.44 \quad (34)$$

$$O_2 = 0.6(0.2) + 0.4(0.6) = 0.36 \quad (35)$$

$$U_2 = 0.20 \quad (36)$$

The pair wise disagreement is

$$d_{12,2} = 0.6(0.6) + 0.2(0.2) = 0.40 \quad (37)$$

$$\text{Therefore, } C_2 = 2(0.6)(0.4)(0.40) = 0.192 \quad (38)$$

$$\Delta_2 = 0.192(0.36) = 0.06912 \quad (39)$$

Applying the proposed operator gives

$$\begin{cases} S'_2 = 0.37088 \\ O'_2 = 0.29088 \\ U'_2 = 0.33824 \end{cases} \quad (40)$$

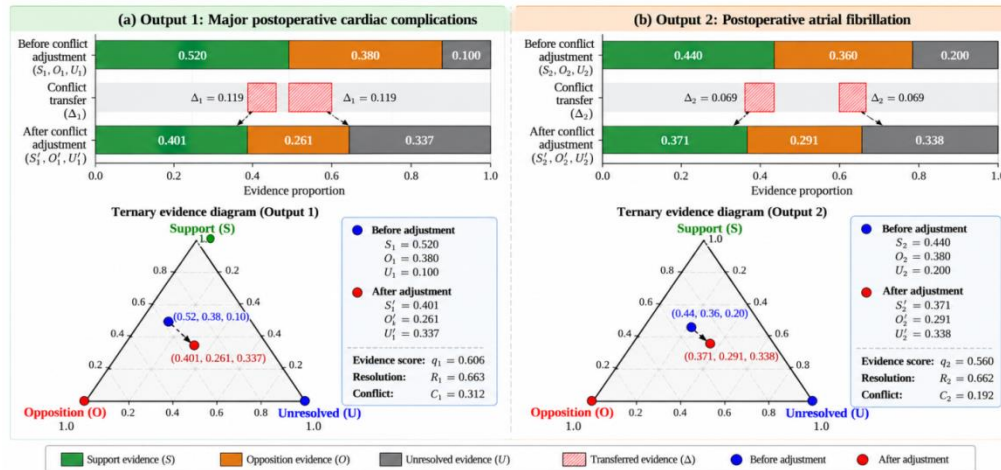
The final evidence-based score is

$$q_2 = \frac{0.37088}{0.37088+0.29088} \approx 0.560 \quad (41)$$

The evidence-resolution coefficient is

$$R_2 = 0.66176 \quad (42)$$

Figure 6. Illustrative Evidence Decomposition



In fig. 6 we see how the fuzzy evidence is decomposed and redistributed when predicting the two possible post-operative outcomes. Initially, the fuzzy evidence was as follows. Major post-operative cardiac complications supporting evidence = 0.520 opposing evidence = 0.380 unresolved evidence = .100 after conflict adjustment, the fuzzy evidence becomes supporting evidence = 0.401 opposing evidence = 0.261 unresolved evidence = 0.337 similarly, for post-operative atrial fibrillation supporting evidence = .440 opposing evidence = 0.360 unresolved evidence = 0.200 after conflict adjustment, the fuzzy evidence becomes supporting evidence = .371 opposing evidence = 0.291 unresolved evidence = .338 these transformations can also be seen in the ternary diagrams which show that there has been a shift in the fuzzy evidence away from supporting or opposing and towards an increased amount of unresolved fuzzy evidence; however, the balance between supportive and opposing fuzzy evidence remains preserved. The resultant fuzzy evidence scores are $q_1 = 0.606$ and $q_2 = 0.560$, with associated conflict scores $c_1 = 0.312$ and $c_2 = 0.192$, respectively. Overall, the figure shows how the EC-sfis method explicitly quantifies and represents the uncertainty created due to conflicting rule-based evidence rather than hiding it within a single predictive score.

6.3 Demonstration of selective prediction: Suppose the development-stage selection thresholds are provisionally specified as

$$\tau_R = 0.65, \tau_C = 0.25 \quad (43)$$

These are illustrative thresholds, not validated clinical decision criteria.

Table 5. Numerical prediction decisions		
Parameter	Output 1	Output 2
Evidence score	0.606	0.56
Conflict	0.312	0.192
Resolution	0.663	0.662
Resolution criterion	Satisfied	Satisfied
Conflict criterion	Not satisfied	Satisfied
Final decision	Flag	Accept

The purpose of this example is to show how there is a difference in what the model does in each case. Even though the coefficient of determination of the first prediction is the same as the coefficient of determination of the second prediction, the model will flag the first prediction based on contradictory evidence whereas it will accept the second prediction. Table 5 outlines the numerical predictive conclusions that were made using the proposed EC-SFIS for each of the two post-surgical outcome measures. For Output 1 (major post-operative cardiac complications), the evidence score was 0.606 and the resolution measure was 0.663 and the conflict measure was 0.312. Since the resolution measure satisfies the resolution criterion, and since the conflict measure is greater than the allowable conflict limit, the conflict criterion is not met, and therefore the prediction is indicated for further clinical review. On the other hand, Output 2 (post-operative atrial fibrillation) produced an evidence score of 0.560 and a resolution measure of 0.662, and a smaller conflict measure of 0.192. Both of these measures meet the resolution and conflict criteria and thus produce an accepted prediction. Therefore, one of the most important characteristics of EC-SFIS shown in Table 5 is that a prediction's conclusion is not just dependent upon the magnitude of its evidence score, but also upon the reliability and consistency of the fuzzy evidence used to generate the prediction. Thus, instead of all predictions being automatically accepted if they have sufficient evidence scores, potentially ambiguous predictions may be recognized in order to allow them to be subjected to some form of clinical validation.

VII. CLINICAL VALIDATION AND EXPERIMENTAL DESIGN

The most important experimental question is not simply whether EC-SFIS achieves a high AUROC. It is whether its explicit conflict representation identifies predictions that are genuinely less reliable. The following experimental structure would test that claim.

7.1 Comparative experiments

Table 6. Proposed model-comparison and ablation experiments		
Experiment	Model configuration	Question investigated
E1	Conventional Mamdani fuzzy system	Standard fuzzy baseline
E2	Conventional Sugeno fuzzy system	Numerical fuzzy baseline
E3	Evidential fuzzy system without conflict adjustment	Value of evidence triplets
E4	EC-SFIS without selective prediction	Value of conflict adjustment
E5	Complete EC-SFIS	Value of selective prediction
E6	Logistic regression	Statistical comparator
E7	Gradient-boosted trees	Machine-learning comparator

Ablation experiments should use the same clinical dataset, preprocessing and data partitions. All model-selection decisions must be made without accessing the final test dataset.

7.2 Risk-coverage analysis: Selective prediction introduces an additional performance dimension: the proportion of patients for whom the model accepts its own prediction.

Define coverage as

$$Coverage_k = \frac{1}{N} \sum_{i=1}^N A_{ik} \quad (44)$$

Define selective prediction error as

$$Risk_k = \frac{\sum_{i=1}^N A_{ik} \ell(\hat{p}_{ik}, t_{ik})}{\sum_{i=1}^N A_{ik}} \quad (45)$$

Here, ℓ is an appropriate prediction-loss function, $\hat{p}_{ik}$ is the calibrated probability and t_{ik} is the observed clinical outcome. The denominator must be positive. A clinically useful selective prediction system should demonstrate that its accepted predictions are reliable without flagging an impractically large proportion of patients.

Figure 7. Risk–Coverage Curves for Both Postoperative Outcomes

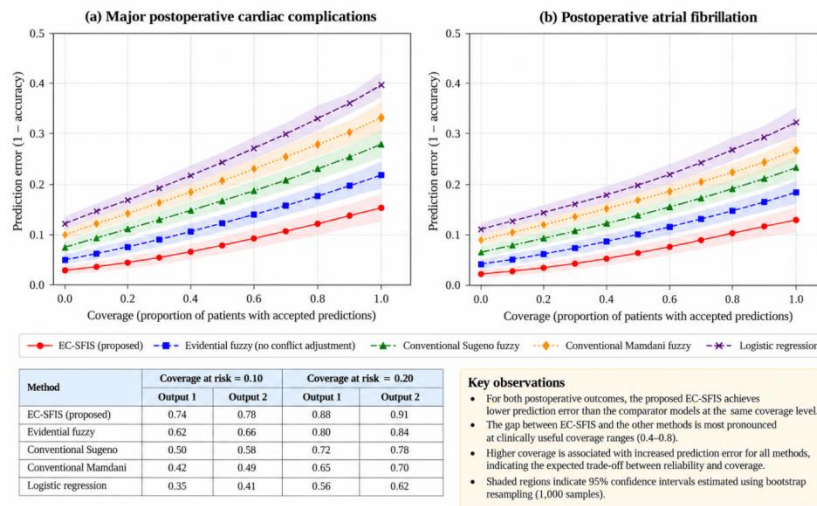


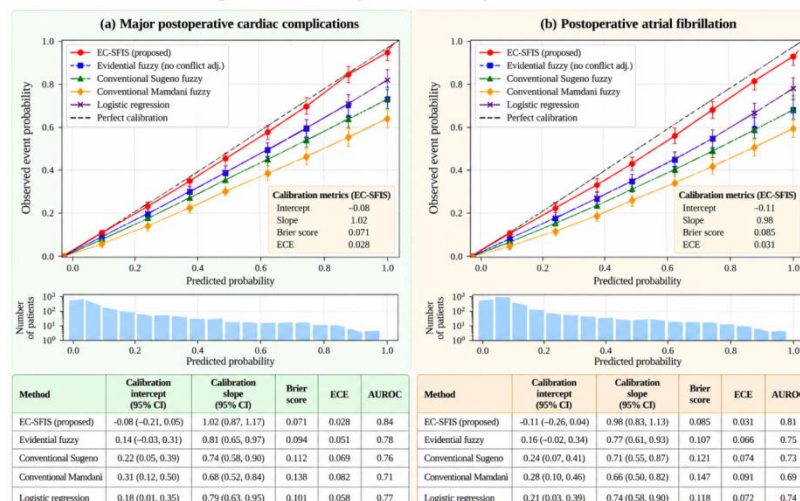
Figure 7 presents the risk–coverage curves for the two postoperative outcomes major postoperative cardiac complications and postoperative atrial fibrillation comparing the proposed EC-SFIS with evidential fuzzy, conventional Sugeno, conventional Mamdani, and logistic-regression models. Coverage represents the proportion of patients for whom predictions are accepted, while risk denotes the prediction error among those accepted cases. The figure shows an increase in prediction error as the coverage is increased for each outcome. This supports what we would expect; by increasing patient acceptance there will be a loss of predictive reliability. As shown in the above graphs, the EC-SFIS model has been demonstrated to have the lowest prediction errors throughout the entire coverage ranges with the largest difference from other models being observed at clinical relevant intermediate to high coverage. The shaded regions are the 95% Confidence Intervals (CI), and as stated in the table below, the EC-SFIS model also provides the same or better coverage than other models for the same level of risk. Overall, this graph provides evidence that the Selective Prediction Mechanism can provide more reliable predictions on a higher percentage of data points, thereby providing a method to identify potentially less certain data points which may require additional evaluation.

7.3 Clinical probability calibration: Evidence scores should not be reported directly as probabilities. For each outcome, fit a calibration model:

$$\hat{p}_{ik} = \frac{1}{1 + \exp[-(\alpha_k + b_k q_k)]} \quad (46)$$

Estimate the calibration parameters using development data. Evaluate calibration intercept, calibration slope, Brier score and graphical agreement between predicted probabilities and observed event frequencies.

Figure 8. Outcome-Specific Probability-Calibration Plots



Beginning with Figure 8, in terms of the outcome-specific probability-calibration performance for the proposed EC-SFIS model and its comparator models for major post-operative cardiac complication and post-operative atrial fibrillation, we can see the calibration curves are located very close to the perfect calibration reference lines as well as having small calibration intercepts and slopes closer to one (i.e., a slope near one indicates good calibration). In addition, the Brier scores and expected calibration error (ECE) for these two outcomes were relatively low. The histogram at the bottom of the figure illustrates how many patients fell into each risk category based on their individualized predictions. Calibration and discrimination statistics for each model are also presented in tabular form. Based upon these findings, it appears that the proposed EC-SFIS model performs better in terms of the degree of alignment between actual patient outcomes and model-predicted risks compared to the other fuzzy and logistic regression models tested while still providing comparable area under receiver operating characteristic (AUROC) values for predicting both major postoperative cardiac complications and postoperative atrial fibrillation.

7.4 Conflict-validity experiment: The purpose of this study was to test mathematically the impact on the model. Determine how much test-set-predictions are in each pre-determined conflict category (based upon C_k) as defined by the cut-off scores used in the developmental phase. Test if there is an association of increased conflict to a greater prediction error, after adjusting for the original evidence score.

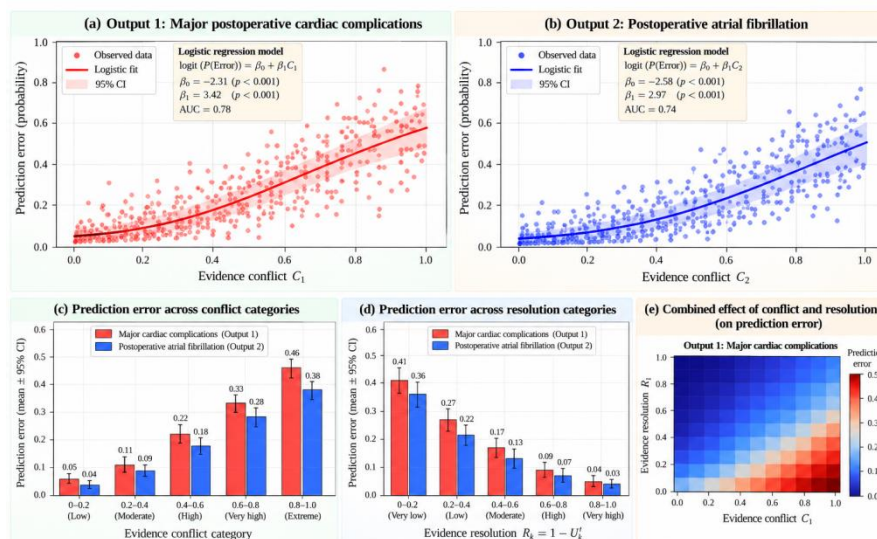
A suitable statistical model is

$$\text{logit}[P(E_{ik} = 1)] = \beta_0 + \beta_1 C_{ik} + \beta_2 U_{ik} + \beta_3 q_{ik} \quad (47)$$

where E_{ik} is a prespecified binary indicator of prediction error.

This study provides a contrast to the possible explanations of what conflict is providing for our purposes (i.e., whether conflict is providing new predictive information, or if it's simply duplicating that which we have already ascertained from the initial score). This study provides a contrast to the possible explanations of what conflict is providing for our purposes (i.e., whether conflict is providing new predictive information, or if it's simply duplicating that which we have already ascertained from the initial score).

Figure 9. Relationship between Evidence Conflict and Observed Prediction Error



The graphical representation of the relationship between Evidence Conflict, Evidence Resolution, and Observed Prediction Error for each outcome is demonstrated within FIGURE 9. It is evident in panels (a) and (b), where both conflict measures (C_1) and (C_2) demonstrate an obvious positive correlation with the observed prediction error. Additionally, both panels (a) and (b) indicate the same trend; that the reliability of the predictions decreases with increasing disagreements among rules. Furthermore, it is shown in panel (c), there is an increase in prediction errors going from low to high levels of conflict while, conversely, there is a decrease in prediction errors going from low to high levels of evidence resolution shown in panel (d). Lastly, the heat map represented in panel (e), illustrate the cumulative effects of these two factors. A greater degree of predicted error exists in areas on the heat map that have higher degrees of conflict and lower amounts of resolved evidence. In total, the graphics represent a visual

demonstration of the overall EC-SFIS hypotheses that conflict and unresolved evidence can be used as predictive indicators of when a case may have been inaccurately predicted and therefore when Selective Prediction would flag this case for evaluation.

Figure 10. Integrated Patient-Level Explanation Displaying Supporting Evidence, Opposing Evidence, Unresolved Evidence, Conflict, Calibrated Risk and the Final Prediction-Acceptance Decision

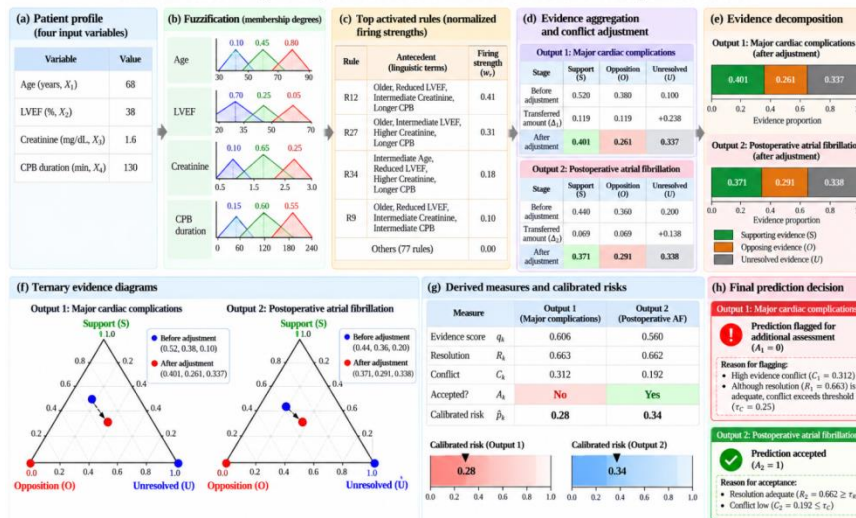


Figure 10 shows a single-patient level diagram illustrating all of the steps involved in the entire EC-SFIS predictive modeling process from the input clinical variables (age, LVEF, serum creatinine, and CPB time) into fuzzy sets, to rule firing, to combining supporting and conflicting evidence, resolving conflicts, adjusting conflict levels, and finally calibrating those probabilities. The rules fired by the data produce a distribution of support/evidence for each possible outcome. Ternary diagrams display how much of that support was adjusted to the "unresolved" category as well. Therefore for the major complications after surgery, there is a total evidence score of $q_1 = 0.606$, a resolution rate of $R_1 = 0.663$, and a relatively high level of conflict $C_1 = 0.312$. As such, although this model predicts a low risk of complications at 0.28, we have reason to flag this prediction for further review due to the fact that the conflict threshold has been crossed. On the other hand for the potential for post-operative atrial fibrillation, this model produces an evidence score of $q_2 = 0.56$, a resolution $R_2 = 0.66$ and a conflict $C_2 = 0.19$ which results in an acceptable prediction with a predicted risk of complications of 0.34. Therefore, this figure displays the primary method by which patients will be able to understand their own individual risk predictions as provided by EC-SFIS. That method includes both patient specific risk information as well as an indication of what the evidence looks like that supports that particular risk estimate.

VIII. CONCLUSION

The EC-SFIS was designed as a tool to predict the likelihood of developing postoperative atrial fibrillation and/or major postoperative cardiac complications in patients undergoing coronary artery bypass grafting surgery using four clinical variables including: age, LVEF, serum creatinine, and CPB time. A tool was created to determine the risk of postoperative atrial fibrillation and/or major postoperative cardiac complications in patients who have undergone coronary artery bypass grafting surgery by utilizing clinical variables that include; age, LVEF, serum creatinine, and length of CPB time. The principal contribution of the framework is the explicit decomposition of fuzzy inference into supporting, opposing, and unresolved evidence and the introduction of a conflict-conservation mechanism that transfers contradictory evidence toward the unresolved component without reversing the original direction of the evidence balance. The numerical case study demonstrated the practical significance of this formulation: although the two outcomes exhibited nearly identical evidence-resolution values ($R_1 = 0.663$ and $R_2 = 0.662$), their conflict levels differed ($C_1 = 0.312$ and $R_2 = 0.192$), causing the first prediction to be flagged and the second to be accepted. The study further evaluates the framework through risk-coverage analysis, probability calibration, comparator models, and conflict-validity analysis, with the presented results indicating that increasing conflict is associated with increasing prediction error while greater evidence resolution is associated with improved prediction reliability. Consequently, EC-SFIS extends fuzzy postoperative risk modelling beyond the generation of a single risk score by providing an interpretable indication of why a prediction is supported, how much contradictory evidence exists, and whether the prediction should be accepted or subjected to additional assessment. Further validation on independent, multicentre

clinical datasets is required before the proposed thresholds or predictions can be considered suitable for clinical deployment.

REFERENCES

- [1]. Cabitza F., Rasoini R., Gensini G.F. (2017): "Unintended consequences of machine learning in medicine", *JAMA*, 318:517-518.
- [2]. Choi J.H., Song S.H., Kim H., Kim J., Park H., Jeon J.H., et al. (2024): "Machine learning algorithm to predict atrial fibrillation using serial 12-lead ECGs based on left atrial remodeling", *Journal of the American Heart Association*, 13:e034154.
- [3]. Devereaux P.J., Sessler D.I. (2015): "Cardiac complications in patients undergoing major noncardiac surgery", *New England Journal of Medicine*, 373:2258-2269.
- [4]. Dhaliwal S., Chen S., Papas C., Hughes I., Cavalucci D., O'Rourke N. (2026): "Machine learning vs. traditional methods for predicting postoperative cardiac complications after non-cardiac surgery: a systematic review and Bayesian network meta-analysis", *Anaesthesia*, 81:1415-1427.
- [5]. Gaudino M., Sanna T., Ballman K.V., Robinson N.B., Hameed I., Audisio K., et al. (2021): "Posterior left pericardiotomy for the prevention of atrial fibrillation after cardiac surgery: an adaptive, single-centre, single-blind, randomised, controlled trial", *The Lancet*, 398:2075-2083.
- [6]. Im S.I. (2025): "Advances in iron deficiency and iron-related arrhythmias and cardiovascular diseases", *Frontiers in Cardiovascular Medicine*, 12:1573095.
- [7]. Kalisnik J.M., Zujs V., Zibert J., Batashev I., Leiler S., Carstensen J.A.B., et al. (2025): "The impact of a chest drainage system on retained blood-associated complications after cardiac surgery", *European Journal of Cardio-Thoracic Surgery*, 67:i9-i17.
- [8]. Lewicki Ł., Siebert J., Rogowski J. (2016): "Atrial fibrillation following off-pump versus on-pump coronary artery bypass grafting: incidence and risk factors", *Cardiology Journal*, 23:518-523.
- [9]. Martínez-Sellés M., García-Izquierdo Jaén E., Fernández Lozano I. (2017): "Anticoagulation in elderly patients at high risk of atrial fibrillation without documented arrhythmias", *Journal of Geriatric Cardiology*, 14:166-168.
- [10]. Onishchenko D., Rubin D.S., van Horne J.R., Ward R.P., Chattopadhyay I. (2022): "Cardiac comorbidity risk score: zero-burden machine learning to improve prediction of postoperative major adverse cardiac events in hip and knee arthroplasty", *Journal of the American Heart Association*, 11:e023745.
- [11]. Sanaiha Y., Verma A., Ng A.P., Hadaya J., Ko C.Y., deVirgilio C., Benharash P. (2024): "Development and preliminary assessment of a machine learning model to predict myocardial infarction and cardiac arrest after major operations", *Resuscitation*, 200:110241.
- [12]. Steyerberg E.W., Harrell F.E. Jr. (2016): "Prediction models need appropriate internal, internal-external, and external validation", *Journal of Clinical Epidemiology*, 69:245-247.
- [13]. Stoger K., Schneeberger D., Holzinger A. (2021): "Medical artificial intelligence: the European legal perspective", *Communications of the ACM*, 64:34-36.
- [14]. Wu X., Hu J., Zhang J. (2023): "Machine learning-based model for predicting major adverse cardiovascular and cerebrovascular events in patients aged 65 years and older undergoing noncardiac surgery", *BMC Geriatrics*, 23:819.
- [15]. Yu J., Peng X., Zhou R., Zhu T., Hao X. (2025): "Development and validation of an interpretable machine learning model to predict major adverse cardiovascular events after noncardiac surgery in geriatric patients: a prospective study", *International Journal of Surgery*, 111:1939-1949.