Asian Journal of Current Research in Clinical Cancer
ISSN: 3062-4444
2024, Volume 4, Issue 2, Page No: 106-122

Available online at: www.galaxypub.co/page/journals

Galaxy Publication

Galaxy Publication

A Hybrid Fuzzy-ML Clinical Decision Support System for Breast Cancer Risk Stratification: Design, Symbolic/Statistical Risk Fusion, and Proof-of-Concept Validation

M. Al-Dosari[1*], F. Al-Marri[1], A. Al-Sulaiti[1]

[1]Department of Oncology, School of Medicine, University of Doha, Doha, Qatar.

***E-mail** ✉ doha.onc.65@gmail.com

Received: 16 August 2024; **Revised:** 29 October 2024; **Accepted:** 08 November 2024

ABSTRACT

Breast cancer remains the most commonly identified malignant condition worldwide and represents the principal cause of death among women. To address this issue, population-based screening initiatives—primarily mammographic exams—began to appear in the 20th century. Their introduction has substantially lowered mortality and enhanced outcomes for individuals diagnosed with this illness. Even so, reading mammograms involves a certain level of inconsistency and depends heavily on the expertise and background of the clinicians who interpret them. Seeking to assist the assessment of mammographic images and strengthen the diagnostic workflow, this study introduces the conception, development, and proof of concept of an innovative intelligent clinical decision support system based on two predictive strategies executed in parallel. The first strategy uses several expert modules driven by fuzzy inference mechanisms designed to process the attributes linked to the chief mammographic findings. From this, a collection of indicators—termed Symbolic Risks—can be derived, each tied to the likelihood of breast cancer based on the detected features. The second strategy relies on a machine-learning classification model that, using both mammography-related descriptors and general patient data, computes another metric, called Statistical Risk, likewise associated with the probability of developing breast cancer. These metrics are then integrated to generate a combined measure, the Global Risk. This value may subsequently be adjusted using a weighting element derived from the BI-RADS category assigned by the medical specialists. The resulting Corrected Global Risk can then be interpreted to determine a patient's status and issue tailored recommendations. The proof-of-concept evaluation and software implementation were performed using a dataset of 130 patients obtained from a database at the University of Wisconsin–Madison School of Medicine and Public Health. The initial outcomes were promising, suggesting the system's potential utility, though thorough clinical testing in real practice settings is still required. Its integration into existing clinical information systems may further streamline diagnostic procedures and enhance patient outcomes.

Keywords: Cancer, Breast cancer, Design; clinical decision support system, Intelligent system, Expert system, Machine learning, Decision making, Medical algorithm, Design science research

How to Cite This Article: Al-Dosari M, Al-Marri F, Al-Sulaiti A. A Hybrid Fuzzy-ML Clinical Decision Support System for Breast Cancer Risk Stratification: Design, Symbolic/Statistical Risk Fusion, and Proof-of-Concept Validation. Asian J Curr Res Clin Cancer. 2024;4(2):106-22. https://doi.org/10.51847/Q2DsJjMLMM

Introduction

At present, breast cancer represents the most frequently encountered malignant disease globally, corresponding to one in eight cancer diagnoses in the general population. Among women, one in four oncologic cases is attributed to this pathology [1-3], which accounted for 15.5% of cancer-related deaths in 2020 [2]. Because of the considerable burden associated with this disease, extensive efforts throughout the latter part of the 20th century and onward have focused on introducing early-detection practices. When paired with prompt diagnosis and adequate therapy, these initiatives have markedly improved patient outcomes and reduced mortality [3]. Such early-detection programs—typically structured as periodic mammographic screening for high-risk groups, especially those over 40 years old—aim to identify any indications suggestive of the disease. Individuals with

suspicious findings are then directed to more specific evaluations to verify or dismiss the potential diagnosis. In this setting, the Breast Imaging Reporting and Data System (BI-RADS) [4] is routinely employed to ensure standardized terminology for describing mammographic observations and to classify exams into seven categories, ranging from no concern to definite malignancy. This common framework supports both documentation and decision-making. However, interpreting mammograms remains challenging; outcomes can vary depending on the practitioner's training and experience [5-7]. Consequently, unnecessary or insufficient follow-up tests may be ordered, posing potential disadvantages to patients.

Given these challenges, the availability of tools capable of supporting complex clinical decisions is vital. Many such systems—often driven by artificial-intelligence methods—are now increasingly used in healthcare, with a wide variety of proposals documented in the literature and typically implemented as clinical decision support systems [8–25]. In the domain of breast cancer, several approaches employ symbolic inference techniques [21, 25], whereas the growing prevalence of connectionist models has led to numerous solutions incorporating large neural networks for tumor detection [26–32]. These technologies have contributed to more efficient use of clinical resources, improved service quality, and lower operational costs. Following this trajectory, the present work offers an adaptation, advancement, and proof-of-concept implementation of the system introduced by Casal-Guisande *et al*. (2022) [25]. That earlier study proposed a diagnostic framework built upon a sequence of expert systems that gathered and formalized patient-related clinical information. The resulting outputs were analyzed to derive covariance elements forming a knowledge base with which to train a statistical classifier responsible for the final predictive stage. Although the proposed sequential structure helped reduce uncertainty in both data processing and interaction, it also constrained the opposing interaction of inferential processes. In particular, statistical inference occurred only after the symbolic stage, relying on its outputs as classifier inputs. This arrangement imposed a dependency of the statistical layer on the symbolic layer, thereby limiting the breadth of knowledge representation and, importantly, the range of correlation patterns that the machine-learning model could uncover.

The enhanced system described in this study is derived from the previously presented model and seeks not only to surpass the earlier reported outcomes but also to restructure the movement of information and knowledge so as to broaden the range of inferences that the models can generate. For this purpose, an intelligent framework composed of two parallel-working modules was developed [12, 13, 17, 25-27]. The first module incorporates expert systems powered by fuzzy-logic inference mechanisms, dedicated to analyzing the characteristics observed in mammographic studies—namely masses, calcification patterns, asymmetries, and architectural distortions. The second module applies a machine-learning classifier that processes combined data from mammograms—excluding the BI-RADS score—alongside general patient information. Running these components at the same time allows the system to compute two types of indicators: Symbolic Risks and the Statistical Risk. These outputs are subsequently merged to generate a percentage estimate of Global Risk, which reflects the probability of developing breast cancer. This value may then be modified using the BI-RADS level assigned by healthcare professionals, producing a Corrected Global Risk which, once interpreted, supports the creation of alerts regarding the patient's status and provides tailored suggestions. Operating both inferential paths concurrently enables cross-comparison and unification of results, while the analytical strategies applied, as explained later, help diminish uncertainty inherent in such architectures.

This paper is divided into five main parts. Section 2 introduces the conceptual framework of the intelligent model, explaining its stages and information routes. It is followed by a detailed discussion of the system's software implementation and a summary of the proof-of-concept findings. Section 3 presents a real-world example demonstrating the system's behavior. Section 4 examines the underlying architecture, and Section 5 outlines key conclusions and potential future enhancements.

Materials and Methods

Definition of the system

Database description

For this investigation, a dataset from the School of Medicine and Public Health at the University of Wisconsin–Madison was utilized, containing records from 130 patients. Out of these 130 individuals, 21 received a confirmed diagnosis of breast cancer following comprehensive diagnostic procedures. Every participant underwent a biopsy, which in each instance either verified or excluded the presence of malignancy from a histopathological standpoint. The dataset included basic demographic and clinical information (age, personal and family cancer history),

typically stored in electronic medical systems, along with mammographic descriptors following BI-RADS nomenclature [28], including the assigned BI-RADS category.

A condensed overview of the key characteristics of the studied population appears in **Table 1**, while **Table 2** and **Table 3** present additional details regarding the variables considered.

Table 1. Summary of the dataset.

Parameter	Value
Total Patients	130
Biopsied Patients	130
Confirmed Cancer Cases	21
Healthy Controls	109
Mean Age	55.2
Number of Criteria	13 Mass (shape, margins, density), calcifications (type, shape, distribution), asymmetries (type), distortion (type), breast tissue density, BI-RADS category, age, personal history, family history
Criteria Type	Quantitative and qualitative

Table 2. Overview of general patient characteristics.

Variable	Type	Notes
Age	Continuous	-
Personal history of cancer	Categorical	Yes / No / N/A
Family history of cancer	Categorical	None / Minor / Major / N/A

Table 3. Summary of mammographic findings.

Subgroup	Data	Type of Data	Commentary
Mass	Shape		
Shape	Margins	Categorical	None, oval, round, lobulated, irregular
Margins	Density	Categorical	None, circumscribed, obscured, micro-lobulated, indistinct, spiculated
Density	Type	Categorical	None, equal density, low density, high density
Calcifications	Shape		
Type	Distribution	Categorical	None, primary, associated
Shape	Type of asymmetry	Categorical	None, skin, vascular, coarse or “popcorn-like”, large rod-like, round, rim, dystrophic, milk of calcium, suture, amorphous, coarse heterogeneous, fine pleomorphic, fine linear or fine linear branching
Distribution	Type of distortion	Categorical	None, diffuse, regional, grouped, linear, segmental
Asymmetries and distortions	-		
Type of asymmetry	-	Categorical	None, missing, focal, developing
Type of distortion		Categorical	None, primary, associated
Breast tissue density	-	Categorical	Missing, fatty, scattered areas of fibro glandular, heterogeneously dense, extremely dense
BI-RADS category	-	Categorical	0, 1, 2, 3, 4.a, 4.b, 4.c, 5, 6

Conceptual system design

A flow diagram representing the intelligent clinical decision support system for breast cancer risk assessment is provided in **Figure 1**.

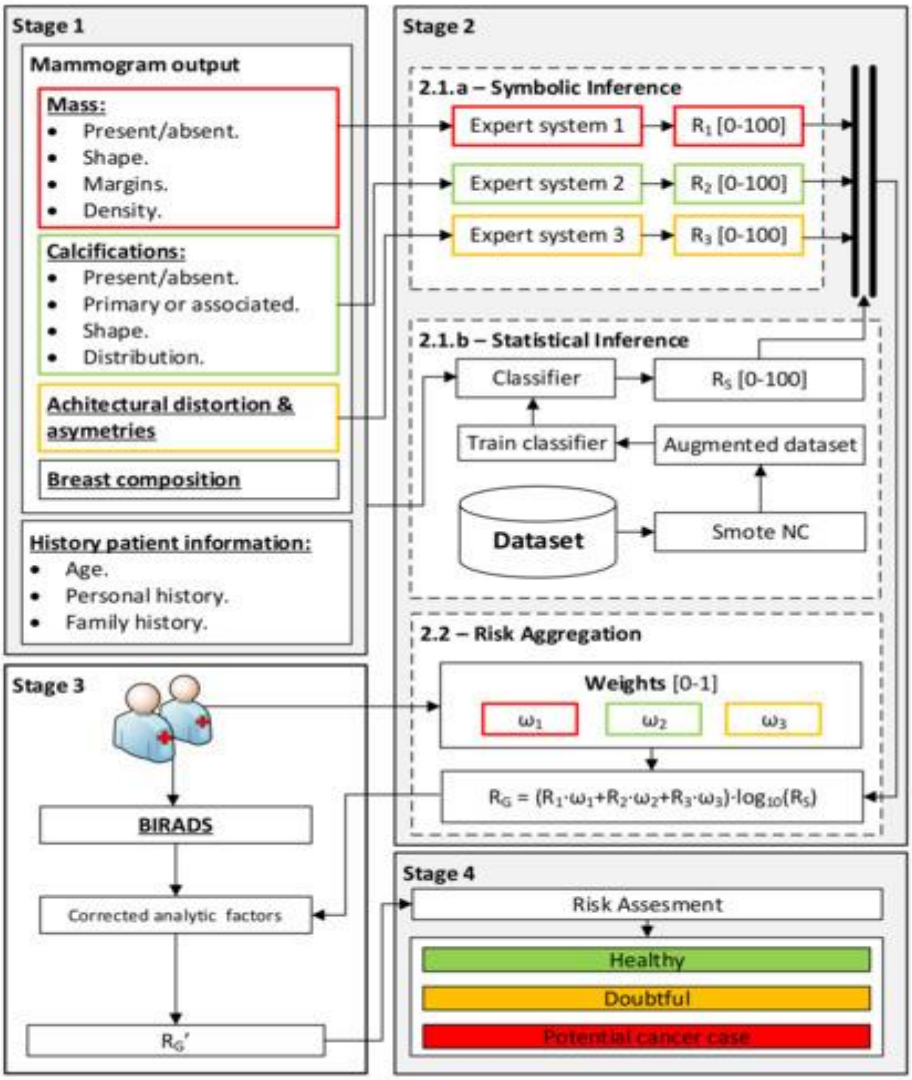


Figure 1. Flow diagram of the intelligent clinical decision support system for estimating breast cancer risk. The figure illustrates how information moves across the system's stages: Stage 1 involves data acquisition; Stage 2 is divided into Stage 2.1.a (symbolic inference), Stage 2.1.b (statistical inference), and Stage 2.2 (risk combination). Stage 3 computes the Corrected Global Risk using the BI-RADS value. Stage 4 issues alerts and supports decision-making.

Gathering and interpretation of patient information

The initial step of the intelligent model consists of collecting general patient details—normally retrieved from electronic health records **(Table 2)**—as well as examining mammogram findings using BI-RADS® terminology [28], as recorded by a clinical specialist **(Table 3)**. For each descriptor, an accompanying field specifying the data type is included, distinguishing numerical inputs from categorical ones.

Data processing

After the information gathered in Stage 1 has been organized, it is fed into the intelligent system, which analyzes it through two simultaneous sub-stages—2.1.a and 2.1.b—composed of several expert models and a machine-learning classifier operating in parallel [12, 13, 17, 25-27]. These two branches produce multiple indicators associated with the likelihood of developing breast cancer. In Stage 2.2, these indicators are combined to produce a single Global Risk value that reflects the estimated probability that a patient may have breast cancer. Each of the sub-stages is explained below:

- Stage 2.1.a — Symbolic Risk estimation:
  Once the mammographic findings captured in Stage 1 are available, they are processed by a set of concurrent expert modules [12, 13, 17, 25-27] constructed using Mamdani-style fuzzy inference

approaches [29-31]. Each expert module is responsible for analyzing one subset of features—masses, calcifications, asymmetries, and architectural distortions. Their outputs are a series of indicators, termed Symbolic Risks (R1, R2, and R3), each ranging from 0 to 100, and each expressing a risk level associated with potential breast cancer development.

- Stage 2.1.b — Statistical Risk estimation:
  Running at the same time as Stage 2.1.a, this sub-stage processes the complete dataset **(Table 2 and Table 3)**, excluding the BI-RADS score assigned by the clinician, using a classification algorithm from the machine-learning domain [32]. Depending on the structure and quality of the inputs, normalization and, if needed, synthetic rebalancing of the dataset may be applied [25]. The model is trained using the dataset described in Section 2.1.1, where every record is labeled either "cancer" or "non-cancer". Because all individuals underwent biopsies, the ground truth for each case is definitively known, minimizing epistemic and interaction-related uncertainty in the training material. After training, the classifier receives the data of a new subject and produces a percentage output—referred to as the Statistical Risk (Rs), also between 0 and 100—intended to reflect the probability that the patient has or may develop breast cancer.

- Stage 2.2 — Aggregation and Global Risk computation:
  With the Symbolic Risks (R1, R2, R3) and the Statistical Risk (Rs) available, they are merged through the formula presented in Eq. 1, yielding the Global Risk (RG). This equation incorporates the weighted sum of the Symbolic Risks and the base-10 logarithm of the Statistical Risk. The weighted sum condenses the Symbolic indicators into a single measure according to the importance assigned by medical experts to each category of mammographic findings (masses, calcifications, asymmetries, and architectural distortions). The logarithmic term then acts multiplicatively, heightening the combined risk when the patient's pattern resembles that of cases used to build the statistical classifier. It should be noted that if any finding category is not present—resulting in one of the Symbolic Risk terms being zero—weights are redistributed proportionally across the remaining categories or reassigned following expert recommendation. Also, by definition, the Global Risk is capped between 0 and 100; any value exceeding this bound, even after the multiplicative adjustment, is limited to a maximum of 100.

$$R_G = (\omega_1 \cdot R_1 + \omega_2 \cdot R_2 + \omega_3 \cdot R_3) \cdot \log_{10} R_s \quad (1)$$

Global risk correction

As explained in the earlier stage, several individual risk indicators were generated and combined to derive the Global Risk. In this phase, the Global Risk can be adjusted by incorporating the BI-RADS score assigned by the clinical team based on their review of the mammograms. The Global Risk represents an estimate of the patient's likelihood of developing breast cancer, derived from statistical categorical inputs along with symbolic variables associated with mammographic findings. Although this metric is useful, the BI-RADS classification—an internationally recognized and standardized reference—is generally preferred. It is currently regarded as the primary diagnostic framework for evaluating breast tumors. However, applying the BI-RADS scale and selecting the correct category for a suspicious lesion demands substantial clinical skill and precision. Its proper use requires familiarity and experience, which can be challenging for medical teams that are less accustomed to interpreting the scale. Even so, its significance in breast cancer staging is undeniable, and it must be considered a core diagnostic component.

In this study, the BI-RADS index is intentionally excluded from the initial calculation of the Global Risk in order to promote a structured representation of knowledge in both statistical and symbolic forms. This design reduces the dependency on highly specialized expertise and instead relies on broader, formalized patient information to generate a credible and robust risk estimate, minimizing epistemic variability linked to BI-RADS interpretation. Nonetheless, the BI-RADS score remains indispensable; therefore, once the Global Risk is computed, it is systematically adjusted using the BI-RADS category assigned by the physicians. To apply this adjustment, a weighted-order procedure was incorporated: depending on the patient's BI-RADS value, scaling factors are applied to the Global Risk. This approach helps reduce uncertainty associated with the BI-RADS category while

reinforcing the alignment expected between the Global Risk output and the BI-RADS assessment. Eq. 2 provides the formula for obtaining the Corrected Global Risk (RG'), where *Fp* denotes the weighting factor.

$$R_G' = R_G \cdot F_p \quad (2)$$

Table 4 lists a proposed set of weighting coefficients corresponding to each BI-RADS category. These factors were established following an empirical evaluation of the relationships and causal links between confirmed cancer cases, the assigned BI-RADS levels, and the resulting Global Risk scores. The objective was to devise a mathematical adjustment allowing the Global Risk to shift according to the BI-RADS category, thereby improving diagnostic performance. The intention is both to refine the predictive capability and to align it more closely with the standard clinical interpretation grounded in BI-RADS, which is often decisive when confirming or rejecting suspected disease. Since its function is purely analytical, the weighting factor does not incorporate additional statistical inference; it simply scales the Global Risk by multiplication or division in accordance with what the BI-RADS score represents. Finally, even after applying the weighting, the Corrected Global Risk is capped at a maximum of 100.

Table 4. Weighting factors according to the BI-RADS index order.

BI-RADS	Weighting Factor (Fp)
1	$F_P = \frac{1}{k} f(R_G)$ $With\ k = -BIRADS_{level} + (3.5)$ $f(R_G) = \frac{10}{R_G + 10}$
2	
3	$F_C = \frac{100}{R_G};\ F_P = t + f(R_G);\ F_P = \begin{cases} t + f(R_G)\ if\ F_C > F_P \\ F_C\ if\ F_C \leq F_P \end{cases}$ $With\ t = BIRADS_{level} - (1.5)$ $f(R_G) = \frac{10}{R_G + 10}$
4A	
4B	
4C	
5	$F_P = \frac{100}{R_G}$
6	

Generation of warnings and decision making

Once the Corrected Global Risk (RG') has been obtained, the system interprets this value and assigns a set of corresponding statuses and recommendations for each patient:

• Healthy case: The patient should be referred for standard follow-up;

• Dubious case**:** The patient's information should be reviewed again, potentially complemented with additional tests, and the patient should be scheduled for another appointment within a defined time frame;

• Possible breast cancer case**:** Conduct confirmatory examinations.

This assessment is simply a suggested categorization intended to support explicit diagnostic decisions. Although the process relies on the Corrected Global Risk, it may also use the uncorrected Global Risk if clinicians deem the adjustment unnecessary.

Implementation of the system

This section covers the implementation of the intelligent clinical decision support platform for breast cancer diagnosis. It describes the developed software application and includes comments on results obtained from the proof-of-concept stage. The design adheres to the guidelines presented by Hevner *et al*. [33, 34], enabling future integration into hospital information systems if required.

MATLAB© (R2021b, MathWorks©, Natick, MA, USA) served as the main development environment. **Table 5** lists the MATLAB® Toolboxes utilized in this work.

Table 5. MATLAB® Toolboxes used in this study.

Toolbox	Description
App Designer [35]	Employed for creating and building the graphical user interface of the software prototype.
Fuzzy Logic Toolbox [36]	Utilized to develop fuzzy logic-based inference engines.
Classification Learner [37]	Applied for training machine learning classification models. Enables efficient parallel evaluation of multiple algorithms to identify the optimal performer.

Alongside these resources, Python (version 3.9.12) was also employed as a complementary tool to perform data augmentation via the SMOTE-NC algorithm provided by the imbalanced-learn library [38].

Figure 2 displays the main interface of the developed application. The Stage #1 area is designed for entering the initial information, such as general patient data and findings extracted from mammography interpretation. The Stage #2 area covers the risk computation phase (Methodology Stage 2), divided into three sections: the Symbolic Reasoning panel (Stage 2.1.a), where Symbolic Risks tied to mammographic findings are computed; the Statistical Reasoning panel (Stage 2.1.b), where the Statistical Risk is obtained; and the Global Risk panel (Stage 2.2), where both risks are integrated. Next, the Stage #3 section adjusts the Global Risk according to the BI-RADS score. Finally, Stage #4 contains the module corresponding to Stage 4 of the methodology, responsible for generating alerts and supporting decisions.

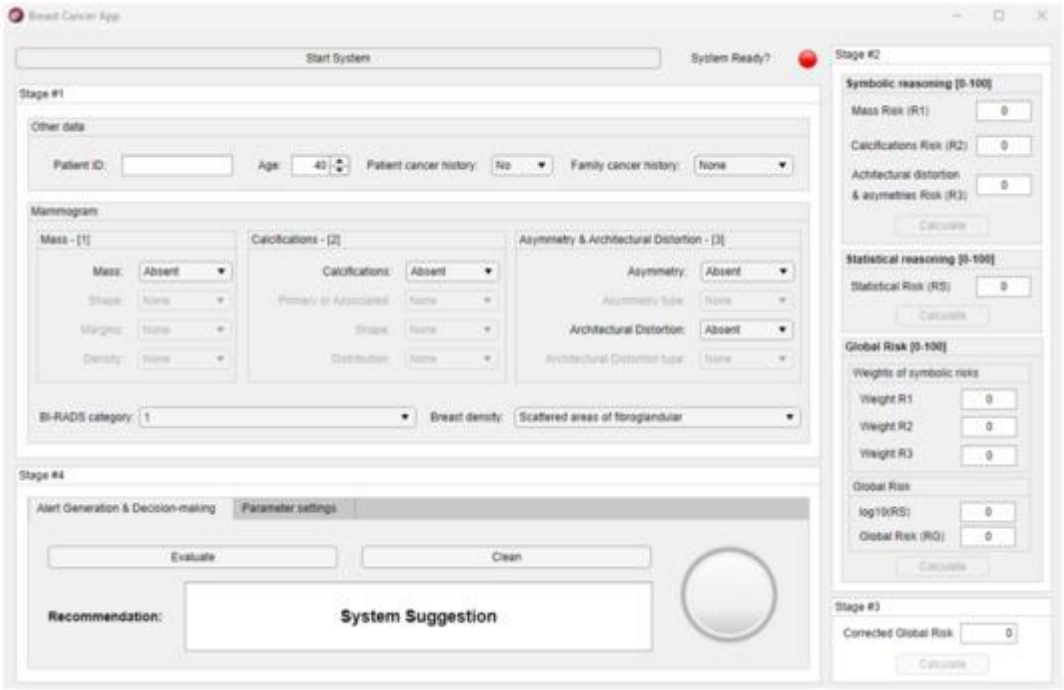


Figure 2. Main graphical interface of the intelligent clinical decision support system for breast cancer risk estimation. Stage #1 corresponds to data entry; Stage #2 performs risk calculations; Stage #3 applies the BI-RADS-based correction; Stage #4 handles alert generation and decision support.

Data collection

Once patient information is prepared for analysis, it must be entered into the software through the fields provided in Stage #1 **(Figure 2)**. Two zones are available: the "Other data" area, containing general patient information, and the "Mammogram" area, which captures the findings observed in the mammograms and the BI-RADS classification assigned by the medical staff after image review. Care must be taken when typing this information, as mistakes or missing data may distort the output and increase uncertainty in the system.

Data Processing

After entering the required inputs, the system processes the information. The computation is carried out in the Stage #2 panel, where the relevant processing modules are located. As explained in Section 2.1.2, the workflow

involves two main components operating concurrently [12, 13, 17, 25-27]. The first comprises several expert systems—also running in parallel—dedicated to determining the Symbolic Risks. The second component employs a machine learning classifier responsible for computing the Statistical Risk. Both risk estimates are then combined to produce the Global Risk.

The following subsections describe the purpose of each component and outline the calculations used to obtain the various risk indicators.

Determination of the symbolic risks

As previously described, several expert systems operating in parallel are used to compute the Symbolic Risks [12, 13,17, 25-27]. These systems process the different categories of mammographic findings—masses, calcifications, asymmetries, and architectural distortion—listed earlier in **Table 3**.

For this study, Mamdani-style fuzzy inference mechanisms [29, 30, 31, 39] were adopted, comparable to those implemented in the first stage of the cascade approach by Casal-Guisande *et al*. [25] and other related works [12, 13, 17-27].

Table 6 outlines the antecedents and consequents included in each expert system. Regarding membership functions, Ross's guidelines [39] were followed, employing normal, convex, and symmetric shapes. For antecedents, singleton membership functions were chosen, while the consequents—representing Symbolic Risks—were modeled with triangular membership functions covering the interval from 0 to 100. Singleton functions were used for antecedents because the associated data represent a single categorical value (e.g., a specific mass shape). Conversely, triangular functions were chosen for consequents because these variables may take membership values between zero and one, yet only one point corresponds to full membership. The risks, understood here as consequent variables, were constructed assuming the existence of a unique value within the measurement scale that reflects the maximum risk, in alignment with previous studies by the authors [25].

Table 6. Summary of the antecedents and consequents used in each expert system.

Expert System	Antecedents	Consequents [0, 100]
Expert System 1 — Masses	Present/absent Shape Margins Density	**R1**
Expert System 2 — Calcifications	Present/absent Type Shape Distribution	**R2**
Expert System 3 — Architectural Distortion and Asymmetries	Asymmetry present/absent Type of asymmetry Distortion of the architecture present/absent Type of distortion	**R3**

A general overview of the configuration adopted for the inference engines in these expert systems is presented in **Table 7**.

Table 7. Overall configuration of the inference engines used.

Inference Engine Component	Details
Type	Fuzzy structure
Mamdani-type	
Defuzzification method	Centroid [39]
Implication method	MIN
Aggregation method	MAX

After completing the symbolic inference stage, three indicators are generated—R1, R2, and R3. Each corresponds to the likelihood of breast cancer based on the group of findings used for its computation. Higher values imply higher estimated risk.

Determination of the statistical risk

Alongside the symbolic analysis [12, 13, 17, 25-27], the Statistical Risk is computed using a machine learning classification model. This classifier relies on the dataset described in Section 2.1.1, excluding the BI-RADS value assigned by clinicians. Most variables are categorical [40, 41], as shown in **Tables 2 and 3**, with age being the only numerical variable. Age is normalized using the Min–Max method, whose formula appears in Eq. 3:

$$t' = \frac{t_i - \min(t)}{\max(t) - \min(t)} \quad (3)$$

Inspection of the dataset reveals a major imbalance between the "cancer" and "non-cancer" labels, which may impair classifier performance and its ability to generalize. To address this issue and following common practice in medical decision-support applications, a controlled data augmentation technique was applied to enhance the performance of binary classifiers [25, 42]. Specifically, the SMOTE-NC method, an extension of the synthetic minority over-sampling technique (SMOTE) [42, 43], was used. This variant supports datasets that contain both numerical and categorical attributes after transformation [25]. The augmentation procedure involved adding synthetic cases until both classes (cancer and non-cancer) reached 200 records each, using k = 5 neighbors [25]. This process produced a consistent and representative dataset suitable for training and optimizing the machine learning classifier. Several candidate models were evaluated through the MATLAB© Classification Learner app, employing a k-fold cross-validation strategy [44] with k = 5. After extensive testing and examining ROC curves [25], the bagged trees algorithm emerged as the top performer. Nonetheless, any other classifier could be substituted if it achieves superior ROC-based metrics. **Figure 3** presents the ROC validation curve for the "cancer" class using the selected bagged trees model, with an AUC of 0.98, indicating excellent performance.

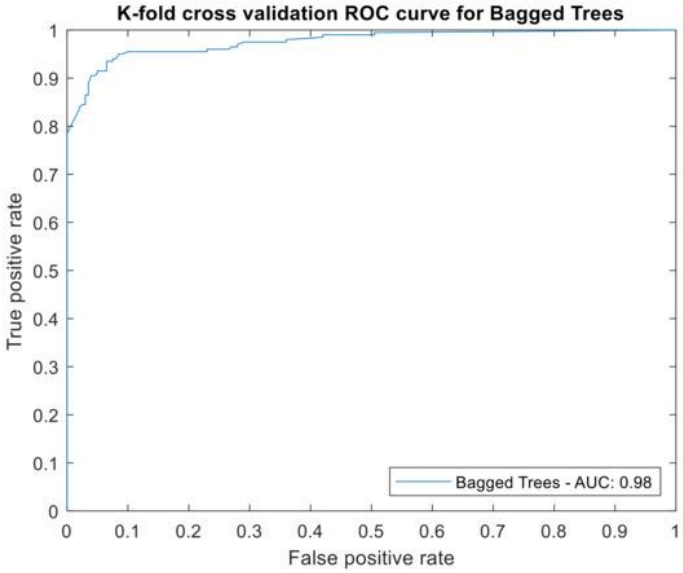


Figure 3. ROC validation curve for the cancer class generated by the bagged trees classifier.

Once the model is finalized, the classifier outputs the Statistical Risk, an estimate of a new patient's probability of having breast cancer. Although internally defined between zero and one, it is rescaled for convenience to span from zero to one hundred. Higher values correspond to a greater estimated likelihood of breast cancer.

Determination of the global risk

After obtaining both the Symbolic Risk and the Statistical Risk, these values are combined through Eq. 1, as explained in the conceptual outline of the system. This combination produces a new indicator—referred to as the Global Risk—that integrates both risk components. Its scale extends from zero to one hundred and, as with the previous metrics, it reflects the probability of developing breast cancer. Unlike individual indicators, this aggregated value incorporates patient information from multiple angles, merging symbolic and statistical assessments into a unified estimate.

Determination of the corrected global risk

As stated earlier, the BI-RADS score was intentionally excluded from the initial risk computation. Nevertheless, it remains a clinically significant element, and therefore, once the Global Risk has been generated, it is adjusted according to the BI-RADS category assigned by the clinicians. The purpose of this adjustment is not to enhance the intrinsic predictive accuracy—already addressed by the Global Risk itself—but rather to align the final output with the conventional diagnostic workflow that relies heavily on BI-RADS. Since medical referrals frequently

hinge on the BI-RADS classification, the system incorporates this influence by producing a corrected risk value that more closely reflects everyday clinical decision-making, even when it differs from the original prediction. To achieve this, a set of empirically derived formulas was established, enabling the Global Risk to be modified in a way that reflects the diagnostic tendencies associated with BI-RADS categories. **Table 4** illustrates the correspondence between BI-RADS levels and the expressions used to compute the Corrected Global Risk.

Generation of warnings and decision making

Once the patient data are processed and both the Global Risk and the Corrected Global Risk have been determined using the BI-RADS level assigned, the system proceeds to assess the risk in order to define the patient's status and generate recommendations to support clinical decision-making. Three possible outcomes are considered, each with corresponding threshold intervals, as summarized in **Table 8**. It should be emphasized that these criteria may be updated in future revisions, depending on clinical validation results and the judgment of medical experts.

Table 8. System states and threshold levels.

State	Threshold
Healthy case	Corrected Global Risk < 40
Dubious case	40 ≤ Corrected Global Risk < 60
Potential case	Corrected Global Risk ≥ 60

It is important to highlight that the thresholds provided here are illustrative, meant to guide an explicit diagnostic step and accommodate the optional Global Risk adjustment. As previously noted, the core functionality of the system already includes classification, based on the Global Risk, separating patients into two groups: potentially cancerous or non-cancerous. However, once a BI-RADS-based correction is introduced, this interpretation becomes less sharply defined, serving instead to align the system's output with the thought process of clinicians. For this reason, maintaining suggested risk levels is considered beneficial, as it better mirrors the physician's perspective than the internal inferential classification method.

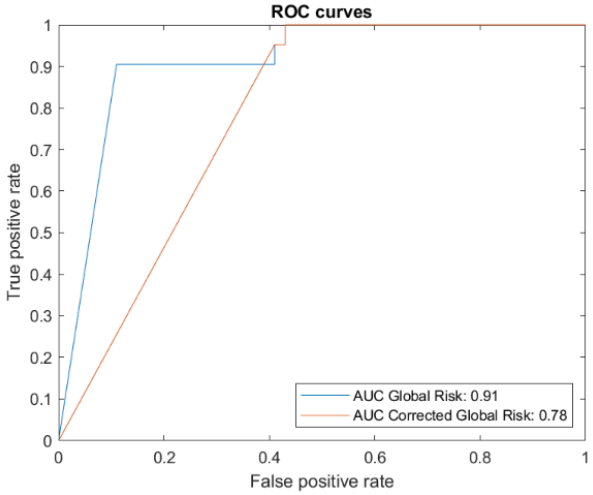
Analysis of results

After implementing the system, a proof of concept [45] was conducted using the labeled dataset to verify the proper operation of the software prototype and to estimate its diagnostic performance. Functionally, the system behaves as a binary classifier [46, 47] that assigns patients to one of two classes—"cancer" or "non-cancer." This classification should be evident both in the Global Risk computation and in the Corrected Global Risk estimation, recognizing that the two metrics rely on different principles: the former is derived from inferential processing of patient data, while the latter is an analytical adjustment aligned with clinical practice using the BI-RADS scale.

To evaluate the model's performance [46, 47], standard metrics were applied, including sensitivity (the ability to correctly identify individuals with the disease) and specificity (the ability to correctly identify individuals without the disease). Additionally, overall precision was quantified using the Matthews correlation coefficient (Mcc) [48-50]. Complementary measures were also incorporated, such as the false negative rate (cases incorrectly predicted as free of disease) and the false positive rate (cases incorrectly predicted as having the disease). Other metrics—like sensitivity per lesion, reflecting the proportion of detected tumors—were encompassed within the indicators previously mentioned. All results are compiled in **Table 9**.

Table 9. Results.

Metric	Global Risk	Corrected Global Risk
Sensitivity	90.5%	100%
False negative rate	9.52%	0%
Specificity	89.81%	60.19%
False positive rate	10.19%	39.81%
Mcc	0.7	0.44
AUC	0.91	0.78



Examining the results listed in **Table 9**, several observations can be drawn. For the Global Risk computed over the study dataset, both sensitivity (90.5%) and specificity (89.81%) show high values, which correspond to an elevated Matthews correlation coefficient (0.7). These findings suggest that, within this proof of concept, the system exhibits strong and distinctive predictive behavior, owing to the combined use of symbolic and statistical inference. The capacity to fuse knowledge-based reasoning with statistical modeling clearly enhances the classifier's effectiveness.

On the other hand, when the Corrected Global Risk is used to obtain performance metrics, the system produces a sensitivity of 100%**,** a specificity of 60.19%**,** and a Matthews correlation coefficient of 0.44**.** These values indicate a decline in the classifier's discriminative power. This reduction is not unexpected: as previously clarified, the correction step was never designed to outperform the original predictive output but to reshape it so that the system behaves in a manner consistent with conventional diagnostic routines. In practical terms, the intention is for the tool to act in a way that clinicians can naturally interpret. With this adjustment, the model's decisions essentially follow the BI-RADS scheme, guiding subsequent evaluations and ensuring that every genuine case is identified with 100% sensitivity—even if this increases the number of patients subjected to additional, ultimately unnecessary, confirmatory procedures. The obtained results clearly demonstrate this behavior.

Accordingly, the intelligent system displays two distinct yet clinically valuable qualities. First, it achieves strong predictive performance when applied to the study dataset. Second, it can reshape those predictions so that they mirror the standard approach of breast-cancer diagnostic teams, ensuring that no potentially suspicious case is overlooked. The proof-of-concept implementation, therefore, confirms both its usability and its intended purpose. Still, it must be emphasized that these outcomes represent only an initial demonstration based on the study dataset. As would reasonably be expected, integrating new data—with greater uncertainty and lower consistency—may reduce predictive accuracy, although the correction method maintains the guarantee of detection.

Results and Discussion

The goal of this section is not to validate the system, but to provide an illustrative example showing how it operates in practice, highlighting both its possible clinical utility and its uncomplicated workflow. Because this work remains at the proof-of-concept stage, the following example simply walks through the procedure using information from a new patient whose data were **not** part of the model-training process.

Data ollection

Table 10 lists the main mammographic observations and the patient's general information. For context, this individual was later confirmed—through diagnostic testing—to have breast cancer, which allows comparison between the system's output and the actual diagnosis.

Table 10. Patient's data to be analyzed.

Mammogram	Type of finding	Characteristic	Value
Mass			

	Present/absent	Present
	Shape	Irregular
	Margins	Spiculated
	Density	Homogeneous
Calcifications		
	Present/absent	Present
	Primary/associated	Associated
	Shape	Coarse heterogeneous
	Distribution	Grouped
Asymmetry		
	Present/absent	Absent
	Type	-
Architectural Distortion		
	Present/absent	Absent
	Primary/associated	-
Breast density		Heterogeneously dense
BI-RADS category		4B
Other data		
	Age	65
	Patient history	No
	Family history	No

After entering the information into the application, the system processed it automatically.

Data processing

Once submitted, the application calculated the Symbolic Risk indicators and the Statistical Risk indicator. The Symbolic Risk values for R1**,** R2, and R3 were 89.97%**,** 99.98%**,** and 0%, respectively, and the Statistical Risk corresponded to 25.61%**.**

These values were then combined according to Eq. 1. All Symbolic Risks were treated as equally weighted, except for the third one, which contributed nothing; thus, the weights were $\omega_1 = \omega_2 = 0.5$ and $\omega_3 = 0$. Eq. 4 presents the computation of the Global Risk, which, due to surpassing its upper bound, resulted in a final value of 100:

$$R_G = (0.5 \cdot 89.97 + 0.5 \cdot 99.98 + 0) \cdot \log_{10} 25.61 \geq 100 \rightarrow R_G = 100 \quad (4)$$

Figure 4 shows the Stage #2 screen of the software with the risk values displayed.

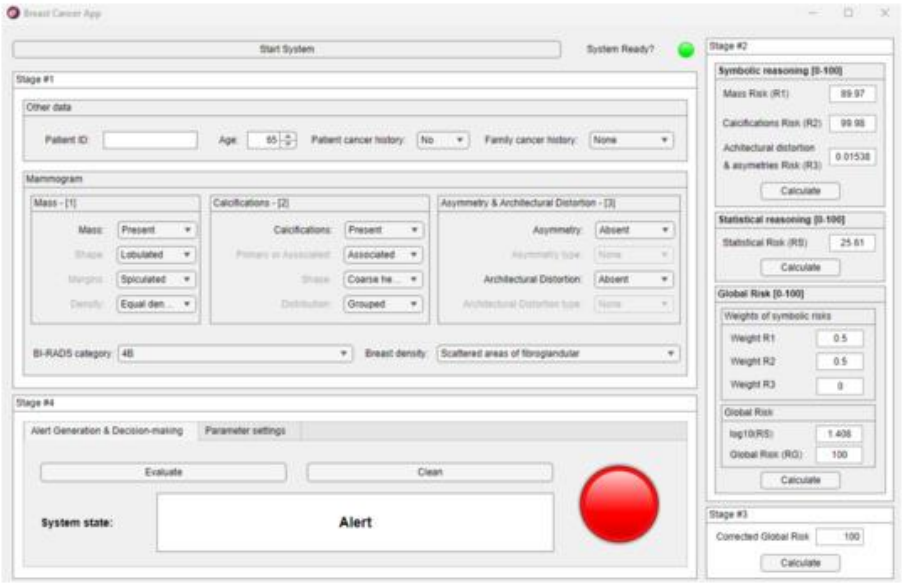


Figure 4. Screenshot of the software tool for the case study.

Global risk correction

The Global Risk obtained in the previous step was then corrected in Stage #3 using the BI-RADS level determined by a specialist. Since the assigned category was 4B and the Global Risk had already reached the maximum allowable score, the Corrected Global Risk also remained at 100. This value appears in the Stage #3 section of **Figure 4**.

Warning generation and decision making

The system then performed the final risk evaluation. Because the Corrected Global Risk exceeded the second threshold, the application issued a maximum-alert status, as shown in the Stage #4 display of **Figure 4**. The medical team was advised to conduct confirmatory testing. In this particular case, the correction step simply reinforced what the inferential process had already indicated: the patient should be considered at high risk for breast cancer. Although the BI-RADS-based correction was not strictly necessary here, it supports alignment with established diagnostic practice.

Ultimately, the system's recommendation matched the patient's confirmed clinical condition.

Discussion

Breast cancer has become the most frequently diagnosed malignant disease worldwide, surpassing lung cancer and remaining a principal cause of mortality among women. Detecting it early—primarily through large-scale screening initiatives—is essential for reducing its overall burden. These programs rely on mammographic examinations targeting populations at elevated risk so that possible tumors can be identified and treated promptly, thereby improving diagnostic outcomes and lowering fatality rates. During the interpretation of mammograms, clinicians typically employ the BI-RADS framework, which provides standardized terminology and a set of categories describing the likelihood of malignancy. Still, the final assessment inevitably shows variability tied to the training, expertise, and subjective interpretation of the radiologists involved, impacting diagnostic decisions and subsequent clinical actions.

This study introduces an updated intelligent clinical decision-support tool for breast cancer, building upon the system architecture previously presented by the authors in Casal-Guisande *et al*. [25]. The goal was to reshape the existing design so that it aligns more closely with the workflow of healthcare teams, thereby easing future clinical validation and refining system effectiveness. The underlying knowledge bases were reorganized to suit clinicians' needs, streamlining the processing stages and adapting the tool to real medical practice.

Diagnosing breast cancer requires integrating several variables to determine whether the patient is affected. Traditional solutions—common in contemporary research—typically rely on a single reasoning paradigm, either symbolic or statistical. Consistent with the authors' recent publications [17, 25, 27], this work instead embraces the joint use of heterogeneous reasoning techniques. Specifically, several fuzzy inference modules operate alongside a machine-learning classifier, each capturing complementary perspectives of the diagnostic process. Unlike the earlier architecture proposed in Casal-Guisande *et al*. [25], where symbolic and statistical components worked sequentially, the redesigned system runs these inferential mechanisms simultaneously. With this concurrent configuration, multiple independent risk indicators—namely, the Symbolic Risks and the Statistical Risk—can be obtained in parallel, each estimating the probability that the patient has breast cancer.

This dual-path inferential design represents a substantial shift from earlier approaches, where all knowledge was fed through a single reasoning chain. While each risk could be analyzed independently to form an impression of a patient's condition, relying exclusively on isolated engines would mean privileging one inferential perspective over the others. The concurrent strategy used here prioritizes the integration of symbolic and statistical outputs by means of an empirical formulation, Eq. 1, which produces a combined metric referred to as the Global Risk**.** This value results from a weighted aggregation of the Symbolic Risks—associated with groups of mammographic findings such as masses, calcifications, asymmetries, and architectural distortions—allowing clinical teams to modulate how much influence each category should exert on the final assessment. The ability to tailor these weightings is one of the system's most flexible and distinctive characteristics, enabling adaptation to different clinical preferences or institutional protocols.

In addition, the Statistical Risk plays a multiplicative role: it amplifies the aggregated Symbolic Risks and raises the danger level for cases that exhibit patterns similar to those seen in the machine-learning model's cancer-positive training examples. Through this mechanism, the system unifies distinct risk types in a straightforward

manner and embeds an extra diagnostic layer into the reasoning model, thereby elevating the risk level in situations that mirror past malignant cases.

A major strength of these inferential frameworks is their capacity for ongoing knowledge expansion. New information—whether codified explicitly or provided in the form of datasets—can be incorporated into the knowledge base, analyzed, and subsequently applied in the prediction stage.

It is worth noting that the BI-RADS category assigned by clinicians in this study does not participate directly in the symbolic or statistical inferential engines, unlike in the authors' prior work. Instead, the BI-RADS level is employed afterward to adjust the previously computed Global Risk through a dedicated analytical function. This enables increases or decreases in the final risk estimate depending on clinical judgment.

Consequently, the approach proposed here is not merely a modification of previously published diagnostic methodologies; it represents a complete reconceptualization. Transitioning from a stepwise propagation of information to a concurrent interpretation of knowledge marks a considerable advance. While sequential inferential pipelines ensure that information is conserved throughout the reasoning flow, concurrent engines — all driven by the same initial data — yield a broader interpretive spectrum while still maintaining information preservation. Any inferential model that aims to produce a diagnosis ought to produce stable, reproducible outputs when given the same inputs; large fluctuations would imply uncontrolled uncertainty. In sequential architectures, such variance may be mitigated through certain computational strategies, but concurrent reasoning requires that the multiple outputs be coherently interpreted as part of a consolidated decision-making process.

Said interpretation—subject to unavoidable variation—must initially rely on analytical formulations capable of aligning predicted values with their actual counterparts. Using such analytical constructs tempers the influence of broad AI-driven reasoning by simplifying, and in some situations nearly linearizing, the deductive stages involved. However, the intention here is not to derive an analytic predictor, but to formulate an expression that enables the different predictive components to be merged. Subsequent empirical evaluation attempts to confine the uncertainty generated by this variability by applying elementary mathematical relationships, which may appear counterintuitive at first glance. Although probabilistic schemes, Bayesian techniques, or fuzzy-based uncertainty treatment could slightly strengthen the stability of these formulations, they would not fundamentally change the situation. This is because realistic associations between the certainty levels of the explanatory variables—in this case, the risks—and the resulting outcome (presence or absence of cancer) cannot be reliably established.

Only after amassing a large collection of labeled triads (variables and corresponding diagnostic labels) could multivariate examinations be carried out to uncover latent causal structures. Such an endeavor, however, is infeasible within the conceptual framework presented here. This characteristic represents both an innovation and one of the clearest limitations of the current model.

All previously discussed circumstances highlight a decisive departure from the authors' earlier work and simultaneously constitute an initial contribution in this research domain.

Beyond these considerations, the conceptual foundation of the intelligent system—built on symbolic inference (expert systems) and statistical inference (a classification algorithm)—enhances and streamlines decision-making. Typically, this type of knowledge is embedded within specialized clinical teams; the tool, therefore, helps unify diagnostic procedures among professionals with differing backgrounds. Additionally, the system implicitly handles uncertainty in its epistemic and random components and acknowledges ambiguity and interaction effects [51, 52]. It achieves this through the combined use of the symbolic and statistical frameworks mentioned earlier, an increasingly widespread strategy in modern intelligent-system architectures.

Apart from these advantages, implementing an intelligent tool of this kind yields notable benefits in medical diagnostics. Producing a percentile-based breast-cancer risk measurement provides clinicians with a meaningful indicator that aids the challenging task of early case identification. This, in turn, lowers false-positive and false-negative counts during screening, thereby reducing costs and enhancing service quality. While minimizing misclassifications helps prevent unnecessary examinations and the associated psychological and logistical burdens on patients, the additional step of applying the Global Risk correction gives the system a distinctive empirical dimension. Thus, although high predictive accuracy is central to any decision-support system, incorporating analytic and empirical adjustments brings the model closer to how real clinical reasoning unfolds. When required, medical teams can choose to prioritize sensitivity by elevating detection rates—even when the system itself categorizes a patient as healthy but the BI-RADS assessment suggests otherwise. The corrective mechanism does not refine the classifier but rather increases clinicians' confidence, since their evaluative criteria are reflected in the final outcome, even when these criteria adopt a conservative stance. Minor inconsistencies may appear in

borderline cases due to this correction, but the medical team can always review the original prediction, compare both values, and reflect on the reasoning process.

At the same time, the system faces limitations beyond the diversification constraints introduced by the empirical correction. The foremost challenge lies in articulating and constructing the knowledge base needed for expert-system rules, which can become a significant hurdle when formulating the declarative structures required for fuzzy inference. Additionally, the concurrent use of symbolic and statistical reasoning—although less reliant on explicit knowledge formalization—means the resulting predictions are difficult to justify through an interpretable chain of reasoning. While hybrid or mixed-model inference can substantially increase diagnostic accuracy, the transparency provided by explicit, structured knowledge remains a key requirement in many clinical contexts.

Conclusion

This work has presented a new intelligent system intended to assist in the diagnostic workflow for breast cancer. The system integrates symbolic reasoning (expert systems) and statistical reasoning (a machine-learning classifier) operating simultaneously to compute a percentage-based risk estimate reflecting the probability of developing breast cancer. Its novelty lies not only in the joint deployment of both inferential paradigms but also in the manner in which they are merged. This fusion is accomplished using an empirical formulation that accommodates the characteristics of the underlying data, along with a subsequent adjustment of the risk metric through a BI-RADS-based weighting factor, ensuring clinical judgment is incorporated into the recommendation. The system has been implemented as a software application using data from the School of Medicine and Public Health at the University of Wisconsin–Madison. A practical case demonstrated its usability, straightforward operation, and potential for broader deployment once fully validated. At present, the system is undergoing refinement and adaptation, with the goal of completing clinical validation in the near future to confirm its reliability and diagnostic strength.

Acknowledgments: None

Conflict of Interest: Yamai reports honoraria for lectures from Taiho Pharmaceutical and Yakult Honsha. Ikezawa honoraria for lectures from Taiho Pharmaceutical, Yakult Honsha, Ono Pharmaceutical, MSD and Incyte Biosciences Japan, and research funding from ASKA Pharmaceutical. Sugimoto reports honoraria for lectures from Chugai Pharm, Daiichi Sankyo, Ono and Eli Lilly, and research funding from Chugai Pharm, Daiichi Sankyo, MSD Pfizer, and Seagen. Takada reports honorarias for lecture from Hisamitsu Pharmaceutical, Novartis and TEIJIN PHARMA. Kunimasa reports honoraria for lectures from Chugai Pharma, and Novartis Pharma. Ohkawa reports honoraria for lectures from Eisai, Chugai Pharmaceutical, Yakult Honsha, Incyte Biosciences Japan, Takeda, Gilead and Hisamitsu, and research grants from Towa Pharmaceutical and Sumitomo Chemical. The other authors have no conflict of interest.

Financial Support: None

Ethics Statement: This study was conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the Ethical Review Committee of the Osaka International Cancer Institute (No. 20148-3, approved on 12 October 2020).

The requirement for informed consent was waived by the opt-out method of our hospital's website. A waiver of informed consent was granted by institutional review board of the Institutional Review Board of Osaka International Cancer Institute.

References

1. Arnold M, Morgan E, Rumgay H, Mafra A, Singh D, Laversanne M, et al. Current and future burden of breast cancer: global statistics for 2020 and 2040. Breast. 2022;66(1):15–23.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates. CA Cancer J Clin. 2021;71(2):209–49.
3. Wilkinson L, Gathani T. Understanding breast cancer as a global health concern. Br J Radiol. 2021;95(1130):20211033.

4. Sickles EA, D'Orsi CJ, Bassett LW, Appleton CM, Berg WA, Burnside ES, et al. BI-RADS del ACR: mamografía. In: Atlas BI-RADS del ACR. Reston (VA): American College of Radiology; 2013.
5. Elmore JG, Jackson SL, Abraham L, Miglioretti DL, Carney PA, Geller BM, et al. Variability in interpretive performance at screening mammography. Radiology. 2009;253(3):641–51.
6. Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in radiologists' interpretations of mammograms. N Engl J Med. 1994;331(22):1493–9.
7. Beam CA. Variability in the interpretation of screening mammograms by US radiologists. Arch Intern Med. 1996;156(2):209–16.
8. Anooj PK. Clinical decision support system: risk prediction of heart disease using fuzzy rules. J King Saud Univ Comput Inf Sci. 2012;24(1):27–40.
9. Montgomery AA, Fahey T, Peters TJ, MacIntosh C, Sharp DJ. Evaluation of computer-based CDSS for hypertension. BMJ. 2000;320(7236):686–90.
10. Hermsen ED, VanSchooneveld TC, Sayles H, Rupp ME. Implementation of a CDSS for antimicrobial stewardship. Infect Control Hosp Epidemiol. 2012;33(4):412–5.
11. Uzoka FME, Osuji J, Obot O. CDSS in malaria diagnosis: comparison of soft computing methods. Expert Syst Appl. 2011;38(2):1537–55.
12. Casal-Guisande M, Comesaña-Campos A, Cerqueiro-Pequeño J, Bouza-Rodríguez JB. Expert-system-based methodology for pressure ulcer treatment. Diagnostics. 2020;10(9):614.
13. Comesaña-Campos A, Casal-Guisande M, Cerqueiro-Pequeño J, Bouza-Rodríguez JB. Expert-system methodology for early detection of hypoxemia. Int J Environ Res Public Health. 2020;17(23):8644.
14. Casal-Guisande M, Cerqueiro-Pequeño J, Comesaña-Campos A, Bouza-Rodríguez JB. Methodology for diabetic foot treatment. In: Proc Int Conf TEEM. 2020. p. 491–5.
15. Berkan Sesen M, Nicholson AE, Banares-Alcantara R, Kadir T, Brady M. Bayesian networks for CDSS in lung cancer. PLoS One. 2013;8(12):e82349.
16. Wagholikar KB, MacLaughlin KL, Henry MR, Greenes RA, Hankey RA, Liu H, et al. CDSS with text processing for cervical cancer screening. J Am Med Inform Assoc. 2012;19(5):833–9.
17. Casal-Guisande M, Torres-Durán M, Mosteiro-Añón M, Cerqueiro-Pequeño J, Bouza-Rodríguez JB, Fernández-Villar A, et al. Intelligent CDSS for obstructive sleep apnea diagnosis. Int J Environ Res Public Health. 2023;20(4):3627.
18. Sesen MB, Peake MD, Banares-Alcantara R, Tse D, Kadir T, Stanley R, et al. Lung Cancer Assistant: hybrid CDSS. J R Soc Interface. 2014;11(91):20140534.
19. Cooley ME, Blonquist TM, Catalano PJ, Lobach DF, Halpenny B, McCorkle R, et al. Feasibility of algorithm-based CDSS in lung cancer. J Pain Symptom Manage. 2015;49(1):13–26.
20. Fernandes AS, Alves P, Jarman IH, Etchells TA, Fonseca JM, Lisboa PJG. CDSS for breast cancer patients. IFIP Adv Inf Commun Technol. 2010;314(1):122–9.
21. Ferreira P, Dutra I, Salvini R, Burnside E. Interpretable models to predict breast cancer. In: IEEE BIBM. 2016. p. 1507–11.
22. Alaa AM, Moon KH, Hsu W, van der Schaar M. ConfidentCare: CDSS for personalized BC screening. IEEE Trans Multimed. 2016;18(10):1942–55.
23. Jiang X, Wells A, Brufsky A, Neapolitan R. CDSS learned from data to prevent metastasis. PLoS One. 2019;14(3):e0213292.
24. Skevofilakas MT, Nikita KS, Templaleksis PH, Birbas KN, Kaklamanos IG, Bonatsos GN. DSS for BC treatment. IEEE Eng Med Biol. 2005;7(1):2429–32.
25. Casal-Guisande M, Comesaña-Campos A, Dutra I, Cerqueiro-Pequeño J, Bouza-Rodríguez JB. Intelligent CDSS for breast cancer risk. J Pers Med. 2022;12(2):169.
26. Cerqueiro-Pequeño J, Comesaña-Campos A, Casal-Guisande M, Bouza-Rodríguez JB. Expert-system methodology for radon risk prevention. Int J Environ Res Public Health. 2020;18(1):269.
27. Casal-Guisande M, Bouza-Rodríguez JB, Cerqueiro-Pequeño J, Comesaña-Campos A. Hybrid intelligent DSS for forest fire detection. Forests. 2023;14(2):172.
28. American College of Radiology. Breast imaging reporting & data system. 2021.
29. Mamdani EH. Advances in linguistic synthesis of fuzzy controllers. Int J Man Mach Stud. 1976;8(6):669–78.

30. Mamdani EH. Application of fuzzy logic to approximate reasoning. IEEE Trans Comput. 1977;C-26(12):1182–91.
31. Mamdani EH, Assilian S. Experiment in linguistic synthesis with fuzzy controller. Int J Man Mach Stud. 1975;7(1):1–13.
32. Wasserman L. All of Statistics: A Concise Course in Statistical Inference; Springer: New York, NY, USA, 2004; Volume 26, ISBN 978-0-387-21736-9.
33. Hevner AR, March ST, Park J, Ram S. Design Science in Information Systems Research. MIS Q. 2004;28:75–105.
34. Hevner AR, Chatterjee S. Design Research in Information Systems: Theory and Practice; Springer: New York, NY, USA, 2010; ISBN 978-1-4419-6107-5.
35. MATLAB App Designer—MATLAB & Simulink. Available online: https://es.mathworks.com/products/matlab/app-designer.html (accessed on 10 August 2022).
36. Fuzzy Logic Toolbox—MATLAB. Available online: https://www.mathworks.com/products/fuzzy-logic.html (accessed on 1 November 2022).
37. Classification Learner. Available online: https://www.mathworks.com/help/stats/classificationlearner-app.html (accessed on 18 October 2022).
38. Imbalanced-Learn. Available online: https://imbalanced-learn.org/dev/index.html (accessed on 18 October 2022).
39. Ross TJ. Fuzzy Logic with Engineering Applications, 3rd edit.; John Wiley & Sons, Ltd.: Chichester, UK, 2010; ISBN 9781119994374.
40. Agresti A. Categorical Data Analysis; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2002; ISBN 0471360937.
41. Powers D, Xie Y. Statistical methods for categorical data analysis. Emerald Group Publishing; 2008.
42. Mohammed AJ, Hassan MM, Kadir DH. Improving classification using SMOTE. Int J Adv Trends Comput Sci Eng. 2020;9(4):3161–72.
43. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE. J Artif Intell Res. 2002;16(1):321–57.
44. Refaeilzadeh P, Tang L, Liu H. Cross-validation. Encycl Database Syst. 2009;53(2):532–8.
45. U.S. House of Representatives Committee on Science and Technology. From the lab bench to the marketplace. Washington (DC): Government Printing Office; 2010.
46. Kumari R, Srivastava S. Machine learning: review on binary classification. Int J Comput Appl. 2017;160(1):11–5.
47. Duda RO, Hart PE, Stork DG. Pattern classification. Hoboken: Wiley; 2000.
48. Boughorbel S, Jarray F, El-Anbari M. Optimal classifier for imbalanced data using MCC. PLoS One. 2017;12(6):e0177678.
49. Chicco D, Jurman G. Advantages of MCC over F1. BMC Genomics. 2020;21(1):6.
50. Guilford JP. Psychometric methods. New York: McGraw-Hill; 1954.
51. Cox DR. Principles of statistical inference. Cambridge: Cambridge University Press; 2006.
52. Thunnissen DP. Propagating uncertainty in multidisciplinary systems design. Pasadena (CA): California Institute of Technology; 2005.