

## НЕЧІТКА БАЗА ЗНАНЬ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ ДІАГНОСТУВАННЯ ВИДІВ РАКУ МОЛОЧНОЇ ЗАЛОЗИ

*У статті розроблено правила діагностування раку молочної залози на основі аналізу цитологічних і гістологічних зображень, які описують форму, колір, текстуру окремих клітин, щільність розміщення клітин, форму груп клітин. В роботі запропонована база діагностичних знань на основі нечіткої логіки та механізм виводу висновку про діагноз при злоякісних новоутвореннях. Розроблені алгоритми можна застосувати в медичних системах підтримки прийняття рішень.*

*Ключові слова: діагностика, нечітка логіка, цитологія, гістологія, рак молочної залози.*

O.M. BEREZSKY, G.M. MELNYK, K.M. BEREZSKA

Ternopil National Economic University

### FUZZY KNOWLEDGE BASE OF BREAST CANCER DIAGNOSIS INTELLECTUAL SYSTEM

*Abstract - This paper describes the process of building of knowledge base of a fuzzy rule-based system for cancer diagnosis.*

*The production rules for cancer diagnosing are based on visual observation of cytological images. Those rules describe the shape, colour and texture of individual cells, density of cells, and form of cells groups. Visual image interpretation knowledge modelling and creation of diagnostic features is complex process. It consists of the following steps: description of microobjects and their qualitative characteristics, determination of numerical features of microobjects, determination of fuzzy variables and membership function, formalization of diagnosing rules, construction of fuzzy rules knowledge base.*

*The diagnostic knowledge base is developed. It based on fuzzy logic and knowledge derivation mechanism for the cancer diagnosis. Developed algorithms can be applied in medical decision support systems.*

*Keywords: diagnosis, fuzzy logic, fuzzy variables, region geometric features.*

### Вступ

В Україні у 2011 р. за даними Національного канцер-реєстру (НКР) [1] було зареєстровано 169 030 нових випадків захворювання на злоякісні новоутворення (ЗН). Загальний показник захворюваності для жіночого населення в цьому році зріс на 2,7% в порівнянні з попереднім. Загальний показник смертності серед населення зріс на 0,3%. Перше місце в структурі як захворюваності так і смертності серед жіночого населення продовжує посідати рак молочної залози (РМЗ) із питомою вагою в 20%.

Одним з найпоширеніших застосувань систем штучного інтелекту є системи медичної діагностики. Їх використання істотно спрощує роботу лікаря, допомагає йому уникнути власних помилок. Однак, оскільки комп'ютерна програма не може нести юридичної відповідальності за отримане з її допомогою рішення, то завжди вважається, що система діагностики виступає тільки в якості консультанта.

Досить велика кількість робіт присвячена використанню нечіткої логіки для постановки діагнозу РМЗ [2, 3, 4, 8], для формалізації термінів [5], інтерпретації набору правил діагностування [9].

У роботах [11, 12] досліджено програму OncoDoc. Це система підтримки прийняття рішень, яка дозволяє оперувати базою знань (БЗ) у вигляді дерева рішень.

Одним із напрямків діагностування є використання штучних нейронних мереж (ШНМ). У статті [6] застосовано самоорганізаційні карти Кохонена для вибору моделі, зменшення кількості змінних, визначення параметрів, а також для оцінки адекватності системи клінічних вимірювань. Достовірність діагностичних правил перевіряється шляхом попередньої обробки даних [13], навчанням ШНМ на основі даних сімейного амнезису раку [14], використанням методів інтелектуального аналізу даних (ШНМ і дерев рішень), а також статистичних методів [15], еволюційного підходу, базованого на Парето-диференціальній еволюції і доповненого алгоритмом локального пошуку для прогнозування РМЗ [16], пришвидшенням вибору правил за допомогою ШНМ [17]. У роботі [18] запропоновано гібридний підхід до побудови автоматизованих систем опрацювання біомедичних зображень, який полягає у поєднанні методів і алгоритмів аналізу зображень із методами інтелектуального опрацювання отриманих візуальних даних – побудови діагностичних баз знань і механізму логічного висновку.

### Постановка задачі

Метою статті є формалізація діагностичних знань про рак молочної залози на основі цитологічних і гістологічних зображень, розробки структури нечіткої бази знань, алгоритму нечіткого логічного виводу та проведення комп'ютерних експериментів.

#### 1. Формалізація діагностичних знань

У клінічній практиці основним методом раннього виявлення РМЗ є рентгенівське дослідження – мамографія. Наступним видом діагностування є забір клітин і тканин ураженого органу для цитологічного дослідження за допомогою мікроскопа, висновком якого є передопераційний діагноз, що служить основою для розробки тактики лікування пацієнта. Мікрооб'єктами на цитологічних зображеннях (ЦЗ) є клітини та їх складові.

Критерії цитологічної діагностики ЗН ґрунтуються насамперед на геометричних ознаках зображень клітини та їх ядер. Іншим критерієм діагностування є оцінка не однієї окремо взятої клітини, а сукупності клітин та фону препарату.

Правила діагностування раку на основі візуального спостереження зображення описують форму, колір, текстуру окремих клітин, їх розміщення одна відносно одної, щільність розміщення, форму груп клітин. Для опису форми груп клітин застосовуються спеціальні терміни, що походять від назв предметів подібних до них

візуально. Характерним для опису препарату в спеціальній літературі є вживання якісних термінів, що описують частоту появи певної ознаки, наприклад: "часто буває", "звичайно буває", "характерно", "деколи" та ін. Типове правило розміщення виглядає так [7]: клітини розташовуються переважно в скупченнях, щільних, пухких. Правило опису клітин може мати такий вигляд [7]: відзначається рясний клітинний склад, неправильна форма ядер, нерівні і нечіткі контури ядерної мембрани.

Моделювання знань на основі візуальної інтерпретації зображень і створення БЗ діагностичних ознак складний багатокроковий процес. Він має такі кроки: опис мікрооб'єктів та їх якісних ознак, визначення числових ознак мікрооб'єктів, визначення нечітких змінних та побудова функцій належності, формалізація правил діагностування ЗН, побудова бази нечітких правил.

Для опису мікрооб'єктів та їх якісних ознак потрібно формалізувати знання (досвід) експерта морфолога: логічні дерева рішень, мову опису клітин та діагнозу. Описові знання можна представити у вигляді спеціальної онтології опису клітин, яка містить всю інформацію про морфологію ракових клітин. Структурована та ієрархічна онтологія складається з усіх значних морфологічних характеристик, які розглядаються окремо. Для кожної характеристики (наприклад, розмір ядра) необхідно зафіксувати якісні значення категорій, які можна присвоїти цій ознаці (наприклад, дуже малий, малий, середній або великий). Морфологічні характеристики і значення повинні інтерпретуватися і "вимірюватися" інтуїтивно, так щоб онтологія застосовувалась і для автоматичного і для ручного морфологічних досліджень.

Онтологія ракових клітин і відповідні бази даних є результатом співпраці між групою експертів в даній предметній області, що гарантує правильність і відповідність структури термінології та еталонних зображень (навчальної вибірки). Як правило, робота по створенню онтології вимагає декілька ітерацій.

Для організації процесу створення навчальної вибірки розроблено інтерфейс за допомогою якого можна швидко створювати якісний опис клітин. Зображення клітин послідовного надаються експерту, який повинен вибрати відповідну якісну категорію для кожної ознаки. Такий інструмент використовується також під час перевірки системи. Оскільки кожний експерт має свої погляди на інтерпретацію якісних термінів, то потрібний крок узгодження їх знань з іншими фахівцями та медичною літературою.

## 2. Формалізація знань

Для формалізації знань з діагностики ракових процесів спершу потрібно навести перелік мікрооб'єктів на зображеннях та визначення термінів, що позначають об'єкти на зображеннях та їх якісні та кількісні ознаки.

### 2.1 Описові знання

Основними мікрооб'єктами інтересу на ЦЗ є клітина певної тканини органу, її складові частини, а також групи клітин. До другорядних мікрооб'єктів можна віднести клітини крові, макрофаги та інші міжклітинні включення (слиз).

Критерії цитологічної діагностики ЗН [7] ґрунтуються насамперед на ознаках клітини і ядра. Ознаки злоякісного процесу відображені в ознаках клітини так:

- збільшена в розмірах, іноді гігантська, рідко близька до норми;
- зміна форми клітинних елементів в порівнянні з нормальною клітиною;
- порушення співвідношення площі ядра і цитоплазми у бік збільшення частки ядра.

Ознаки злоякісного процесу відображені в ознаках ядра так:

- збільшення розміру в порівнянні з нормою;
- нерівномірний малюнок хроматину;
- найбільш постійна ознака - нерівність контурів.

Отже, одними з основних ознак, що потрібні для діагностування ЗН є ознаки розміру і форми клітини.

Мікрооб'єктами інтересу на гістологічних зображеннях (ГЗ) є: поперечний переріз протоки молочної залози, поперечний переріз дольки молочної залози. Дані мікрооб'єкти представлені структурами, які мають у своєму складі епітеліальні клітини. Злоякісний процес характеризується різким збільшенням розмірів та форми клітини, нагромадженням однотипних, атипичних клітин та їх низькою диференціацією.

### 2.2 Узагальнена структура нечіткої бази знань

Узагальнена структура нечіткої бази знань представлена на рис. 1. Вона складається із блоку обчислення числових характеристик ознак (БОЧХО), бази даних числових ознак (БДЧО), блоку фазифікації (БФ), бази правил діагностування (БПД), блоку функцій належності (БФН), блоку нечіткого логічного виводу (БНЛВ), блоку дефазифікації (БДФ) та блоку постановки діагнозу (БПДЗ).

Нечітка база знань (НБЗ) працює в двох режимах:

1. Накопичення знань,
2. Постановка попереднього діагнозу.

В режимі накопичення знань у БПД заносяться правила діагностування в нечіткій формі, які отримані від експертів (цитолога та гістолога). Крім цього БФН заповнюється конкретними параметрами для відповідних функцій належності. У режимі постановки діагнозу НБЗ працює так. На вхід БОЧХО поступають цитологічні

$\mathbf{Im}^C$  та гістологічні  $\mathbf{Im}^G$  зображення, в якому обчислюються відповідно цитологічні  $\bar{\mathbf{F}}^C$  та гістологічні  $\bar{\mathbf{F}}^G$  ознаки, що утворюють множину вхідних ознак  $\bar{\mathbf{F}}$ , яка записується в БДЧО. На основі числових характеристик ознак ідентифікуються функції належності  $\mu_{\bar{\mathbf{F}}}(u)$ . Використовуючи множину числових ознак  $\bar{\mathbf{F}}$  і їх функції належності, БФ утворює нечітку множину вхідних ознак  $\tilde{\mathbf{F}}$ . БНЛВ, спираючись на множину правил  $\bar{\mathbf{R}}$  та множину ознак  $\tilde{\mathbf{F}}$ , утворює нечітку множину діагнозів  $\tilde{\mathbf{D}}$ . БДФ переводить нечітку множину діагнозів  $\tilde{\mathbf{D}}$  у

множину діагностичних ознак  $\bar{F}_{\bar{D}}$  із відповідними ваговими коефіцієнтами кожного правила  $W_{\bar{D}}$ .

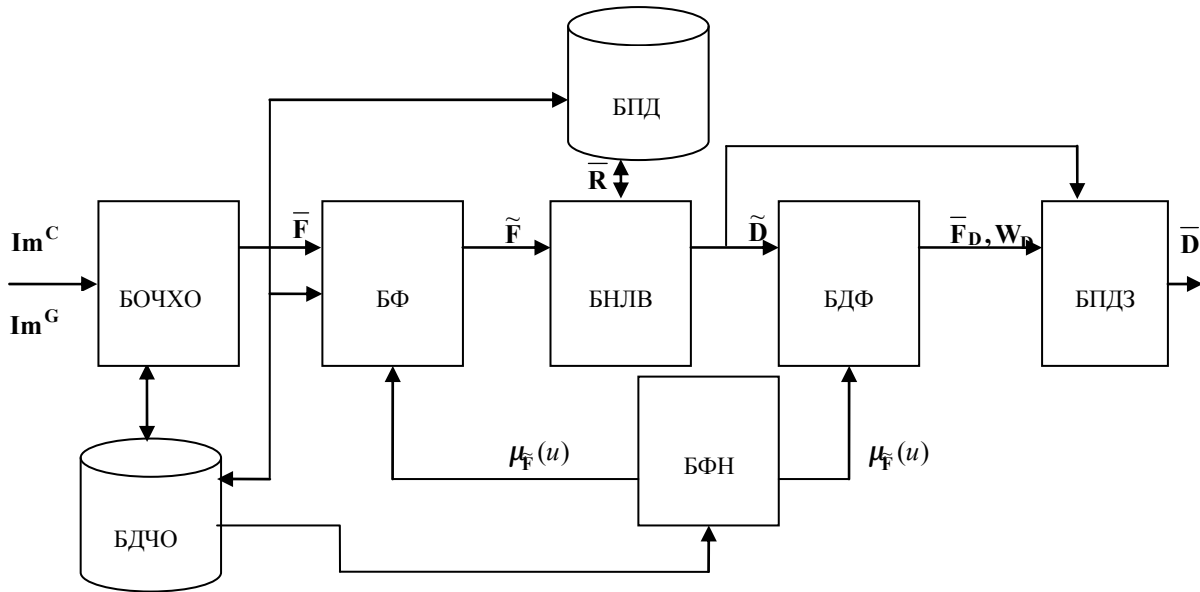


Рис. 1. Узагальнена структура нечіткої бази знань

- $\bar{F}$  – множина вхідних ознак;
- $\tilde{F}$  – нечітка множина вхідних ознак;
- $\bar{R}$  – множина правил;
- $\tilde{D}$  – нечітка множина діагнозів;
- $\bar{D}$  – множина діагнозів;
- $\bar{F}_{\bar{D}}$  – множина діагностичних ознак, що характеризують попередні діагнози;
- $W_{\bar{D}}$  – множина вагових коефіцієнтів діагнозів  $\bar{D}$ .

Отже, маючи множину нечітких діагнозів  $\tilde{D}$  і множини діагностичних ознак  $\bar{F}_{\bar{D}}$  з ваговими коефіцієнтами  $W_{\bar{D}}$ , БПД дає можливі варіанти діагнозів  $\bar{D}$ .

### 2.3 Продукційна модель представлення діагностичних знань

Для представлення діагностичних знань використаємо продукційну модель.

Як відомо, [19] в загальному випадку продукцією називається вираз такого вигляду:

$$(i): Q; V; P \Rightarrow D, N,$$

де  $i$  – ім'я продукції;  $Q$  – сфера застосування продукції.

$P \Rightarrow D$  – ядро продукції у вигляді правила IF  $P$  THEN  $D$ .

$V$  – умова застосування ядра продукції – предикат.

$N$  – постумова продукції.

В загальному випадку ядро продукції може мати складнішу конструкцію:

$$\text{IF } P \text{ THEN } D_1 \text{ ELSE } D_2.$$

Розглянемо випадки, коли  $P_1, P_2, \dots, P_n$  є нечіткими множинами. При цьому можливі такі конструкції побудови ядра продукції:

1. Складний кон'юнктивний антецедент:

$$\text{IF } X \text{ is } P_1 \text{ and } P_2 \text{ and } \dots \text{ and } P_n \text{ THEN } Y \text{ is } D^s.$$

Сумарна нечітка множина  $P^s$  рівна:  $P^s = P_1 \cap P_2 \cap \dots \cap P_n$ , а функція належності визначається:

$$\mu_{P^s}(x) = \min[\mu_{P_1}(x), \mu_{P_2}(x), \dots, \mu_{P_n}(x)].$$

2. Складний диз'юнктивний антецедент:

$$\text{IF } X \text{ is } P_1 \text{ or } P_2 \text{ or } \dots \text{ or } P_n \text{ THEN } Y \text{ is } D^s.$$

Сумарна нечітка множина  $P^s$  рівна:  $P^s = P_1 \cup P_2 \cup \dots \cup P_n$ , а функція належності визначається:

$$\mu_{P^s}(x) = \max[\mu_{P_1}(x), \mu_{P_2}(x), \dots, \mu_{P_n}(x)].$$

Ядро продукції у випадку нечіткої бази знань представимо так:

$$\text{IF } P_1 \text{ LO } P_2 \text{ LO } \dots \text{ LO } P_n \text{ THEN } D_1 \text{ LO } D_2 \text{ LO } \dots \text{ LO } D_k,$$

де  $P = \{P_1, P_2, \dots, P_n\}$  – множина елементних пропозицій;

$LO = \{I, АБО, НЕ\}$  – множина логічних операцій;

$\bar{D} = \{D_1, D_2, \dots, D_k\}$  – множина пропозицій, що утворюють рішення.

Будемо розглядати множину діагнозів як функцію ознак, що їх характеризують

$\bar{D} = g(f_1, f_2, \dots, f_m)$ , де  $f_1, f_2, \dots, f_m$  – ознаки діагнозу.

Дану залежність представимо у вигляді нечіткої бази знань [20].

$$\begin{aligned} & \text{IF } \left[ (f_1 = a_1^{k1}) \text{ AND } (f_2 = a_2^{k1}) \text{ AND } \dots \text{ AND } (f_m = a_m^{k1}) \right] (w_{k1}); \\ & \text{OR } \left[ (f_1 = a_1^{k2}) \text{ AND } (f_2 = a_2^{k2}) \text{ AND } \dots \text{ AND } (f_m = a_m^{k2}) \right] (w_{k2}); \\ & \quad \vdots \\ & \text{OR } \left[ (f_1 = a_1^{kp_k}) \text{ AND } (f_2 = a_2^{kp_k}) \text{ AND } \dots \text{ AND } (f_m = a_m^{kp_k}) \right] (w_{kp_k}). \\ & \text{THEN } D = d_k, \text{ для всіх } k = \overline{1, n}, \end{aligned}$$

- де  $n$  – кількість нечітких значень вихідної змінної  $D$ ;  
 $a_i^{kl}$  – нечіткий терм для  $f_i$  в правилі з номером  $kl$ ,  $k = \overline{1, n}$ ,  $l = \overline{1, p_k}$ ;  
 $d_k$  – нечіткий терм для оцінки вихідної змінної  $D$ ;  
 $p_k$  – кількість правил, в яких вихід оцінюється термом  $d_k$ ,  $k = \overline{1, n}$ ;  
 $w_{kl}$  – ваговий коефіцієнт  $kl$ -го правила,  $w_{kl} \in [0, 1]$ .

### 2.4 Нечіткий логічний вивід

Нечіткий логічний вивід ґрунтується на множині нечітких правил  $\bar{R}$  та множині ознак  $\tilde{F} = \langle \bar{F}^C, \bar{F}^G \rangle$ . В системі реалізований прямий і зворотний вивід.

Прямий вивід (рис.2) формується на основі правил  $\bar{R}^C$ , які базуються на ознаках цитологічних зображень  $\bar{F}^C$ . В результаті подачі конкретних значень ознак  $\bar{F}^C$  спрацьовують правила і вибирається діагноз з найбільшим ваговим коефіцієнтом  $W$  і таким чином формується діагноз  $D_1$ .

Цей діагноз підтверджується або заперечується на основі зворотного виводу. На вхід графу формування діагнозів поступає діагноз  $D_1$  відшукуються правила із множини правил  $\bar{R}^G$ , які підтверджують даний діагноз. Правила, які спрацьовують для підтвердження діагнозу  $D_1$  впливають із множини ознак гістологічних зображень  $\bar{F}^G$ . У разі відсутності правила  $\bar{R}_i^G$  даний діагноз ігнорується і коректуються правила  $\bar{R}^C$ , які привели до невірного діагнозу.

Степінь належності  $F^C = (f_1, f_2, \dots, f_n)$  до класу  $d_k$ ,  $d_k \in \bar{D}$  визначається так:

$$\mu_{d_k}(F^C) = \max_{l=1, p_k} w_{kl} \min_{i=1, m} (\mu_{kl}(f_i)), \quad k = \overline{1, n}.$$

Тоді наймовірнішим діагнозом буде діагноз з максимальним ступенем належності:

$$d_k = \arg \max_{\{d_1, d_2, \dots, d_m\}} (\mu_{d_1}(F^C), \mu_{d_2}(F^C), \dots, \mu_{d_m}(F^C)).$$

### 2.5 Аналіз структурної атипії гістологічних зображень

Для аналізу структурних патологічних змін у тканинах розроблено методи і алгоритми аналізу ГЗ на основі теорії кристалографічних груп [18]. Аналіз ГЗ складається із таких етапів: визначення відповідних точок на контурах мікрооб'єктів, обчислення афінних перетворень між мікрооб'єктами, визначення групи симетрії для групи мікрооб'єктів (частини зображення).

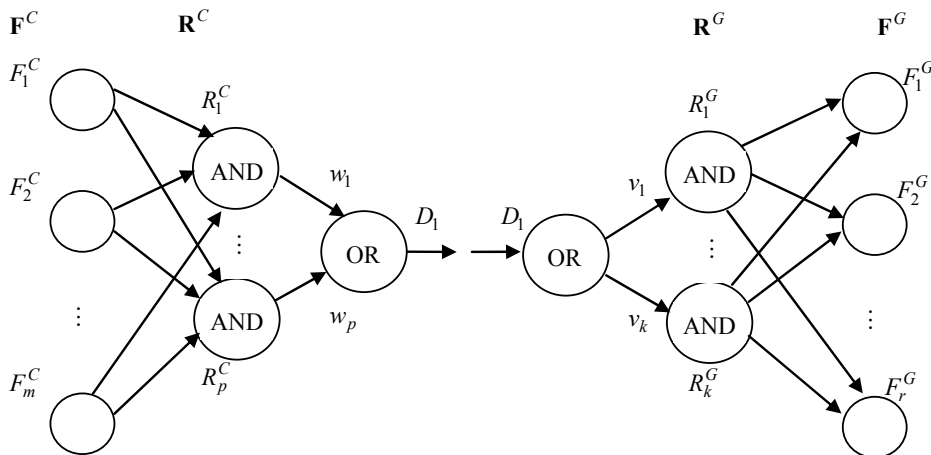


Рис. 2. Граф прямого і зворотного виводу діагнозу

Відповідні точки визначаються на основі дискретних центральних моментів області: обчислюється кут нахилу головної осі, координати центру інерції об'єкта. Для області визначаються три відповідні контурні точки, на основі яких визначаються коефіцієнти афінних перетворень.

Проведено експериментальне дослідження коефіцієнта структурної атипії, шляхом класифікації тренувальної вибірки ГЗ тканин епітеліального шару із відомим класом зображення, обчислено показники чутливості та специфічності (табл. 1). При цьому використано БД обсягом 400 зображень.

Таблиця 1

Коефіцієнт структурної атипії тканин			
Класи зображень	Коефіцієнт структурної атипії	Чутливість	Специфічність
Важка дисплазія	0,54-0,67	0,96	0,93
Помірна дисплазія	0,42-0,49	0,93	0,92
Нормальна тканина	0,25-0,41	0,95	0,94

### 3. Комп'ютерні експерименти

#### 3.1 Нечіткі змінні та функції належності

У співдружності з експертами кафедри патологічної анатомії з секційним курсом судової медицини Тернопільського державного медичного університету імені І.Я. Горбачевського сформована база даних гістологічних і цитологічних зображень РМЗ (близько 1500 зображень), на основі якої проведені комп'ютерні експерименти.

Дослідники [10] виділяють такі інформативні кількісні геометричні ознаки окремих клітин на ЦЗ:

- радіус (середнє значення відстаней від центру до точок контуру),
- периметр,
- площа,
- гладкість (локальне відхилення радіусів),
- компактність,
- увігнутість (величина точок увігнутості контуру),
- кількість точок увігнутості контуру,
- симетрія,
- фрактальна розмірність.

В якості нечітких змінних виділимо форму та розмір клітини. Для розміру клітини [7] задаємо таку термножину: дрібну, збільшену, гігантську. Для форми клітини [7] задаємо таку термножину: циліндричну, кубічну, округлу. Для оцінки розміру клітини використаємо площу області клітини. Для оцінки форми області використаємо ознаку округлості.

Для створення функцій належності на основі експертних знань та числових ознак мікрооб'єктів здійснено вимірювання розміру та ознак форми для нормальних тканин та двох видів раку. Дослідження проведені на навчальній БД загальним обсягом 1500 зображень. Метод фарбування зразків – гематоксилін-еозин, кратність збільшення об'єктива –  $\times 20$ . Для ілюстрації ідентифікації параметрів нормального закону розподілу значень числових ознак на рис. 3. приведено графіки густини функцій розподілу ймовірностей для ознаки розміру.

Результати вимірювань геометричних ознак мікрооб'єктів, виділених на початковій вибірці ЦЗ, підсумовано в табл. 2.

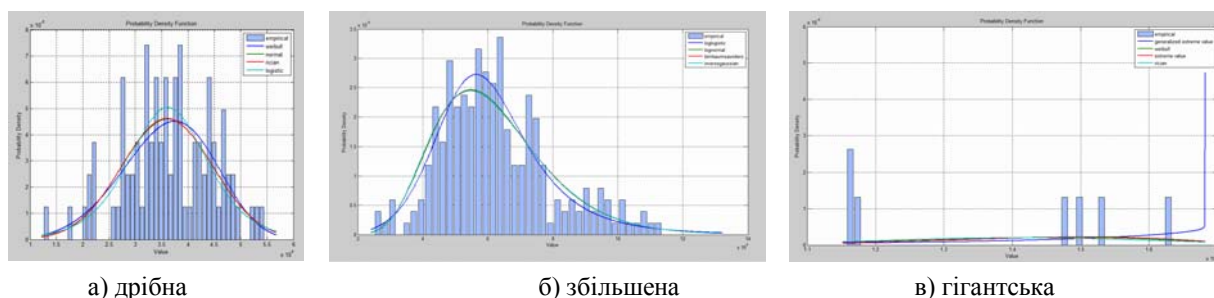


Рис. 3. Функції розподілу ймовірностей ознаки розміру

Таблиця 2

Ознака	Ознаки мікрооб'єктів		
	Терм-множина		
Розмір (піксели)	дрібна (small)	збільшена (enlarged)	гігантська (giant)
Мінімальне значення площі	12147	24026	115217
Максимальне значення площі	56727	132280	168202
Параметри нормального розподілу, $M, \sigma$	3.6072e+004, 8.6834e+003	6.1816e+004, 1.7714e+004	1.4098e+005, 2.1748e+004
Форма	циліндрична	кубічна	округла
Мінімальне значення округлості	0,384	0,424	0,763
Максимальне значення округлості	1,041	0,876	1,085
Параметри нормального розподілу, $M, \sigma$	0.2072, 0.006	0.71, 0.07	0.94, 0.11

Таблиця 3

## Параметри нечітких множин

Нечітка множина	Кількісна ознака області	Носій	Терм-множина	Ядро
Розмір клітини	Площа	[12147, 168202]	Дрібна	[12147, 24026]
			Збільшена	[56727, 115217]
			Гігантська	[132280, 168202]
Форма клітини	Ознака форми	[0.384, 1.085]	Циліндрична	[0.384, 0.512]
			Кубічна	[0.721, 0.74]
			Округла	[0.953, 1.085]

Побудовано функції належності. На основі експериментальних даних (табл. 2), за допомогою засобу Fuzzy Logic Toolbox програмного комплексу Matlab побудовано модель із обраними нечіткими змінними та функціями належності, що зображено на рис. 4.

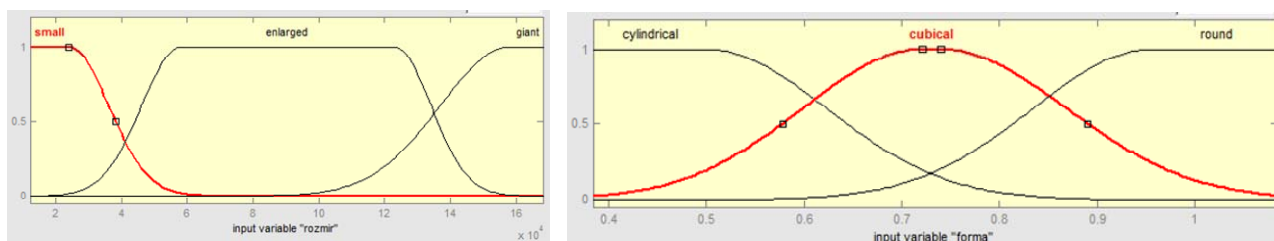


Рис. 4. Функції належності розроблених нечітких змінних

## 3.2 Правила діагностування

Для діагностування ЗН на основі ЦЗ знання експертів формалізовано у вигляді правил виводу про наявність злоякісного процесу на основі розглянутих ознак розміру та форми клітини. Результат формалізації представлено у табл. 4.

Таблиця 4

## Моделювання правил діагностування злоякісних новоутворень

Зображення	Розмір клітини	Форма клітини
Нормальна тканина	дрібний	кубічна циліндрична
Рак узагальнений	збільшений іноді гігантський	циліндрична: часто кубічна: переважно

Приведемо приклад експериментальної БЗ діагностичних правил. Вихідна змінна правила вказує на наявність раку (злоякісного процесу). У запропонованій БЗ вихідна змінна складається із двох нечітких множин "нормальна тканина" (normal), "рак" (cancer). Для вихідної змінної параметри функції належності наведено у табл. 5.

Таблиця 5

## Вихідна змінна "діагноз" (diagnosis)

Значення	Носій	Ядро
нормальна тканина (normal)	[0, 1]	[0, 0.24]
рак (cancer)		[0.8, 1]

Приведемо фрагменти правила діагностичних ознак на основі досліджень ЦЗ.

Фрагмент правила для класифікації нормальних клітин виглядає так:

IF розмір\_клітини=дрібний AND форма\_літини=кубічна TO діагноз=нормальна

OR розмір\_клітини=дрібний AND форма\_літини=кругла

THEN діагноз =норма.

Фрагмент правила для класифікації раку такий:

IF розмір\_клітини= збільшений AND форма\_літини= циліндрична

OR розмір\_клітини= збільшений AND форма\_літини= кубічна

OR розмір\_клітини= гігантський AND форма\_літини= кубічна

OR розмір\_клітини= гігантський AND форма\_літини= циліндрична

OR розмір\_клітини= гігантський AND форма\_літини= округла

THEN діагноз =рак.

На рис. 5 наведено приклад моделювання виводу висновку на основі створеної БЗ у середовищі Matlab.

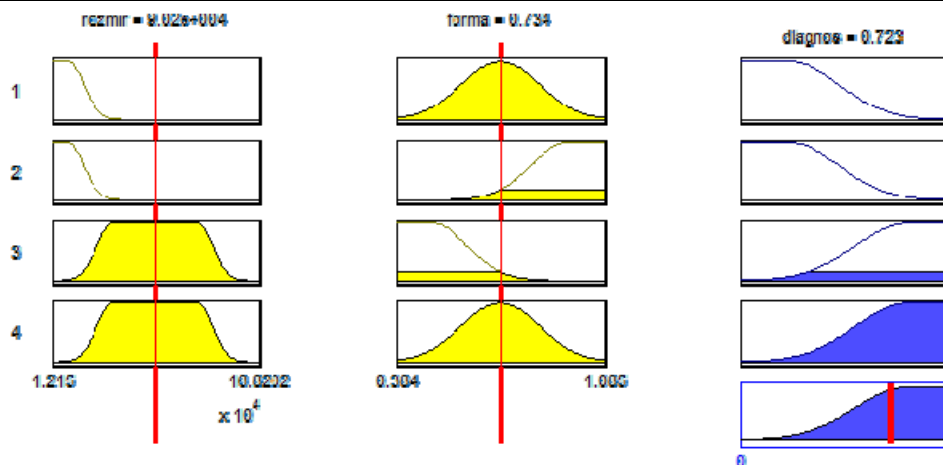


Рис. 5. Моделювання бази знань нечітких правил

Гістологічні дослідження дали змогу сформувані такі фрагменти правил.

Фрагмент правила для класифікації нормальної тканини виглядає так:

IF просвіт\_протоки=вільний\_на 70%

OR епітелій\_стінки\_протоки=однорядний

THEN діагноз = норма.

Фрагмент правила для класифікації раку такий:

IF епітелій\_стінки\_протоки=багаторядний AND ЯЦВ=більше 0,5

OR просвіт\_протока=вільний AND епітелій\_стінки\_протоки= багаторядний

OR просвіт\_протоки=повністю\_заповнений\_солідною\_пробкою

THEN діагноз = рак.

### Висновки

В роботі приведені результати експериментальних досліджень видів раку молочної залози на основі аналізу гістологічних і цитологічних досліджень, здійснена формалізація якісних знань експертів, розроблена структура бази знань на основі нечіткої продукційної моделі, запропоновано алгоритм логічного виводу діагнозу та проведено комп'ютерні експерименти. Робота виконана в межах державної бюджетної теми "Інтелектуальна система для діагностування різних форм раку молочної залози на основі аналізу гістологічних та цитологічних зображень", державний реєстраційний номер 0112U000736.

### Література

1. Рак в Україні, 2011-2012: Бюлетень Національного канцер-реєстру – № 14 – 203 с.
2. Pena-Reyes C.A. A fuzzy-genetic approach to breast cancer diagnosis / Carlos Andrés Pena-Reyes, Moshe Sipper Moshe Sipper // Artificial intelligence in medicine – № 17. – 1999. – P. 131–155.
3. Mahfouf M. A survey of fuzzy logic monitoring and control utilisation in medicine / M. Mahfouf, M.F Abbod, D.A Linkens // Artificial intelligence in medicine – № 21. – 2001. – P. 27–42.
4. Soria D. A quantifier-based fuzzy classification system for breast cancer patients / Daniele Soria, Jonathan M. Garibaldi, Andrew R. Green // Artificial intelligence in medicine – № 58. – 2013. – P. 175–184.
5. Kovalerchuk B. Fuzzy logic in computer-aided breast cancer diagnosis: analysis of lobulation / B. Kovalerchuk, E. Triantaphyllou, J. F. Ruiz // Artificial intelligence in medicine – № 11. – 1997. – P. 75–85.
6. West D. Model selection for a medical diagnostic decision support system: a breast cancer detection case / David West, Vivian West // Artificial intelligence in medicine – № 20. – 2000. – P. 183–204.
7. Шабалова И. П. Цитологический атлас. Диагностика заболеваний молочной железы / И.П. Шабалова, Т.В. Джангирова, Н.Н. Волченко, К.К.Пугачев – М. – Тверь: ООО «Издательство «Триада», 2005. – 119 с.
8. Gadaras I. An interpretable fuzzy rule-based classification methodology for medical diagnosis / Ioannis Gadaras, Ludmil Mikhailov // Artificial intelligence in medicine – № 47. – 2009. – P. 25–41.
9. Soria D. A quantifier-based fuzzy classification system for breast cancer patients / Daniele Soria, Jonathan M. Garibaldi, Andrew R. Green, Desmond G. Powe, Christopher C. Nolan, Christophe Lemetre, Graham R. Ball, Ian O. Ellis // Artificial intelligence in medicine – № 58. – 2013. – P. 175–184.
10. Wolberg W. H. Computer-derived nuclear features distinguish malignant from benign breast cytology / W. H. Wolberg, W. N. Street, D. M. Heisey, O. L. Mangasarian // Human Pathology – 1995. – № 26. – P. 792–796
11. Séroussi B. OncoDoc: a successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer / B. Séroussi, J. Bouaud, E.-C. Antoine // Artificial intelligence in medicine – № 18. – 2000. – P. 205–219.
12. Séroussi B. Using OncoDoc as a computer-based eligibility screening system to improve accrual onto breast cancer clinical trials / Brigitte Séroussi, Jacques Bouaud // Artificial intelligence in medicine – № 29. – 2003. – P. 153–167.
13. Setiono R. Generating concise and accurate classification rules for breast cancer diagnosis / Rudy Setiono //

Artificial intelligence in medicine – № 18. – 2000. – P. 205–219.

14. Ronco A. Use of artificial neural networks in modeling associations of discriminant factors: towards an intelligent selective breast cancer screening / A. Ronco // Artificial intelligence in medicine – № 16. – 1999. – P. 299–309.
15. Delen D. Predicting breast cancer survivability: a comparison of three data mining methods / Dursun Delen, Glenn Walker, Amit Kadam // Artificial intelligence in medicine – № 34. – 2005. – P. 113–127.
16. Abbass H. A. An evolutionary artificial neural networks approach for breast cancer diagnosis / Hussein A. Abbass // Artificial intelligence in medicine – № 25. – 2002. – P. 265–281.
17. Setiono R. Extracting rules from pruned neural networks for breast cancer diagnosis / Rudy Setiono // Artificial intelligence in medicine – № 8. – 1996. – P. 37–51.
18. Berezsky O. Histological images processing automation / O. Berezsky, G. Melnyk, V. Valkiv // Proceedings of the 12-th International Conference "The Experience of Designing and Application of CAD Systems in Microelectronics" (CADSM'2009), 19-23 February 2013, Polyana Svalyava (Zakarpattia), 2013. – P. 150–151.
19. Искусственный интеллект: в 3-х кн. Кн. 2. Модели и методы: справочник; под ред. Д. А. Поспелова. – М.: Радио и связь, 1990. – 304 с.
20. Штовба С. Д. Проектирование нечетких систем средствами MATLAB. – М.: Горячая линия - Телеком, 2007. – 288 с.
21. Мельник Г.М. Інформаційна технологія опрацювання гістологічних зображень / Г.М. Мельник // Вісник Хмельницького національного університету. Технічні науки. – 2012. – № 5. – С. 154–161.

## References

1. Rak v Ukraini, 2011-2012: Byuleten Nacionalnogo kantser-reyestru – № 14 – 203 p.
2. Pena-Reyes C.A. A fuzzy-genetic approach to breast cancer diagnosis / Carlos Andrés Peña-Reyes, Moshe Sipper Moshe Sipper // Artificial intelligence in medicine – № 17. – 1999. – Pages 131–155.
3. Mahfouf M. A survey of fuzzy logic monitoring and control utilisation in medicine / M. Mahfouf, M.F Abbod, D.A Linkens // Artificial intelligence in medicine – № 21. – 2001. – Pages 27–42.
4. Soria D. A quantifier-based fuzzy classification system for breast cancer patients / Daniele Soria, Jonathan M. Garibaldi, Andrew R. Green // Artificial intelligence in medicine – № 58. – 2013. – P. 175–184.
5. Kovalerchuk B. Fuzzy logic in computer-aided breast cancer diagnosis: analysis of lobulation / Boris Kovalerchuk, Evangelos Triantaphyllou, J. F. Ruiz // Artificial intelligence in medicine – № 11. – 1997. – P. 75–85.
6. West D. Model selection for a medical diagnostic decision support system: a breast cancer detection case / David West, Vivian West // Artificial intelligence in medicine – № 20. – 2000. – P. 183–204.
7. Shabalova Y.P. Tsytolohychesky atlas. Dyahnostyka zabolevanyi molochnoi zhelezy / Y. P. Shabalova, T. V. Dzhanyrova, N. N. Volchenko, K.K. Puhachev – M. – Tver: OOO «Yzdatelstvo «Tryada», 2005. – 119 p. [in Russian]
8. Gadaras I. An interpretable fuzzy rule-based classification methodology for medical diagnosis / Ioannis Gadaras, Ludmil Mikhailov // Artificial intelligence in medicine – № 47. – 2009. – Pages 25–41.
9. Soria D. A quantifier-based fuzzy classification system for breast cancer patients / Daniele Soria, Jonathan M. Garibaldi, Andrew R. Green, Desmond G. Powe, Christopher C. Nolan, Christophe Lemetre, Graham R. Ball, Ian O. Ellis // Artificial intelligence in medicine – № 58. – 2013. – P. 175–184.
10. Wolberg W. H. Computer-derived nuclear features distinguish malignant from benign breast cytology / W. H. Wolberg, W. N. Street, D. M. Heisey, O. L. Mangasarian // Human Pathology – 1995. – № 26. – P. 792–796.
11. Seroussi B. OncoDoc: a successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer / B. Séroussi, J. Bouaud, É.-C. Antoine // Artificial intelligence in medicine – № 18. – 2000. – P. 205–219.
12. Seroussi B. Using OncoDoc as a computer-based eligibility screening system to improve accrual onto breast cancer clinical trials / Brigitte Séroussi, Jacques Bouaud // Artificial intelligence in medicine – 2003. – P. 153–167.
13. Setiono R. Generating concise and accurate classification rules for breast cancer diagnosis / Rudy Setiono // Artificial intelligence in medicine – № 18. – 2000. – P. 205–219.
14. Ronco A. Use of artificial neural networks in modeling associations of discriminant factors: towards an intelligent selective breast cancer screening / Alvaro L. Ronco // Artificial intelligence in medicine – № 16. – 1999. – P. 299–309.
15. Delen D. Predicting breast cancer survivability: a comparison of three data mining methods / Dursun Delen, Glenn Walker, Amit Kadam // Artificial intelligence in medicine – № 34. – 2005. – P. 113–127.
16. Abbass H. A. An evolutionary artificial neural networks approach for breast cancer diagnosis / Hussein A. Abbass // Artificial intelligence in medicine – № 25. – 2002. – P. 265–281.
17. Setiono R. Extracting rules from pruned neural networks for breast cancer diagnosis / Rudy Setiono // Artificial intelligence in medicine – № 8. – 1996. – P. 37–51.
18. Berezsky O. Histological images processing automation / O. Berezsky, G. Melnyk, V. Valkiv // Proceedings of the 12-th International Conference "The Experience of Designing and Application of CAD Systems in Microelectronics" (CADSM'2009), 19-23 February 2013, Polyana Svalyava (Zakarpattia), 2013. – P. 150–151. [in Ukrainian]
19. Yskusstvennyy yntellekt: v 3-kh kn. Kn. 2. Modely y metody: spravochnyk; pod red. D. A. Pospelova. – M.: Radyo y sviaz, 1990. – 304 p. [in Russian]
20. Shtovba S. D. Proektyrovanye nechetkyh system sredstvamy MATLAB. – M.: Goryachaya linia – Telekom, 2007. – 288 p. [in Russian]
21. Melnyk G.M. Informacijna tehnologiya opracyuvannya gistologichnyh zobrazhen / G.M. Melnyk // Visnyk Khmelnytskogo nacionalnogo universytetu. Tekhnichni nauky. – 2012. – N 5. – P. 154–161. [in Ukrainian]

Рецензія/Peer review : 6.7.2013 р. Надрукована/Printed : 23.11.2013 р.

Рецензент: Николайчук Я. М., д.т.н., проф. зав. каф. спеціалізованих комп'ютерних систем Тернопільського національного економічного університету