

УДК 512.816

**М.А. Омаров¹, Р.И. Цехмистро², А.А. Трубицин³**^{1,3}ХНУРЭ, г. Харьков, Украина²УКРДУЗТ, г. Харьков, Украина**ПОИСК СПОСОБОВ ОЦЕНИВАНИЯ ТЯЖЕСТИ ЗАБОЛЕВАНИЯ АТОПИЧЕСКИМ ДЕРМАТИТОМ С ИСПОЛЬЗОВАНИЕМ МОДЕЛИ НЕЧЕТКОГО ВЫВОДА**

В статье рассматривается подход к способам создания экспертных оценок для установления степени заболевания человека атопическим дерматитом путем использования вероятностного подхода на основе Байесовских процессов в сравнении с подходом на основе теории нечетких множеств.

АТОПИЧЕСКИЙ ДЕРМАТИТ, БАЙЕСОВСКАЯ ОЦЕНКА, БАЗА ДАННЫХ, ЭКСПЕРТНАЯ СИСТЕМА, ЭКСПЕРТНАЯ ОЦЕНКА, НЕЧЕТКОЕ МНОЖЕСТВО

Введение

Проблема прогнозирования течения атопического дерматита является актуальной, ей посвящено немалое количество работ связанных как с диагностикой, так и моделированием процесса заболевания [1]. Одни из сложностей построения эффективных диагностических (экспертных) систем состоит в необходимости учета более 80 данных пациента, необходимости принимать решения о стадии заболевания как на основании объективных так и субъективных данных. Не существует однозначного подхода к построению диагностической оценочной экспертной системы для прогнозирования дерматита. Экспертные системы, как известно, состоят из базы данных, базы знаний и решающего модуля. Решающий модуль, используя исходные данные из рабочей памяти и знания из базы знаний, формирует такую последовательность правил, которые, будучи примененными к исходным данным, приводят к решению задачи. Выбор и создание последовательности правил и составляет одну из главных проблем построения медицинских экспертных систем. Эта проблема непосредственно связана с созданием и эффективным использованием баз данных из которых берутся исходные данные для выбора и запуска того или иного правила. Правила представляют собой один или несколько алгоритмов, моделирующих процесс диагностики заболевания.

Одним из направлений построения экспертных систем является использование Байесовских сетей доверия в сравнении с использованием теории нечеткой логики [2-4]. Это приобретает особую актуальность в связи с необходимостью эффективного поиска информации в базах данных, что в немалой степени используется в алгоритмах моделирующих процессы диагностики и прогнозирования заболеваний [3-5].

Например, возникает важный вопрос какой метод (сети Байеса или нечетких множеств по Мамдани) предъявляет менее жесткие требования как к наполняемости баз данных так и к их поиску без потери эффективности. Использование реляционных баз данных позволяет использовать язык SQL-запросов, который позволяет практически мгновенно

выводить данные из таблиц с количеством символов более 10 000 и взаимодействовать между таблицами. Системы подобные PostgreSQL позволяют еще и проводить вычисления в таблицах используя встроенный язык программирования. Однако подобные средства разработки (Oracle, PostgreSQL и т.д.) созданы исключительно для работы с базами данных и непригодны для построения экспертных систем. Анализ существующих подходов к проектированию баз данных указывает большие возможности объектно-реляционных баз данных, в которых используется языки объектно-ориентированного программирования, в сочетании с технологией SQL.

Однако создание реальных экспертных систем, несмотря на имеющийся прогресс в области проектирования баз данных, по-прежнему сопряжен с трудностями, связанными как с визуализацией полученных результатов (придания наглядности) так и с необходимостью привлечения квалифицированных разработчиков и экспертов. Коммерческие средства разработки экспертных систем недостаточно адаптированы под конкретные отечественные условия и несколько дорогостоящи.

Следовательно, разработкой конкретной медицинской экспертной системы под конкретное заболевание необходимо с точки зрения программирования начинать с методов моделирования течения заболевания.

В оценке состояния больного, страдающего атопическим дерматитом (АД) и для принятия решений, связанных с выбором дальнейшей тактики лечения, важную информацию для лечащего врача дает анализ субъективных факторов заболевания.

На сегодняшний день для характеристики тяжести АД наиболее часто используется система SCORAD (scoring of atopic dermatitis). Индекс SCORAD представляет собой результат суммирования баллов, полученных при субъективной оценке выраженности и распространенности проявления заболевания.

Например, площадь поражения кожи, оценка интенсивности морфологических элементов дерматита, сумма баллов субъективных признаков (зуд+нарушение сна).

Сложность объективной оценки перечисленных факторов является причиной использования методов, ориентированных на построение моделей, учитывающих неточность исходных данных. Именно в таких случаях применение математического аппарата нечетких множеств (НМ) является наиболее эффективным.

Целью исследования является выработка рекомендаций для разработки математической модели поддержки принятия решений, связанных с диагностикой больного и принятия решений относительно дальнейшей стратегии лечения.

Выработка практических рекомендаций для построения экспертной системы оценки состояния больного на основе базы знаний и алгоритмов теории нечетких множеств путем анализа базы знаний.

1. Постановка задачи

В исследовании использовались данные пациентов, полученные на базе кафедры пропедевтики №2 ХНМУ.

В рамках задачи оценки состояния больного АД в качестве лингвистических переменных использовались такие качественные показатели АД как: оценка площади поражения кожи; оценка морфологических элементов сыпи; оценка субъективных симптомов.

Лингвистическая переменная «оценка площади поражения» была представлена как:

$$\langle b1, S, X1, G1, M1 \rangle,$$

где $b1$ – площадь поражения(в %); $S = \{“S_{\min}$ – малая область поражения”, “ $S_{\text{сред}}$ – средняя область поражения”, “ $S_{\text{знач}}$ – значительная область поражения”}; $X1 = [0, 100]$ – количество термов (элементов множества); $G1$ – процедура образования новых термов с помощью логических связок “И”, “ИЛИ” и модификаторов типа “очень”, “НЕ”, “слегка” и др.; $M1$ – процедура семантические правила, задающие функции принадлежности нечетких термов, порожденных синтаксическими правилами $G1$.

Лингвистическая переменная «оценка морфологических проявлений сыпи» была представлена как:

$$\langle b2, K, X2, G2, M2 \rangle,$$

где $b2$ – оценка морфологических проявления сыпи; $K = \{“0$ – отсутствует”, “ $K_{\text{лег}}$ – легкая”, “ $K_{\text{сред}}$ – средняя”, “ $K_{\text{тяж}}$ – тяжелая”}; $X2 = [0, 18]$ – количество термов (элементов множества); $G2$ – процедура образования новых термов с помощью логических связок “И”, “ИЛИ” и модификаторов типа “очень”, “НЕ”, “слегка” и др.; $M2$ – процедура семантические правила, задающие функции принадлежности нечетких термов, порожденных синтаксическими правилами $G2$.

Лингвистическая переменная «оценка субъективных симптомов», оценивающая зуд и нарушение сна в баллах по 10 бальной шкале для каждого из признаков, представлена как:

$$\langle b3, C, X3, G3, M3 \rangle,$$

где $b3$ – оценка субъективных симптомов; $T3 = \{“C_{\text{лег}}$ – незначительное беспокойство”, “ $C_{\text{умер}}$ – умеренное беспокойство”, “ $C_{\text{сред}}$ – среднее беспокойство”, “ $C_{\text{знач}}$ – значительно беспокойство”, “ $C_{\text{умер}}$ – сильное беспокойство”, “ $C_{\text{очень знач}}$ – очень сильное беспокойство”}; $X3 = [0, 20]$ – количество термов (элементов множества); $G3$ – процедура образования новых термов с помощью логических связок “И”, “ИЛИ” и модификаторов типа “очень”, “НЕ”, “слегка” и др.; $M3$ – процедура семантические правила, задающие функции принадлежности нечетких термов, порожденных синтаксическими правилами $G3$.

Лингвистическая переменная «степень тяжести состояния больно» была представлена как:

$$\langle b4, T4, X4, G4, M4 \rangle,$$

где $b4$ – степень тяжести состояния больного; $T4 = \{“легкая$ степень тяжести”, “средняя степень тяжести”, “тяжелое состояние”, “очень тяжелое состояние”}; $X4 = [0, 10]$; $G4$ – процедура образования новых термов с помощью логических связок “И”, “ИЛИ” и модификаторов типа “очень”, “НЕ”, “слегка” и др.; $M4$ – процедура аналогичная $M3$ для правил $M4$.

Нечеткая база знаний, представляющая совокупности лингвистических высказываний имеет вид (1).

$$\mu(b1, b2, b3, S, K, C) = \left\{ \begin{array}{l} d_{\text{лег}} \text{ if } ((0 < b1 \leq S_{\min}) \text{ and } (0 < b2 \leq K_{\text{лег}}) \text{ and } (0 < b3 \leq C_{\text{лег}})) \\ \text{or } ((S_{\min} < b1 \leq S_{\text{сред}}) \text{ and } (0 < b2 \leq K_{\text{лег}}) \text{ and } (C_{\text{лег}} < b3 \leq C_{\text{умер}})) \\ d_{\text{сред}} \text{ if } ((S_{\min} < b1 \leq S_{\text{сред}}) \text{ and } (C_{\text{лег}} < b2 \leq K_{\text{сред}}) \text{ and } (C_{\text{умер}} < b3 \leq C_{\text{сред}})) \\ d_{\text{тяжел}} \text{ if } ((0 < b1 \leq S_{\min}) \text{ and } (C_{\text{лег}} < b2 \leq K_{\text{сред}}) \text{ and } (C_{\text{сред}} < b3 \leq C_{\text{знач}})) \\ d_{\text{очень тяжел}} \text{ if } \left(\begin{array}{l} (S_{\text{сред}} < b1 \leq S_{\text{знач}}) \text{ and } (C_{\text{лег}} < b2 \leq K_{\text{сред}}) \\ \text{and } (C_{\text{знач}} < b3 \leq C_{\text{очень знач}}) \end{array} \right) \\ \text{or } ((S_{\text{сред}} < b1 \leq S_{\text{знач}}) \text{ and } (C_{\text{сред}} < b2 \leq K_{\text{тяжел}}) \text{ and } (C_{\text{умер}} < b3 \leq C_{\text{знач}})) \\ \text{or } ((S_{\text{сред}} < b1 \leq S_{\text{знач}}) \text{ and } (C_{\text{лег}} < b2 \leq K_{\text{сред}}) \text{ and } (C_{\text{знач}} < b3 \leq C_{\text{очень знач}})) \end{array} \right. \quad (1)$$

В данном соотношении присутствует пять переменных, не считая констант на которые раскладываются переменные S, K, C , что затрудняет построение функции принадлежности в явном виде.

2. Аналитический анализ

Аналитические зависимости между термом d_i и переменными b_1, b_2, b_3 устанавливается в случае объединения множеств S, K, C , либо рассмотрения по отдельности каждого множества. Например, введя в рассмотрение $a = 0, b = S_{\min}, c = S_{\text{сред}}, b = S_{\text{знач}}$ получим трапецидальную функцию принадлежности

$$\mu(b1) = \begin{cases} d1 = \left(1 - \frac{S_{\min} - b1}{S_{\min}}\right), (0 < b1 \leq S_{\min}) \\ d2 = 1, (S_{\min} < b1 \leq S_{\text{сред}}) \\ d3 = \left(1 - \frac{b1 - S_{\text{сред}}}{S_{\text{тяжел}} - b1}\right), (S_{\text{сред}} < b1 \leq S_{\text{тяжел}}) \\ d4 = \text{"другие варианты"} \end{cases} \quad (2)$$

Аналогичные треугольные или трапецидальные принадлежности можно представить и для множеств K и C если рассматривать эти факторы по отдельности, однако это не будет описывать реальную базу знаний для диагностики заболевания, к тому же этот вариант усложнит или сделает невозможным процесс автоматизации (возможность построения экспертной системы).

Однако выражение (1) можно упростить, используя функции принадлежности, введенные в рассмотрение Болдвинном [5-9].

Применяя квантификаторы, получим функции принадлежности термов применительно к соотношению (1) в предположении наличия одного нечеткого множества (признака заболевания), например проявление сыпи:

$$\begin{aligned} -\mu_{\text{сред}}(X2) &= d_{\text{сред}}(b2) = d_{\text{лег}}(b2)^2; \\ -\mu_{\text{тяж}}(X2) &= d_{\text{тяж}}(b2) = 1 - d_{\text{лег}}(b2). \end{aligned} \quad (3)$$

Данные соотношения представлены для нечеткого множества K , для нечеткого множества S с константами ($S_{\min}, S_{\text{сред}}, S_{\text{знач}}$) функции принадлежности имеют вид:

$$\begin{aligned} -\mu_{\text{сред}}(X1) &= d_{\text{сред}}(b1) = d_{\min}(b1)^2; \\ -\mu_{\text{тяж}}(X1) &= d_{\text{тяж}}(b1) = 1 - d_{\min}(b1). \end{aligned} \quad (4)$$

Для нечеткого множества C определяем следующие функции принадлежности принимая во внимание большее число констант характеризующих границы степеней заболевания:

$$\begin{aligned} -\mu_{\text{умер}}(X3) &= d_{\text{умер}}(b3) = d_{\text{лег}}(b3)^{1/2}; \\ -\mu_{\text{сред}}(X3) &= d_{\text{сред}}(b3) = d_{\text{знач}}(b3)^{1/2}; \\ -\mu_{\text{знач}}(X3) &= d_{\text{знач}}(b3) = 1 - d_{\text{лег}}(b3); \\ -\mu_{\text{оченьзнач}}(X3) &= d_{\text{оченьзнач}}(b3) = d_{\text{знач}}(b3)^2. \end{aligned} \quad (5)$$

Таким образом мы имеем дело с тремя нечеткими множествами S, K, C с разным значени-

ем конечных элементов(термов) $b1=X1=\{0...100\}, b2=X2\{0...18\}, b3=X3\{0..20\}$ в каждом множестве и с разным количеством функций принадлежности, которые обуславливаются медицинскими потребностями в рассматриваемой ситуации.

Для эффективной оценки состояния больного мы не можем уменьшить количество функций принадлежности и безосновательно пренебрегать количеством функций принадлежности данных на рассмотрение нечетких множествах.

Рассмотрим нечеткие отношения, следующие из модели базы знаний (1) на основании импликации и t -нормам, предварительно установив, что множества S, K, C являются входными, а множество D – выходным. Отношение с нечеткими множествами строятся на определенных правилах, наиболее важными из которых являются объединение и пересечения (правило ассоциативности для трех множеств):

$$\begin{aligned} -\mu_{S \cup K \cup C} &= \max(\mu_S(x), \max(\mu_K(x), \mu_C(x))), \\ -\mu_{S \cap K \cap C} &= \max(\mu_S(x), \max(\mu_K(x), \mu_C(x))). \end{aligned} \quad (6)$$

Введем в рассмотрение множество KC путем объединения множеств K и C для которого:

$$\begin{aligned} -\mu_{K \cup C} &= \max(\mu_K(x), \mu_C(x)), \\ -\mu_{K \cap C} &= \max(\mu_K(x), \mu_C(x)), \end{aligned}$$

и выражения для отрицаний и импликаций:

$$\begin{aligned} \mu_{\overline{KC}}(X) &= 1 - (\mu_{KC}(x)), \\ \mu_{S=KC}(X) &= \max(1 - \mu_S(x), \mu_{KC}(x)). \end{aligned}$$

Введенные выше операции над нечеткими множествами основаны на использовании операций \max и \min , что является удобным для их реализации в программных пакетах научного назначения Mathcad и Matlab. Эти соотношения реализуемы и в средах разработки пригодных для создания экспертных систем и баз данных (они есть в языке SQL). В теории нечетких множеств разрабатываются вопросы построения обобщенных, параметризованных операторов пересечения, объединения и дополнения, позволяющих учесть разнообразные смысловые оттенки соответствующих им лингвистических связей естественного языка «и», «или», «не» [7]. Один из подходов к операторам пересечения и объединения заключается в их определении, при помощи нечетких операторов, т.н.треугольных норм. Следует обратить внимание на то, что представленные выше операции пересечения $\min(\mu_A(x), \mu_B(x))$ и объединения $\max(\mu_A(x), \mu_B(x))$, использующиеся как самостоятельно, так и при введении операций разности, симметрической разности и дизъюнктивной суммы – это только один из возможных вариантов определения данных операций, введенный основоположником теории нечетких множеств Л.Заде.

Треугольной нормой (t -нормой) называется двуместная действительная функция T , ото-

бражающая две функции принадлежности нормальных нечетких множеств $\mu_A(x)$, $\mu_B(x)$ в одну функцию принадлежности нормального нечеткого множества и удовлетворяющая следующим условиям: Рассмотрим вначале дизъюнктивную сумму множеств S и K , отметим что данные множества являются не доминируемыми:

$$S \otimes K = (S - K) \cup (K - S) = (S \cap \bar{K}) \cup (\bar{K} \cap S), \quad (7)$$

с функцией принадлежности

$$\mu_{S-K}(X) = \max \left\{ \begin{array}{l} (\min(\mu_S(x), 1 - \mu_K(x))); \\ (\min(1 - \mu_S(x), \mu_K(x))) \end{array} \right\},$$

$$\mu_{A-B}(X) = \max \left\{ \begin{array}{l} (\min(\mu_A(x), 1 - \mu_B(x))); \\ (\min(1 - \mu_A(x), \mu_B(x))) \end{array} \right\}.$$

Результаты расчетов в системе Mathcad представлены на рис.1, на котором представлены две кривые; верхняя соответствует дизъюнктивной сумме $S \oplus K$ (REZ^2), а нижняя $\mu_{S-K}(x)$ (μREZ^2) аргументом обеих функций являются элементы множества S , заданные в таблице 4.

Из рисунка видно, что зависимости построенные на реальных практически важных данных даже для небольшого числа элементов множеств гораздо сложнее чем трапецидальных зависимостей.

Этот подталкивает нас к другим методам экспертных оценок для установления более конкретных зависимостей.

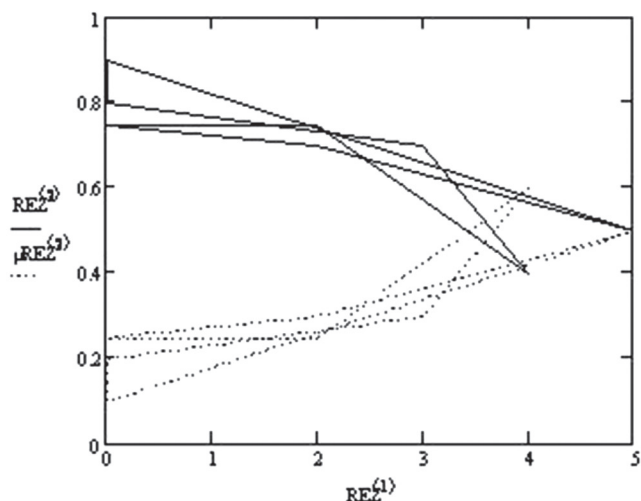


Рис. 1. Функции принадлежности множества S и K

2. Численный анализ

Предшествующие рассуждения, несмотря на относительную несложность базы знаний (1) не привели к конкретной наглядной концепции получения методов оценок. Перейдем к численным методам.

На основании результатов медицинского обследования пациента по трем признакам: оценка площади поражения кожи; оценка морфологических элементов сыпи; оценка субъективных симптомов рассмотрим таблицу1 диагностики 10 (де-

сяти) пациентов обследованных на atopический дерматит.

Представим их результаты в упрощенном виде при котором количество значащих элементов одинаково, т.е. соответствует 10, однако если одно и то же значение множества встречается более одного раза количество переменных уменьшается.

Таблица 1

Данные 10 пациентов больных дерматитов

Номер больного	SCORAD	Площадь поражения кожи (баллы) – S (100макс.)	Субъективные признаки (зуд) – C (10макс.)	Сумма баллов объективных признаков – K (эритема, отек, сухость) (18 макс.)
1	38	5	5	8
2	61	95	5	9
3	28	10	2	6
4	21	20	0	5
5	20	4	2	5
6	60	12	4	15
7	24	20	3	5
8	15	5	0	4
9	7	0	0	2
10	20	2	5	8

Например, в множестве K значение 5 встречается у 4-ого, 5 и 7 пациента, таким образом увеличивается их степень принадлежности значения 5 составляет 0,3.

Таблица 2

Степень присутствия признака S в группе больных

Степень принадлежности элементов множества S	0,2	0,1	0,1	0,2	0,1	0,1	0,1
Значение элементов множества S	5	95	10	20	4	12	2

Таблица 3

Степень присутствия признака K в группе больных

Степень принадлежности элементов множества K	0,2	0,1	0,1	0,3	0,1	0,1	0,1
Значение элементов множества K	8	9	6	5	4	15	2

Таблица 4

Степень присутствия признака C в группе больных

Степень принадлежности элементов множества C	0,3	0,2	0,3	0,1	0,1
Значение элементов множества C	5	2	0	4	3

Рассмотрим матрицу оценок заболевания девяти пациентов конкретным врачом.

Матрица оценок из таблицы 1 имеет вид (8):

$$M = \begin{pmatrix} 5 & 5 & 8 \\ 95 & 5 & 9 \\ 10 & 2 & 6 \\ 20 & 0 & 5 \\ 4 & 2 & 5 \\ 12 & 4 & 15 \\ 20 & 3 & 5 \\ 5 & 0 & 4 \\ 0 & 0 & 2 \end{pmatrix}. \tag{8}$$

Целесообразно представить все три качества в системе с одинаковым максимальным количеством баллов для каждой группы. Однако по количеству баллов, которыми пользуются медики для каждого фактора можно судить о степени важности каждого фактора. На рис. 2 представлена зависимость $Cr_i = f(i)$ – усредненное графическое отображение оценок больных врачом по каждому фактору, соотношению (9):

$$Cr_i = \frac{1}{m} \sum_{j=0}^{m-1} M_{j,i}. \tag{9}$$

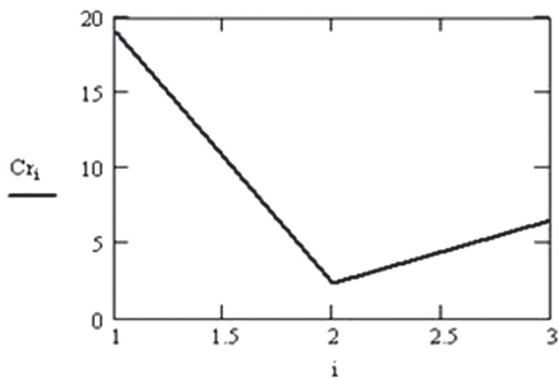


Рис. 2 Усредненное графическое отображение оценок больных по каждому фактору

На рис. 3 представлена зависимость $Eksprt_j = f(j)$ усредненное графическое отображение общей оценки состояния конкретного больного по соотношению (10):

$$E_{ksprt_j} = \frac{1}{n} \sum_{j=0}^{n-1} (M^T)_j. \tag{10}$$

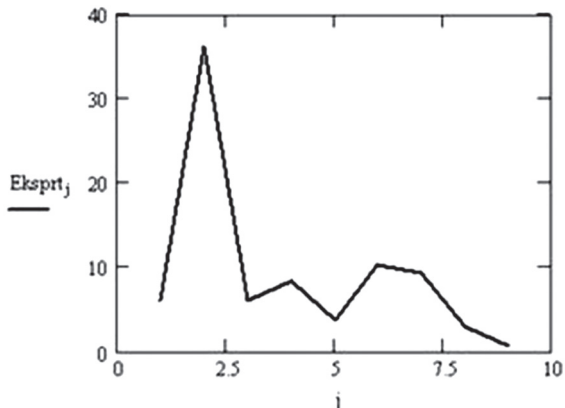


Рис. 3. Общая оценка состояния конкретного (j-) больного

Указанные оценки имеют как количественную, так и качественную составляющую.

Из сопоставления зависимостей вытекает вывод о том, что разброс абсолютных значений по первому графику составил 10 раз, по второму около 20 раз.

Расчет весовых коэффициентов по результатам обследования больных проводится по соотношению:

$$\alpha_i = \frac{1}{n} \sum_{j=1}^n \left(\frac{M_{j,i}}{\sum_{i=1}^m M_{j,i}} \right), \tag{11}$$

$$\alpha = \begin{pmatrix} 0.503 \\ 0.095 \\ 0.402 \end{pmatrix},$$

$$\alpha^T = (0.503 \quad 0.095 \quad 0.402).$$

Расчет средневзвешенной оценки результатов обследования больных в баллах:

$$K_m = \frac{1}{n} \cdot \sum_{j=1}^n \sum_{i=1}^m (\alpha_i \cdot M_{j,i}) = 12,413. \tag{12}$$

Весовые коэффициенты i (строка) фактора заболевания (переменной) в общей оценке заболевания каждого пациента j (столбец):

$$\alpha_{i,j} = \frac{M_{j,i}}{\sum_{i=1}^m M_{j,i}}, \tag{13}$$

$$\alpha = \begin{pmatrix} 0,276 & 0,872 & 0,556 & 0,8 & 0,364 & 0,387 & 0,714 & 0,556 & 0 \\ 0,278 & 0,046 & 0,111 & 0 & 0,182 & 0,129 & 0,107 & 0 & 0 \\ 0,444 & 0,083 & 0,333 & 0,2 & 0,455 & 0,484 & 0,179 & 0,444 & 1 \end{pmatrix}.$$

Оценка результирующего вывода о заболевании пациентов при равной значимости весовых коэффициентов:

$$K_m^1 = \frac{1}{n} \cdot \sum_{j=1}^n \sum_{i=1}^m \left(\frac{M_{j,i}}{n} \right) = 3,099. \tag{14}$$

Численные значения весовых коэффициентов для трех признаков.

В числителях находятся коэффициенты, величина которых зависит от приоритетности признаков. Признак $S = 0.95$ для (α_1) , $K = 0.8$ для (α_2) , $C = 0.5$ для (α_3) . Поскольку приоритеты уже учтены баллах отводимые на каждый признак:

$$\alpha_1 = \frac{1,15}{n}, \alpha_2 = \frac{0,95}{n}, \alpha_3 = \frac{0,95}{n}.$$

Оценка степени заболевания пациента, при разном значении весовых коэффициентов учитывающих приоритеты признаков представлена как:

$$\alpha_1 = \frac{1,15}{n}, \alpha_2 = \frac{0,95}{n}, \alpha_3 = \frac{0,95}{n}, \tag{15}$$

$$K1 = 9.462.$$

Данное оценивание нельзя назвать абсолютно четким показало, что средневзвешенная оценка заболевания пациентов врачом составило 12,5 усл. баллов. При предположении равнозначности весовых коэффициентов оценка заболевания боль-

ных пациентом падает до 3,1 усл. баллов. При учете приоритета признаков оценка заболевания составляет 9,5 усл. баллов. Данный метод экспертного оценивания основывается на гипотезе равной значимости весовых коэффициентов, значения которых четко зависят от максимального количества баллов, которые определяются медиками по определению. К тому же его нельзя назвать методом четкого оценивания, поскольку есть необходимость учета приоритета признаков.

Для установления влияния приоритетности, ввиду их важности признаков проведем дополнительное Байесовское оценивание.

Рассмотрим определения важности одного из трех приведенных диагностических признаков. Разделим приведенные диагностические признаки в зависимости от их важности (табл. 5). Важность определим следующим образом: средняя важность – K , существенная важность – S , малосущественный признак – C . Таким образом рассматривая конкретный признак существует три варианта. Признак попадает в одну из трех групп. Примем, что 50% признаков имеет среднюю важность, 20% признаков имеют существенную важность и 30% признаков не существенно важны. Априорные вероятности в данном случае $P(B_1)=0,5$, $P(B_2)=0,3$, $P(B_3)=0,2$. Одними из критериев важности того или иного признака является использование его врачами и указывание на него пациентами. Существенный признак используют 80% процентов врачей. 50% процентов врачей используют признак средней важности, и 20% процентов врачей используют также малосущественные признаки. Укажем также на то, что на существенный признак отмечается 80 % больных $P(A_2/B_1)=0,8$ у которых данный признак явно присутствует, на несущественный указывает 40% больных (среди тех у которых он явно указывает на дерматит), $P(A_2/B_3)=0,8$ на признак средней важности указывает 20 % больных $P(A_2/B_1)=0,2$ (однако по обследованиям он полностью подтверждает дерматит).

Вероятность того что признак S для себя приоритетным считают указывает 80 проц. врачей $P(A_1/B_1)=0,8$. Признак “ K ” приоритетным для себя считают 50 проц. врачей. И 20 проц. врачей считают необходимым учитывать признак – C .

Таблица 5

Распределение вероятностей по признакам важности

Вероятность	1 средняя важность признака	2 низкая важность признака	3 высокая важность признака
$P(B_i)$	0,5	0,3	0,2
$P(A_1/B_i)$	0,8	0,5	0,25
$P(A_2/B_i)$	0,8	0,2	0,4

$$P\left(\frac{B_i}{A_i}\right) = \frac{P\left(\frac{A_i}{B_i}\right) \cdot P(B_i)}{\sum_{k=1}^3 P\left(\frac{A_i}{B_k}\right) \cdot P(B_k)}, \quad (16)$$

$$P\left(\frac{B_1}{A_1}\right) = \frac{0,8 \cdot 0,5}{0,8 \cdot 0,5 + 0,5 \cdot 0,2 + 0,25 \cdot 0,2} = \frac{0,4}{0,6} = 0,66,$$

$$P\left(\frac{B_2}{A_2}\right) = \frac{0,5 \cdot 0,3}{0,8 \cdot 0,5 + 0,5 \cdot 0,3 + 0,25 \cdot 0,2} = \frac{0,15}{0,6} = 0,25,$$

$$P\left(\frac{B_3}{A_3}\right) = \frac{0,25 \cdot 0,2}{0,8 \cdot 0,5 + 0,5 \cdot 0,3 + 0,25 \cdot 0,2} = \frac{0,05}{0,6} = 0,083.$$

После того как жалобы больных на возможное присутствие деоматита полностью подтвердилось, доверия к предположениям B_2 и B_3 потеряло важность, т. е. признаки средней и меньшей значимости подтверждаются как мнением врачей так и жалобами пациентов с меньшей вероятностью. Событие A_1 – использование врачом признака как главного A_2 – жалоба пациента на признак заболевания. B_1 – диагностический признак имеет существенную важность, B_2 – диагностический признак имеет среднюю важность, B_3 – признак имеет малую важность.

Если независимость событий A_2 и A_3 объективна, то при верности гипотез B_1 , B_2 , B_3 проведем оценки по формуле Байеса в следующем виде [8-10]:

$$P\left(\frac{B_i}{A_1 A_2}\right) = \frac{P\left(\frac{A_1}{B_i}\right) \cdot P\left(\frac{A_2}{B_i}\right) \cdot P(B_i)}{\sum_{k=1}^3 P\left(\frac{A_1}{B_k}\right) \cdot P\left(\frac{A_2}{B_k}\right) \cdot P(B_k)}, \quad (17)$$

$$P\left(\frac{B_1}{A_1 A_2}\right) = \frac{0,8 \cdot 0,5 \cdot 0,8}{0,8 \cdot 0,8 \cdot 0,5 + 0,5 \cdot 0,3 \cdot 0,2 + 0,25 \cdot 0,2 \cdot 0,4} = 0,86,$$

$$P\left(\frac{B_2}{A_1 A_2}\right) = \frac{0,8 \cdot 0,5 \cdot 0,8}{0,8 \cdot 0,8 \cdot 0,5 + 0,5 \cdot 0,3 \cdot 0,2 + 0,25 \cdot 0,2 \cdot 0,4} = 0,86,$$

$$P\left(\frac{B_3}{A_1 A_2}\right) = \frac{0,3 \cdot 0,5 \cdot 0,2}{0,8 \cdot 0,8 \cdot 0,5 + 0,5 \cdot 0,3 \cdot 0,2 + 0,25 \cdot 0,2 \cdot 0,4} = 0,08,$$

$$P\left(\frac{B_3}{A_1 A_2}\right) = \frac{0,2 \cdot 0,25 \cdot 0,4}{0,8 \cdot 0,8 \cdot 0,5 + 0,5 \cdot 0,3 \cdot 0,2 + 0,25 \cdot 0,2 \cdot 0,4} = 0,05.$$

Данные оценки подтверждают оценки по предыдущему соотношению с использованием экспертного оценивания.

Выводы

Построенная модель базы знаний позволяет проводить экспертную оценку степени заболевания (или его отсутствия) пациента атопическим дерматитом. Численный анализ функций принадлежности термов от их значений при разном количестве термов показал реальные результаты качественно сопоставимые с результатами медицинского об-

следования. Использование разного количества термов, т.е. разные размеры исходных множеств, в предложенном методе не существенно повлияли на усложнение алгоритмов численного анализа базы знаний благодаря использованию матричного представления матричного представления введенных в рассмотрение нечетких множеств. Данный подход является менее наглядным, но практически более универсальным, нежели использование функций одной и более переменных для определения функций принадлежности, с помощью которых строится база знаний. К тому же он более функционален для построения экспертных систем, поскольку легче реализуются в средствах разработки связанных с базами данных [11, 12].

Проведенное экспертное оценивание с использованием средневзвешенных оценок и Байесовское оценивание четко показали необходимость учета приоритетности признаков заболевания при построении базы знаний и необходимости установления взаимосвязи между количеством баллов отводимых на оценку каждого признака с их приоритетностью.

Учет приоритетности признаков заболевания также важен для скорейшего определения диагноза, поскольку больные могут указывать на различные признаки. Приоритетность признаков заболевания также может варьироваться при разных стадиях заболевания.

На основании предложенного подхода в дальнейшем появляется реальная возможность разработки методик получения поправок к признаку SCORAD.

Список литературы: 1. Бых А.И. Метод анализа иерархии при построении мультифакторных моделей дифференциальной диагностики атопического дерматита / А.И. Бых, Е.В. Высоцкая, И.И. Ключник, А.А. Трубицин, А.П. Порван. // Восточно-Европейский журнал передовых технологий 2009 г. 5-8. 2. Клименко В.А. Информационная система диагностики и прогнозирования атопического дерматита у детей / В. А. Клименко, А. П. Порван, А. А. Трубицин. // Системы обработки информации. – Х.:ХВУ, 2004. – Вып. 8 (47). – С. 27-33. 3. Кораблев Н.М. Сравнительный анализ методов определения абсолютных приоритетов признаков при нечеткой исходной информации / Н.М. Кораблев, А.С. Непокупный, Алзин Ферас // Системы обработки информации. – Х.:ХВУ, 2005. – Вып. 9 (49). – С. 75-83. 4. Бых А.И. Разработка модели базы данных информационной системы зберігання та обробки інформації про пацієнтів із захворюваннями різними дерматозами / Бых А.И., Панфьорова І.Ю., Висоцька О.В., Жуков В.І., Кириченко Ю.В. // Вісник НТУ “ХПІ”. Тематичний

випуск: Інформатика і моделювання. – Харків: НТУ “ХПІ”. – 2010. – № 31. – С. 23 – 29. 5. Поворознюк А.И. Модель нечеткой экспертной системы прогноза риска развития профессионально обусловленных заболеваний / А.И. Поворознюк, Н.А. Чикина, И.В. Антонова // Системы обработки информации, 2010, вып. 5(86). 6. Baldwin J.F. Fuzzy logic and fuzzy reasoning. - London, Academic Press, 1981.] 7. Кораблев Н.М. Дифференциальная диагностика аллергодерматозов с использованием адаптивной модели нечеткого вывода / Н.М. Кораблев, И.В. Сорокина, А.Э. Макогон // Системы обработки информации, 2009, вып. 3(77). 8. Штовба С.Д. Проектирование нечетких систем средствами MATLAB / С.Д. Штовба. – М.: Горячая линия-Телеком, 2007. – 288с. 9. Брежнев Е.В. Модель прогнозирования риска с использованием нечеткого вывода / Е.В. Брежнев, А.А. Адаменко // Системы обработки информации, 2009, вып. 2(76). 10. Леоненков А. Нечеткое моделирование в среде MATLAB и fuzzyTECH / Леоненков А. СПб.: БХВПетербург, 2005. – 736 с. 11. Цехмистро Р.И. Розробка транслятора простої мови програмування. / Цехмистро Р.И., Цимбал А.М., // Радиоелектроника и информатика №1(32) 2006. – С. 57-65. 12. Thekhmistro R.I. The comparative analysis of a multiprobe microwave multimeters with involvement of processing by the kalman filtering and the least-squares methods with regard for re-reflection of probes O.B. Zaichenko, I.I. Klyuchnik, M.A. Miroshni, R.I Thekhmistro - Telecommunications and Radio Engineering, 2015 p.79-86.

Поступила в редколлегию 2.06.2015

УДК 512.816

Пошук способів оцінювання тяжкості захворювання атопічний дерматит з використанням моделі нечіткого виведення / М.А. Омаров, Р.И. Цехмистро, А.А. Трубицин // Біоніка інтелекту: наук.-техн. журнал. – 2015. – № 2 (85). – С. 74–80.

Проведене експертне оцінювання з використанням середньозважених оцінок і Байєсова оцінювання чітко показали необхідність врахування пріоритетності ознак захворювання при побудові бази знань і необхідності встановлення взаємозв'язку між кількістю балів відводяться на оцінку кожної ознаки з їх пріоритетністю. Табл. 5. Іл. 3. Бібліогр.: 12 найм.

UDC 512.816

Search sposobov otsenyvaniya gravity with atopic dermatitis disease using the fuzzy logic model / М.А. Omarov, R.I. Tsekhmistro, А.А. Trubitsin // Bionics of Intelligense: Sci. Mag. – 2015. – № 2 (85). – P. 74–80.

Of expert evaluation using a weighted average of estimates and Bayesian estimation clearly showed the need to consider the priority of signs of disease in the construction of knowledge and the need to establish the relationship between the number of points allocated to the assessment of each characteristic to their priority .

Tab. 5. Fig. 3. Ref.: 12 items.