

И.В. АНТОНОВА, канд. техн. наук, доц., НТУ «ХПИ»

НЕЧЕТКАЯ ЭКСПЕРТНАЯ СИСТЕМА ПРОГНОЗА РИСКА В СТРАХОВАНИИ ПРОФЗАБОЛЕВАНИЙ

Предложена экспертная система с нечеткими решающими правилами для прогноза индивидуального риска развития профессионально обусловленных аллергодерматозов на основе анализа факторов риска, оценки адаптационных возможностей и защитных свойств организма в условиях вредного производства. Определены функциональные характеристики ее модулей, проведено компьютерное моделирование и параметрическая оптимизация в среде MatLab.

Ключевые слова: экспертная система, нечеткие решающие правила, прогноз индивидуального риска, факторы риска, компьютерное моделирование, параметрическая оптимизация.

Постановка проблемы. Ухудшение социальной и производственной среды за последние 20-30 лет привело к увеличению риска развития профессиональных и профессионально обусловленных заболеваний (ПОЗ). Особенно это проявилось на предприятиях химической и химико-фармацевтической отрасли, где к профессиональным рискам относится воздействие активных химических веществ, вызывающих различные кожные заболевания. Наличие полиса медицинского страхования является одной из форм гарантий получения при заболевании качественной медицинской помощи. Поэтому актуальной является разработка автоматизированных информационных систем поддержки принятия решений в страховании профзаболеваний. Основой для принятия решений в медицинском страховании является максимально индивидуализированная количественная оценка риска развития заболеваний.

По результатам исследований, проводимых в рамках научно-исследовательских работ Харьковского НИИ Дерматологии и Венерологии совместно с НТУ «ХПИ» на предприятиях химико-фармацевтической промышленности, были выявлены групповые и внутренние факторы риска развития ПОЗ, проведена оценка адаптационно-компенсаторных возможностей в группах риска [1-4]. Наличие выявленных закономерностей развития ПОЗ дает возможность построить экспертную систему прогноза риска профзаболеваний.

Анализ литературы. Появление и совершенствование автоматизированных систем для обследований групп населения привело к трансформации экспертных систем для дифференциальной диагностики заболеваний с похожей симптоматикой, невероятно популярных еще 30-40 лет назад, в автоматизированные информационные системы поддержки принятия решений. Системы поддержки принятия решений осуществляют выбор наилучшего решения из

множества возможных либо упорядочение возможных решений по предпочтительности.

Построение моделей, использующих способность человека принимать решения в условиях неполной и неопределенной информации, и их реализация в компьютерных системах представляет собой одну из важнейших научно-технических проблем. Предложенный Л. Заде [5] математический аппарат нечеткой логики явился одним из самых мощных инструментов решения подобных проблем. Это стало толчком к созданию особого класса экспертных систем (ЭС) – *нечетких ЭС*.

Логико-лингвистические методы описания нечетких систем основаны на том, что входные и выходные параметры системы являются лингвистическими переменными, а качественное описание процесса задается совокупностью простейших правил следующего вида:

$$L_j : \text{ЕСЛИ } \tilde{A}_j, \text{ ТО } \tilde{B}_j, \quad (j = \overline{1, k_0}),$$

где \tilde{A}_j и \tilde{B}_j – нечеткие подмножества, заданные на декартовом произведении универсальных множеств входных и выходных лингвистических переменных соответственно. К достоинствам нечетких ЭС следует отнести тот факт, что в их основе лежит *композиционное правило Заде*, позволяющее вычислять суперпозиции влияний различных процессов на окончательный результат, что дает возможность учесть совместное влияния целого ряда факторов в решении задачи прогноза риска развития ПОЗ.

В качестве модели нечеткого логического вывода разработанной системы выбрана *модель Мамдани* (Mamdani). Нечеткий вывод типа Мамдани используется в том случае, когда необходимо извлечь знания из имеющихся данных либо при поиске закономерностей во множестве данных [6]. Легко интерпретируемые лингвистические решающие правила являются безусловным достоинством систем такого вида. К недостаткам моделей Мамдани следует отнести высокий уровень вычислительных затрат при получении правила нечеткого вывода.

Целью статьи является описание структуры нечеткой ЭС прогноза риска развития ПОЗ у рабочих предприятий химико-фармацевтической промышленности и ее возможных применений в страховании профзаболеваний.

Разработка модели экспертной системы. Структура рассматриваемой нечеткой ЭС проектировалась на основе выявленных в ходе предыдущих исследований факторов риска [1-3].

Пусть для описания множества Ω , состоящего из m объектов, используется n признаков: $\overline{X} = (X_1, X_2, \dots, X_n)$. Каждому объекту из множества Ω соответствует некоторое значение целевого признака X_0 . Признаки X_i ($i = \overline{0, n}$)

измерены в порядковой или номинальной шкале. Целевой признак X_0 имеет k_0 градаций, т. е. принимает k_0 различных значений.

Структурная схема нечеткой ЭС прогноза риска развития ПОЗ представлена на рис. 1.

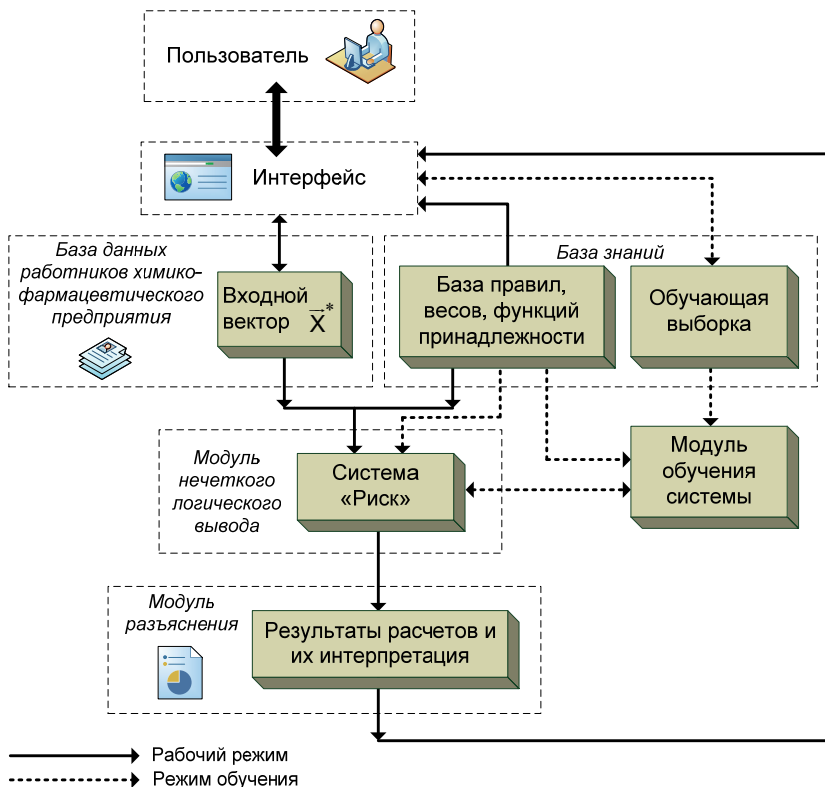


Рис. 1 – Структурная схема нечеткой ЭС прогноза риска развития ПОЗ.

В качестве переменных, описывающих состояние объекта исследования, выбраны выявленные индивидуальные [1] и групповые [2-4] факторы риска X_i ($i = \overline{0, n}$) развития ПОЗ. Внутренние факторы риска выявлены в ходе анализа анамнестических данных у рабочих с проявлениями ПОЗ. При построении групповых факторов риска применялся *метод корреляционной адаптометрии* [2], анализ кривых оценки резистентности гомеостаза [3] и *цепи Маркова* [4].

Применение метода корреляционной адаптометрии дало возможность выявить стажные группы рабочих с максимумом компенсаторной реакции организма (*эффект группового стресса*).

С помощью анализа кривых оценки резистентности гомеостаза решалась задача о влиянии на развитие профессионально обусловленных алергодерматозов наличия в анамнезе обследованных рабочих различных соматических заболеваний и кожных заболеваний неаллергического генеза. Построение кривых проводилось по *методу Каплана-Мейера* [7].

В качестве формализованного метода прогноза риска заболевания ПОЗ применялся математический аппарат цепей Маркова. Представление процесса эволюции системы под действием внешних и внутренних факторов в виде последовательности переходов между состояниями $S_i(t_k)$ и $S_j(t_{k+1})$, относящимися к смежным моментам времени t_k и t_{k+1} , является эффективным способом исследования динамических систем с дискретным временем. Задача прогнозирования риска развития ПОЗ в момент времени t_{k+1} при фиксированных внешних условиях сводится к отысканию вероятности p_{ij} перехода системы из состояния $S_i(t_k)$ в состояние $S_j(t_{k+1})$ страхового случая.

Обоснованием для предоставления той или иной индивидуальной формы корпоративного страхового полиса станут *матрицы условных вероятностей* p_{ij} , учитывающие влияние индивидуальных и групповых факторов риска развития профессиональных заболеваний у рабочих предприятий данной промышленной отрасли.

Согласно [8] *лингвистическая переменная* описывается набором $\{X, T(X), U, G, M\}$, в котором X – название этой переменной; $T(X)$ – термножество X , т.е. совокупность ее лингвистических значений; U – универсальное множество; G – синтаксическое правило, порождающее термы множества $T(X)$; M – семантические правила, задающие функции принадлежности нечетких термов, порожденных синтаксическими правилами G .

В процессе структурной идентификации модели для каждой переменной X_i ($i = \overline{0, n}$) задавалось ее имя, термножества (градации) T_i , соответствующие им универсальные множества U_i .

Для всех градаций признаков X_i ($i = \overline{1, n}$) построены функции принадлежности $\mu^{jp}(X_i)$ входного признака X_i нечеткому терму X_i^{jp} .

Например, график функции принадлежности фактора риска $X_2 =$ «Стаж работы на предприятии» представлен на рис. 2.

На основе результатов предыдущих исследований [1-3] был синтезирован набор k_0 правил вида:

L_j : ЕСЛИ $(X_1 = X_1^{j1})$ И $(X_2 = X_2^{j1})$ И ... И $(X_n = X_n^{j1})$ (с весом ω_{j1})

ИЛИ $(X_1 = X_1^{j2})$ И $(X_1 = X_1^{j2})$ И ... И $(X_1 = X_1^{j2})$ (с весом ω_{j2})

.....

ИЛИ $(X_1 = X_1^{jq_j})$ И $(X_2 = X_2^{jq_j})$ И ... И $(X_n = X_n^{jq_j})$ (с весом ω_{jq_j}),

ТО $X_0 = X_0^j$,

где L_j – решающее правило с номером j ($j = \overline{1, k_0}$); X_i^{jp} – градация входной переменной X_i в правиле с номером jp ($j = \overline{1, k_0}, p = \overline{1, q_j}$); q_j – количество правил, в которых выходная переменная X_0 оценивается термом X_0^j . На основе базы нечетких правил L_j ($j = \overline{1, k_0}$) была получена модель Мамдани нечеткого логического вывода для прогноза риска развития ПОЗ.

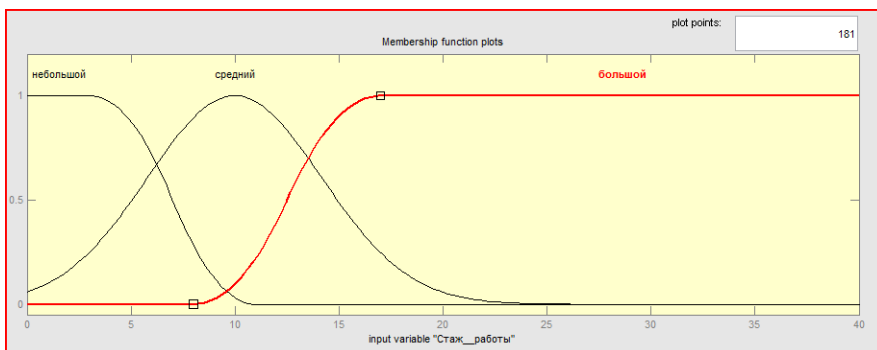


Рис. 2 – Функции принадлежности признака «Стаж работы на предприятии».

Согласно механизму Мамдани нечеткий вывод в ЭС осуществляется следующим образом:

1. На этапе фаззификации определяются *степени истинности*, т.е. значения функций принадлежности $\mu^{jp}(X_i)$ ($i = \overline{1, n}, j = \overline{1, k_0}, p = \overline{1, q_j}$) для левой части правила L_j , состоящей из jp подправил.

2. Для текущих входных значений $\vec{X}^* = (X_1^*, X_2^*, \dots, X_n^*)$ *минимаксным методом* определяется степень истинности целевого признака, т.е. значения

$$\mu_0^j(X_0) = \bigvee_{p=1, q_j} (\omega_{jp} \bigwedge_{i=1, n} \mu^{jp}(X_i^*)), \quad j = \overline{1, k_0},$$

где $\bigvee(\bigwedge)$ – операция s -нормы (t -нормы).

3. На этапе дефаззификации четкое значение целевого признака X_0 , соответствующее входному вектору $\vec{X}^* = (X_1^*, X_2^*, \dots, X_n^*)$, определяется *методом нахождения центра тяжести*.

Пример применения разработанной системы. Проверка адекватности и оптимизация модели нечеткой ЭС проводилась средствами пакета Fuzzy Logic среды MatLab на случайно сформированной выборке из имеющейся базы, со-

держашей информацию о 430 обследованных рабочих профильных предприятий.

На рис. 3 и 4 представлены примеры окон вывода модуля разъяснений нечеткой ЭС.

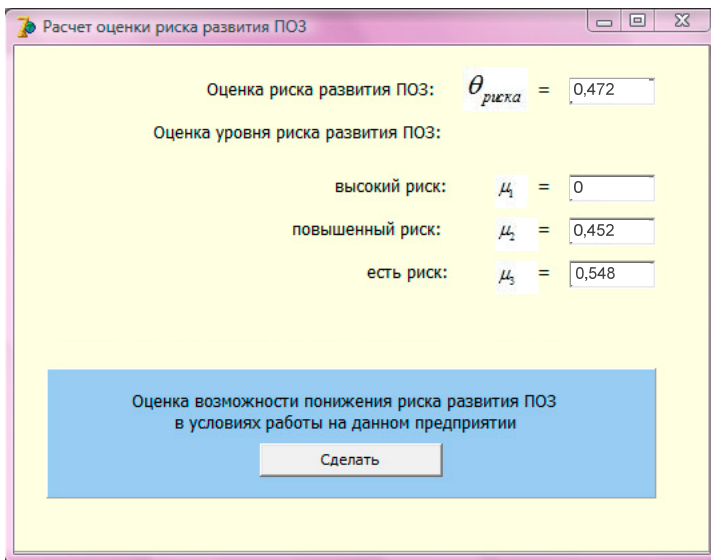


Рис. 3 – Окно вывода «Расчет оценки риска развития ПОЗ».

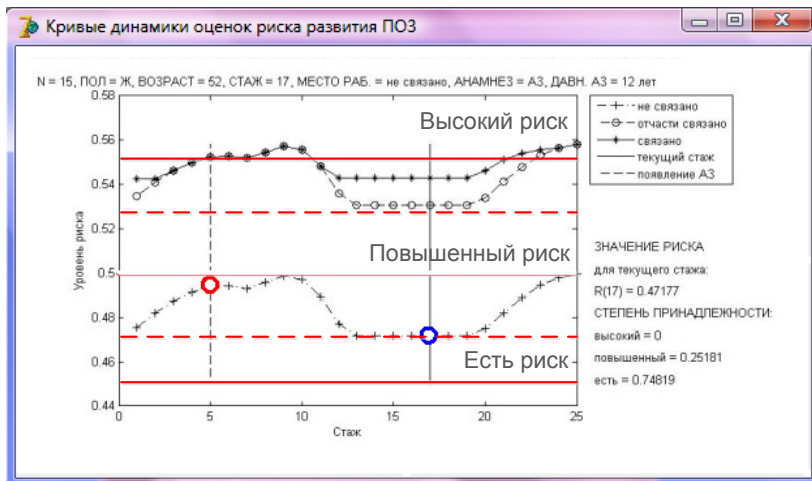


Рис. 4 – Окно вывода кривых динамики оценок риска развития ПОЗ при различных значениях признака «Место работы».

Выводы. В работе представлена нечеткая ЭС прогноза риска развития профессионально обусловленных алергодерматозов и определены функциональные характеристики ее модулей. Оптимизация нечеткой ЭС осуществлялась путем ее параметрической идентификации по обучающей выборке. Прогнозное значение риска – $\theta_{\text{риска}}$ развития ПОЗ – является определяющим при выборе вида страхового полиса в системе корпоративного медицинского страхования. Результаты работы данной ЭС могут быть использованы на предприятии химико-фармацевтической промышленности при проведении профилактических медицинских осмотров в качестве рекомендаций по мерам профилактики профзаболевания.

Список литературы: 1. Чикина Н.А. Изучение влияния внутренних факторов риска на развитие алергодерматозов у рабочих химико-фармацевтических предприятий / Н.А. Чикина, И.В. Антонова // Международной науч.-техн. конф. «MicroCAD – 2007». – Харьков: НТУ «ХПИ». – 2007. – Ч.8. – С. 195–200. 2. Чикина Н.А. Математические модели адапционных процессов у рабочих предприятий химико-фармацевтической промышленности / Н.А. Чикина, И.В. Антонова // Информационные технологии и компьютерная инженерия. – Винница: ВНТУ, 2009. – № 2 (15). – С. 45–50. 3. Антонова И.В. Дифференциация уровней риска развития профессионально обусловленных заболеваний / И.В. Антонова, Н.А. Чикина. – Системи обробки інформації. Збірник наукових праць ХУПС. – Харків, 2013. – Випуск 5 (82). – С. 180–184. 4. Чикина Н.А. Математическая модель прогнозирования риска в системе медицинского страхования профпатологий / Н.А. Чикина, И.В. Антонова // Вестник НТУ «ХПИ». Тематический выпуск: Информатика и моделирование. – Харьков, НТУ «ХПИ». – 2013. – № 40. – С. 196–201. 5. Zadeh L.A. Fuzzy sets // Information and Control, 1965. – № 8 (3). – P. 338–353. 6. Штовба С.Д. Проектирование нечетких систем средствами MATLAB – М.: Горячая линия-Телеком, 2007. – 288 с. 7. Гланц С. Медико-биологическая статистика: [Пер. с англ.] / Стентон Гланц. – М.: Практика, 1999. – 459 с. 8. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений – М.: Мир, 1976. – 165 с.

Поступила в редколлегию 20.05.2013

УДК 681.518

Нечёткая экспертная система прогноза риска в страховании профзаболеваний / И. В. Антонова // Вісник НТУ «ХПІ». Серія: Математичне моделювання в техніці та технологіях. – Харків: НТУ «ХПІ», 2013. – №37 (1010). – С. 10 – 16. Бібліогр.: 8 назв.

Запропонована експертна система з нечіткими вирішальними правилами для прогнозу індивідуального ризику розвитку професійно обумовлених алергодерматозів на основі аналізу факторів ризику, оцінки адаптаційних можливостей та захисних властивостей організму в умовах шкідливого виробництва. Визначені функціональні характеристики її модулів, проведено комп'ютерне моделювання й параметрична оптимізація в середовищі MatLab.

Ключові слова: експертна система, нечіткі вирішальні правила, прогноз індивідуального ризику, комп'ютерне моделювання, параметрична оптимізація.

In the article the fuzzy expert system of risk prognosis of development of the professionally conditioned allergodermatoses is worked out on the basis of analysis of risk factors, estimation of adaptation possibilities and protective properties of organism in the conditions of harmful production. Functional descriptions of her modules are certain, a computer modeling and optimization of parameters are conducted in the environment of MatLab. Figs.: 4. Refs.: 8 titles.

Key words: expert system, fuzzy decision rules, prognosis of individual risk, risk factors, computer modeling, optimization of parameters.