

УДК 004.891.3

**МЕДИЦИНСКАЯ ДИАГНОСТИКА ЗДОРОВЬЯ РАБОТНИКОВ  
ГАЗОВОЙ ПРОМЫШЛЕННОСТИ НА ОСНОВЕ СИНТЕЗА  
КОМБИНИРОВАННОГО РЕШАЮЩЕГО ПРАВИЛА**

**Бисалиев Рафаэль Валерьевич**

доктор медицинских наук

**Джамбеков Азамат Матифулаевич**

аспирант

Астраханский государственный технический университет, Астрахань

*author@apriori-journal.ru*

**Аннотация.** В работе сделан обзор алгоритмов и методов синтеза решающих правил, которые применяются в имеющихся на сегодняшний день медицинских экспертных системах диагностики. С целью повышения эффективности принятия решений при диагностике здоровья работников газовой промышленности предложен вариант комбинированного решающего правила, которое включает описание структуры симптомокомплексов и вероятностный метод Вальда.

**Ключевые слова:** метод Вальда; комбинированное решающее правило; поддержка принятия решений; симптомокомплекс.

# MEDICAL DIAGNOSIS OF HEALTH WORKERS GAS INDUSTRY BASED ON COMBINED SYNTHESIS DECISION RULES

**Bisaliev Rafael Valeryevich**

doctor of medicine

**Dzhambekov Azamat Matifulaevich**

post-graduate student

Astrakhan State Technical University, Astrakhan

**Abstract.** The paper provides an overview of the methods of synthesis algorithms and decision rules that apply to currently available medical diagnostic expert systems. To improve the efficiency of decision-making in the diagnosis of the health of workers of the gas industry offered the option of a combined decision rule, which includes a description of the structure and symptom probabilistic method Wald.

**Keywords:** Wald method; combined decision rule; decision support; symptom.

Современные информационно-вычислительные комплексы способны в некоторой степени заменить работу специалистов разных областей или упростить их работу. В частности, это относится к области медицинской диагностики, где возникает множество плохо формализуемых задач. Постановка верного диагноза работникам различных отраслей промышленности является одной из приоритетных задач для медицинских учреждений здравоохранения предприятий.

В работах отечественных и зарубежных исследователей указано наличие у работников нефтегазовой промышленности повышенной частоты распространения разных хронических заболеваний, в том числе

патологий бронхолегочной и сердечно-сосудистой систем, заболеваний верхних дыхательных путей и желудочно-кишечного тракта, болезней кожи, костей и мышц, а также репродуктивных нарушений [1-4].

*Целью исследования* является анализ имеющихся на сегодняшний день методов синтеза решающего правила, которое является ядром подсистемы диагностического комплекса, а также разработка методики синтеза комбинированного решающего правила для повышения эффективности процессов постановки диагноза работникам газовой промышленности.

Среди существующих на сегодняшний день медицинских систем поддержки принятия решений можно отметить: PUFF – диагностика легочных заболеваний; MUCIN – диагностика инфекционных заболеваний; AI/COAG – диагностика заболеваний крови; АСПОН – автоматизированный комплекс для профилактического осмотра населения [5] и другие.

Формирование и принятие решений в данных системах осуществляется на основе следующих алгоритмов и методов.

1. *Метод Байеса*. На основе формулы Байеса производится вычисление вероятностей возможных диагнозов на основе вероятностей возникновения симптомокомплексов и априорных вероятностей самих диагнозов. При превышении полученного значения вероятности некоторого порогового значения делается вывод о наличии диагноза.

2. *Метод Вальда*. Производится вычисление отношения величины правдоподобия для совокупности признаков и его сравнение со значением, полученным при обработке выборки обучения. Для системы признаков  $X$  и двух диагнозов  $D_q$  и  $D_w$  получаем следующее выражение для отношения правдоподобия:

$$\Omega = \frac{P(X / D_q)}{P(X / D_w)}, \quad (1)$$

при

$$P(D_q) + P(D_w) = 1, \quad (2)$$

где  $P(X/D_k)$  – условная вероятность наличия признаков  $X$  при диагнозе  $D_k$ .

Далее, производится сравнение полученного значения  $\Omega$  с пороговыми:

$$A = \frac{1-\beta}{\alpha}; \quad (3)$$

$$B = \frac{\beta}{1-\alpha}; \quad (4)$$

где  $\alpha$  – вероятность совершения ошибки при принятии диагноза  $D_w$  при наличии у больного диагноза  $D_q$ ;  $\beta$  – вероятность совершения ошибки при принятии диагноза  $D_q$  при верном диагнозе  $D_w$ .

Если  $\Omega > A$ , то заключают о наличии диагноза  $D_q$ . Если  $\Omega > B$ , то принимают диагноз  $D_w$ . В противном случае производят отказ от выполнения постановки диагноза.

3. *Методы распознавания образов.* Здесь объект представляется точкой в пространстве признаков. Среди данных методов существует большое разнообразие алгоритмов, имеющих различное вычисляемое расстояние (расстояния Минковского, Эвклида, Камберра и др.) [6], а также различные критерии их оценки (алгоритмы голосования, сравнение с прототипом и др.). Также, сюда можно отнести методы, основанные на нейронных сетях.

4. *Детерминистические методы, основанные на применении симптомокомплексов.* Для каждого диагноза существует набор симптомов (признаков диагноза), которые в совокупности образуют симптомокомплекс заболевания (может быть представлен вектором дихотомических признаков). Неформализованное представление информации о симптомах заболеваний размещено в разных медицинских справочниках [7]. Данная информация служит экспертной оценкой заболевания, которая выработана врачами многих поколений.

Выделяют следующие типы симптомов, на основе которых формируются симптомокомплексы: патогномонические симптомы (однозначно

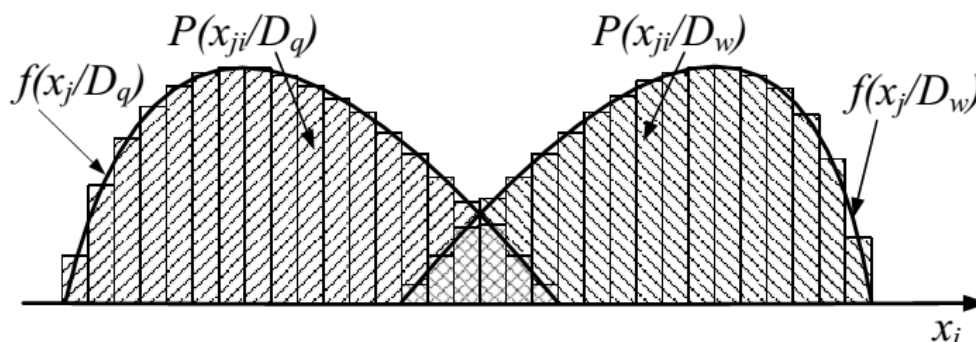
определяют заданное заболевание), специфические симптомы (неоднозначно определяют заданные заболевания), неспецифические симптомы (возможны при заданном заболевании).

Процесс диагностики состоит в сравнении вектора признаков объекта диагностики с симптомокомплексами для известных заболеваний. Критерием сходства может, в частности, служить расстояние Хэмминга [8].

На основе анализ недостатков и достоинств существующих методов и алгоритмов можно сделать вывод о перспективности применения методики синтеза комбинированного решающего правила, которая основана на объединении детерминистического и вероятностного подходов.

Вероятностный подход, базирующийся на методе Вальда, состоит в вычислении априорных объективных условных вероятностей  $P(x_j / D_k)$  на основе анализа обучающей выборки. Осуществляется разбивка числовых признаков  $x_j$  на непересекающиеся интервалы  $\tau_i (i = \overline{1, n})$ , на основе которых строят гистограммы, аппроксимирующие аналитический закон распределения  $f(x_j / D_k)$  рассматриваемого признака, для двух диагнозов  $D_w$  и  $D_q$  на одной координатной плоскости.

На рис.1 изображены некоторые аналитические законы распределения вероятностей  $f(x_j / D_q)$  и  $f(x_j / D_w)$  и аппроксимация гистограммами данных законов распределения.

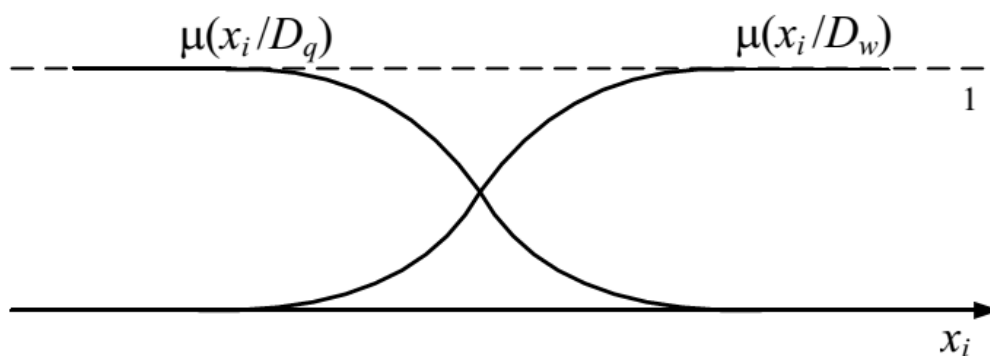


**Рис. 1. Аналитические законы распределения вероятностей и аппроксимация гистограммами данных законов распределения**

Другой подход, базирующийся на методах теории нечеткой логики [9], связан с мнениями экспертов. Здесь вычисляется значение непрерывной функции  $\mu(x_j / D_k)$  – функции принадлежности, полученной на основе суждений специалиста, которая отражает степень соответствия значения численного признака  $x_j$  некоторой лингвистической переменной, описывающей рассматриваемый признак (например, «высокая температура» или «повышенное артериальное давление»).

На рис. 2 приведены функции принадлежности  $\mu(x_i / D_q)$  и  $\mu(x_i / D_w)$ . Для этих функций имеют место соотношения:

$$\begin{aligned} \mu(x_i / D_q) &= \bar{\mu}(x_i / D_w) = 1 - \mu(x_i / D_w) \\ \mu(x_i / D_w) &= \bar{\mu}(x_i / D_q) = 1 - \mu(x_i / D_q) \end{aligned} \quad (5)$$



**Рис. 2. Функции принадлежности  $\mu(x_i / D_q)$  и  $\mu(x_i / D_w)$**

Помимо этого, для каждого признака  $x_j$  обозначается экспертная оценка, отражающая его вес в симптомокомплексе. Данные оценки служат для взвешивания значения  $\mu(x_i / D_k)$  и принимают значения  $e_0, e_1, e_2, e_3$ , где  $e_0$  – вес патогномонических симптомов;  $e_1$  – вес специфических симптомов;  $e_2$  – вес неспецифических симптомов;  $e_3$  – вес показателей, не входящих в симптомокомплекс данного заболевания.

При этом справедливо неравенство:

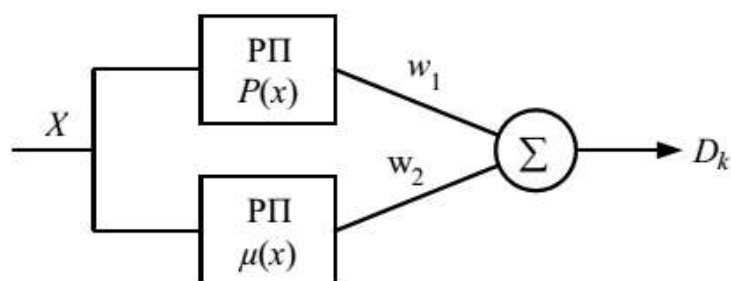
$$e_0 \geq e_1 \geq e_2 \geq e_3, \quad \sum_{i=0}^3 e_i = 1. \quad (6)$$

Таким образом, учитывая оценки  $e_i$ , функция принадлежности принимает вид:

$$\mu_1(x_j / D_k) = e_i \cdot \mu(x_j / D_k). \quad (7)$$

Итак, разные подходы (вероятностный подход, и подход на основе нечеткой логики) используются для решения одной и той же задачи классификации. Вполне естественным является предположение, что совместное использование данных подходов перспективно, поэтому можно предложить следующие варианты их совместного использования:

1. *Коллектив решающих правил.* Данный метод приведен на рис. 3.



**Рис. 3. Структурная схема коллектива решающих правил**

Производится анализ симптомокомплекса  $X$  с помощью каждого из РП со степенью доверия  $w_i$ . На основании взвешенных результатов работы всех решающих правил блок логического вывода формулирует с последующей дефаззификацией окончательный диагноз  $D_k$ .

2. *Суммирование оценок.* Так как условные вероятности  $P(x_j / D_k)$  и значения функции принадлежности  $\mu(x_j / D_k)$  являются нормированными величинами, то можно выполнить их усреднение с учетом весовых коэффициентов  $k_i$ . Тогда выражение (7) преобразуется в

$$\mu(x_j / D_k) = k_1 P(x_{ji} / D_k) + k_2 \mu_1(x_j / D_k), \quad (8)$$

где  $k_i > 0, \sum k_i = 1, i = \overline{1,2}$ , а отношение правдоподобия приобретает вид:

$$\Omega = \prod_{j=1}^m \frac{h(x_j / D_q)}{h(x_j / D_w)}. \quad (9)$$

Таким образом, в работе был осуществлен анализ существующих методов и алгоритмов синтеза решающих правил, применяемых в медицинских диагностических системах. Также, с целью повышения эффективности диагностики здоровья работников газовой промышленности предложен вариант комбинированного решающего правила, который позволяет учесть не только объективные вероятности, но и субъективные оценки экспертов.

### **Список использованных источников**

1. Валеева Э.Т. Профессиональная и производственно-обусловленная патология у работающих в современных нефтехимических производствах / Э.Т. Валеева, Л.К. Каримова, Г.Г. Гимранова // Агрэкологическая безопасность в условиях техногенеза: международный симпозиум: сб. научных докладов. Ч. 1. (21-23 июня 2006 г.). Казань, 2006. С. 129-132.
2. Сивочалова О.В. Критерии оценки профессионального риска репродуктивного здоровья / О.В. Сивочалова, М.А. Фесенко // Профессия и здоровье: материалы IV Всерос. конгресса. М., 2005. С. 136-138.
3. Macdonald, E. B. Occupational medicine in Europe: Evolution of the profession / E. B. Macdonald. WHO/ECEH. Bilthoven 29 March. 1999. P. 31.
4. Takala J. Global estimates of traditional occupational risks: Int. Conference on Occupational Health Services (Helsinki, 25-27 Jan., 2005) / Jukka Takala // Scand. J. Work Environ. and Health. 2005. V. 31. № 1. P. 62-67.
5. Ахутин В.М. Формальная модель автоматизированной системы профилактических осмотров населения / В.М. Ахутин, В.В. Шаповалов // Медицинская техника. 2002. № 1. С. 3-7.
6. Дюк В.А. Компьютерная психодиагностика / В.А. Дюк. СПб.: Братство, 1994. 364 с.



7. Справочник практического врача / под ред. А.А. Михайлова, Л.И. Дворецького. М.: Новая Волна, 2001. 528 с.
8. Поворознюк А.И. Компьютерные системы медицинской диагностики: лабораторный практикум: в 2 ч. Ч. 2. / А.И. Поворознюк, А.Е. Филатова, А.Н. Шеин. Харьков: НТУ «ХПИ», 2007. 96 с.
9. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилиньский, Л. Рутковский. М.: Горячая линия – Телеком, 2006. 452 с.