

ОЦЕНКА СОСТОЯНИЯ ЗДОРОВЬЯ МЕДИЦИНСКИХ РАБОТНИКОВ С ПРИМЕНЕНИЕМ МНОГОМЕРНОГО СТАТИСТИЧЕСКОГО АНАЛИЗА

Татьяна Анатольевна Ермолина^{1*}, Надежда Аркадьевна Шилова¹,
Наталья Алексеевна Мартынова¹, Алексей Генрихович Калинин¹,
Сергей Валентинович Красильников²

¹Северный (Арктический) федеральный университет, г. Архангельск,

²Первая городская клиническая больница им. Е.Е. Воловевич, г. Архангельск

Реферат

Цель. Выявление наиболее важных дифференциально-диагностических критериев из иммунологических показателей с помощью дискриминантного анализа для оценки уровня здоровья медицинского персонала онкологического диспансера.

Методы. С целью моделирования обобщённых показателей уровня здоровья было исследовано состояние здоровья врачей и среднего медицинского персонала, а также проведено одномоментное исследование иммунного статуса у медицинских работников онкологического диспансера. Прежде всего было изучено содержание Т- и В-лимфоцитов с различным фенотипом.

Модель «уровень здоровья медицинских работников» построена на основании выборки из 96 наблюдений (группа медицинского персонала, связанного с работой излучающей аппаратуры). Контрольную группу составили 98 человек того же лечебного учреждения, относящихся к кабинетным работникам.

Результаты. Состояние здоровья можно описывать с помощью макропараметров. При этом под состоянием здоровья в самом общем случае понимают конкретное значение иерархических взаимосвязанных параметров аддитивной модели здоровья. Проведение дифференциальной диагностики на основе оценки клинико-лабораторных данных, функциональных показателей и их значимости даёт возможность прогнозировать уровень здоровья медицинских работников.

Вывод. Информативными признаками для проведения дифференциальной диагностики служат работа в отделении лучевой диагностики, работа в хирургических отделениях, работа с излучающей аппаратурой, наличие у работника среднего или высшего профессионального образования, возраст сотрудника, пол и стаж работы в отделении; наиболее информативные признаки — работа в отделении лучевой диагностики ($F=22,292$), в хирургических отделениях ($F=7,890$), а также с излучающей аппаратурой ($F=3,985$).

Ключевые слова: медицинские работники, заболеваемость, моделирование, прогнозирование, уровень здоровья, дискриминантный анализ.

MODELLING THE HEALTH CONDITION OF MEDICAL WORKERS USING A MULTIVARIATE STATISTICAL ANALYSIS T.A. Ermolina¹, N.A. Shilovskaya¹, N.A. Martynova¹, A.G. Kalinin¹, S.V. Krasilnikov². ¹Northern (Arctic) Federal University, Arkhangelsk, Russia, ²First City Clinical Hospital named after E.E. Volosevich, Arkhangelsk, Russia. **Aim.** To identify the most important differential diagnostic criteria of the immunological parameters with the help of discriminant analysis in order to assess the level of health of medical personnel of an oncology dispensary. **Methods.** With the aim to model the generalized indicators of the health level, evaluated was the health of physicians and nurses, and a cross-sectional study of the immune status of health workers of the oncology dispensary was conducted. Firstly the content of T- and B-lymphocytes with different phenotypes was studied. The model «level of health of medical workers» was built on the basis of a sample of 96 observations (the group of medical personnel, the work of whom is associated with emitting apparatuses). The control group consisted of 98 people of the same medical institution, who are mainly office staff. **Results.** The health status can be described using macro parameters. At the same time under the state of health in the most general sense a specific value of the hierarchically interrelated parameters of the additive model of health is understood. Carrying out a differential diagnosis based on the assessment of clinical and laboratory data, functional parameters and their significance makes it possible to predict the level of health of medical workers. **Conclusion.** Informative features for differential diagnosis are the work at the department of radiology, working in surgical departments, working with the emitting apparatuses, secondary professional or higher professional education of the worker, the age of the employee, gender of the employee and the employee's length of service for the department; the most informative features — work at the department of radiology ($F=22,292$), at the surgical departments ($F=7,890$), as well as work with the emitting apparatuses ($F=3,985$). **Keywords:** medical workers, morbidity, modelling, predicting, level of health, discriminant analysis.

В настоящее время в медико-биологических исследованиях широко используют математическое моделирование [1]. С целью выявления наиболее значимых иммунологических показателей мы применили дискриминантный анализ для оценки уровня здоровья медицинского персонала онкологического диспансера.

Для моделирования обобщённых показателей уровня здоровья обследованы врачи и средний медицинский персонал онкологического диспансера с исследованием иммунного статуса (прежде всего — содержания субпопуляций Т- и В-лимфоцитов).

Модель «уровень здоровья медицинских работников» построена на основании выборки из 96 наблюдений (группа медицинского персона-

ла, связанного с работой излучающей аппаратуры). Контрольную группу составили 98 человек того же лечебного учреждения, относящихся к кабинетным работникам.

Для иммунологического обследования использовали периферическую венозную кровь (10 мл), взятую из локтевой вены утром натощак. Гематологические исследования включали определение общего содержания лейкоцитов и лейкограмму для последующей оценки концентрации лимфоцитов с различным фенотипом. Общее количество лейкоцитов определяли в камере Горяева с применением объектива $\times 20$ и окуляра $\times 7$, вычисляли по формуле абсолютное содержание клеток в 1 мл крови. Количество лимфоцитов, моноцитов, нейтрофилов подсчитывали в мазках крови, окрашенных по методу Романовского-Гимзы. Для регистрации клеток использовали клавишные лабораторные счётчики, в каждом случае подсчитывали 100 клеток. Субпопуляции лимфоцитов определяли методом непрямой иммунопероксидазной реакции с использованием моноклональных антител производства НПЦ «МедБиоСпектр» (г. Москва). Содержание цитокинов (интерлейкинов-10 и -6, фактора некроза опухолей альфа, растворимого рецептора фактора некроза опухолей) определяли методом «конкурентного» иммуноферментного анализа с реактивами «Biosource» (США) на фотометре серии «Мультискан» (Финляндия). Результаты получали в пикограммах на миллилитр. Количественное содержание в сыворотке крови раково-эмбрионального антигена оценивали твёрдофазным неконкурентным методом, основанным на прямой «сэндвич»-технологии с помощью иммуноферментного набора фирмы «Cap Ag» (Швейцария). Фагоцитарную активность нейтрофилов определяли с помощью тест-набора химической компании «Раекомплекс» (Чита). Оценку результатов проводили по фагоцитарному показателю (процент фагоцитирующих клеток из числа сосчитанных нейтрофилов), а также по фагоцитарному числу (количество частиц, поглощённых одним активным нейтрофилом).

Соблюдали этические принципы, предьявляемые Хельсинской декларацией Всемирной медицинской ассоциации (World Medical Association Declaration of Helsinki) от 1964 г., с последним пересмотром в 2000 г.

Для статистической обработки материала использовали программу «Statistica 6.0». Серии сравнивали между собой по дисперсии с помощью критерия Стьюдента и U-критерия Манна-Уитни, включая проверку гипотез о равенстве средних и дисперсии групп.

Проводили дискриминантный анализ (метод многомерного статистического анализа), который позволил оценить качество и точность группирования объектов, а также выделить наиболее информативные признаки их деления на группы [2].

Проверку точности представленных групп проводили с помощью линейных классифика-

ционных функций, информативность признаков оценивали по F-критерию Фишера. Критический уровень значимости при проверке статистических гипотез $p < 0,05$.

В ходе дискриминантного анализа рассчитывали ряд характеристик, необходимых для интерпретации групп и их различий, — линейные классификационные функции, канонические дискриминантные функции, факторная нагрузка канонических дискриминантных функций. Дискриминантный анализ дал возможность построения графиков положения групп в признаковом пространстве.

При проведении дискриминантного анализа выявлены достоверные различия в уровне здоровья между группами медицинских работников, работающих в разных отделениях, с излучающей аппаратурой и без её использования, а также в зависимости от возраста, пола и стажа работы в отделениях. Все представленные в модели переменные — информативные параметры с уровнем статистической достоверности от $p < 0,05$ до $p < 0,001$.

Предполагаем, что на состояние здоровья влияют факторы, которые характеризуются двумя типами данных:

1) качественными данными, связанными с социальным статусом работников и их профессиональной деятельностью (работа в отделе лучевой диагностики, работа в отделении хирургии, наличие у работника высшего профессионального образования, наличие у работника среднего профессионального образования, работа с излучающей аппаратурой, возраст, пол, стаж работы);

2) количественными данными, то есть иммунологическими показателями крови — содержанием клеток CD8, CD10, CD25, CD71, CD HLA DR, CD16, CD4.

После определения факторных и зависимых переменных была построена простейшая линейная вероятностная модель для выявления влияния факторных переменных на уровень здоровья.

Простейшая линейная вероятностная модель имеет следующий вид:

$$\hat{x}_9 = 0,69 - 0,41 \cdot x_1 - 0,51 \cdot x_2 + 0,57 \cdot x_3 + 0,22 \cdot x_4 - 0,12 \cdot x_5 + 0,004 \cdot x_6 - 0,06 \cdot x_7 - 0,01 \cdot x_8$$

$\begin{matrix} 1,85 & -1,9 & -2,22 & 2,81 & 1,32 & -0,65 \\ 0,62 & -0,40 & -1,68 \end{matrix}$

где: x_1 — работа в отделе лучевой диагностики; x_2 — работа в отделении хирургии; x_3 — наличие у работника высшего профессионального образования; x_4 — наличие у работника среднего профессионального образования; x_5 — работа с излучающей аппаратурой; x_6 — возраст сотрудника; x_7 — пол сотрудника; x_8 — стаж работы.

Факторы x_1 , x_2 , x_5 , x_7 и x_8 снижают уровень здоровья работника, так как коэффициенты перед ними имеют отрицательное значение, а факторы x_3 , x_4 и x_6 повышают уровень здоровья работника, поскольку коэффициенты перед ними имеют положительное значение.

Сравнивая коэффициенты модели с табличным $t_{табл} = 1,96$, заметим, что значимыми в модели являются свободный коэффициент 0,69 (наибольшее численное значение) и коэффициенты при переменных x_1 (работа в отделе лучевой диагностики), x_2 (работа в отделении хирургии), x_3 (наличие у работника высшего профессионального образования), x_5 (работа с излучающей аппаратурой), x_6 (возраст сотрудника), x_7 (пол сотрудника), x_8 (стаж работы).

Далее определяли коэффициент детерминации, который представляет собой долю объяснённой дисперсии отклонений зависимой переменной от её среднего значения. Зависимая переменная объясняется (прогнозируется) с помощью функции от объясняющих переменных. В частном случае R^2 является квадратом коэффициента корреляции между зависимой переменной и её прогнозными значениями с помощью объясняющих переменных. Коэффициент детерминации, вычисленный по выборочным данным, невелик ($R^2=0,22$), но с учётом числа степеней свободы его значение увеличивается ($R^2_{множ} = 0,47$), и можно предположить, что качество модели удовлетворительное. Наблюдаемое значение критерия Фишера $F=3,01$ превосходит табличное значение $F_{табл} = 2,04$, что позволяет утверждать, что модель пригодна для прогнозирования (уравнение регрессии статистически значимо). Причиной статистической незначимости отдельных коэффициентов модели может быть мультиколлинеарность факторов.

Выявлена сильная корреляция между факторами x_2 (работа в отделении хирургии) и x_5 (работа с излучающей аппаратурой) ($r_{25} = -0,69$), x_3 (наличие у работника высшего профессионального образования) и x_7 (пол сотрудника) ($r_{37} = 0,58$), x_6 (возраст сотрудника) и x_8 (стаж работы) ($r_{68} = 0,75$).

Для выявления влияния факторных переменных на уровень здоровья строили логистическую регрессию. На первом шаге анализа оценена значимость влияния отдельных предикторов на зависимую переменную без учёта влияния остальных. Для переменной x_3 приписаны баллы 14,058, для x_4 — 5,964, для x_8 — 0,873, для x_6 — 0,814, для x_1 — 0,474, для x_2 — 0,306.

На втором шаге оценены значимые коэффициенты логистической регрессии. Уравнение принимает следующий вид:

$$\hat{x}_9 = \frac{1}{1 + e^{-0,475} + e^{1,728 \cdot x_3}}$$

При проверке модели получено 67,7% правильных предсказаний. Можно сделать вывод, что, согласно модели логистической регрессии, на уровень здоровья наибольшее влияние оказывает факторная переменная x_3 (наличие у работника высшего профессионального образования).

Для выявления влияния факторных переменных на уровень здоровья проводили дискриминантный анализ. Исходными данными для классификации были объекты (объём выборки

$n=96$), разделённые на две группы (1 — здоров, 0 — нездоров), причём каждый объект относится только к одной группе.

Каждый объект наблюдения характеризовался рядом признаков или дискриминантных переменных: x_1 (работа в отделе лучевой диагностики), x_2 (работа в отделении хирургии), x_3 (наличие у работника высшего профессионального образования), x_4 (наличие у работника среднего профессионального образования), x_5 (работа с излучающей аппаратурой), x_6 (возраст сотрудника), x_7 (пол сотрудника), x_8 (стаж работы).

В качестве обучающих выборок последовательно используются x_9 (уровень здоровья), x_{10} (CD8), x_{11} (CD10), x_{12} (CD25), x_{13} (CD71), x_{14} (CD HLA DR), x_{15} (CD16), x_{16} (CD4).

Оценка степени влияния дискриминантных переменных на зависимую переменную x_9 (уровень здоровья) представлена в табл. 1.

Таблица 1

Оценка степени влияния дискриминантных переменных на зависимую переменную x_9 (уровень здоровья)

Факторные переменные	F-включения
x_1 — работа в отделе лучевой диагностики	6,399*
x_2 — работа в хирургических отделениях	8,476*
x_3 — наличие у работника высшего профессионального образования	16,126**
x_4 — наличие у работника среднего профессионального образования	4,637
x_5 — работа с излучающей аппаратурой	3,985
x_6 — возраст сотрудника	3,448*
x_7 — пол сотрудника	3,009
x_8 — стаж работы сотрудника в отделении	5,524*

Примечание: * $p < 0,05$; ** $p < 0,001$. Наиболее информативные признаки — параметры x_1 , x_2 и x_3 .

Вычисленное значение критерия Пирсона $\chi^2=14,804$ значительное (чем выше его значение, тем качественнее дискриминантная функция различает группы).

Каноническая дискриминантная функция, в соответствии с которой предсказывались значения зависимой переменной, имела следующий вид:

$$\hat{x}_9 = -1,431 + 2,687 \cdot x_3$$

Наибольшее влияние на уровень здоровья оказывает дискриминантная переменная x_3 (наличие у работника высшего профессионального образования).

Для решения диагностической задачи по каноническим уравнениям определяли канонические линейные дискриминантные функции (КЛДФ), обобщающие все признаки, включённые в модель, по всем обследуемым медицинским работникам. Вклад КЛДФ в меж-

Расчётные значения коэффициентов канонических линейных дискриминантных функций для рассматриваемых факторов, влияющих на уровень здоровья медицинских работников

Фактор	Работа в отделе лучевой диагностики	Работа в хирургических отделениях	Наличие высшего профессионального образования	Наличие среднего профессионального образования	Работа с излучающей аппаратурой	Возраст	Пол	Стаж работы сотрудника в отделе
CD4	–	2,668	2,019	–	–	–	–	–
CD8	–	12,056	0,055	–	8,484	0,387	–	–
CD10	1,388	–	–	–	–	–	–	–
CD16	–	–	–	–	–	–	–	0,153
CD25	14,27	12,827	–	–	–	–	–	–
CD HLA DR	–	–	–	3,374	–	0,279	2,61	–
CD71	–	–	–	2,369	–	–	–	–

групповую дисперсию признаков оценивали по критерию Пирсона χ^2 . Этот вклад признавали достоверным при уровне статистической значимости $p < 0,05$.

Для выявления наибольшего влияния предикторов на зависимые переменные исследовали влияние факторных переменных отдельно на содержание в крови клеток CD4, CD8, CD10, CD16, CD25, CD HLA DR, CD71.

Коэффициенты КЛДФ представлены в табл. 2. Они показывают относительный вклад каждой переменной в значение дискриминантной функции с учётом влияния всех остальных переменных. Чем больше абсолютное значение коэффициента, тем больше относительный вклад данной переменной в значение дискриминантной функции, разделяющей классы.

По приведённым в табл. 2 данным можно оценить вклады различных факторов, влияющих на уровень здоровья, и кумулятивный вклад. С первой КЛДФ (содержание в крови клеток CD4) более тесно связаны такие переменные, как работа в хирургических отделениях и наличие высшего профессионального образования, со второй КЛДФ (содержание в крови клеток CD8) более тесно связаны следующие переменные: работа в хирургических отделениях, наличие высшего профессионального образования, работа с излучающей аппаратурой, возраст, и т.д.

Данные о факторной нагрузке КЛДФ можно использовать для оценки коэффициентов «весомости» переменных при решении диагностической задачи.

ВЫВОДЫ

1. Информативные признаки для проведения дифференциальной диагностики – работа в отделении лучевой диагностики, работа в хирургических отделениях, работа с излучающей аппаратурой, наличие среднего профессионального образования, наличие высшего профессионального образования, возраст, пол и стаж работы сотрудника в отделении.

2. Наиболее информативные признаки – работа в отделении лучевой диагностики ($F=22,292$), в хирургических отделениях ($F=7,890$), а также работа с излучающей аппаратурой ($F=3,985$).

3. Точность диагностики по решающим правилам в среднем имеет достоверность 95,3%.

4. Состояние здоровья можно описывать с помощью макропараметров. При этом под состоянием здоровья в самом общем случае понимаются конкретное значение иерархических взаимосвязанных параметров аддитивной модели здоровья. Проведение дифференциальной диагностики на основе оценки клинко-лабораторных данных, функциональных показателей и их значимости даёт возможность прогнозировать уровень здоровья медицинских работников.

ЛИТЕРАТУРА

1. Токмачев М.С. Здоровье населения региона: модели и управление // Проблем. управл. – 2010. – № 6. – С. 45–52.
2. Фаткуллина И.Б., Протопопова Н.В., Михалевич И.М. Дискриминантный анализ как метод проведения дифференциальной диагностики артериальной гипертензии при беременности // Вестн. нов. мед. технол. – 2011. – Т. XVIII, №1. – С. 134–135.