

УДК 612:001.891+504.05

БИОХИМИЧЕСКАЯ ОЦЕНКА СОСТОЯНИЯ ЗДОРОВЬЯ ЧЕЛОВЕКА**И.П. ТАРАСОВА***(Витебский государственный университет им. П.М. Машерова, Витебск)*

Рассмотрены экспертные системы, которые позволили бы на биохимическом уровне зарегистрировать ранние, доклинические изменения метаболизма и тем самым различить здоровый и больной организм. Математически рассчитана производная величина, отражающая «идеальный» уровень здоровья индивида. Полученные экспертные системы могут быть использованы для индивидуальной диагностики людей.

В современных условиях высокой технической оснащенности и психофизической интенсивности труда одним из главных лимитирующих факторов становится фактор здоровья. Длительное нарушение баланса между работой и отдыхом, хроническое действие неблагоприятных экологических и производственных факторов приводит к истощению компенсаторно-приспособительных механизмов организма. Неправильный выбор поля деятельности, когда с возрастом начинают проявляться противоречия между требованием профессии, с одной стороны, и индивидуально-отологическими (конехитуниональными) свойствами организма, реальными показателями здоровья - с другой, является причиной снижения производительности труда, уменьшения продолжительности жизни и т.д.

Загрязнения всех жизненных сред и продуктов питания приводит к возникновению целого ряда социальных проблем, среди которых и рост заболеваемости [1]. Под действием «химического пресса» и ряда других экологически неблагоприятных факторов происходят изменения и в биохимическом составе крови [2]. Действующие в организме биологические системы, обеспечивающие гомеостаз, в этих условиях не в состоянии их полностью восстановить. Следствием этого является возникновение патологических изменений. Такая перестройка происходит задолго до явного и устойчивого нарушения гомеостаза и свидетельствует о скрытом изменении «уровня здоровья» индивидуума.

Если проанализировать весь спектр заболеваний, то окажется, что медицина сама по себе имеет к ним отношение примерно на 10 % [3]. В остальном болезни в значительной степени зависят от неблагоприятных экологических и производственных факторов (20 %), образа жизни (50 %), распываемой наследственности (20 %). При этом все большее внимание должно уделяться не только явным, лежащим на поверхности и легко поддающимся анализу потерям труда в результате временной нетрудоспособности, но и скрытым, являющимся следствием хронического недомогания и предболезненного состояния.

Особую актуальность в связи с этим приобретает разработка и реализация комплексных оздоровительно-профилактических мероприятий, направленных на снижение уровня профессиональных заболеваний, выделение основных факторов риска, снижение функциональной напряженности на различных этапах жизнедеятельности. Используя достижения математики и вычислительной техники, можно значительно упростить и ускорить процесс распознавания патологических образов. Комплекс средств, обеспечивающих массовое прогностическое обследование здоровья населения, предполагает использование автоматизированной системы донозологической диагностики, позволяющей с помощью пакета программ экспресс-методами оценить уровень адаптационного резерва различных функциональных систем организма.

Развитие и внедрение современных медицинских технологий предусматривают применение информационных систем, позволяющих улучшить качество диагностического и лечебного процессов. Данный подход предусматривает создание экспертных систем (ЭС), под которыми понимают системы, способные выносить суждения и принимать решения без участия человека-эксперта [4]. Для построения экспертной системы могут быть использованы вероятностные и детерминистские подходы. В основе идеологии создания ЭС лежат системный подход, концепция различия и статистические методы распознавания образов.

В связи с вышесказанным актуальным является внедрение и использование новых способов анализа полученной информации с помощью компьютера. Одним из подходов в разработке новой системы диагностики может служить классификационная система, использование которой с высокой долей вероятности отнесет исследуемого индивида к одной из нескольких групп, каждая из которых характеризуется наличием патологии или ее отсутствием.

Таким образом, целью настоящей работы явилось создание экспертных систем, которые позволили бы на биохимическом уровне зарегистрировать ранние, доклинические изменения метаболизма и тем самым различить здоровый и больной организм.

Материал и методы. Для анализа нами была использована база данных Республиканского липидного лечебно-диагностического центра метаболической терапии (г. Витебск). Она включала в себя результаты комплексного биохимического обследования 129 практически здоровых людей (средний воз-

раст $36,68 \pm 1,04$) и 5071 - с различными патологиями (средний возраст $51,96 \pm 0,17$). Объектом анализа явились следующие биохимические показатели крови: ТГ; общий холестерин (ОХС); ХС ЛПВП (холестерол липопротеидов высокой плотности); ХС ЛПНП (холестерол липопротеидов низкой плотности); ХС ЛПОНП (холестерол липопротеидов очень низкой плотности); ИА (индекс атерогенности липопротеидов); мочевины; билирубин; глюкоза; общий белок; альбумин; мочевиная кислота; аспаратаминотрансфераза (АСТ); аланинаминотрансфераза (АЛТ), щелочная фосфатаза (ЩФ) и у-глутамилтранспептидаза (ГГТ).

Математико-статистический анализ проводили с помощью прикладных компьютерных пакетов Microsoft Excel 2000, а также универсального статистического пакета STATGRAPHICS Plus 2.1 для Windows [5]. На основании полученной информации были созданы исходные базы данных.

Для решения поставленной задачи нами был использован дискриминантный анализ, так как он позволяет изучить различия между двумя и более группами объектов по отношению к нескольким переменным одновременно.

Результаты и обсуждения. Первоначально полученные результаты обследований были подвергнуты стандартному статистическому анализу: расчет среднего значения, стандартной ошибки, медианы, моды, стандартного отклонения, дисперсии выборки, коэффициентов эксцесса и асимметрии. Далее исходная база данных была подвергнута корреляционному анализу. После анализа матрицы коэффициентов корреляции несколько уменьшился объем материала, подлежащего дальнейшему рассмотрению. Во-первых, содержание ОХС сильно коррелировано ($r = 0,95$; $p < 0,0001$ для двух групп) с содержанием ХС ЛПНП. Во-вторых, обнаружена функциональная корреляционная зависимость между триацетилглицеринами сыворотки крови и холестерином ЛПОНП у практически здоровых людей ($r = 1,00$; $p < 0,0001$) и весьма тесная взаимосвязь этих же показателей у больных ($r = 0,99$; $p < 0,0001$). Так как в дискриминантный анализ [6] может быть взят только один член пары признаков, для которой коэффициент корреляции близок к 1. поэтому ХС ЛПОНП и ХС ЛПНП не были использованы в качестве входных параметров. Кроме того, был исключен и индекс атерогенности как производный показатель, вычисляемый из значений ОХС и ХС ЛПВП, что сделано также в соответствии с этими требованиями. Все остальные биохимические показатели линейно независимы, что подтверждается результатами корреляции, и могут быть использованы в дискриминантном качественном анализе.

При проведении дискриминантного анализа в качестве классифицирующего признака мы использовали перекодированные ряды, которые рассматривались как входные данные: 1 - больные, 2 - практически здоровые люди. В качестве входных параметров мы использовали пол (X_1), возраст (X_2), а также перекодированные биохимические показатели крови: X_3 - ОХС; X_4 - ХС ЛПВП; X_5 - ТГ; X_6 - мочевины; X_7 - билирубин; X_8 - глюкоза; X_9 - общий белок; X_{10} - мочевиная кислота; X_{11} - ГГТ; X_{12} - ЩФ; X_{13} - АСТ; X_{14} - АЛТ; X_{15} - альбумин.

Все модели построены по единому алгоритму, с различиями в выборе входных параметров. Для удобства понимания логики построения отдельных моделей опишем общую структуру всех построенных уравнений.

Модель № 1. Входят все биохимические показатели, а также учитывался возраст и пол обследованных.

Модель № 2. Пол, возраст, ОХС, ХС ЛПВП, ХС ЛПНП, ТГ, мочевины, билирубин, глюкоза, общий белок, мочевиная кислота и альбумин.

Модель № 3. Пол, возраст, ОХС, ХС ЛПВП, ХС ЛПНП, ТГ, ГГТ, ЩФ, АСТ и АЛТ.

Модель № 4. Пол, возраст, мочевины, билирубин, глюкоза, общий белок, мочевиная кислота, альбумин, ГГТ, ЩФ, АСТ и АЛТ.

Модель № 5. Пол, возраст, ОХС, ХС ЛПВП, ХС ЛПНП и ТГ.

Модель № 6. Пол, возраст, мочевины, билирубин, глюкоза, общий белок, мочевиная кислота и альбумин.

Модель № 7. Пол, возраст, ГГТ, ЩФ, АСТ и АЛТ.

Модель № 8. Пол, возраст, ОХС, ХС ЛПВП, ХС ЛПНП, ТГ, ГГТ, глюкоза и общий белок.

Модель № 9. Пол, возраст, ОХС, ХС ЛПНП, ТГ, ГГТ, глюкоза и общий белок.

Модель № 10. Пошаговый дискриминантный анализ, позволяющий выбрать наиболее значимые показатели, которыми оказались возраст и глюкоза.

По коду проведения дискриминантного анализа нами оценивались: 1) безошибочность ЭС; 2) чувствительность и специфичность ЭС; 3) коэффициент канонической корреляции.

Исходя из полученных данных установлено, что для проведения анализа необходима одна дискриминантная функция, которая объясняет 100 % суммарной дисперсии.

По результатам эксперимента нами предложено 10 моделей ($p < 0,001$), которые имеют процент общих совпадений по классам 78-91 % (табл. 1 - 3), коэффициенты канонической корреляции находятся в пределах $r = 0,381 - 0,679$ ($p < 0,001$).

Таблица 1

Эффективность отнесения к классам и процент общих совпадений в экспертных системах, построенных для различения больных ИБС и практически здоровых людей

№ варианта	Каноническая корреляция	% общих совпадений по классам				% общих совпадений
		Класс 1		Класс 2		
		А	В	А	В	
1	0.679	90,61	9,39	93,94	6,06	91,12
2	0.599	89,41	10,59	90,24	9,76	89,90
3	0.660	90,06	9,94	93,94	6,06	90,65
4	0.672	89,50	10,50	93,94	6,06	90,19
5	0.381	78,30	21,70	72,97	27,03	78,20
6	0.583	88,16	11,84	90,24	9,76	88,40
7	0.653	88,40	11,60	93,94	6,06	89,25
8	0.462	90,41	9,59	92,68	7,32	90,54
9	0.462	89,41	10,59	92,68	7,32	89,59
10	0.658	88,95	11,05	93,94	6,06	89,72

Условные обозначения: класс 1 соответствует группе больных; класс 2 - соответствует группе практически здоровых людей; А - процент обследованных людей, правильно отнесенных к классу; В - процент обследованных людей, ошибочно отнесенных к классу.

Таблица 2

Коэффициенты дискриминантных функций в экспертных системах, построенных для различения практически здоровых и больных людей

Показатель	Номер варианта экспертной системы									
	1	2	3	4	5	6	7	8	9	10
X ₁	0,227	-0,294	0,239	0,236	-0,033	-0,361	0,236	0,323	0,338	11
X ₂	0,083	0,085	0,088	0,085	0,081	0,088	0,09	-0,077	-0,077	н
X ₃	0,034	0,038	0,019	н	0,051	н	н	-0,067	-0,065	0,084
X ₄	-0,109	-0,123	-0,017	н	-0,296	н	н	0,194	н	н
X ₅	0,15	0,237	0,18	н	0,154	н	н	-0,13	-0,144	н
X ₆	-0,055	-0,044	н	-0,057	н	-0,052	н	н	н	н
X ₇	0,02	-0,004	н	0,016	н	-0,014	н	н	н	н
X ₈	0,251	0,18	н	0,269	н	0,204			-0,081	0,271
X ₉	-0,009	-0,023	н	-0,008	н	-0,023	н	0,046	0,047	н
X ₁₀	-0,0004	-0,0002	н	0,0002	н	0,001	н	н	н	н
X ₁₁	-0,006	н	0,006	0,006	н	н	0,007	н	н	н
X ₁₂	0,002	н	0,001	0,002	н	н	0,001	н	н	н
X ₁₃	-0,007	н	-0,01	-0,008	н	н	-0,011	н	н	н
X ₁₄	0,005	н	0,003	0,008000	н	н	0,006	н	н	н
X ₁₅	0,006	0,005	н	0,007	н	0,005		н	н	н
Константа	-6,194	-3,481	-5,408	-6,354	-4,41	-3,275	-5,214	0,746	0,937	-5,983

Таблица 3

Координаты групповых центроидов в экспертных системах, построенных для различения практически здоровых и больных людей

№ варианта	Групповой центроид		№ варианта	Групповой центроид	
	№ 1	№2		№ 1	№2
1	0,393	-2,154	6	0,256	-2,002
2	0,266	-2,08.5	7	0,366	-2,007
3	0,374	-2,049	8	-0,126	2,151
4	0,386	-2,115	9	-0,126	2,146
5	0,025	-1,331	10	0,371	-2,035

Не рассматривая подробно результаты дискриминантного анализа, представленные в табл. 1 – 3, перейдем к характеристике наиболее удачных и практически значимых, на наш взгляд, моделей.

Самой информативной оказалась ЭС, где в качестве входных параметров использовались все биохимические показатели, а также учитывались пол и возраст. Она позволяет получить самый высокий процент общих совпадений по классам – 91,21 %. Данная система способна выявить патологические изменения, происходящие в организме, на биохимическом уровне с точностью 90,61 %. Практически здоровые люди при этом правильно диагностируются в 93,94 % случаев. Коэффициент канонической корреляции данной экспертной системы свидетельствует о взаимосвязи средней ($r = 0,679$; $p < 0,0001$) тесноты между биохимическими показателями двух групп.

Функциональная модель № 1 выглядит следующим образом:

$$Y_1 = -6,194 + 0,227X_1 + 0,083X_2 + 0,034X_3 - 0,109X_4 + 0,15X_5 - 0,055X_6 + 0,02X_7 + 0,251X_8 - 0,009X_9 - 0,0004X_{10} + 0,006X_{11} + 0,002X_{12} - 0,007X_{13} + 0,005X_{14} + 0,006X_{15}.$$

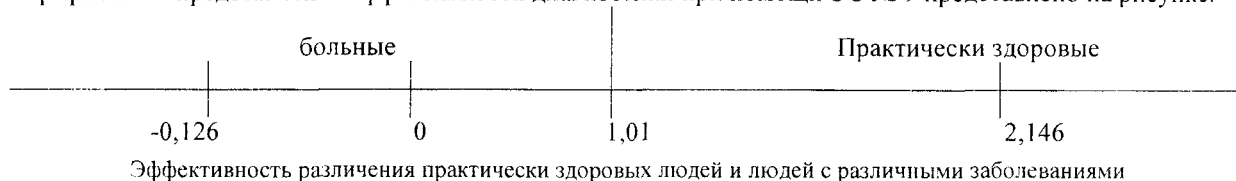
Координаты групповых центроидов данной экспертной системы следующие: (0,393) – для больных; (-2,154) – для практически здоровых людей.

Нами разработана экспертная система № 8, в которой в качестве входных параметров использовались следующие: пол, возраст и всего 5 биохимических показателей. Безошибочность этой экспертной системы составила 90,54 %. При этом больные правильно диагностируются на 90,41 %, а здоровые на 92,68 %, коэффициент канонической корреляции равен 0,462 ($p < 0,0001$).

Модель № 9. Безошибочность ЭС составляет 89,59 %. Чувствительность экспертной системы составляет 90,41 %, специфичность – 92,68 %. Коэффициент канонической корреляции равен 0,462 ($p < 0,0001$). Уравнение дискриминантной функции № 9 имеет вид:

$$Y_9 = 0,937 + 0,338X_1 - 0,077X_2 - 0,065X_3 - 0,144X_5 + 0,081X_8 + 0,047X_9.$$

Координаты групповых центроидов: (-0,126) – для больных; (2,146) – для практически здоровых. Графическое представление эффективности диагностики при помощи ЭС № 9 представлено на рисунке.



С учетом значений групповых центроидов, критической точкой является 1,01 (см. рисунок); ее превышение свидетельствует о принадлежности к группе практически здоровых людей, а цифры ниже критической точки указывают на наличие биохимических изменений, характерных для патологических сдвигов.

Конкретные значения интегральных показателей состояния организма (Y), вычисленные с помощью приведенного выше уравнения для всех представителей, размещаются на непрерывной шкале состояний. При этом величины у лиц 1-й группы (больные) концентрируются в левой части шкалы, в области отрицательных значений, а величины у лиц 2-й группы (практически здоровые) размещаются в правой части шкалы, в зоне её положительных значений.

Значительный интерес представляет возможность использования полученных экспертных систем для индивидуальной диагностики людей. Например, если при вычислении интегрального показателя состояния организма полученная величина приближается к центроиду группы № 1, то можно говорить об отрицательном воздействии какого-либо фактора среды на организм человека. Если рассчитанное значение Y находится в правой части шкалы и приближается к центроиду группы № 2 (практически здоровых людей), то говорят об отсутствии риска развития какого-либо заболевания.

Таким образом, можно сказать, что центроид группы практически здоровых людей – это некая производная величина, отражающая «идеальный» уровень здоровья индивида. Любые отклонения от него в сторону центроида больных свидетельствуют о нарушении гомеостаза в организме.

Данные экспертные системы могут иметь важное практическое значение, так как применение их на практике позволит зарегистрировать ранние, доклинические изменения метаболизма и организовать раннюю профилактику выявленных нарушений. Кроме того, возможно использовать полученные экспертные системы для оценки влияния неблагоприятных экологических факторов на здоровье человека.

ЛИТЕРАТУРА

1. Смычек В.Б. Здоровье населения Республики Беларусь: современное состояние и ближайшие перспективы // Экологическая антропология: Ежегодник. - Мн., 2003. - С. 8 - 11.
2. Маршал В.Дж. Клиническая биохимия / Пер. с англ. - М.: БИНОМ, 2000. - С. 368.
3. Казни Э.М., Блинова Н.Г., Литвинова Н.А. Основы индивидуального здоровья человека: введение в общую и прикладную валологию. - М.: ВЛАДОС. - 2000. - 192 с.
4. Кейлор К. Как построить свою экспертную систему. - М.: Энергоатомиздат, 1991. - 288 с.
5. Дюк В. Обработка данных на ПК в примерах. - СПб.: Питер, 1997. - 240 с.
6. Факторный, дискриминантный и кластерный анализ / Под ред. И.С. Енюкова. - М.: Финансы и статистика. - 1989. - 215 с.