



И.М. Прищепа, И.П. Тарасова, Э.А. Доценко

## Оценка влияния экологически неблагоприятных факторов среды на организм человека при помощи экспертных систем

Современные представления о возникновении патологии базируются на нескольких моделях возникновения болезней [1]. Основной медицинской моделью является экологическая [1], в соответствии с которой болезни вызываются внешними повреждающими факторами в результате чего хронические заболевания становятся основными. Доминирование их в среднем и пожилом возрасте объясняется тем, что по мере увеличения длительности жизни повышается вероятность возникновения болезней, вызываемых внешними факторами, так как возрастает время экспозиции этих факторов. Согласно этой модели, для изменения уровня заболеваемости, надо установить природу внешних патогенных факторов и «факторов» риска, которые повышают вероятность развития болезней, и ликвидировать или внешние патогенные причины, или устойчивое их действие. Другой моделью возникновения болезней является генетическая, в основе которой лежит сбой выполнения генетической программы. Имеет место и аккумуляционный механизм возникновения болезней, основанный на том, что источником повреждений, прежде всего, становятся нормальные процессы метаболизма.

Загрязнения всех жизненных сред и продуктов питания приводит к возникновению целого ряда социальных проблем, среди которых и рост заболеваемости [2-4]. Одной из самых важных и в тоже время лабильных систем человека является сердечно-сосудистая система. Она никогда не является индифферентной [5, 6], даже если токсический агент обладает более выраженным тропизмом к нервной системе, печени, почкам и другим органам и тканям, так как кровеносные сосуды являются неотъемлемым компонентом структуры любого органа. Под действием «химического пресса» и ряда других экологически неблагоприятных факторов происходят изменения и в биохимическом составе крови [6, 7]. При стрессовом воздействии факторов внешней среды биологические системы, обеспечивающие гомеостаз организма не в состоянии полностью восстановиться. Следствием этого является возникновение патологических изменений.

Переход от состояния полного здоровья к различным заболеваниям происходит постепенно, сопровождаясь отклонениями биохимических показателей крови от нормальных величин. Не вызывает сомнений то обстоятельство, что для каждого патологического состояния имеет место уникальное сочетание биохимических показателей [8], характерное для той или иной болезни. Более того, формирование образа на биохимическом уровне может происхо-

доть задолго до появления клинических признаков болезни и сопровождается как прогрессирующими отклонениями некоторых показателей от нормальных величин, так и увеличением числа измененных признаков. Используя достижения математики и вычислительной техники можно значительно упростить и ускорить процесс распознавания патологических образцов. В связи с этим возникает необходимость в создании экспертных систем, которые позволили бы проводить компьютерное различение здоровых и патологических образцов и компьютерную диагностику процессов, происходящих в организме.

Исходя из выше сказанного, целью настоящей работы явилось создание экспертных систем, которые на биохимическом уровне способны различить здоровый организм и больной, а так же оценить влияние экологических факторов на сердечно-сосудистую систему.

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 и 2863 больных ИБС. Объектом анализа явились следующие биохимические показатели крови: триацилглицерины (ТГ), общий холестерол (ХС), холестерол липопротеинов высокой плотности (ХС ЛПВП), холестерол липопротеинов низкой плотности (ХС ЛПНП), холестерол липопротеинов очень низкой плотности (ХС ЛПОНП), индекс атерогенности липопротеинов (ИА), мочеви́на, билирубин, глюкоза, общий белок, альбумин, мочева́я кислота, креатинкина́за, амила́за, щелочная фосфота́за, аспаратами́нотрансфера́за (АСТ), аланина́минотрансфера́за (АЛТ) и  $\gamma$ -глутамилтранспептида́за (ГГТ).

Математико-статистический анализ проводили с помощью прикладных компьютерных пакетов Microsoft Excel 2000, а так же универсального статистического пакета STATGRAPHICS Plus 2.1 для Windows [9]. На основании полученной информации были созданы исходные базы данных. Для решения поставленной задачи нами был использован дискриминантный анализ.

Первоначально полученные результаты обследований были подвергнуты стандартному статистическому анализу. Почти все средние значения биохимических показателей у больных ИБС достоверно ( $p < 0,001$ ) превышают эти же значения у практически здоровых людей. Исключения составляют – ХС ЛПВП, общий белок, альбумин – средние значения которых выше у практически здоровых людей. Различия в содержании АСТ и билирубина статистически не достоверны.

Далее исходная база данных была подвергнута корреляционному анализу. После анализа матрицы коэффициентов корреляции несколько уменьшился объем материала, подлежащего дальнейшему рассмотрению. Во-первых, содержание общего ХС сильно коррелировало с содержанием ХС ЛПНП. Коэффициент корреляции для больных ИБС был равен ( $r = 0,95$ ,  $p = 0$ ), а для практически здоровых людей ( $r = 0,97$ ,  $p = 0$ ). Во-вторых, обнаружена функциональная корреляционная зависимость между ТГ сыворотки крови и ХС ЛПОНП у практически здоровых людей ( $r = 1,00$ ,  $p = 0$ ) и весьма тесная взаимосвязь этих же показателей у больных ИБС ( $r = 0,99$ ,  $p = 0$ ). Коэффициенты корреляции равны 1,00 и 0,99 соответственно. Так как в дискриминантный анализ может быть взят только один член пары признаков [10], 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 ИА как производный показатель, вычисляемый из значений общего ХС и ХС ЛПВП, что сделано также в соответствии с этими требованиями. Все остальные биохимические показатели линейно независимы, что подтвержда-

ются результатами корреляции, и могут быть использованы в качестве входных параметров при проведении анализа.

При проведении дискриминантного анализа в качестве классифицирующего признака мы использовали перекодированные ряды, которые рассматривались как входные данные: 1 – больные ИБС, 2 – практически здоровые люди. В качестве входных параметров мы использовали пол ( $X_1$ ), возраст ( $X_2$ ), а так же перекодированные биохимические показатели крови:  $X_3$  – ХС,  $X_4$  – ХС ЛПВП,  $X_5$  – ТГ,  $X_6$  – мочевины,  $X_7$  – билирубин,  $X_8$  – глюкоза,  $X_9$  – общий белок,  $X_{10}$  – мочевая кислота,  $X_{11}$  – креатинкиназа,  $X_{12}$  – амилаза,  $X_{13}$  – ГГТ,  $X_{14}$  – щелочная фосфатаза,  $X_{15}$  – АСТ,  $X_{16}$  – АЛТ,  $X_{17}$  – альбумин.

В результате дискриминантного анализа нами оценивались: 1) эффективность отнесения к классам; 2) процент общих совпадений по классам; 3) коэффициент канонической корреляции. Исходя из полученных данных нами установлено, что для проведения анализа необходима одна дискриминантная функция, которая объясняет 100% суммарной дисперсии. В предложенных моделях в качестве входных данных мы использовали от 5 до 17 биохимических показателей крови. В процессе эксперимента нами предложено 17 моделей ( $r = 0,313 - 0,867$ ), которые имеют общие совпадения по классам в пределах 83,23 - 94,80%. Из предложенных экспертных систем в 15 случаях мы получили эффективность общих совпадений выше 92% ( $r = 0,691 - 0,867$ ).

Первоначально мы построили модели, используя все биохимические показатели крови. Затем из моделей попробовали поочередно исключить липиды, ферменты и органические вещества крови (без липидов). Полученные модели обладают высокими статистическими характеристиками:  $r = 0,850-1,867$ , эффективность общих совпадений по классам составляет 93,68-94,80%. Затем мы получили экспертные системы, в которые входят только липиды, или ферменты, или органические вещества крови (без липидов), а также учитывается возраст и пол. Диагностические возможности системы построенной на основе липидов снижаются до 83,23%, а в двух остальных случаях до 92,63-91,91% соответственно.

На следующем этапе исследования с целью уменьшения количества переменных, использованных в построении моделей, и улучшения статистических характеристик мы попробовали поочередно удалять биохимические показатели из общего набора. При построении экспертных систем за основу мы взяли пол, возраст, липиды, глюкозу и общий белок. Полученная модель вышла удачной, она позволяет определить диагноз ИБС в 94,94% случаев, при этом практически здоровые правильно определяются в 92,68%. Общая эффективность экспертной системы – 94,68%, а коэффициент канонической корреляции равен 0,694 ( $p < 0,001$ ), что говорит о высоких статистических характеристиках системы. Далее мы из вышеназванных входных параметров исключили ХС ЛПВП, при этом результаты прогнозирования фактически не изменились (общая эффективность – 94,4%,  $r = 0,691$ ). С нашей точки зрения эта модель заслуживает внимания, так как она обладает высокими статистическими характеристиками и минимальным количеством входных параметров.

Уравнение дискриминантной функции выглядит следующим образом:

$$Y = 1,134 + 0,655 \cdot X_1 - 0,089 \cdot X_2 - 0,122 \cdot X_3 - 0,153 \cdot X_5 - 0,014 \cdot X_8 + 0,048 \cdot X_9$$

Конкретные значения интегральных показателей состояния организма ( $Y$ ), вычисленные с помощью приведенного выше уравнения для всех представителей размещаются на непрерывной шкале состояний. При этом величины  $Y$  больных ИБС концентрируются в левой части шкалы, в области ее отрицательных значений (-0,343 – групповой центростой большой ИБС), а величины  $Y$

практически здоровых людей в зоне ее положительных значений (2,646 – групповой центроид здоровых). Используя коэффициенты дискриминантных функций и координаты групповых центроидов можно вычислить величину, близость которой к какому либо центроиду свидетельствует о принадлежности точки в пространстве вектор-признаков к данной совокупности.

Данные экспертные системы имеют важное практическое значение, так как применение их на практике позволит врачу не только правильно диагностировать заболевание, но и зарегистрировать ранние, доклинические изменения метаболизма, а также организовать раннюю профилактику выявленных нарушений. Кроме того, возможно, использование полученных экспертных систем для оценки влияния неблагоприятных экологических факторов на здоровье человека.

## ЛИТЕРАТУРА

1. Дильман В.М. Четыре модели медицины. Л., 1987. С. 288.
2. Манак Н.А., Русецкая В.Г. Динамика показателей заболеваемости системы кровообращения в Республике Беларусь за 10 лет (1990-1999) // Достижения медицинской науки Беларуси. Вып. VI. Мн., 2001. С. 106-107.
3. Пилипцевич Н.Н., Ломать Г.Н., Гальбурт Г.Н. и др. Основные показатели здоровья населения Беларуси в 1990-е гг. // Достижения медицинской науки Беларуси. Вып. VI. Мн., 2001. С.93-94.
4. Отчет о состоянии здоровья в мире, 1997 год: победа над страданиями, обогащение человечества. Женева, 1997. С. 206.
5. Трахненбург И.М., Факторов И.В., Верич Г.Е. Химическое загрязнение производственной и окружающей среды как факторы риска сердечно-сосудистых заболеваний // Гигиена и санитария, 1984, № 9. С. 61-66.
6. Алексеев С.В., Хаймович М.Л., Кадыскина Е.Н., Суворов Г.А. Производственный шум. Л., 1991. С. 136.
7. Трахненбург И.М., Колесников В.С., Луковенко В.П. Тяжелые металлы во внешней среде: Современные гигиенические и токсикологические аспекты. Мн., 1994. С. 285.
8. Маршал В. Дж. Клиническая биохимия / Пер. с англ. М.-СПб., 2000. С. 368.
9. Дюк В. Обработка данных на ПК в примерах. СПб., 1997. С. 240.
10. Факторный, дискриминантный и кластерный анализ / Под ред. И.С. Енюкова. М., 1989. С. 215.

## S U M M A R Y

*The present paper is aimed at creating expert systems, which could give possibility to distinguish the healthy organism and the organism suffering from ischemic heard disease. The work is of great practical importance as it affords opportunity of using the obtained expert systems for people's individual diagnosing. And what is more important, such systems will help to register the earlier pre-clinical changes of metabolism and to organize the early prophylaxis of the revealed disorders.*

*Поступила в редакцию 26.06.2003*