

С.С. САДЫКОВ,
И.А. ОРЛОВ,
А.С. БЕЛЯКОВА

**Разработка и исследование
математических моделей ССЗ:
нарушения ритма и проводимости**

УДК 004.67, 004.891.3

Муромский институт
(филиал) ФГБОУ ВПО
«Владимирский
государственный
университет имени
А.Г. и Н.Г. Столетовых»

В статье рассматриваются этапы построения математических моделей нарушения ритма и проводимости, включая выделение наиболее значимых факторов, оценку взаимосвязи между ними, расчет коэффициентов регрессионной модели и их анализ.

In article stages of construction of mathematical models of disturbance of a rhythm and conductivity, including allocation of the most significant factors, an interrelation estimation between them, calculation of factors regression models and their analysis are surveyed.

Целью работы является исследование применения методов математической статистики для построения математических моделей сердечно-сосудистых заболеваний: нарушений ритма и проводимости.

В статье рассматриваются основные этапы и методы решения задачи построения и исследования математических моделей сердечно-сосудистых заболеваний, предоставляющих возможности: получение объективной оценки функционального состояния организма; определение факторов, оказывающих наибольшее влияние; прогнозирование дальнейших изменений состояния.

В течение последних 40—50 лет сердечно-сосудистые заболевания (ССЗ) являются основной причиной в структуре смертности большинства европейских популяций, обуславливая 49% всех смертей и 30% смертей лиц в возрасте моложе 65 лет. В отличие от

стран Запада, где благодаря разработке и внедрению общенациональных программ по профилактике, отмечается тенденция к снижению смертности от ССЗ, в России абсолютные показатели смертности продолжают расти, превышая в 6—8 раз аналогичные показатели в развитых странах Европы, Японии и США. Сегодня участились случаи инфаркта миокарда у людей младше 30 лет и у мужчин, которые интенсивно работают. Именно поэтому диагностирование, лечение, а также прогнозирование заболеваний кардиологии являются на сегодняшний день наиболее актуальными[1,2].

Параллельно с разработкой прогностической модели необходимо налаживать регламентированный сбор статистической информации в реальном времени, так как для расчета прогноза на будущее необходимо знать всю информацию о прошлом и настоящем. На данном этапе неопределимую роль играет определение состояния исследуемой системы, а также изучение факторов влияющих на рассматриваемый показатель.

Для достижения цели по улучшению качества диагностирования, прогнозирования и лечения в работе решается задача реализации этапа сбора информации, а также задача разработки и исследования математических моделей ССЗ[3].

Разрабатываемые математические модели должны положительным образом отразиться на прогнозировании и профилактике заболеваний кардиологии в целом, что и определяет актуальность настоящей работы[4].

Решение задачи построения математических моделей ССЗ, а также моделей прогноза дальнейших изменений состоит из последовательного выполнения следующих этапов:

1. Систематизация статистических данных, полученных в ходе массовых обследований и анкетирования пациентов;
2. Определение показателей, обеспечивающих получение оценки функционального состояния организма;
3. Присвоение веса значениям каждого параметра в соответствии со значениями нормы;
4. Определение факторов, оказывающих наибольшее влияние, с использованием факторного анализа;
5. Проведение дискриминантного анализа;

6. Проведение регрессионного анализа с целью сформировать уравнение, описывающее каждую болезнь сердца, представленную в выборке. Оценка уровня значимости полученного уравнения;

7. Разработка моделей прогноза дальнейших изменений на основе полученных моделей.

Особенности и методы используемых статистических анализов

1. Факторный анализ[5]

Факторный анализ, анализ главных компонент – анализ, главной целью которого является сокращение числа переменных и определение структуры взаимосвязей между ними. Сокращение достигается за счет выделения скрытых общих факторов, объясняющих связи между наблюдаемыми признаками объекта, т.е. вместо исходного набора переменных анализируются данные по выделенным факторам. Факторный анализ также позволяет выявить факторы, оказывающие наибольшее влияние.

Факторный анализ необходимо начинать с вычисления корреляционной матрицы. Ее анализ позволит оценить степень коррелированности переменных между собой. И если эта степень окажется высокой, то данные переменные можно объединять в один фактор.

2. Дискриминантный анализ

Дискриминантный анализ - это метод многомерной статистики, применяемый для решения задач классификации (распознавания образов) и позволяющий отнести объект с определенным набором признаков (симптомов) к одному из известных классов

Результат анализа представляет собой линейные классификационные функции в виде линейных уравнений, выработанным методами дискриминантного анализа на основе обучающей информации.

Медицинская диагностика с применением дискриминантного анализа выполняется в три этапа.

1) На первом этапе формируется обучающая информация. Определение видов заболеваний для диагностики, диагностических признаков осуществляется врачом специалистом.

2) На втором этапе вырабатываются решающие правила и дается оценка их информативности. Решающие правила представля-

ются в виде линейных классификационных функций (ЛКФ) и канонических линейных дискриминантных функций (КЛДФ).

3) На третьем этапе непосредственно решается задача медицинской диагностики по выработанным решающим правилам. После обследования больного определяются количественные значения симптомов, включенных в ЛКФ или КЛДФ, рассчитываются эти функции и по их величинам дается решение об отнесении больного к той или иной группе заболеваний.

3. Регрессионный анализ

В настоящее время множественная регрессия — один из наиболее распространенных методов в эконометрике, медицине, биологии, экологии и генетике. Основная цель множественной регрессии — построить модель с большим числом факторов, определив при этом влияние каждого из них в отдельности, а также совокупное их воздействие на моделируемый показатель.

Для оценки коэффициентов уравнения множественной регрессии наиболее часто используется метод наименьших квадратов (МНК).

Построение математических моделей ССЗ

Сбор статистических данных осуществляется путем анкетирования и анализа индивидуальных карт пациентов терапевтического и кардиологического отделений МУЗ МГБ №3. Анализируемый информационный массив включает в себя результаты обследования 221 человека.

В качестве объясняющих переменных использовались показатели, являющиеся наиболее традиционными для диспансерных обследований в кардиологическом отделении:

1. Возраст;
2. Пол;
3. Конституция;
4. ЧСС;
5. ЭКГ: PQ, P, QRS, QT;
6. Ритм.

Исходные данные приведены к общей системе координат, где значению каждого параметра присвоена количественная определенность или степень изменения в баллах. Полученные значения

преобразованы следующим образом: а) значение нормы соответствует 100; б) значения параметров (в зависимости от процентного соотношения с нормой) соответственно – меньше или больше 100.

Проведение факторного анализа исходной выборки позволило на основе критерия Кайзера и критерия «каменистой осыпи» выделить 4 фактора, оказывающие наибольшее влияние.

Из анализа полученной таблицы факторных нагрузок (рис. 1) следует, что Factor 1 имеет высокие факторные нагрузки по переменным PQ и P и низкие по переменным QRS и QT, а Factor 2 — наоборот: низкие по переменным PQ и P и высокие по переменным QRS и QT. Factor 3 имеет высокие факторные нагрузки по переменной Pulse (ЧСС), а Factor 4 - по переменной Sex (пол). Полученные значения факторных нагрузок позволяют сделать вывод, что выделенные 4 фактора наилучшим образом характеризуют данные, а что общее состояние организма, в основном, определяется величинами P, PQ, QRS, QT и ЧСС.

Таблица 1

Факторные нагрузки (поворот осей)

	Фактор 1	Фактор 2	Фактор 3	Фактор 4
Возраст	0.502777	-0.053467	-0.371021	-0.564175
Пол	0.060246	0.044572	-0.084995	0.879579
Конст. Тип	0.210888	0.034160	0.502669	0.233472
Пульс (ЧСС)	-0.03196	0.032784	0.859465	-0.055473
PQ	0.786650	0.227927	0.012045	-0.003201
P	0.761287	-0.311458	0.013341	0.048750
QRS	0.146023	-0.834426	-0.301964	0.092417
QT	0.144244	0.772084	-0.327478	0.239187

Анализ полученной на этапе факторного анализа корреляционной матрицы позволяет оценить степень коррелированности переменных между собой и выделить параметры с выраженной взаимосвязью:

- PQ и P;
- QRS и QT.

В ходе работы проведены факторный, дискриминантный и регрессионный анализы и получены уравнения множественной регрессии, описывающие:

- нарушение сердечного ритма (брадикардия);
- увеличение частоты сердечных сокращений (тахикардия);
- нарушение процесса реполяризации;

- диффузное нарушение процесса реполяризации;
- замедление внутрисердечной проводимости.

Для проведения статистических анализов из исходной выборки выделены случаи каждого отклонения в отдельности, а также случаи нормального функционального состояния организма.

При проведении факторного анализа для случаев нарушения сердечного ритма (брадикардия) критерий Кайзера и критерий каменной осыпи позволили выделить 4 фактора, оказывающих наибольшее влияние:

- фактор 1 в большей степени отвечает за QT;
- фактор 2 - за пол;
- фактор 3 - за возраст и конституцию;
- фактор 4 - за ЧСС.

Параметры с выраженной взаимосвязью:

- зубец P - ЧСС;
- возраст - конституция.

При проведении факторного анализа в качестве зависимой переменной выступает величина, определяющая состояние организма человека, значение которой представляет собой количественную определенность или степень изменения в баллах: 1 – норма; 2 - нарушение.

Окно результатов регрессионного анализа представлено на рис. 1.

```

Multiple Regression Results (Step 6)

Dependent: Result          Multiple R = ,97071531      F = 125,1774
                          R² = ,94228822      df = 6,46
No. of cases: 53          adjusted R² = ,93476060      p = 0,000000
                          Standard error of estimate: ,104577047
Intercept: -,347616068 Std.Error: ,2365590 t( 46) = -1,469 p = ,1485

-----
Pulse beta=,946           PQ beta=,128           P beta=,082
Sex beta=,118            Age beta=,069         QRS beta=,040
  
```

Рис. 1. Окно результатов регрессионного анализа

В первой части окна содержится основная информация о результатах оценивания, во второй — высвечиваются значимые стандартизованные регрессионные коэффициенты.

Из приведенных результатов анализа следует, что зависимость между откликом и предикторами сильная (R^2 , т.е. коэффициент детерминации $> 0,75$); построенная линейная регрессия адекватно

описывает взаимосвязь между откликом и предикторами, свободный член статистически значим.

Таблица результатов с анализа представлена в таблице 2

Таблица 2

Результаты регрессионного анализа

	Beta	Ст. ош. Beta	B	Ст. ош. B
Свободный член			-0,347	0.237
Возраст	0.069	0.039	0,117	0.069
Пульс (ЧСС)	0.946	0.036	0,504	0.019
Пол	0.118	0.039	0,196	0.064
PQ	0.128	0.038	0,197	0.057
P	0.082	0.038	0,064	0.029
QRS	0.040	0.037	0,119	0.111

Таблица 2 содержит стандартизованные (Beta) и нестандартизованные (B) регрессионные коэффициенты (веса), их стандартные ошибки и уровни значимости. Величины Beta-коэффициентов позволяет сравнить вклады каждого предиктора в предсказание отклика.

Таким образом, в зависимую переменную больший вклад вносит переменная Pulse (ЧСС), а меньший — величины QRS, возраст, величина зубца P. Параметры, не включаемые в результирующее уравнение регрессии: конституция и QT.

Полученное уравнение регрессии:

$$\begin{aligned} \text{Результата} = & -0,348 + 0,504\text{Пульс} + 0,197\text{PQ} + \\ & + 0,064\text{P} + 0,1\text{Пол} + 0,117\text{Возраст} + 0,120\text{QRS} \end{aligned}$$

Факторный и регрессионный анализ, проведенные для тахикардии, нарушения процесса реполяризации, диффузного нарушения процесса реполяризации, а также замедления внутрипредсердной проводимости, позволили получить следующие результаты, представленные в таблице 1, где ФА – факторный анализ, РА – регрессионный анализ.

Таблица 1

Результаты факторного и регрессионного анализа отдельных ССЗ

Отклонение	Результаты анализа
Тахикардия	Факторы, оказывающие наибольшее влияние: - фактор 1 - ЧСС и величина зубца P; - фактор 2 - пол;

Отклонение	Результаты анализа
	<p>- фактор 3 - значения PQ и QRS. Параметры с выраженной взаимосвязью: - зубец P - ЧСС; - возраст - конституция.</p> <p>Уравнение регрессии: <i>Результат</i> = 0,456 + 0,429 · ЧСС + 0,059 · P + 0,037 · QT</p>
Нарушение процесса реполяризации	<p>Факторы, оказывающие наибольшее влияние: - фактор 1 - ЧСС и величина зубца P; - фактор 2 - величина PQ; - фактор 3 - величина PQ; - фактор 4 - значение QRS. Параметры с выраженной взаимосвязью: - зубец P - ЧСС.</p> <p>Уравнение регрессии: <i>Результат</i> = -0,909 + 0,287 · ЧСС + 0,357 · P + 0,348 · QRS + 0,240 · Пол</p>
Диффузное нарушение процесса реполяризации	<p>Факторы, оказывающие наибольшее влияние: - фактор 1 - ЧСС и величина QRS; - фактор 2 - возраст и величина P; - фактор 3 - значение PQ; - фактор 4 - значения P и пол. Параметры с выраженной взаимосвязью: - QRS - ЧСС; - возраст - зубец P.</p> <p>Уравнение регрессии: <i>Результат</i> = 0,728 + 0,156 · Пульс + 0,365 · P - 0,222 · Пол</p>
Замедление внутри-предсердной проводимости	<p>Факторы, оказывающие наибольшее влияние: - фактор 1 - пол и величина PQ; - фактор 2 - величина P; - фактор 3 - пол и величина PQ; - фактор 4 - возраст. Параметры с выраженной взаимосвязью: - зубец P - ЧСС; - пол - величина PQ.</p> <p>Уравнение регрессии: <i>Результат</i> = 0,394 + 0,233 · Пульс + 0,333 · P</p>

Таким образом, в ходе проведения ряда анализов решены следующие задачи:

- определены основные показатели, обеспечивающие получение оценки функционального состояния организма;
- решена задача классификации, позволяющая отнести определенный набор признаков (симптомов) к одному из известных классов;
- получены уравнения множественной регрессии, описывающие ССЗ.

Заключение

В работе решена задача построения и исследование математических моделей сердечно-сосудистых заболеваний.

Для построения модели произведен анализ статистических данных, полученных в ходе массовых обследований и анкетирования пациентов, а также определены основные показатели, обеспечивающие получение объективной оценки состояния исследуемой системы.

Построенные модели ориентированы на людей, имеющих проблемы сердечно-сосудистой системы и системы кровообращения.

В работе произведена оценка уровня значимости полученных уравнений и математических моделей, а также разработана модель прогноза дальнейших изменений, что должно положительным образом отразиться на прогнозировании и профилактике заболеваний в кардиологии.

Литература

1. Садыков С.С., Белякова А.С. Математические модели некоторых сердечно-сосудистых заболеваний. Информационные технологии. №12, 2011, с. 59-63.
2. Садыков С.С., Белякова А.С. Регрессионные модели стенокардии и зависимость их информативности от количества параметров работы сердца. Системы управления и информационные технологии. №3.1(45), 2011, с.190-194.
3. Белякова А.С. Основные признаки оценки состояния сердечно-сосудистой системы. Алгоритмы, методы и системы обработки данных. №14, 2009, с.14-31.
4. Суворов А.В. Клиническая электрокардиография. – Нижний Новгород: Изд-во НМИ, 1993. – 124 с.
5. Юнкеров В.И., Григорьев С.Г. Математико-статистическая обработка данных медицинских исследований. – СПб.: ВМедА, 2002. – 266 с.