

С.С. САДЫКОВ,  
А.С. БЕЛЯКОВА, Е.А. ПУГИН

**Математическая модель  
диагностики инфекционного  
эндокардита**

УДК 004.67, 004.891.3

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

*Статья посвящена построению и исследованию регрессионной модели инфекционного эндокардита. Выявлены наиболее важные характеристики заболевания.*

*Article is devoted construction and research regression to model of an infectious endocarditis. The most important characteristics of disease are taped.*

Диагностика сердечно-сосудистых заболеваний (ССЗ) является одной из важнейших проблем современной медицины. ССЗ являются причиной смертей в 56% летальных исходов в России [1-3]. Поэтому своевременное определение вида заболевания является важным и актуальным. Особенно при диагностике редко диагностируемых болезней сердца.

Одним из таких заболеваний является эндокардит - воспаление внутренней оболочки сердца, эндокарда. При локализации на створках клапанов сердца эндокардит является основой формирования большинства приобретенных пороков сердца и относится к ведущим проявлениям поражения сердца при ревматизме и других диффузных болезнях соединительной ткани. Ежегодная заболеваемость инфекционным эндокардитом составляет 38 случаев на 100 тысяч населения, причем чаще заболевают лица трудоспособного возраста[4].

Инфекционный эндокардит диагностируется по результатам проведения осмотра и расспроса пациента, эхокардиографии и биохимического анализа крови. Параметры обследования при этом представлены количественными (ЭхоКГ и анализ крови) и каче-

ственными значениями (осмотр). Врач-кардиолог должен учитывать все эти значения и их сочетания.

Математические модели диагностики ССЗ предназначены для помощи врачу при постановке диагноза пациенту, а также для увеличения степени его объективности. При этом необходимо принимать во внимание имеющиеся взаимосвязи между параметрами и их значимость для данного заболевания[2].

Целью работы является построение и исследование математических сердечно-сосудистых заболеваний на примере эндокардита, оценка наиболее значимых параметров, связей между ними.

Наиболее распространенными симптомами эндокардита являются: лихорадка, озноб, потливость, головные боли [3]. Эти данные собираются врачом при осмотре и являются качественными.

Обозначим количественные параметры, оцениваемые при диагностике эндокардита следующим образом:

$x_1$  - возраст, лет;

$x_2$  - пол: 1 – женщины, 2 – мужчины;

$x_3$  - гемоглобин, г/л;

$x_4$  - скорость оседания эритроцитов (СОЭ), мм/ч;

$x_5$  - фибриноген, г/л;

$x_6$  - альбумины, снижение, г/л;

$x_7$  - гамма – глобулины, увеличение, г/л;

$x_8$  - вегетации на клапанах, мм.

Исходная матрица данных содержит  $M = 400$  записей результатов обследований пациентов: 200 пациентов больные, остальные – здоровые (таблица 1).

Таблица 1

Исходная матрица данных

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$
1	46	1	9,5	51	3	57,7	11,8	7
2	30	1	12,4	60	2,1	54,9	11,1	6
3	25	2	11,8	59	3	52,3	11	4
...	...	...	...	...	...	...	...	...
399	66	2	157	4	3,6	57,3	13,1	0
400	94	1	153	9	3,5	59,3	14,7	1

Входные параметры  $x_{ij}, j=1,2,\dots,m, m=8$  влияют на исход заболевания  $y_i, i=1,2,\dots,M$ . Связь между входными параметрами и выходным значением  $y_i$  описывается регрессионным уравнением[2]:

$$\hat{y}_i = v_0 + v_1 x_{i1} + v_2 x_{i2} + \dots + v_m x_{im} \quad (1)$$

где  $v_0, v_1, \dots, v_m$  - коэффициенты регрессии,  $\hat{y}$  - значение выходного параметра, предсказанное по уравнению регрессии.

Для оценки силы связей между параметрами были рассчитаны коэффициенты их корреляции по:

$$r_{x_i x_j} = \frac{\sum_{i=1}^m (x_{ij} - \bar{x}_i)(x_{ji} - \bar{x}_j)}{\sqrt{\sum_{i=1}^m (x_{ij} - \bar{x}_i)^2 \sum_{i=1}^n (x_{ji} - \bar{x}_j)^2}} \quad (2)$$

Коэффициент  $r_{x_i x_j}$  является относительной величиной  $-1 \leq r_{x_i x_j} \leq 1, i, j=1,2,\dots,m$ . При этом сила связи между параметрами

оценивается следующим образом:

$$\begin{cases} |r_{x_i x_j}| < 0.3 - \text{слабая} \\ 0.3 \leq |r_{x_i x_j}| \leq 0.7 - \text{умеренная} \\ |r_{x_i x_j}| > 0.7 - \text{сильная.} \end{cases} \quad (3)$$

В результате анализа исходных данных получена следующая корреляционная матрица (таблица 2):

Таблица 2

Коэффициенты корреляции между параметрами

	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$
$x_1$	1,000							
$x_2$	-0,057	1,000						
$x_3$	0,028	-0,017	1,000					
$x_4$	-0,027	-0,000	<b>-0,972</b>	1,000				
$x_5$	-0,026	-0,050	<b>0,515</b>	<b>-0,513</b>	1,000			
$x_6$	-0,059	-0,009	<b>0,780</b>	<b>-0,758</b>	<b>0,393</b>	1,000		
$x_7$	-0,005	-0,023	<b>0,536</b>	<b>-0,512</b>	0,217	0,440	1,000	
$x_8$	-0,019	-0,028	<b>-0,912</b>	<b>0,893</b>	<b>-0,461</b>	<b>-0,691</b>	<b>-0,497</b>	1,000

Жирным шрифтом в таблице 2 выделены коэффициенты корреляции параметров, между которыми наблюдаются сильные и средние связи.

Коэффициенты регрессии получены методом наименьших квадратов и представлены в таблице 3. Значения коэффициентов регрессии получены при уровне значимости  $p \leq 0.05$

Таблица 3

**Коэффициенты регрессии**

$V_0$	$V_1$	$V_2$	$V_3$	$V_4$	$V_5$	$V_6$	$V_7$	$V_8$
0,925	0	0	0,006	-0,002	0	0,002	0,002	-0,004

Таким образом, получено уравнение диагностики инфекционного эндокардита.

$$y_i = 0.925 + 0.006x_3 - 0.002x_4 + 0.002x_6 + 0.002x_7 - 0.004x_8 \quad (4)$$

Оценим адекватность и информативность полученного регрессионного уравнения. Для этого рассчитаем коэффициент детерминации, показывающий силу связи между входными параметрами и откликом или информативность полученного уравнения:

$$D = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (5)$$

Коэффициент детерминации оценивается по формуле:

$$\left\{ \begin{array}{l} 0.1 \leq D < 0.3 - \text{слабая} \\ 0.3 \leq D < 0.5 - \text{умеренная} \\ 0.5 \leq D < 0.7 - \text{заметная} \\ 0.7 \leq D < 0.9 - \text{высокая} \\ 0.9 \leq D < 1 - \text{весьма высокая.} \end{array} \right. \quad (6)$$

По ( ) информативность уравнения составила  $D = 0.97$  и является высокой. Также рассчитаем смещенный коэффициент детерминации:

$$D_n = 1 - (1 - D) \frac{N - 1}{N - m' - 1} \quad (7)$$

Смещенная оценка коэффициента детерминации:  $D_n = 0.96$ , что также подтверждает высокую адекватность модели.

Следовательно, полученная математическая модель инфекционного эндокардита является адекватной и информативно и может использоваться при диагностике заболеваний.

При тестировании полученной модели эндокардита на контрольной выборке данных объемом 150 записей: 94 пациента здоровые, 56 имеют данную патологию получено: у 91 пациента из 94 определено наличие заболевания, остальные пациенты согласно результатам моделирования здоровы. Следовательно, в 97% случаев модель дала верный результат.

Таким образом, математические модели диагностики ССЗ могут использоваться для определения наличия заболеваний, учитывают взаимосвязи между параметрами и предназначены для поддержки принятия решения кардиологом при определении болезни сердца.

### 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. Садыков С.С., Сафиулова И.А., Ткачук М.И. Полная автоматизированная обработка последовательности ультразвуковых снимков сердца. Известия высших учебных заведений. Приборостроение. 2010. Т. 53. № 9. С. 27-33.