

ПРОГНОЗИРОВАНИЕ ВОЗНИКНОВЕНИЯ ФАКТОРОВ РИСКА ИЗВИТОСТИ КОРОНАРНЫХ АРТЕРИЙ

Ключевые слова: бинарная логистическая регрессия, дискриминантный анализ, коронарные артерии, извитость артерий.

ВВЕДЕНИЕ

При ангиографических исследованиях у больных с явными признаками ишемии и клиникой стенокардии часто не наблюдают значимых органических сужений просвета коронарных артерий (КА), а диагностируют только наличие их извитости.

Извитость КА [1] можно рассматривать как проявление ее избыточного роста. В результате проведенных исследований было установлено, что длина КА превышает длину участка ее расположения в 1,5–2 раза [2]. Избыточная длина КА способствует пропорциональному увеличению сопротивления движению крови, при этом не учитывается сопротивление, обусловленное изгибом. Иначе говоря, извитость КА создает дополнительное сопротивление, вызывает локальные перегрузки ее стенки, травмы эндотелия и изменение его функции. Нарушение кровотока в извитом сосуде в ряде случаев приводит к ишемии миокарда или к рефлекторному спазму сосуда.

Для более глубокого понимания механизмов необходимо исследовать факторы, влияющие на образование извитости КА, чтобы в дальнейшем, по возможности, разработать профилактические и терапевтические мероприятия.

В последние годы в НИССХ им. Н.М. Амосова НАМНУ ведется работа по определению, может ли извитость КА рассматриваться в качестве самостоятельной патологии.

Цель настоящей статьи — исследовать факторы, приводящие к развитию извитости КА, с помощью статистических моделей.

МАТЕРИАЛЫ И МЕТОДЫ

В описанном в статье исследовании участвовало 130 пациентов, которым первичная коронарорентрикулография (КВГ) выполнялась в НИССХ им. Н.М. Амосова НАМНУ. В зависимости от наличия извитости и атеросклеротического поражения КА пациентов разделили на две группы: в первую вошли больные с извитостью и атеросклеротическими поражениями коронарных сосудов ($n=66$, где n — общее количество больных в группе), которым в дальнейшем проводилось коронарное шунтирование (КШ), во вторую — пациенты с атеросклеротическими поражениями коронарных сосудов и без признаков их выраженной извитости ($n=64$).

Описание каждого пациента включало 47 показателей (переменных):

- гендерные (возраст, пол);
- гемодинамические (частота сердечных сокращений (ЧСС), систолическое артериальное давление (АД), фракция выброса (ФВ), конечно-систолический индекс (КСИ), конечно-диастолический индекс (КДИ) и др.);
- биохимические (глюкоза, гемоглобин и др.);

— сопутствующие патологии (заболевания дыхательной, эндокринной и сердечно-сосудистой систем), вредные привычки (курение), предварительно проведенное лечение (антигипертензивная (АГ) терапия).

Статистическая обработка материала была выполнена с использованием пакета программ IBM Statistics 20.0.

Построение моделей проводилось на основе метода бинарной логистической регрессии (БЛР) [3] и дискриминантного анализа.

Для проверки гипотез при сравнении качественных (категориальных) данных между двумя независимыми группами использовался критерий χ^2 Пирсона. Различие между группами считали достоверным, если $p \leq 0,05$, где p — вероятность ошибки первого рода.

Для изучения статистической взаимосвязи признаков использовали параметрический корреляционный анализ Спирмена. Значение достоверности коэффициента корреляции трактовали следующим образом: если $p < 0,05$, то полученный коэффициент корреляции является достоверным.

Различие средних значений определяли с помощью t -критерия Стьюдента для нормально распределенных переменных и с использованием критерия Манна–Уитни для остальных. Для проверки гипотезы о нормальном распределении переменных применяли критерий Колмогорова–Смирнова.

При выполнении корреляционного анализа из исследуемых показателей отобрали факторы, которые имели наиболее весомое влияние на извитость КА. Для построения моделей логистической регрессионной и дискриминантного анализа пользовались методом принудительного включения, т.е. в математическую модель были принудительно включены факторы, которые наиболее сильно коррелировали с развитием извитости КА.

Для каждой модели рассчитывали параметры чувствительности и специфичности, а также площадь под ROC-кривой.

ПОСТРОЕНИЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ ИЗВИТОСТИ КА НА ОСНОВЕ МЕТОДА БЛР

Средний возраст пациентов в группах больных с неизвитыми и с извитыми КА со склерозом составил 58 ± 9 лет и 61 ± 8 лет соответственно. Подавляющее большинство пациентов — женщины. Так, в группе с извитыми КА со склерозом 51 женщина и 15 мужчин, а с неизвитыми КА со склерозом — 50 женщин и 14 мужчин.

По результатам корреляционного анализа были отобраны семь категориальных и девять количественных переменных, для которых $p < 0,05$.

В табл. 1 приведены: средние значения и стандартные отклонения для количественных переменных с нормальным распределением и медиана; минимальное и максимальное значения (в круглых скобках), а также II и III квартили (в квадратных скобках) — для распределений других типов, которые отличаются от нормальных, для групп пациентов с извитыми и неизвитыми КА со склерозом; статистическая значимость p для переменных в группах с извитыми и неизвитыми КА со склерозом.

Наиболее значимые различия ($p \leq 0,01$) были выявлены у пяти переменных: КДИ, КСИ, ФВ, протромбиновый индекс (ПТИ) и уровень глюкозы. К переменным, для которых $p < 0,05$, относятся: возраст, систолическое АД и гемоглобин.

В табл. 2 приведены независимые категориальные переменные, которые использовались для построения регрессионной модели, а также отношение шансов (OR) с 95%-м доверительным интервалом (95% CI), которое показывает соотношение вероятностей извитости КА при положительном и отрицательном значении изучаемого фактора риска.

Таблица 1

Переменная	Средние значения показателей пациентов		P
	с неизвитыми КА со склерозом (n = 64)	с извитыми КА со склерозом (n = 66)	
Возраст	58±9	61±8	0,031
Систолическое АД	127,5 (100–190) [120–145]	140 (90–240) [127,5–155]	0,041
КДИ	87,2 (52,1–231,5) [70,8–104]	67,2 (43,0–108,6) [60,2–77,9]	0,000
КСИ	41,2 (14,7–186,2) [27,8–60]	28 (13,4–61,4) [22,1–33,4]	0,000
ФВ	52 (18,0–27,8) [42,3–58]	58,5 (35–74) [50,0–63,3]	0,000
ПТИ	100 (45–100) [89–100]	85 (29–100) [80–93]	0,000
Гемоглобин	149 (109–208) [141–157]	142 (89–180) [135–153,3]	0,043
Гематокрит	0,45±0,05	0,44±0,04	0,327
Глюкоза	6 (4,3–19,3) [5,4–6,6]	5,5 (3,9–12,8) [4,9–6,5]	0,009

Таблица 2

Факторы риска	Градация значений факторов риска	Категориальные переменные для пациентов		OR (95 % CI)
		с неизвитыми КА со склерозом (n = 64)	с извитыми КА со склерозом (n = 66)	
Вредность профессии	Нет	58	65	1,0
	Есть	6	1	0,15 (0,02–1,27)
Заболевания дыхательной системы	Нет	52	60	1,0
	Есть в анамнезе	12	4	0,29 (0,09–0,95)
Воспалительные заболевания сердца	Нет	62	56	1,0
	Есть в анамнезе	2	8	4,43 (0,9–21,7)
Заболевания щитовидной железы	Нет	57	63	1,0
	Есть в анамнезе	7	1	0,13 (0,02–1,08)
Вредные привычки (курение)	Нет	54	62	1,0
	Да	10	2	0,17 (0,04–0,83)
АГ терапия	Не проводилась	37	19	1,0
	Лечилась	27	45	3,25 (1,56–6,74)
Тромбоциты	Ниже нормы	10	40	1,0
	Норма	50	20	0,1 (0,04–0,24)
	Выше нормы	3	2	0,17 (0,02–1,14)

Из данных табл. 2 следует, что наибольшие шансы развития извитости КА были выявлены у пациентов, которые перенесли воспалительные заболевания сердца (OR = 4,43, 95% CI 0,9–21,7), а также получили АГ терапию (OR = 3,25, 95% CI 1,56–6,74).

Для построения модели на основе метода БЛР вошли 15 статистически значимых ($p < 0,05$) факторов риска развития извитости КА в группе, где сравниваются пациенты с неизвитыми и извитыми КА со склерозом (табл. 3). Гематокрит, у которого $p > 0,1$ не был включен в регрессионный анализ.

Математическая модель прогнозирования извитости КА в группе больных, у которых сравнивались случаи с неизвитыми и извитыми КА со склерозом, имеет следующий вид:

$$z_1 = -x_1 \cdot 5,9 + x_2 \cdot 1,3 + x_3 \cdot 4,6 - x_4 \cdot 7,0 - x_5 \cdot 1,6 + x_6 \cdot 1,9 - x_7 \cdot 3,9 + x_8 \cdot 0,04 + x_9 \cdot 0,017 - x_{10} \cdot 0,15 + x_{11} \cdot 0,16 + x_{12} \cdot 0,17 - x_{13} \cdot 0,07 - x_{14} \cdot 0,06 - x_{15} \cdot 0,54 + 22,76,$$

где z_1 — степень влияния факторов риска на возникновение извитости КА; a_0 — свободный член уравнения логистической регрессии; $a_{1...n}$ — весовые коэффициенты уравнения регрессии; $x_{1...n}$ — показатели, влияющие на возникновение извитости КА.

Таблица 3

Факторы риска	Условное обозначение	Коэффициент корреляции	p	Коэффициент уравнения регрессии, $a_{1...n}$, $a_0 = 22,76$
Вредность профессии	x_1	-0,174	0,047	-5,9
Заболевания дыхательной системы	x_2	-0,189	0,033	1,3
Воспалительные заболевания сердца	x_3	0,175	0,048	4,6
Заболевания щитовидной железы	x_4	-0,194	0,028	-7,0
Вредные привычки (курение)	x_5	-0,214	0,015	-1,6
АГ терапия	x_6	0,242	0,002	1,9
Тромбоциты	x_7	-0,474	0	-3,9
Возраст	x_8	0,189	0,031	0,04
Систолическое АД	x_9	0,18	0,041	0,017
КДИ	x_{10}	-0,42	0	-0,15
КСИ	x_{11}	-0,361	0	0,16
ФВ	x_{12}	0,342	0	0,17
ПТИ	x_{13}	-0,449	0	-0,07
Гемоглобин	x_{14}	-0,182	0,043	-0,06
Глюкоза	x_{15}	-0,234	0,009	-0,54

Анализ чувствительности и специфичности логистической регрессионной модели базируется на построении ROC-кривой или кривой ошибок, которую используют для представления результатов бинарной классификации.

Для верификации выявленных закономерностей применялась перекрестная проверка (cross validation test). Две группы пациентов (исходный массив наблюдений) были разделены на две равных подгруппы (блока): в первой проводилось обучение модели (обучающая выборка), а вторая (тестовая выборка) — использовалась для проверки.

В табл. 4 показана площадь под ROC-кривой, правильные отнесения, чувствительность и специфичность моделей в процентах при проверке разбивкой исходного массива, в котором сравниваются больные с извитыми и неизвитыми КА со склерозом, на выборки обучающую ($N_1 = 65$, где N_1 — количество больных в обучающей выборке) и тестовую ($N_2 = 65$, где N_2 — количество больных в тестовой выборке).

Таблица 4

Выборка данных	Площадь под ROC-кривой auROC	Правильные отнесения Perf, %	Специфичность Sp, %	Чувствительность Se, %
Полная	0,979	86,8	86	87,5
Обучающая	0,978	91,2	93,1	89,3
Тестовая	0,842	71,4	61,9	78,6

Из табл. 4 следует, что при прогнозировании развития извитости КА в общей группе, в которой сопоставлялись больные с извитыми и неизвитыми КА со склерозом методом БЛР, чувствительность модели на обучающей выборке составляет 89,3 %, а специфичность — 93,1 % при суммарной точности 91,2 %. При проверке по тестовой выборке результаты оказались ниже: чувствительность модели составила 78,6 %, а специфичность — 61,9 % при суммарной точности 71,4 %.

ПОСТРОЕНИЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ ИЗВИТОСТИ КА НА ОСНОВЕ ДИСКРИМИНАНТНОГО АНАЛИЗА

Описание каждого пациента включало 47 показателей. После логического и корреляционного анализа для дальнейшего исследования использовались 15 переменных, которые характеризовались статистически значимой ($p \leq 0,05$) корреляционной связью с извитостью КА.

Отбор признаков в окончательную модель выполнялся методом принудительного включения дискриминантного анализа в ППП SPSS 20.0.

В табл. 5 представлены статистически значимые ($p < 0,05$) факторы риска развития извитости КА в группе, в которой сравниваются больные с неизвитыми КА со склерозом и с извитыми КА со склерозом.

Отнесение объекта (больного) к определенному классу (извитые или неизвитые КА) было выполнено в соответствии с набором его факторов риска на основе расчета линейных дискриминантных функций (ЛДФ) по следующим формулам:

$$\text{ЛДФ}_1 = -314,83 + 27,9 \cdot x_1 - 3,3 \cdot x_2 + 18,2 \cdot x_3 + 5,1 \cdot x_4 + 20,1 \cdot x_5 + 1,7 \cdot x_6 + 13,8 \cdot x_7 + 1,3 \cdot x_8 + 0,17 \cdot x_9 - 0,13 \cdot x_{10} + 1,5 \cdot x_{11} + 3,2 \cdot x_{12} + 0,8 \cdot x_{13} + 0,75 \cdot x_{14} + 2,7 \cdot x_{15};$$

$$\text{ЛДФ}_2 = -300,51 + 26 \cdot x_1 - 2,2 \cdot x_2 + 22,1 \cdot x_3 + 0,6 \cdot x_4 + 18,6 \cdot x_5 + 3,4 \cdot x_6 + 10,1 \cdot x_7 + 1,3 \cdot x_8 + 0,19 \cdot x_9 - 0,28 \cdot x_{10} + 1,7 \cdot x_{11} + 3,4 \cdot x_{12} + 0,7 \cdot x_{13} + 0,7 \cdot x_{14} + 2,4 \cdot x_{15}.$$

После обследования больного в формулы линейных дискриминантных функций подставляются значения переменных, которые включены в модель, далее проводится расчет уравнений. Конкретного пациента необходимо отнести в ту группу, значение ЛДФ для которой наибольшее.

Дискриминантная модель прогноза извитости и неизвитости КА со склерозом состоит из 15 признаков, имеет достаточно большое число правильных отнесений (87,7 %) и есть статистически значимой ($p \leq 0,05$). Чувствительность модели составляет 89,3 %, специфичность — 86 %. Качество модели допустимо для прогнозирования извитости и неизвитости КА со склерозом

В табл. 6 представлены расчеты площади под ROC-кривой, правильные отнесения, чувствительность и специфичность моделей в процентах при проверке разбивкой исходного массива на обучающую ($N_1 = 65$) и тестовую ($N_2 = 65$) выборки в общей группе, в которой сравниваются пациенты с извитыми и неизвитыми КА со склерозом.

При прогнозировании развития извитости КА в группе, в которой сравнивались больные с извитыми и неизвитыми КА со склерозом методом дискриминантного анализа, чувствительность модели на обучающей выборке составила 88,2 %, а специфичность — 90,9 % при суммарной точности 89,6 %. При проверке по тестовой выборке результаты оказались ниже: чувствительность модели составила 83,3 %, а специфичность — 87,1 % при суммарной точности 85,2 %.

Таблица 5

Факторы риска	Градация признака с кодами	Условное обозначение	Значение коэффициентов ЛДФ ₁ (неизвитая КА со склерозом)	Значение коэффициентов ЛДФ ₂ (извитая КА со склерозом)	<i>p</i>	Коэффициент корреляции
Вредность профессии	Нет — 1 Есть — 2	x_1	27,9	26	0,047	- 0,174
Заболевания дыхательной системы	Нет — 1 Есть — 2	x_2	- 3,3	- 2,2	0,033	- 0,189
Воспалительные заболевания сердца	Нет — 1 Есть — 2	x_3	18,2	22,1	0,048	0,175
Заболевания щитовидной железы	Нет — 1 Есть — 2	x_4	5,1	0,6	0,028	- 0,194
Вредные привычки (курение)	Нет — 1 Да — 2	x_5	20,1	18,6	0,015	- 0,214
АГ терапия	Не принимал — 1 Лечился — 2	x_6	1,7	3,4	0,002	0,242
Тромбоциты	Ниже нормы — 1 Норма — 2 Выше нормы — 3	x_7	13,8	10,1	0	- 0,474
Возраст		x_8	1,3	1,3	0,031	0,189
Систолическое АД		x_9	0,17	0,19	0,041	0,18
КДИ		x_{10}	- 0,13	- 0,28	0	- 0,42
КСИ		x_{11}	1,5	1,7	0	- 0,361
ФВ		x_{12}	3,2	3,4	0	0,342
ПТИ		x_{13}	0,8	0,7	0	- 0,449
Гемоглобин		x_{14}	0,75	0,7	0,043	- 0,182
Глюкоза		x_{15}	2,7	2,4	0,009	- 0,234
Константа		—	- 314,83	- 300,51	—	—

Таблица 6

Выборка данных	Площадь под ROC-кривой auROC	Правильные отнесения Perf, %	Специфичность Sp, %	Чувствительность Se, %
Полная	0,970	86,8	86	87,5
Обучающая	0,937	89,6	90,9	88,2
Тестовая	0,864	85,2	87,1	83,3

СРАВНЕНИЕ ПОЛУЧЕННЫХ МОДЕЛЕЙ

При сравнении качества полученных моделей использовались параметры чувствительности, специфичности, а также площадь под ROC-кривой.

Для определения прогноза извитости КА с помощью моделей на основе БЛР и дискриминантного анализа использовали модели, построенные на обучающей выборке 100 % массива.

В табл. 7 представлены расчеты площади под ROC-кривой, правильные отнесения, чувствительность и специфичность моделей в процентах для группы, в ко-

торой сравниваются больные с извитыми и неизвитыми КА со склерозом, построенных с помощью метода БЛР и дискриминантного анализа.

Таблица 7

Тип модели и размер обучающей выборки	Площадь под ROC-кривой auROC	Правильные отнесения Perf, %	Чувствительность Se, %	Специфичность Sp, %
БЛР (весь массив)	0,979	86,8	87,5	86
Дискриминантный анализ (весь массив)	0,970	87,7	89,3	86

На рис. 1 показаны ROC-кривые, построенные с помощью модели БЛР и дискриминантного анализа для общей группы, в которой сравнивались пациенты с извитыми и неизвитыми КА со склерозом, где 1 — логистическая регрессия, 2 — дискриминантный анализ, 3 — опорная линия.

Результаты проверки моделей показали, что они могут точно прогнозировать возникновение извитости КА. Модели на основе БЛР и дискриминантного анализа обеспечили правильный прогноз для 86,8% и 87,8% пациентов соответственно.

Проведенное исследование также показало, что качество регрессионной и дискриминантной моделей приемлемо (см. табл. 7). Площади под ROC-кривыми составили 0,979 для модели БЛР и 0,970 — для модели на основе дискриминантного анализа.

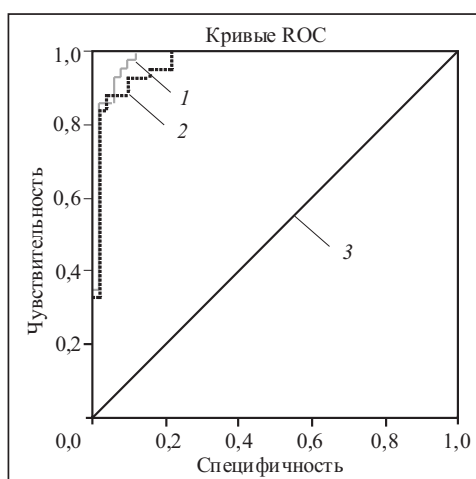


Рис. 1

ЗАКЛЮЧЕНИЕ

При построении моделей на основе БЛР и дискриминантного анализа было использовано 15 независимых переменных. Данные показатели являются значимыми для прогнозирования извитости КА и их изъятие из моделей снизит качество последних. Для сравнения качества построенных моделей использовались правильные отнесения, чувствительность, специфичность в процентах и площадь под ROC-кривой.

Результаты анализа влияющих факторов можно применять для выявления групп риска, а также разработки профилактических и терапевтических мероприятий для рассматриваемых групп пациентов.

СПИСОК ЛИТЕРАТУРЫ

1. Coronary tortuosity: a long and winding road / E.S. Zegers, B.T.J. Meursing, E.B. Zegers, A.J.M. Oude Ophuis // Neth. Heart J. — 2007. — N 15. — P. 191–195.
2. Методика определения истинной длины извитых сосудов / Г.В. Кнышов, Е.А. Настенко, В.Б. Максименко и др. // Щорічник наук. праць Асоціації серцево-судинних хірургів України. — 2011. — Вип. 19. — С. 208–212.
3. Григорьев С.Г., Юнкеров В.И., Клименко Н.Б. Логистическая регрессия. Многомерные методы статистического анализа категориальных данных медицинских исследований: Уч. пособие. — Санкт-Петербург: ВМедА, 2001. — 266 с.

Поступила 28.02.2013