

Прогнозування виникнення факторів ризику склерозу коронарних артерій

А. О. Матвійчук¹, Є. А. Настенко², Є. О. Лебедєва², О. Г. Кисельова¹

¹Національний технічний університет України «Київський політехнічний інститут»

²Національний інститут серцево-судинної хірургії ім. М. М. Амосова НАМН України

Резюме

Методами логістичної регресії і дискримінантного аналізу перевірили існування специфічних чинників ризику розвитку коронарного атеросклерозу у пацієнтів зі звивистими коронарними артеріями. Порівняли якість отриманих моделей. Виявили, що досліджені фактори ризику відповідають загальноновизнаним представленням у світовій літературі.

Ключові слова: бінарна логістична регресія, дискримінантний аналіз, коронарні артерії, коронарний атеросклероз, звивистість артерій.

Клин. информат. и Телемед.
2013. Т.9. Вып.10. с.90–96

Вступ

Атеросклероз коронарних артерій є однією з найбільш частих причин виникнення ішемії міокарда. Атеросклероз коронарних судин, при виникненні ішемічної хвороби серця (ІХС) проявляється стенокардією, розвитком інфаркту міокарда, різними порушеннями провідності і аритміями серця.

До теперішнього часу встановлено цілий ряд факторів різної природи, що сприяють розвитку і прогресуванню ІХС: артеріальна гіпертензія, надлишкова маса тіла, куріння, гіподинамія, цукровий діабет та ін. Як показали результати одного з найбільших міжнародних досліджень MONICA (Multinational Monitoring of Trends and Determinants in Cardiovascular Disease), класичні фактори ризику розвитку атеросклерозу не можуть повністю пояснити розвиток серцево-судинних ускладнень. У даній статті зроблена спроба знайти інші фактори в розвитку та прогресуванні склерозу коронарних артерій (КА).

Мета роботи: перевірити існування специфічних чинників ризику розвитку коронарного атеросклерозу у пацієнтів зі звивистими КА.

Матеріал і методи дослідження

В дослідження було включено 121 пацієнт, яким первинно коронарорентрикулографія (КВГ) виконувалась в НІССХ ім. М. М. Амосова НАМНУ. В залежності

від наявності звивистості та атеросклеротичного ураження КА пацієнти були розподілені на 2 масива даних. Перший масив включав в себе дані пацієнтів зі звивистими інтактними КА (n=55), яким КВГ виконувалась для діагностики ішемічної хвороби серця. До другого масиву увійшли хворі зі звивистістю та атеросклеротичними ураженнями коронарних судин (n=66), яким у подальшому проводилось коронарне шунтування (КШ).

З використанням методу «кожний за себе» порівнювались наступні вибірки: звивисті КА зі склерозом та звивисті КА без склерозу – для виявлення факторів, які впливають на розвиток склеротичних уражень КА.

Статистична обробка матеріалу проводилась з використанням програми IBM Statistics 20.0.

З метою оцінки впливу факторів ризику на перебіг атеросклерозу у пацієнтів зі звивистими КА для перевірки гіпотез при порівнянні якісних (категоріальних) даних між двома незалежними групами використовувався критерій χ^2 Пірсона. Був прийнятий довірчий інтервал 95%, при $p < 0,05$ різницю ознак слід вважати статистично значущою. З масиву досліджуваних ознак було відібрано ряд факторів, які мали найбільш вагомий вплив на прогресування коронарного склерозу. Для вивчення статистичного взаємозв'язку ознак використовували параметричний кореляційний аналіз Спірмена.

Для обробки та визначення системи впливових чинників використано розвідувальний статистичний аналіз з використанням методу бінарної логістичної регресії та дискримінантного аналізу. За допомогою даних методів було створено математичні моделі, в яких вплив кожного показника визначався його ваговою оцінкою або відповідним коефіцієнтом, які використовувались

для оцінки інформативності та прогностичної значущості кожного з факторів прогресування коронарного склерозу у хворих зі звивистими КА.

В процесі кореляційного аналізу з масиву досліджуваних ознак було відібрано ряд факторів, які мали найбільш вагомий вплив на склероз КА.

В математичну модель були примусово включені чинники, що найбільш тісно корелювали з розвитком склерозу КА.

Для оцінки якості побудованих моделей розраховували параметри чутливості і специфічності та площу під ROC-кривою.

Результати і обговорення

По результатам кореляційного аналізу були відібрані 7 категоріальних змінних та 1 неперервна (кількісних) змінна, для яких $p < 0,05$.

Табл. 1. Середні значення показників пацієнтів зі звивистими КА без склерозу [код = 1] та звивистими КА зі склерозом [код = 2].

Група хворих	Хворі з звивисті КА без склерозу (n=55)	Хворі із звивистими КА зі склерозом (n=66)	p
Показник	M ± SD, або Медіана (мін.-макс.) [II-III квантилі]	M ± SD, або Медіана (мін.-макс.) [II-III квантилі]	
Вік	55 (18-76) [47-62]	61 (34-73) [55-68]	0,001

n – загальна кількість хворих в групі.

В табл. 1 відображені середні значення і стандартні відхилення по кількісним змінним для груп пацієнтів зі звивистими та незвивистими КА зі склерозом, а також статистична значимість p. Для змінної вік $p = 0,001$.

Для побудови регресійної моделі використовувались також дискретні (категоріальні) змінні (табл. 2). Всі змінні були включені в модель по результатам кореляційного аналізу.

В табл.2 наведено відношення шансів (OR), яке виражає співвідношення ймо-

вірностей склерозу у звивистих КА при позитивному та негативному значенні вивчаємого фактору.

Із даних таблиці 2 випливає, що найбільші шанси виявити склероз КА мають пацієнти, які перенесли запальні захворювання серця (OR=7,7, 95% CI (0,9-63,8)), в яких спостерігається артеріальна гіпертензія (OR=2,57, 95% CI (0,68-9,63)), цукровий діабет 2-го типу (OR=3,6, 95% CI (0,9-13,64)). Також, впливають на склероз КА такі фактори, як холестерин (OR=1,33, 95% CI

Табл. 2. Фактори ризику, включені в модель прогнозу розвитку звивистості КА для БЛР в групі, в якій порівнюються звивисті КА без склерозу [код = 1] та звивисті КА зі склерозом [код = 2].

Групи по факторам ризику		Звивисті КА без склерозу, n=55	Звивисті КА зі склерозом, n=66	OR (95%CI)*
Фактори ризику	Розбиття значень факторів ризику			
Пол	Чоловічий	26	51	1.0 (reference)
	Жіночий	29	15	0,26 (0,12-0,58)
Артеріальна гіпертензія по історії хвороби (АГ по ІХ)	Норма	7	10	1.0 (reference)
	Гіпертонічна хвороба (ГХ) 1ст	4	2	0,35 (0,05-2,47)
	ГХ 2ст	38	32	0,59 (0,2-1,73)
	ГХ 3ст	6	22	2,57 (0,68-9,63)
Триглицериди	Норма	26	15	1.0 (reference)
	Вище норми	29	40	2,39 (1,08-5,3)
Холестерин	Нижче норми	10	10	1.0 (reference)
	Норма	36	33	0,9 (0,34-2,48)
	Вище норми	9	12	1,33 (0,39-4,57)
Запальні захворювання серця	Немає	54	56	1.0 (reference)
	Є	1	8	7,7 (0,9-63,8)
Захворювання щитовидної залози	Немає	50	63	1.0 (reference)
	Є	5	1	0,16 (0,02-1,4)
Цукровий діабет 2 типу	Немає	52	53	1.0 (reference)
	Є	3	11	3,6 (0,9-13,64)

Примітки:

*) позначення: OR – відношення шансів, CI – 95% довірчий інтервал для OR.

(0,39–4,57)), триглицериди (OR=2,39, 95% CI (1,08–5,3)) і пол (OR=0,26, 95% CI (0,12–0,58)). Тобто, у жінок склероз КА спостерігається майже в чотири рази частіше, ніж у чоловіків.

В загальний список параметрів для побудови моделі на основі бінарної логістичної регресії (БЛР), увійшли 8 змінних (табл. 3).

Модель БЛР в групі пацієнтів, в якій порівнювались випадки зі звивистими

КА без склерозу та звивистими КА зі склерозом має наступний вигляд:

$$z_1 = x_1 \times 1,7 - x_2 \times 0,002 - x_3 \times 0,11 - x_4 \times 0,17 + x_5 \times 2,4 - x_6 \times 3,1 + x_7 \times 1,81 + x_8 \times 0,08 - 8,4$$

Аналіз коефіцієнтів кореляції (стовпець 3, табл. 3) в групі, в якій порівнювались випадки зі звивистими КА без склерозу та звивистими КА зі склерозом

показав, що найбільш значимими факторами є пол та вік.

Порівняння тестової і навчальної вибірок в групі, в якій були порівняні звивисті КА без склерозу із звивистими КА зі склерозом представлено в табл. 4.

Порівняння тестової і навчальної вибірок в групі, в якій порівнюються звивисті зі склерозом КА та без склерозу представлено на рис. 1.

Табл. 3. Статистично значимі ($p < 0,05$) фактори ризику розвитку звивистості КА для моделі БЛР в групі, в якій порівнюються звивисті КА без склерозу [код = 1] та звивисті КА зі склерозом [код = 2].

Фактор ризику	Умове позначення	Коефіцієнт кореляції	Статистична значущість кореляції, p	Коефіцієнт рівняння регресії, $a_{1..n}$
1	2	3	4	5
Пол	X_1	-0,31	0,00	1,7
АГ по IX	X_2	0,18	0,02	-0,002
Триглицериди	X_3	-0,07	0,00	-0,11
Холестерин	X_4	-0,13	0,01	-0,17
Запальні захворювання серця	X_5	0,20	0,03	2,4
Захворювання щитовидної залози	X_6	-0,17	0,06	-3,1
Цукровий діабет 2 типу	X_7	0,18	0,05	1,81
Вік	X_8	0,303	0,001	0,08
$a_0^* = -8,4$				

Примітки:

*) a_0 – вільний член рівняння БЛР.

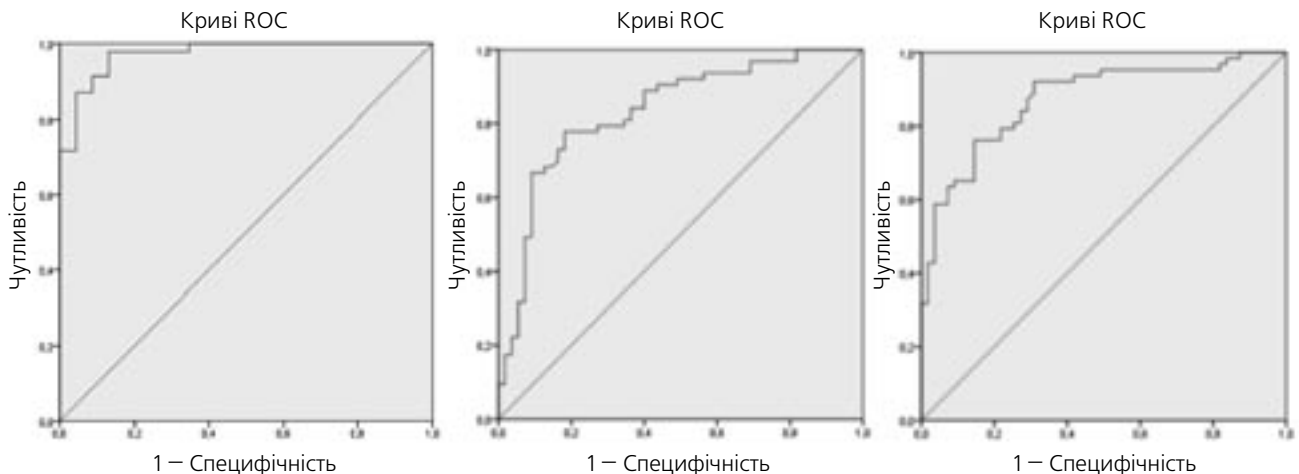


Рис. 1. Криві – ROC рівняння z_1 : а) повна вибірка даних; б) вибірка даних, яка використовується для навчання моделі; в) перехресна вибірка даних в групі, в якій порівнюються звивисті КА без склерозу із звивистими КА зі склерозом.

При прогнозуванні розвитку звивистості КА у групі, в якій були порівняні звивисті КА без склерозу із звивистими КА зі склерозом методом БЛР чутливість моделі на навчальній вибірці становить 80%, а специфічність 75,8% при сумарній точності 77,8%. А при перевірці по тестовій вибірці результати виявилися вищими – чутливість моделі становить 80%, а специфічність 86,7% при сумарній точності 83,6%.

На наступному етапі виконувати побудову моделі на основі дискримінантного аналізу.

Матриця вихідних даних містить 47 ознак. Після логічного та кореляційного аналізу вихідних даних в навчальній матриці для подальшого дослідження були залишені 8 ознак, які мали статистично значущий

Табл. 4. Відсоток правильних віднесень (Perf), чутливість (Se) і специфічність (Sp) моделей при перевірці розбиттям масиву на навчальну (N=61) і тестову (N=60) вибірки в групі, в якій порівнюються звивисті КА без склерозу із звивистими КА зі склерозом.

	auROC	Perf (%)	Sp (%)	Se (%)
Повна вибірка даних	0,974	86,5	91,8	76
Навчальна вибірка	0,833	77,8	75,8	80
Тестова вибірка	0,873	83,6	86,7	80

($p \leq 0,05$) кореляційний зв'язок зі склерозом звивистих КА.

Відбір ознак в остаточну модель здійснювався методом примусового включення дискримінантного аналізу в ППП SPSS 20.0.

Лінійні дискримінантні функції (ЛДФ) розраховуються за формулами:

$$\begin{aligned} \text{ЛДФ}_1 = & -60,3 + 14,57 \times x_1 + 2,6 \times x_2 - \\ & - 1,76 \times x_3 + 4,8 \times x_4 + 13,7 \times x_5 + \\ & + 23 \times x_6 + 13,1 \times x_7 + 0,56 \times x_8 \end{aligned}$$

Табл. 5. Результати досліджень для групи, в якій порівнюються звивисті КА без склерозу [код=1] із звивистими КА зі склерозом [код=2].

Найменування ознаки	Градація ознаки	Код ознаки	Значення коефіцієнтів ЛДФ ₁ * (звивисті без склерозу)	Значення коефіцієнтів ЛДФ ₂ * (звивисті зі склерозом)	Значення НККДФ*	Статистична значущість кореляції, r*	Коефіцієнт кореляції
Пол	Жіночий – 1	X ₁	14,57	16,4	0,66	0	-0,31
	Чоловічий – 2						
Артеріальна гіпертензія по ІХ	Норма – 1	X ₂	2,6	2,6	0,01	0,02	0,18
	ГХ 1 ст. – 2						
	ГХ 2 ст. – 3						
	ГХ 3 ст. – 4						
Триглицериди	Норма – 1	X ₃	-1,76	-1,9	-0,09	0	-0,07
	Вище норми – 2						
Холестерин	Нижче норми – 1	X ₄	4,8	4,6	-0,8	0,01	-0,13
	Норма – 2						
	Вище норми – 3						
Запальні захворювання серця	Немає – 1	X ₅	13,7	15,8	0,42	0,03	0,2
	Є – 2						
Захворювання ЦЗ	Немає – 1	X ₆	23	21,3	-0,29	0,06	-0,17
	Є – 2						
Цукровий діабет 2-го типу	Немає – 1	X ₇	13,1	14,8	0,41	0,05	0,18
	Є – 2						
Вік		X ₈	0,56	0,65	0,61	0,001	0,303
Константа			-60,3	-69,9			

$$\text{ЛДФ}_2 = -69,9 + 16,4 \times x_1 + 2,6 \times x_2 - 1,9 \times x_3 + 4,6 \times x_4 + 15,8 \times x_5 + 21,3 \times x_6 + 14,8 \times x_7 + 0,65 \times x_8$$

За результатами обстеження хворого в формули лінійних дискримінантних функцій підставляються значення тих ознак, які включені в модель, і проводиться розрахунок рівнянь. Конкретного хворого необхідно віднести в ту групу хворих, значення ЛДФ для яких найбільше.

Модель є статистично значущою ($p \leq 0,05$) і має досить високу прогностичну здатність (79,2%). Класифікаційна матриця, в якій представлені результати класифікації хворих за даними навчальної інформації та співвідношення цієї класифікації зі спостережуваними результатами звивистості КА, представлені в табл. 6.

На основі аналізу даних таблиці можна зробити висновок, що в групі пацієнтів зі звитими КА без склерозу запропонована модель забезпечує співпадіння прогнозованого результату з реальним

результатом в 76% випадків (співпадіння результатів у 19 із 25 пацієнтів), у групі звивисті КА зі склерозом – співпадіння прогнозованого результату із реальними результатами склало 80,9% (співпадіння результатів у 38 із 47 пацієнтів зі звивистими КА зі склерозом).

Таким чином, дискримінантна модель прогнозу звивистості КА зі склерозом і без склерозу, заснована на 18 ознаках, володіє класифікаційною здатністю в межах норми (79,2%) і є статистично значущою ($p \leq 0,05$). Чутливість моделі становить 76%, специфічність – 80,9%. Якість моделі допустима для прогнозування звивистості КА зі склерозом і без.

Порівняння тестової і навчальної вибірок, в якій порівнюються звивисті КА без склерозу та звивисті КА зі склерозом, представлено в табл. 7.

При прогнозуванні розвитку звитості коронарних артерій у групі, в якій порівнюються звивисті КА зі склерозом із звивистими КА без склерозу методом дискримінантного аналізу чутливість мо-

делі на навчальній вибірці складає 100%, а специфічність 96,4% при сумарній точності 98,1%. При перевірці по тестовій вибірці результати виявилися вищими – чутливість моделі становить 84%, а специфічність 81,3% при сумарній точності 82,5%.

Порівняння отриманих моделей

При порівнянні якості отриманих моделей було використано в першу чергу параметр чутливості. Для визначення прогнозу за допомогою моделей на основі бінарної логістичної регресії і на основі дискримінантного аналізу використовуються моделі, побудовані на навчальній вибірці 100% масиву.

При оцінці якості моделей при розпізнаванні всіх випадків навчальної вибірки найкращий результат має модель бінарної логістичної регресії.

Табл. 6. Класифікаційна матриця.

Звивистість			Передбачена приналежність до групи		Всього
			Звивисті без склерозу	Звивисті зі склерозом	
Вихідні	Частота	Звивисті без склерозу	19	6	25
		Звивисті зі склерозом	9	38	47
	%	Звивисті без склерозу	76,0	24,0	100,0
		Звивисті зі склерозом	19,1	80,9	100,0

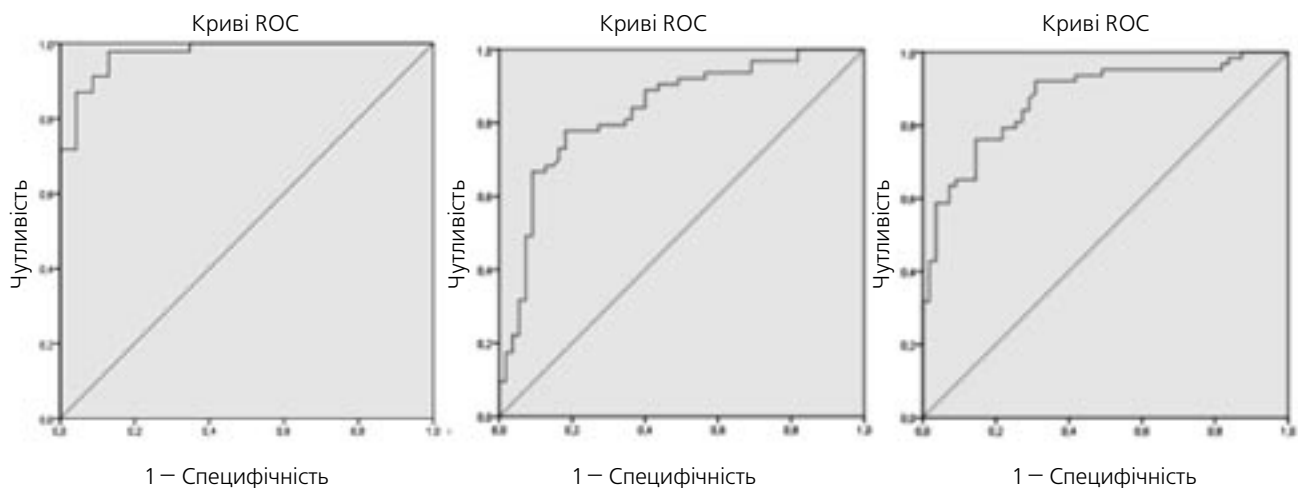


Рис. 2. Криві – ROC рівняння: а) повна вибірка даних; б) вибірка даних, яка використовується для навчання моделей; в) перехресна вибірка даних в групі, в якій порівнюються звивисті КА без склерозу зі звивистими КА зі склерозом.

Табл. 7. Відносна кількість правильних віднесень (Perf), чутливість (Se) і специфічність (Sp) моделей при перевірці розбиттям масиву на навчальну (N = 61) і тестову (N = 60) вибірки в групі, в якій порівнюються звивисті КА зі склерозом із звивистими КА без склерозу.

	auROC	Perf (%)	Sp (%)	Se (%)
Повна вибірка даних	0,950	86,5	91,8	76
Навчальна вибірка	0,988	98,1	96,4	100
Тестова вибірка	0,851	82,5	81,3	84

Табл. 8. Площа під ROC-кривою, відсоток правильних віднесень, чутливість і специфічність моделей для групи, в якій порівнюються звивисті КА зі склерозом із звивистими КА без склерозу, побудованих за допомогою бінарної логістичної регресії і дискримінантного аналізу.

Тип моделі і розмір навчальної вибірки	auROC	Відсоток правильних віднесень	Чутливість	Специфічність
Бінарна логістична регресія (весь масив)	0,974	83,3%	72%	89,4%
Дискримінантний аналіз (весь масив)	0,950	79,2%	76%	80,9%

Висновки

В результаті при побудові моделей за допомогою бінарної логістичної регресії та дискримінантного аналізу було використано 8 незалежних змінних. Всі ці показники є прогностично значущими і не можуть бути вилучені з моделей без серйозного зниження їх якості. Для зіставлення якості побудованих моделей використовувалися відсоток вірних віднесень, чутливість, специфічність і площа під ROC-кривою.

Дане дослідження показало, що якість, як регресійної моделі, так і дискримінантної є цілком прийнятною (табл. 8.).

Застосований метод моделювання дозволив виявити чинники ризику розвитку коронарного атеросклерозу у пацієнтів з ЗКА, які відповідають загальноновизнаним представленням у світовій літературі. Співставлення чинників, які пов'язані з атеросклерозом дозволяють більш чітко виокремити комплекси факторів для обох патологій.

Література

- Амосова Е. Н. Эффективность коррекции факторов риска и различных методов хирургического лечения больных хронической ИБС в отношении предотвращения смерти от инфаркта миокарда: мифы и реальность // Сердце і судини. – 2009. – № 4. – С. 12–24.
- Бокерия Л. А., Алекаян Б. Г., Бузиашвили Ю. И. Стентирование как метод лечения больных при возврате стенокардии после операции коронарного шунтирования // Руководство по рентгеноэндоваскулярной хирургии сердца и сосудов / Под ред. Л. А. Бокерия, Б. Г. Алекаяна. – М.: НЦССХ им. А. Н. Бакулева РАМН, 2008. – Т. 3. – Глава 32. – С. 623–636.
- Григорьев, С. Г. Логистическая регрессия. Многомерные методы статистического анализа категориальных данных медицинских

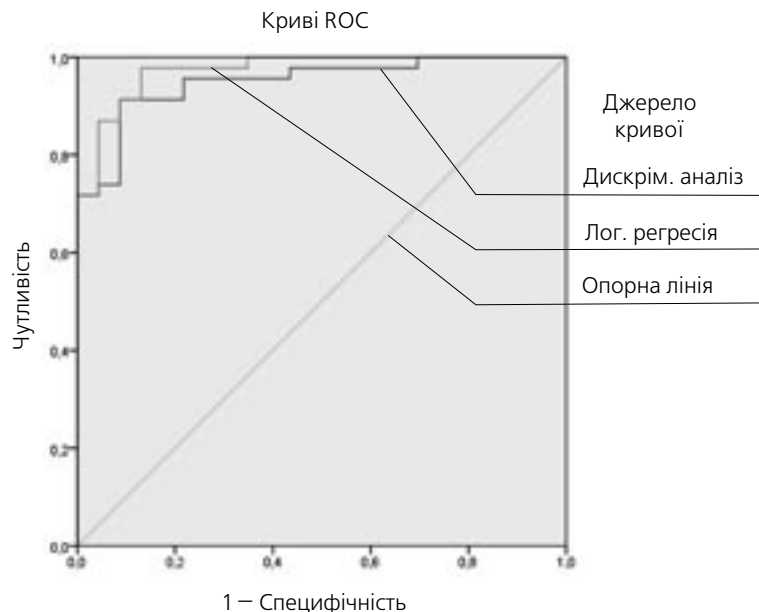


Рис. 3. Криві — ROC для групи, в якій порівнювались звивисті КА зі склерозом із звивистими КА без склерозу, побудованих за допомогою бінарної логістичної регресії і дискримінантного аналізу.

исследований / С. Г. Григорьев, В. И. Юнкеров, Н. Б. Клименко // Уч. пособие. ВМедА. – СПб, 2001. – С. 10–21.

Prediction of risk factors for coronary sclerosis

**A. O. Matviychuk¹, E. A. Nastenko²
E. O. Lebedeva², O. G. Kyselova¹**

¹National Technical University of Ukraine

«Kyiv Polytechnic Institute»

²Amosov National Institute of Cardiovascular Surgery NAMS of Ukraine

Abstract

Risk factors of coronary atherosclerosis in patients with tortuous coronary arteries were identified using logistic regression and discriminant analysis methods. Quality of the models were compared. It was found that examined risk factors correspond to generally accepted views in the world literature.

Key words: binary logistic regression, discriminant analysis, coronary artery disease, coronary atherosclerosis, arterial tortuosity.

Прогнозирование возникновения факторов риска склероза коронарных артерий

**А. О. Матвийчук¹, Е. А. Настенко²
Е. О. Лебедева², О. Г. Киселева¹**

¹Национальный технический университет Украины «Киевский политехнический институт»

²Национальный институт сердечно-сосудистой хирургии им. Н. М. Амосова НАМН Украины

Резюме

Методами логистической регрессии и дискриминантного анализа проверили существование специфических факторов риска развития коронарного атеросклероза у пациентов с извитыми коронарными артериями. Сравнили качество полученных моделей. Обнаружили, что исследованные факторы риска соответствуют общепризнанным представлениям в мировой литературе.

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

Листування

А. О. Матвійчук

Національний технічний університет України «Київський політехнічний інститут»

вул. Космонавта Волкова 24, кв. 69

Київ, 02166, Україна

тел. +38 (097) 127-23-95

ел. пошта: nutius@yandex.ru