

УДК 616.133.3 + 616.133.31-004.6-078.33

ИДЕНТИФИКАЦИЯ ЛАТЕНТНЫХ ИНФОРМАТИВНЫХ НАБОРОВ ПРИЗНАКОВ ПРИ ОЦЕНКЕ НЕСТАБИЛЬНОСТИ АТЕРОСКЛЕРОТИЧЕСКОЙ БЛЯШКИ

¹Розыходжаева Г.А., ²Игнатьев Н.А., ³Усманова З.А.

¹Центральная клиническая больница № 1, Ташкент, e-mail: gulnoradm@inbox.ru;

²Национальный университет Узбекистана имени Мирзо Улугбека, Ташкент;

³Ташкентский институт усовершенствования врачей, Ташкент

Рассмотрены вопросы, связанные с разработкой информационной модели нестабильности атеросклеротической бляшки с учетом активности матричных металлопротеаз и их ингибиторов. Подробно освещены все этапы интеллектуального анализа данных проведенного исследования. Разработана логическая модель нестабильной атеросклеротической бляшки. Получено решающее правило, согласно которому вычислялось значение индекса нестабильности атеросклеротической бляшки.

Ключевые слова: информационная модель, матричная металлопротеиназа-9, тканевой ингибитор металлопротеиназы-1, нестабильная атеросклеротическая бляшка

IDENTIFICATION OF LATENT SET OF INFORMATIVE PARAMETERS IN THE ASSESSMENT OF ATHEROSCLEROTIC PLAQUE INSTABILITY

¹Rozykhodzhaeva G.A., ²Ignatev N.A., ³Usmanova Z.A.

¹Central Clinical Hospital № 1, Tashkent, e-mail: gulnoradm@inbox.ru;

²National University of Uzbekistan named Mirzo Ulugbek, Tashkent;

³Tashkent Institute of Postgraduate Medical, Tashkent

The article is devoted to the questions connected with development of information model of atherosclerotic plaque instability taking into account the activity of matrix metalloproteinases and their inhibitors. Is explained in detail all stages of data minning of researchs. Is developed logical a model of unstable atherosclerotic plaque. Obtained decision rule, according to which was calculated the index of instability of atherosclerotic plaques.

Keywords: information model, matrix metalloproteinase-9, tissue inhibitor of metalloproteinase-1, unstable atherosclerotic plaque

Ранняя диагностика различных заболеваний на основе биологических маркеров в последние годы становится более актуальной [7, 9]. В кардиологии большой интерес вызывает выявление в крови биомаркеров, характеризующих состояние атеросклеротической бляшки (АСБ). В последнее время широко обсуждается роль матричных металлопротеиназ (ММП) и их тканевых ингибиторов (ТИМП) в патогенезе нестабильности АСБ [13]. Анализ современных литературных источников показал, что необходимы дальнейшие исследования, чтобы понять роль ММП и ТИМП, вовлеченных в атерогенез и разрыв бляшки. Настоящая работа выполнена в соответствии с приоритетными направлениями развития науки и технологии Республики Узбекистан в рамках программы ГНТП-9 «Разработка новых технологий профилактики, диагностики, лечения и реабилитации заболеваний человека» в рамках прикладного Государственного научно-технического проекта МЗ РУз АДСС 25 «Разработка метода ранней

прижизненной диагностики атеросклероза магистральных артерий».

Цель исследования – разработка информационной модели нестабильности атеросклеротической бляшки с учетом активности матричных металлопротеаз и их тканевых ингибиторов.

Материалы и методы исследования

Работа по интеллектуальному анализу медицинских данных по атеросклеротическим бляшкам проводилась совместно с сотрудниками кафедры программных и сетевых технологий Национального университета Узбекистана имени М. Улугбека.

По данным клинических, лабораторных, инструментальных исследований сформирована обучающая выборка, которая послужила основой при построении информационной модели для решения задач прогнозирования. Для решения задач математического моделирования с использованием обобщенных оценок мы отобрали данные 73 больных ИБС со стенозом сонной артерии более 50%. Основным условием включения в анализируемую базу данных было отсутствие пропусков в вариационном ряде. В базу данных введено 83 параметра. В работе использовались обобщенные оценки – агрегированные (комбинированные) показатели, которые применялись для отображения отно-

шений между объектами двух групп (классов K_1 , K_2) в разнотипном признаковом пространстве на числовую ось. В K_1 , K_2 включены соответственно пациенты со стабильными и нестабильными АСБ более 50% по результатам предварительного дуплексного сканирования. Для получения комбинированных показателей использовано правило агломеративной иерархической группировки разнотипных признаков с целью нелинейного отображения их значений в описании объектов на числовую ось.

Результаты исследования и их обсуждение

Анализ данных проведенного исследования включал следующие этапы:

1. Разведочный анализ с учетом пропусков (не измеренных значений) в данных.
2. Формирование набора информативных признаков на основе оценки их вкладов в процесс разделения классов.
3. Формирование обучающей выборки путём удаления шумовых объектов.

4. Отбор индивидуальных информативных наборов признаков для каждого объекта.

5. Нелинейное отображение определяемых наборов признаков на числовую ось и формирование решающего правила.

Для разработки информационной модели использовано 9 методов обследования, 83 параметра (признака), 73 из которых измерялись в количественных и 10 в номинальных шкалах измерений.

Использование традиционных методов статистики (например, метода Стьюдента) допускает включение в анализ только количественных признаков. Методы статистики не позволяют оценить значимость качественных показателей и сочетаний количественных и качественных показателей в принятии решений. Статистические характеристики количественных признаков по классам K_1 , K_2 приведены в таблице.

Статистические характеристики количественных признаков по стабильным и нестабильным АСБ

Название признака, единица измерения	K_1 Больные со стабильными атеросклеротическими бляшками	K_2 Больные с нестабильными атеросклеротическими бляшками	Значение p по Стьюденту
1	2	3	4
Возраст, годы	67,9 ± 1,5	62,8 ± 1,2	> 0,05
Мужчины/женщины, n (%)	30/15 (67/33)	25/3 (89/11)*	< 0,05
Рост, м	1,68 ± 0,01	1,70 ± 0,01	> 0,05
Вес, кг	78,31 ± 1,89	78,36 ± 2,4	> 0,05
Индекс массы тела, усл. ед. (кг/м ²)	27,7 ± 0,59	27,2 ± 0,79	> 0,05
Частота сердечных сокращений покоя, уд./мин	73,91 ± 1,88	76,18 ± 1,87	> 0,05
Систолическое артериальное давление, мм рт.ст.	130,78 ± 3,09	132,5 ± 2,45	> 0,05
Диастолическое артериальное давление, мм рт.ст.	80,89 ± 1,65	83,57 ± 1,17	> 0,05
Артериальная гипертензия, n (%)	42 (93,3)	26 (92,9)	> 0,05
Гиперлипидемия, n (%)	22 (48,9)	15 (53,6)	> 0,05
Ожирение, n (%)	19 (42,2)	7 (25)	> 0,05
Сахарный диабет, n (%)	14 (31,1)	14 (50)	> 0,05
Стенокардия напряжения: n (%)			
ФК II	32 (71,1)	23 (82,1)	> 0,05
ФК III	4 (8,9)	3 (10,7)	> 0,05
Инфаркт миокарда в анамнезе, n (%)	6 (13,3)	4 (14,3)	> 0,05
Перенесенный инсульт, n (%)	10 (22,2)	16 (57,1)*	< 0,01
Периферический атеросклероз, n (%)	1 (2,2)	4 (14,3)	> 0,05
Положительный семейный анамнез, n (%)	1 (2,2)	5 (17,9)*	< 0,05
Курение, n (%)	3 (6,7)	7 (25)*	< 0,05
Применение бета-адреноблокаторов, n (%)	32 (71,1)	10 (35,7)*	< 0,02
Применение статинов, n (%)	14 (31,1)	22 (78,6)*	< 0,001
Максимальная толщина комплекса «интима-медиа» общей сонной артерии в стандартной точке измерения, мм	0,99 ± 0,04	1,2 ± 0,05*	< 0,002
Максимальная степень стеноза сонных артерий, %	44,2 ± 3,6	75,4 ± 2,9*	< 0,001

Окончание таблицы

1	2	3	4
Тип атеросклеротической бляшки слева	2,86 ± 0,21	3,32 ± 0,31	> 0,05
Тип атеросклеротической бляшки справа	3,35 ± 0,26	3,6 ± 0,28	> 0,05
Диаметр аорты, см	3,3 ± 0,08	3,17 ± 0,05	> 0,05
Конечный диастолический размер, см	5,22 ± 0,09	4,56 ± 0,08*	< 0,001
Конечный диастолический объем, мл	132,08 ± 5,19	99,21 ± 3,8*	< 0,001
Диаметр левого предсердия, см	3,4 ± 0,07	3,12 ± 0,04*	< 0,001
Конечный систолический размер, см	3,7 ± 0,09	2,94 ± 0,07*	< 0,001
Конечный систолический объем, мл	60,03 ± 3,49	35,04 ± 2,03*	< 0,001
Толщина межжелудочковой перегородки, см	1,18 ± 0,03	1,34 ± 0,06*	< 0,05
Фракция выброса левого желудочка, %	54,18 ± 1,3	64,07 ± 0,9*	< 0,001
Ударный объем, мл	72,03 ± 2,76	61,5 ± 2,62*	< 0,02
Толщина задней стенки левого желудочка, см	1,19 ± 0,03	1,27 ± 0,02*	< 0,05
Толщина стенки правого желудочка, см	2,99 ± 0,08	2,72 ± 0,02*	< 0,002
Гипертрофия межжелудочковой перегородки, <i>n</i> (%)	1 (2,2)	1 (3,6)	> 0,05
Диастолическая дисфункция левого желудочка, <i>n</i> (%)	12 (26,7)	25 (89,3)*	< 0,001
Атриовентрикулярная блокада, <i>n</i> (%)	1 (2,2)	1 (3,6)	> 0,05
Блокада левой ножки пучка Гисса, <i>n</i> (%)	1 (2,2)	1 (3,6)	> 0,05
Сахар в крови, ммоль/л	5,69 ± 0,24	6,68 ± 0,46	> 0,05
Общий холестерин, ммоль/л	5,2 ± 0,16	4,6 ± 0,18	> 0,05
Триглицериды, ммоль/л	1,8 ± 0,12	2,2 ± 0,29	> 0,05
Холестерин липопротеидов высокой плотности, ммоль/л	0,94 ± 0,02	1,04 ± 0,05	> 0,05
Холестерин липопротеидов низкой плотности, ммоль/л	3,7 ± 0,12	3,2 ± 0,15	> 0,05
Индекс атерогенности ≤ 3,0 (ХС ЛНП/ХС ЛВП)	3,98 ± 0,12	3,24 ± 0,16	< 0,001
Коэффициент атерогенности ≤ 3,5 ((ОХС-ЛВП)/ЛВП)	4,59 ± 0,14	3,48 ± 0,15	< 0,001
Матриксная металлопротеиназа-9, нг/мл	200,9 ± 8,5	539,03 ± 57,1	< 0,001
Тканевой ингибитор металлопротеиназы-1, нг/мл	1180,98 ± 47,98	2410,2 ± 123,8	< 0,001
Матриксная металлопротеиназа-9/ тканевой ингибитор металлопротеиназы-1	0,18 ± 0,009	0,24 ± 0,03	> 0,05
Цинк в сыворотке, мкмоль/л	13,7 ± 1,4	15,9 ± 0,62	> 0,05
Цинк в волосах, мкг/г	220,5 ± 17,07	212,5 ± 17,01	> 0,05
Цинк в бляшках, мкг/г	–	82,8 ± 16,8	–

Учитывая сложность постановки диагноза в многопараметрическом пространстве, предлагается в качестве инструментария при формировании математической модели диагностирования пациентов со стабильной и нестабильной АСБ применить технологию синтеза обобщенных оценок. Обобщенные оценки – это агрегированные (комбинированные) показатели, которые используются для отображения отношений между объектами двух групп (классов K_1 , K_2) в разнотипном признаковом пространстве на числовую ось [1].

Комбинированные показатели формируют новое признаковое пространство, размерность которого меньше исходного. Для получения комбинированных показателей

использовано правило агломеративной иерархической группировки разнотипных признаков с целью нелинейного отображения их значений в описании объектов на числовую ось. Результаты нелинейного отображения рассматривались как значения обобщенных оценок (новых признаков) в описании объектов.

При решении задач прогнозирования АСБ применялся метод поиска логических закономерностей в форме нелинейных проекций обобщенных оценок объектов по определяемым наборам из 31-го признака (симптома, синдрома) на выборке из представителей классов K_1 и K_2 . В самую устойчивую закономерность вошли 5 исходных показателей: вес пациента, ТИМП-1, ММП-9, ФВ, СССА. Согласно решающему правилу

с пороговым значением [2] вычислялось значение индекса нестабильности атеросклеротической бляшки (ИНАСБ):

$$\text{ИНАСБ} = 0,2 \cdot (0,99623 \cdot (y_3 - 0,02887) - 0,0061 \cdot (\text{СССА} - 55,0)) + 0,8 \cdot (9,1088603533\text{E} - 03) \cdot (y_3 \cdot \text{СССА} + 1,41804) - 0,0125,$$

где

$$y_1 = 0,3 \cdot (-0,00059 \cdot (\text{ММП-9} - 274,96) - 0,0003 \cdot (\text{ТИМП-1} - 1602,878)) + 0,7 \cdot (-2,6440737712\text{E} - 07) \cdot (\text{ММП-9} \cdot \text{ТИМП-1} - 428481,41184);$$

$$y_2 = 0,2 \cdot (1,00483 \cdot (y_1 + 0,01192) - 0,0091 \cdot (\text{ФВ} \% - 60,0)) + 0,8 \cdot (1,5979164850\text{E} - 02) \cdot (y_1 \cdot \text{ФВ} \% + 0,77465);$$

$$y_3 = 0,6 \cdot (0,00448 \cdot (\text{вес} - 72,0) + 1,06 \cdot (y_2 + 0,00913)) + 0,4 \cdot (1,0333362905\text{E} - 02) \cdot (\text{вес} \cdot y_2 + 0,72105).$$

Если ИНАСБ > 0, то пациент относится к классу стабильной АСБ, при ИНАСБ < 0 – к классу нестабильной АСБ. Пороговое значение для решающего правила из [2] вычислялось по формуле

$$P = \frac{\min K_1 + \max K_2}{2},$$

где $\min K_1$, $\max K_2$ – соответственно минимальное значение обобщенной оценки из класса K_1 и максимальное из K_2 [3].

Математический аппарат интеллектуального анализа данных является наиболее приемлемой альтернативой традиционному статистическому анализу, так как позволяет избежать искажения результатов оценки риска развития клинических проявлений атеросклероза в случаях введения в систему новых данных, появления данных, существенно отличающихся от основной массы наблюдений, и ошибочных измерений [10, 12]. Для извлечения новых знаний (скрытых закономерностей) из базы данных мы использовали современные методы интеллектуального анализа [6].

В отличие от традиционных статистических методов интеллектуальный анализ находит широкое применение там, где необходимо моделировать подобие человеческой интуиции и, как правило, трудно построить явные алгоритмы. Для достижения высокой точности решения прикладных задач данные должны соответствовать набору жестких требований. Из теории и практики вычислений известно, что эти требования часто бывают диаме-

трально противоположными. Алгоритмы методов интеллектуального анализа могут использоваться в качестве инструментария для пополнения баз знаний экспертных систем [1].

Опыт использования системы диагностики доклинического атеросклероза подтвердил эффективность применения медицинской статусметрии в целях своевременной профилактики сердечно-сосудистых заболеваний атерогенной природы. Однако одновременно был выявлен и ряд серьезных недостатков используемой системы, проявляющихся в значительном искажении оценки риска для лиц, показатели которых имели значения, далеко отстоящие от основной массы наблюдений [8].

В другом исследовании [11, 12] искусственные нейронные сети (ИНС) рассматривались в качестве математического аппарата, воспроизводящего условную функцию частоты на основании альтернативных обучающих выборок. Математический аппарат ИНС приведен как наиболее приемлемая альтернатива модели Фишера для оценки риска развития атеросклероза по данным врачебных, инструментальных и лабораторных исследований.

В отличие от перечисленных выше исследований применение метода вычисления обобщенных оценок с помощью иерархической агломеративной группировки в нашей работе было целесообразным по следующим причинам [1]:

– обобщенные оценки образуют новое признаковое пространство, размеры которого меньше исходного;

– решается проблема использования алгоритмов классификации, реализация которых была неэффективна из-за большой размерности признакового пространства, либо возможна при одном типе шкал измерений;

– в процессе группировки происходит последовательный отбор информативных наборов признаков;

– нелинейное отображение описаний объектов на числовую ось по определяемым комбинациям признаков является средством обнаружения устойчивых логических закономерностей (новых знаний) в хранилищах данных [2].

Нелинейное отображение описаний объектов на числовую ось по определяемым комбинациям признаков является средством обнаружения устойчивых логических закономерностей (новых знаний) в хранилищах данных [1, 2, 3, 4, 5]. Статистические методы для этих целей не подходят, так как в процессе вычисления статистических показателей используются усредненные данные.

В работе решалась задача повышения информативности и устойчивости прогнозирования нестабильности АСБ у пациентов с каротидным атеросклерозом за счет вычисления индекса нестабильности атеросклеротической бляшки. Существующие методы прогнозирования (например, способ индивидуальной количественной оценки риска развития клинических проявлений атеросклероза [8]) имеют основной недостаток – большое число исследуемых параметров показателей, полученных в ходе клинического обследования, которые попадают в область равенства прогностических коэффициентов.

В данном исследовании поставленная задача решается с применением решающего правила в форме нелинейных зависимостей между такими пятью исходными показателями как вес пациента, ТИМП-1, ММП-9, ФВ, СССА. Эти исходные показатели образуют латентный (явно не измеряемый) признак. Смысл полученного результата заключается в повышении информативности и устойчивости оценки риска развития нестабильной АСБ и индивидуального риска развития клинических осложнений АСБ. В самую устойчивую закономерность вошли пять исходных показателей: вес пациента, ТИМП-1, ММП-9, ФВ, СССА. Используя данные предварительного обследования и значе-

ний этих пяти показателей (вес пациента, ТИМП-1, ММП-9, ФВ, СССА), которые прямо или косвенно отражают интенсивность развития атеросклеротических процессов на всех стадиях развития, в клинической практике возможно с высокой эффективностью прогнозировать нестабильность атеросклеротической бляшки. Данные верификации разработанной модели диагностики атеросклеротического процесса показали в 92% случаев соответствие диагнозам пациентов, полученным по экспертной информации.

Высокая чувствительность метода к изменению взаимосвязей показателей позволяет спрогнозировать развитие нестабильности АСБ, провести своевременные профилактические и лечебные мероприятия. Последующий анализ динамики риска позволяет также оценить эффективность принятых мер и при необходимости скорректировать профилактику и лечение клинического атеросклероза.

Выводы

1. Разработанный алгоритм автоматизированной диагностики с использованием математических методов способствует повышению информативности и устойчивости прогнозирования нестабильности АСБ у пациентов с каротидным атеросклерозом за счет вычисления индекса нестабильности атеросклеротической бляшки.

2. Выделенные на основе информационного подхода наиболее существенные «весомые» факторы, вошедшие в наиболее устойчивую закономерность (вес пациента, сывороточный уровень ТИМП-1, ММП-9, фракция выброса левого желудочка по данным эхокардиографии и степень стеноза сонных артерий по данным дуплексного сканирования) позволяют оптимизировать лечебно-диагностический процесс при атеросклерозе. Получено решающее правило, согласно которому на основе вышеуказанных факторов можно вычислить значение индекса нестабильности атеросклеротической бляшки.

3. Разработанная логическая модель НАСБ направлена на повышение достоверности результатов диагностики. Данные верификации разработанной модели диагностики атеросклеротического процесса показали в 92% случаев соответствие диагнозам пациентов, полученным по экспертной информации.

Список литературы

1. Игнатъев Н.А. Интеллектуальный анализ данных на базе непараметрических методов классификации и разделения выборок объектов поверхностями. – Ташкент: Национальный университет Узбекистана им. Мирзо Улугбека, 2008. – 108 с.
2. Игнатъев Н.А., Нуржонов Ш.Ю. Выбор параметров регуляризации для повышения обобщающей способности дискриминантных функций // Узбекистон Республикаси Курол Кучлари академиясининг хабарлари. – 2014. – № 1 (14). – С. 81–87.
3. Игнатъев Н.А. Извлечение явных знаний из разнотипных данных с помощью нейронных сетей // Вычислит. технологии. – 2003. – № 2 (8). – С. 69–73.
4. Игнатъев Н.А. О повышении эффективности реализации методов итеративной группировки // Проблемы управления и информатики. – 2002. – № 1. – С. 115–119.
5. Игнатъев Н.А., Мадрахимов Ш.Ф. О некоторых способах повышения прозрачности нейронных сетей // Вычислит. технологии. – 2003. – № 6 (8). – С. 31–37.
6. Колесникова С.И. Методы анализа информативности разнотипных признаков // Вестник Томского государственного университета Управление, вычислительная техника и информатика. – 2009. – № 1 (6). – С. 69–80.
7. Пальцева Е.М. Комплексный подход к молекулярной диагностике ряда социально значимых заболеваний: автореф. дис. ... д-ра мед. наук. – М.: МГУ, 2010. – 47 с.
8. Патент Российская Федерация № 2385668. Способ индивидуальной количественной оценки риска развития клинических проявлений атеросклероза / Бовтюшко В.Г., Бовтюшко П.В., Поддубский Г.А., Юсупов А.Н. // Бюл. – 2010. – № 10.
9. Полонская Я.В., Чернявский А.М., Волков А.М. и др. Корреляции биомаркеров воспаления и деструкции в крови и в сосудистой стенке у мужчин с коронарным атеросклерозом // Бюллетень СО РАМН. – 2011. – № 5. – С. 25–30.
10. Розыходжаева Г.А. Функциональное состояние сердечно-сосудистой системы больных ишемической болезнью сердца пожилого и старческого возраста (патогенетические и клинико-диагностические аспекты): автореф. дис. ... д-ра мед. наук. – Ташкент: РСНПМЦТ и МР, 2006. – 44 с.
11. Юсупов А.Н., Бовтюшко П.В. Применение искусственных нейронных сетей в оценке риска развития клинических проявлений атеросклероза // Вестник Российской Военно-медицинской академии. Приложение. – 2010. – № 2 (30) – С. 193–197.
12. Юсупов А.Н. Применение искусственных нейронных сетей в оценке риска развития клинических проявлений атеросклероза у работников, занятых во вредных условиях труда (на примере нефтеперерабатывающих производств): автореф. дис. ... канд. биол. наук. – М.: ФГБУ «Военно-медицинская Академия имени С.М. Кирова», 2012. – 24 с.
13. Fiotti N., Altamura N., Orlando C. et al. Metalloproteinases-2, -9 and TIMP-1 expression in stable and unstable coronary plaques undergoing PCI // Int. J. Cardiol. – 2008. – Vol. 21. – P. 350–357.