

раженным органам и выделение изоферментного спектра сердца производят на шаге 5 алгоритма. Наиболее вероятной комбинацией в точке пика является сочетание 40% ЛДГ сердца и 60% печени, а в точке $K=2-30$, ЛДГ сердца и 70% печени. В соответствии с данной комбинацией осуществляется построение векторов $\vec{b}^1$ и $\vec{b}^2$, где $\vec{b}^1$ и $\vec{b}^2$ характеризуют влияние сердца и печени на изменение ферментативной активности ЛДГ в сыворотке крови. На основании проведенного анализа можно сделать вывод, что на повышение ферментативной активности после пика в течение двух суток влияет функция $p(i)$ сердца и печени, отражающая реакцию этих органов на гипоксию организма.

Таким образом, разработанный алгоритм на основе построенной математической модели осуществляет дифференциацию активности изоферментов ЛДГ по органам, вовлеченным в процесс гиперферментемии. На примере инфарктов сердца показана возможность использования этого алгоритма в клинической практике при определении степени поражения органов и динамических характеристик патологических процессов.

ЛИТЕРАТУРА

1. Акопян Ж. Н., Зарафян Н. М., Казанчян Г. П. Биолог. ж. Армении. 39, 2, 97, 1986.
2. Вилкинсон Д. Принципы и методы диагностической энзимологии. 256, М., 1981.
3. Зарафян Н. М., Акопян Ж. Н. Вопросы молекулярно-клеточной биологии. 48, Ереван, 1983.
4. Исаков И. И., Коровкин Б. Ф., Маркелов Н. М. Введение в клиническую энзимологию. 80, Л., 1974.
5. Камаров Ф. И., Коровкин Б. Ф., Меньшиков В. В. Биохимические исследования в клинике. 40, Л., 1981.
6. Мосс Д. В., Баттсворт П. Дж. Энзимология и медицина. 14, М., 1978.
7. Zondag H. A. Determination and diagnostic significance of lactate dehydrogenase isoenzymes., 5, Assen, 1963.

Поступило 15.IV 1986 г.

Биолог. ж. Армении, т. 39, № 9, с. 763—768, 1986

УДК 616.7.681.3+621.391.193

НЕКОТОРЫЕ ВОПРОСЫ АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ ВНУТРИ ОПРЕДЕЛЕННЫХ НОЗОЛОГИЙ

Г. С. ПОГОСЯН, К. М. ДЖУГАРЯН

РИВЦ МЗ Армянской ССР, Ереван

Аннотация — С помощью математических методов и ЭВМ проведена классификация объектов биомедицинских исследований с выделением среди них гомогенных групп. Использован иерархический среднесвязывающий кластер-анализ с последующим применением дискриминантного анализа для получения решающих правил отнесения объекта исследования к одной из выделенных гомогенных групп.

Անուագիր—Մաթեմատիկական մեթոդների և էՀՄ-ի միջոցով առաջված է կենսաբժշկական հետազոտությունների ենթարկված օբյեկտների դասակարգումը՝ նրանցում հոմոգեն խմբերի առանձնացմամբ:

Կիրառված է հիերարխիական միջին կապող կլաստեր-անալիզի դրան հաջորդող դիսկրիմինանտ անալիզով որոշող կանոններ առանալու համար. որոնք թույլ են տալիս ուսումնասիրվող օբյեկտի ներդրավումը դասակարգումից առաջված հոմոգեն խմբերից որն է մեկուսի:

Abstract — The classification of biomedical investigations objects with the isolation of homogeneous groups among them is proposed using mathematical methods and computers. Hierarchic average linkage cluster-analysis followed by the discriminant analysis has been used for the development of decisive rules to forecast the object of investigation to one of the isolated homogeneous groups.

Ключевые слова: моделирование биоматематическое, кластер, линейная дискриминантная функция, реклассификация, расстояние Мехалюковиса, плоскость Физнера.

Одной из важнейших задач применения математических методов и ЭВМ является создание принципов построения и использования различных моделей, имитирующих описание объектов биомедицинских исследований, взаимосвязи параметров и протекание реальных процессов. Однако гетерогенность медико-биологических данных требует обоснования необходимости и правомочности применения той или иной модели в конкретных приложениях [2, 6], основным критерием которого является биомедицинская интерпретируемость параметров модели и полученных результатов. В этом отношении биоматематическое моделирование (БММ) [3] является достаточно интересным и результативным подходом. Для решения задач БММ необходимо выделить гомогенные группы, поскольку каждая модель адекватна лишь в определенной области данных. Именно поэтому вопрос объективной классификации (т. е. выделения гомогенных групп) является одним из основных этапов БММ. При этом возможны три подхода к ее решению:

1. Класс может описываться перечнем своих членов; 2. Класс может определяться некоторым общим набором свойств своих членов; 3. Класс может определяться тенденцией к образованию кластеров в пространстве параметров [9].

Очевидно, что третье определение наиболее применимо в различных приложениях, когда число членов класса неизвестно и список общих свойств также не определен.

Цель настоящего исследования состояла в выявлении гомогенных групп больных острым инфарктом миокарда (ОИМ), исследовании их статистической значимости, а также получении решающих правил для отнесения нового больного к одной из групп.

Нами были проанализированы истории болезни больных ОИМ, госпитализированных в НИИ кардиологии МЗ АрмССР с 1976 по 1982 гг. Математической базой настоящего исследования служили методы теории распознавания образов, которые по праву считаются одним из наиболее адекватных при решении подобных задач [7, 8].

На начальном этапе в анализ были включены 56 параметров (данные радикардиограмм, лабораторных анализов, физических исследований и т. д.), характеризующих изучаемые объекты. Для последующего анализа были отобраны лишь объекты, обладающие полным набором исследуемых характеристик. Таким образом, анализируемая выборка была представлена 219 объектами.

Как известно [10, 12], одним из наиболее результативных методов выявления гомогенных групп является кластер-анализ. К сожалению, в настоящее время не существует методологических основ для объективного выбора одного из существующих на сегодняшний день многочисленных методов кластер-анализа [13]. Кроме того, выбор критерия существования и статистической значимости выявляемых кластеров далеко не однозначен [16]. Однако, как свидетельствуют большинство авторов [11, 14, 15], во многих случаях иерархический среднесвязывающий кластер-анализ является одним из лучших. В [4] показано, что лишь на множестве, представленном в основном радиокардиографическими параметрами, удастся достигнуть удовлетворительной картины кластирования.

В свете сказанного в настоящей работе объект исследования представляется вектором X из десяти компонент: МОС, УН, СН, УО, ОЦК1, ОЦК2, пульс, рост, вес и возраст. Результаты проведенного кластирования представлены в виде дендрограммы (рис. 1).

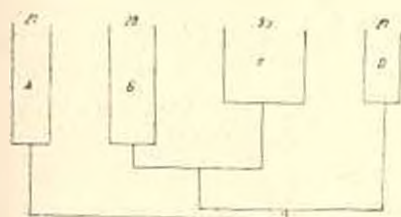


Рис. 1. Результаты проведенного иерархического среднесвязывающего кластер-анализа.

Обратимся теперь к анализу полученной дендрограммы в смысле выявления статистически значимых кластеров. На сегодняшний день не существует однозначных критериев для выявления статистически значимых кластеров, однако большинство авторов сходятся на том, что «время жизни» (промежуток от образования анализируемого кластера до его слияния с другим) и «населенность» (число объединяемых объектов) могут служить объективными критериями. На рис. 1 представлены кластеры, удовлетворяющие указанным двум критериям. Выбор для последующего анализа лишь относительно крупных кластеров обосновывается еще и тем, что, как отмечено [17], малочисленные кластеры приводят к артефактам.

Таким образом, анализ рис. 1 свидетельствует о существовании четырех—А, В, С, Д—кластеров с населенностями 21, 28, 55 и 21 соответственно. Разделимость выявленных четырех кластеров может быть оценена по данным табл. 1.

Таблица 1. Разделимость выявленных кластеров

Кластеры	А	В	С
В	14.79	—	—
С	27.24	6.35	—
Д	6.21	7.28	12.31

Примечание: приведены значения F-критерия [1].

Центры выделенных кластеров различаются статистически более, чем на 99%-ном уровне.

На следующем этапе работы была поставлена задача изучения линейной разделимости полученных кластеров в пространстве радиокардиографических параметров и получения решающих правил для отнесения нового объекта к одной из четырех выделенных групп. Поставленная задача решается в рамках модели дискриминантного анализа, применимость которого для изучения сердечно-сосудистых заболеваний нами была показана [5]. Поскольку представительность выявленных объективных групп в генеральной совокупности больных ОИМ неизвестна, то априорная вероятность появления объекта любого класса принимается равной. Решающие правила основаны на линейных дискриминантных функциях (ЛДФ). Для выявления относительной информативности того или иного параметра применен пошаговый алгоритм на основе F-критерия для ввода параметра в качестве предиктора. Вычисленный F-критерий равен 10,5, что значительно превышает $F_{30,329}$ (99%). Таким образом, полученное кластер-анализом разбиение статистически значимо более чем на 99%-ном уровне в пространстве радиокардиографических параметров.

Реклассификация анализируемого материала проводится по максимуму апостериорной вероятности. Прогностические возможности полученных соотношений оцениваются по методу скользящего контроля. Основные результаты реклассификации и прогноза приведены в табл. 2.

Таблица 2. Результаты реклассификации и прогноза

Наименование классов	Процент верной реклассификации	А	В	С	Д
А	90,5	19	0	0	2
В	79,3	0	23	4	2
С	83,9	0	9	47	0
Д	78,9	3	1	0	18
Общий	83,2	22	33	51	19
Процент верного прогноза					
А	71,4	15	0	2	4
В	62,1	1	18	7	3
С	82,1	0	9	46	1
Д	68,4	5	1	0	13
Общий	73,6	21	28	55	21

Примечание: приведено также количество объектов, реклассифицируемых или прогнозируемых в тот или иной класс.

Как видно из табл. 2, полученные средние значения верной реклассификации и прогноза (83,2 и 73,6% соответственно) достаточно велики. Наилучшая реклассификация ЛДФ достигается для класса А

(90,5%), а наихудшая—для класса *Д* (78,9%). В случае же прогноз-
стических возможностей ЛДФ картина несколько меняется. Наилуч-
ший прогноз достигается для класса *С* (82,1%), наихудший—для клас-
са *В* (62,1%).

Для геометрической интерпретации взаимного расположения вы-
деленных кластеров была получена проекция анализируемого множе-
ства на плоскость первых двух канонических переменных (плоскость
Фишера), дающую наилучшее разделение классов среди всех линейных
проекций [1] (рис. 2).

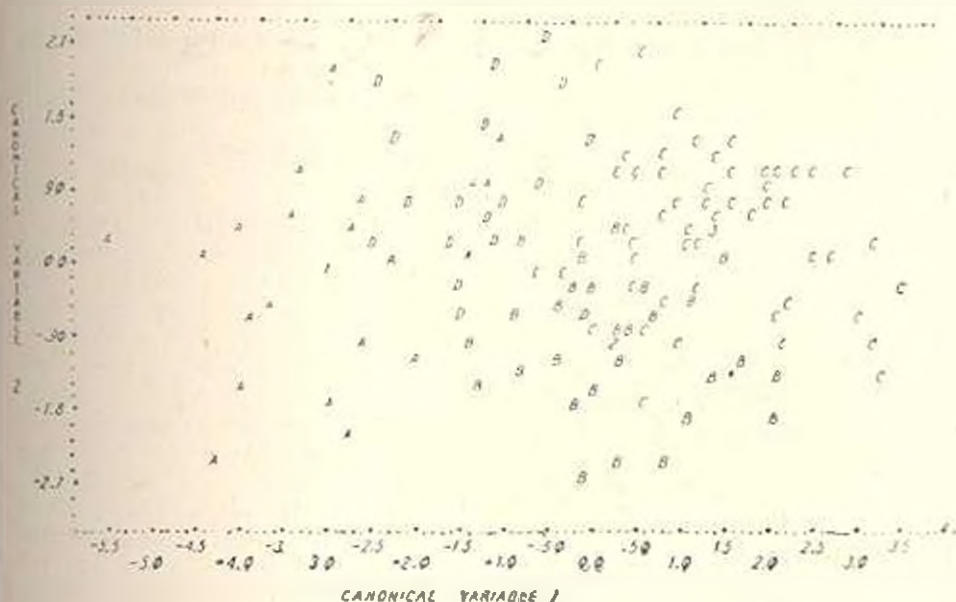


Рис. 2. Проекция анализируемого множества на плоскость Фишера.

Как и следовало ожидать (табл. 1), наиболее далеко отстоящими
классами являются *А* и *С*, тогда как *В* и *Д* занимают промежуточное
положение. Кроме того, можно заметить, что классы *В* и *С* хуже раз-
деляются, чем *А* и *Д*. Классы *А* и *В* вообще линейно не разделяемы.

После полученной плоскостной картины взаимного расположения
выделенных кластеров, т. е. перехода от линейного представления с по-
мощью древа кластеризации к плоскостному (от рис. 1 к рис. 2), воз-
можен визуальный анализ исследуемого многомерного материала.

Для численной оценки разброса объектов по классам были расчи-
таны D^2 —расстояния Махаланобиса.

Таким образом, на основании приведенного визуального и количе-
ственного анализа можно выделить три группы объектов: 1. наиболее
типичные для данного класса; 2. обладающие переходными между
классами свойствами; 3. обладающие экстремальными свойствами.

Естественно, что необходимо дифференцированное отношение к
объекту в зависимости от принадлежности его к одной из выделенных
гомогенных групп. А отнесение, как уже было отмечено, возможно до-
статочно четко на основе полученных ЛДФ.

Л И Т Е Р А Т У Р А

1. Афифи А., Эйзен С. Статистический анализ. Подход с использованием ЭВМ. М., 1982.
2. Ластей Л. Введение в проблему принятия решений и медицине. М., 1971.
3. Ордукхян А. А., Налбандян С. Г. 1-я Всесоюз. конф. «Физиология, патофизиология и фармакология мозгового кровообращения». Тез. докл., 131. Ереван, 1984.
4. Ордукхян А. А., Джугарян К. М., Степанян К. Г. В сб.: Актуальные вопросы кардиологии, 34—35, Ереван, 1982.
5. Ордукхян А. А., Джугарян К. М., Оганесян Н. М., Бабалян А. С., Степанян К. Г. 1-й Всесоюз. биофиз. съезд, 2, 175. М., 1982.
6. Розен Р. Принцип оптимальности в биологии. М., 1969.
7. Сокал Р. В сб.: Классификация и кластер. Под ред. Дж. Вэн Райзинга, 7—18. М., 1980.
8. Соломон Г. В сб.: Классификация и кластер. Под ред. Дж. Вэн Райзинга, 129—117, М., 1980.
9. Ту Дж., Гонсалес Р. Принципы распознавания образов, М., 1978.
10. Хант Э. Искусственный интеллект, М., 1978.
11. Bayne C. et al. *Patl. Rec.*, 12, 2, 51—63, 1980.
12. Bourne L. *Human conceptual behavior*, Boston, 1966.
13. Conlgham K. *Computer J.*, 15, 3, 209—213, 1972.
14. Dubes R., Jain A. *Patl. Rec.*, 8, 4, 247—261, 1976.
15. Dubes R., Sain A. *Patl. Rec.*, 11, 4, 235—255, 1979.
16. Milligan G. *Psychometrika*, 46, 2, 187—199, 1981.
17. Smith S., Dubes R. *Patl. Rec.*, 12, 3, 177—189, 1980.

Биолог. ж. Армении, т. 39, № 9, с. 768—772, 1986

УДК 577.154.5

ОЧИСТКА ПУРИННУКЛЕОЗИДФОСФОРИЛАЗЫ ПОЧЕК КРОЛИКА

А. В. АНАНЬЕВ, Х. О. БЕЗИРДЖЯН, Ж. И. АКОПЯН

Институт экспериментальной биологии АН Армянской ССР,
лаборатория молекулярной энзимологии, Ереван.

Аннотация — Выделена пуриинуклеозидфосфорилаза (КФ—2.4.2.1) из почек кролика. В результате шести стадий очистки с использованием фракционирования сульфатом аммония и колоночной хроматографии на ДЭАЭ-целлюлозе и гидроксилапатите получен электрофоретически гомогенный препарат. Фермент специфичен к нуклеозидам гипоксантина, гуанина, а также к ксантозину, но не расщепляет нуклеозиды аденина.

Անոտացիա — Կատարվել էրիկամներից անջատված է պուրինուկլեոզիդֆոսֆորիլազա (ԿՖ— 2. 4. 2. 1)։ Օգտագործելով մաքրման վեց փուլեր, որոնք ընդգրկում էև բաժանումը ամոնիումի սուլֆատով և զրոմատոգրաֆիան ԴԵԱԷ-ցելլուլոզի ու հիդրոքսիլապատիտի սյուներում, ստացված է էլեկտրաֆորետիկ հոմոգեն պատրաստով։ Ֆերմենտը սպեցիֆիկ է հիպոքսանտինի և գուանինի նուկլեոզիդներին, ինչպես նաև քսանտոզինի նկատմամբ, բայց չի տրոհում ադենինի նուկլեոզիդները։

Abstract — Purine nucleoside phosphorylase (EC—2.4.2.1) from rabbit kidney has been isolated. The six-step purification procedure by $(\text{NH}_4)_2\text{SO}_4$