

«Санкт-Петербургский Государственный Университет»
Факультет Прикладной математики-Процессов управления
Кафедра Диагностики функциональных систем

Харитонов Андрей Андреевич

Выпускная квалификационная работа бакалавра

Дискриминантный анализ иммунных статусов онкологических
больных.

Направление 01.03.02

Прикладная математика и информатика

Научный руководитель,
старший преподаватель
Орехов А.В.

Санкт-Петербург
2019

Содержание

Введение	3
Постановка задачи	5
Глава 1. Дискриминантный анализ и сфера его использования	6
Глава 2. Дискриминантный анализ иммунных статусов он- кобольных	14
Заключение	20
Список литературы	21
Приложение	22

Введение

В настоящее время медицина в своих исследованиях стала все чаще обращаться к достижениям математических наук, с помощью которых изучаются многообразные связи основных клинических параметров и их комбинаций.

Статистическое моделирование основных жизненных функций в норме и в патологии имеет большое теоретическое и практическое значение. Оно позволяет объяснить сущность патологического процесса, выявить взаимосвязи между клиническими, инструментальными и лабораторными показателями, а также правильно оценить прогноз заболевания и разработать эффективные лечебно-профилактические мероприятия.

Как правило в медицине используют качественную оценку результатов лечения, которая может быть положительной либо отрицательной. Но не всегда возможно адекватно сравнить результаты клинических исследований новых препаратов.

Например, в онкологии разработана клиническая классификация степени ответа опухоли, которая основана на изменении линейных размеров новообразования, чаще определяемых лучевыми методами исследования. Сама классификация была принята Всемирной организацией здравоохранения (ВОЗ) в 1981 году, однако в 2000 году она была пересмотрена и представлена как критерий оценки ответа солидных опухолей, в которой стали использовать двухмерные измерения вместо трехмерных и уменьшили количество измеряемых очагов.

Критерии оценки ответа опухолей на лечение (response evaluation criteria in solid tumors — RECIST) разработаны Европейской организацией по исследованию и лечению рака (European Organisation for Research and Treatment of Cancer — EORTC), Национальным институтом злокачественных новообразований США (National Cancer Institute — NCI), а также Национальным онкологическим институтом Канады (National Cancer Institute of Canada) и впервые внедрены в практику в 2000 году. В настоящее время актуальной является вторая редакция — RECIST 1.1, выпущенная в 2009 году, которая

отличается от первоначальных версий как граничными значениями, так и количественными показателями. В то же время увеличилось и количество оцениваемых параметров, их значимость. Игнорирование формальных признаков неэффективности по критериям RECIST позволило продолжить эффективное лечение у ряда пациентов с формальным прогрессированием (псевдо-прогрессирование). Это наиболее стандартизированный и практически удобный метод для оценки ответа и прогрессирования в разделе лечения солидных опухолей.

Адаптация критериев RECIST является новым шагом на пути к оптимизации подхода к оценке эффективности лечения пациентов при иммунотерапии. Существенно сократилось число случаев раннего и необоснованного прекращения лечения пациентов современными препаратами при правильной оценке эффективности иммунотерапевтических препаратов и при смене режимов терапии на менее эффективную и более токсичную программу химиотерапии. Но следует понимать, что широкое применение отграничено отсутствием адаптационного коммерческого программного обеспечения. С другой стороны, критерий RECIST уже входит в практику ряда медицинских учреждений, что дает возможность надеяться на положительные перспективы в клинической онкологии.

Постановка задачи

Цель данного исследования заключается в анализе статистически важных данных среди различных показателей иммунного статуса у больных онкологическими заболеваниями, а также поиск коэффициентов позволяющих оценивать дальнейшее состояние у больных данным заболеванием на основании дискриминантного анализа.

Глава 1. Дискриминантный анализ и сфера его использования

Многомерные статистические методы могут использоваться для однозначного определения информативных признаков, которые характеризуют тот или иной диагностический процесс на основе законов распределения и значений вероятностей присутствия, и, соответственно, построения адекватной модели [1]. Дискриминантный анализ (далее Д. А.) — это метод, который используется для анализа массива данных, когда критерий или зависимая переменная являются категориальными (например, принимают значения: «да» или «нет»), а предиктор или независимая переменная имеет интервальный характер. Термин категориальная переменная означает, что зависимая переменная состоит из нескольких категорий.

Дискриминантный анализ — это раздел математической статистики, содержанием которого является разработка методов решения задач различения (дискриминации) объектов наблюдения по определенным признакам [2].

Роналд Фишер — основоположник дискриминантного анализа, один из представителей школы английской биометрии и создатель дисперсионного анализа. Он ясно понимал, что остаточная (внутригрупповая) дисперсия, связанная с естественным разнообразием биологических объектов, служат препятствием для объективного сопоставления сравниваемых групп. Единственно возможный способ ликвидировать эффекты внутригрупповой изменчивости — её искусственное апостериорное уменьшение, в идеале сведение к нулю, то есть минимизация.

Теоретическая база дискриминантного анализа была заложена Р. Фишером и в дальнейшем получила свое развитие в трудах Х. Аренса и Ю. Лейтера [3], А. Аффифи и С. Эйзена [4], М. Кендалла и А. Стьюарта [5], С.А. Айвазяна с соавторами [6]. Все они подчёркнули, что если наблюдения относятся к нескольким генеральным совокупностям, то для их классификации представляется разумным найти линейную комбинацию признаков, которая называется дискриминантной функцией, в которую исходные признаки

входят с некоторыми постоянными величинами — коэффициентами.

Дискриминация (от лат. *discriminatio*) означает различение, разделение. Основной задачей дискриминантного анализа является различие групп объектов и изучение характера этих групп. Задача дискриминантного анализа — это нахождение значений коэффициентов, которые снижают вероятность ошибочной классификации.

Непосредственно сам анализ принадлежит к многомерным методам статистического анализа, предполагающих описание объектов комплексом признаков. Высокая разрешающая способность таких методов характеризуется не только принципом «больше учтённых признаков — больше информации». Не менее важно, что эти методы могут учитывать концепцию корреляционных свойств. Как и многие другие многомерные методы, дискриминантный анализ базируется на построении линейных комбинаций признаков, то есть функций, в которые каждый из них входит со своим коэффициентом (вкладом). В дискриминантном анализе линейные комбинации называются соответственно дискриминантными функциями:

$$DF = b_1x_1 + \dots + b_ix_i + \dots + b_px_p + C,$$

где DF — значение дискриминантной функции, x_i — численное значение i -го признака, b_i — вклад i -го признака в значение функции, p — число признаков.

Дискриминантный анализ обеспечивает объективное сравнение (разделение) групп за счёт искусственной минимизации внутригруппового разнообразия (дисперсии).

Применяя дискриминантный анализ, можно не только оценивать достоверность межгрупповых различий и оценить «расстояния» между группами, также возможно определить те признаки из числа учтённых, которые в первую очередь обуславливают межгрупповые различия. Более того, когда две (или большее число) категории ранее поделены в дискриминантном анализе, возможно определить принадлежность неизвестного объекта к одной из них. Исходные данные для Д. А. могут выглядеть как совокупность объектов, которые разделены на несколько групп/классов. Число признаков не может быть

выше суммарного числа объектов, по которым описаны объекты и не должно быть меньше двух.

Признаки должны быть количественными, а распределение их значений в каждом классе — нормальным. Включение в комплекс признаков с единичной (полной) корреляцией недопустимо.

Описание объекта по комплексу из p признаков геометрически равнозначно определению его координат в p -мерном пространстве. Единичному объекту в пространстве может соответствовать точка, группе объектов — «облако точек». Два или более «облака точек», отвечающих данным группам объектов, нередко перекрываются, то есть неоднозначно различимы. Практически всегда есть особи с такой ординацией в пространстве, которая не позволяет точно отнести их к одной или другой группе. Цель дискриминантного анализа геометрически формулируется как построение нового пространства, в котором принадлежность объектов к группам определяется конкретно — перекрывание «облаков точек» происходит минимальным.

Координаты объектов в новом ортогональном пространстве определяются значениями дискриминантных функций (DF) — специфических линейных комбинаций признаков, которые находятся на основе двух корреляционных матриц: межгрупповой (корреляции групповых средних) и внутригрупповой (корреляции признаков особей в группах). Выбор таких осей производится по критерию максимума отношения межгрупповой дисперсии к 7 внутригрупповой. Данный математически выбор основан на сравнении матрицы корреляции групповых средних с матрицей корреляции тех же признаков внутри групп.

В дискриминантном анализе различают два типа коэффициентов: это нестандартизованные и стандартизованные. Посредством коэффициентов первого типа можно получить уравнение для каждого объекта, которое позволяет вычислить собственно значение дискриминантной функции. В данное уравнение помимо значений признаков, умноженных на значения коэффициентов, также входит константа. Нормированные коэффициенты приводят к дискриминантным значениям, измеряемым в единицах стандартного квадратического отклонения, то есть можно определить на сколько сигм кон-

кретный объект отклоняется от главного центроида.

Стандартизованные коэффициенты позволяют судить об относительном вкладе конкретной переменной в дискриминантную функцию. То есть анализ абсолютных величин стандартизованных коэффициентов может дать ответ на вопрос, какова же роль каждого признака в межгрупповых различиях. Если абсолютная величина коэффициента для данной переменной у всех значимых дискриминантных функций мала, то эту переменную можно исключить и таким образом сократить число переменных. Данная процедура называется «определение информативного списка признаков».

В качестве основных результатов дискриминантного анализа называют и вычисление структурных коэффициентов. Они представляют собой коэффициенты корреляции между конкретными переменными и дискриминантными функциями, когда абсолютная величина такого коэффициента велика для конкретного признака, то чаще всего считают, что вся информация о дискриминантной функции заключена в данной переменной. Это позволяет интерпретировать дискриминантные функции в соответствии с ролью наиболее тесно скоррелированных с ней переменных.

Можно достаточно точно выявить факторы, которые оказывают существенное влияние на конечный показатель, а также определить силу и направление этого влияния, используя методы регрессионного анализа. Но сами синтезированные модели могут оказаться неточными или, в случае сложных моделей, переобученными, что также может снижать эффективность их использования. Применение методов кластерного анализа позволяет сопоставить объекты моделирования по их качественным характеристикам, агрегировать экспертные оценки текущего и прогнозируемого уровней развития объектов. Использование данного анализа требует значительной затраты сил и времени для проведения нужных расчетов, а сами полученные результаты группирования не всегда могут адекватно интерпретироваться, особенно если эти информативные признаки, которые характеризуют объекты исследования, были достаточно разнородны.

Указанные выше недостатки позволяют учесть Д. А., который дает возможность быстро и качественно классифицировать исследу-

емые объекты и построить адекватную линейную математическую модель. Данный анализ имеет ряд преимуществ, таких как: учитывает изменчивости параметра, рассматривает совокупности всех клинических и параклинических показателей, взятых со своими коэффициентами и отражающие удельный вес влияния каждого показателя на постановку диагноза.

Цель Д. А. — это построение дискриминантных функций, которые представляют собой не что иное, как линейную комбинацию предикторов, которая будет идеально различать категории зависимой переменной. Это позволяет изучить, имеются ли существенные различия между группами с точки зрения зависимых переменных. Он также оценивает точность классификации.

Перед тем как приступить непосредственно к самой процедуре классификации, сперва необходимо определить границу, разделяющую в частном случае две рассматриваемые группы.

Классификация — это разделение данной совокупности объектов или явлений на однородные в определенном смысле группы.

Различают:

- классификацию при наличии обучающих выборок (дискриминантный анализ);
- классификацию без обучения (автоматическая классификация — кластерный анализ).

При решении задачи классификации, то есть отнесении неизвестных объектов в одну из известных групп, применяются классифицирующие функции, представляющие собой уравнения, составленные для каждой группы. Неизвестный объект относится к классу, у которого в результате решения каждого из уравнений значение классифицирующей функции оказывается наибольшим.

Д. А. является одним из методов многомерного статистического анализа. Цель анализа — на основе измерения данных характеристик (признаков, параметров) объекта классифицировать его, то есть отнести к одной из представленных групп (классов) некоторым оптимальным способом. Под оптимальным способом понимается либо минимум математического ожидания потерь, либо минимум вероятности ложной классификации. Этот вид анализа является многомер-

ным, так как измеряется несколько параметров объекта, по крайней мере, больше одного, например, температура, влажность в технологическом процессе, давление, состав крови, температура, общее состояние больного и т.д.

Широкий круг задач, возникающих на практике и связанных с классификацией, можно решить методами Д. А.

Дискриминантный анализ описывается количеством категорий, которыми обладает зависимая переменная. Как и в статистике, все предполагается вплоть до бесконечности, поэтому в этом случае, когда зависимая переменная имеет две категории, тогда используется тип Д. А. для двух групп. Если зависимая переменная имеет три или более трех категорий, то используется тип множественного дискриминантного анализа.

Основное различие между типами Д. А. заключается в том, что для двух групп можно вывести только одну дискриминантную функцию. С другой стороны, в случае множественного дискриминантного анализа может быть вычислено более одной дискриминантной функции.

Рассмотрим виды Д. А.:

1) Пошаговый анализ с включением — в пошаговом анализе дискриминантных функций модель дискриминации строится по шагам. То есть, на каждом шаге просматриваются все предикторы и находится те из них, приносящая наибольший вклад в различие между совокупностями. Эта переменная должна быть включена в модель на данном шаге, и происходит переход к следующему шагу.

2) Пошаговый анализ с исключением — можно также двигаться в обратном направлении. В этом случае все предикторы будут сначала включены в модель, а затем на каждом шаге будут устраняться переменные, которые вносят малый вклад в предсказания. Тогда в качестве результата успешного анализа можно сохранить только важные переменные в модели, то есть те переменные, чей вклад в дискриминацию больше остальных.

3) Расчет на случай — пошаговый Д. А., основанный на использовании статистического уровня значимости. Поэтому по своей природе пошаговые процедуры рассчитывают на случай, так как они

«тщательно перебирают» переменные, включенные в модель для получения максимальной дискриминации. При использовании пошагового метода можно заметить, что используемый при этом уровень значимости не отражает истинного значения альфа, то есть, вероятности ошибочного отклонения гипотезы H_0 (нулевой гипотезы, которая заключается в том, что между совокупностями нет различия).

Есть ряд предположений, при которых Д. А. «работает»:

1) Предположение о том, что наблюдаемые величины (измеряемые характеристики объекта) имеют нормальное распределение. Это предположение нужно проверять. Следует заметить, что умеренные отклонения от этого предположения не являются неизбежными;

2) Предположение об однородности дисперсий анализируемых переменных в разных классах (различия между классами имеется только в средних). Умеренные отклонения от этого предположения также допускаются. Выдвигаются предположения:

1) имеются разные классы объектов;

2) каждый класс имеет нормальную функцию плотности $f(\mu(i), Y)$ от k переменных; где $\mu(i)$ — вектор математических ожиданий переменных размерности k , Y — ковариационная матрица.

В случае, если параметры известны, дискриминацию можно провести таким образом.

Имеются функции плотности нормально распределённых классов. Задана точка в пространстве k измерений. Предполагая, что имеет наибольшую плотность, необходимо отнести точку к i -му классу. Имеется доказательство, что если априорные вероятности для определяемых точек каждого класса одинаковы и потери при неправильной классификации i -й группы в качестве j -й не зависят от i и j , то решающая процедура снижает потери при неправильной классификации.

Методы, которые реализуются в практических задачах, являются линейными. Функции классификации и дискриминантные функции также являются линейными комбинациями исходных величин.

Необходимо сделать важное замечание о проверке гипотез. Дискриминантный анализ может быть проведен, когда основные положения не выполняются (положение о нормальности и однородности

дисперсий). Задача состоит в интерпретации результатов. В конечном счете практика является наиболее важным критерием правильности построенного классификатора. Если окажется, что в результате построен классификатор, который «работает» на практике, то это будет показателем успеха.

Глава 2. Дискриминантный анализ иммунных статусов онкобольных

Анализ взаимосвязи между иммунологическими параметрами и течением заболевания проводился на клиническом материале, предоставленном научным отделом онкоиммунологии НИИ онкологии им. Н. Н. Петрова. У пациентов на распространенных стадиях заболевания, проходивших лечение, проводилась оценка ответа солидных опухолей.

Оценка иммунной системы проводилась по следующим показателям иммунного статуса:

1. CD4/CD8 — отношение клеток регуляторов иммунного ответа к исполнительным клеткам иммунитета.

2. Лимфоциты — главные клетки иммунной системы, обеспечивают гуморальный иммунитет

3. НКТ-клетки — представляют собой субпопуляцию лимфоцитов, экспрессирующих как маркеры НК-клеток, так и Т-клеточные дифференцированные антигены.

4. Т-лимфоциты — обеспечивают распознавание и уничтожение клеток, несущих чужеродные антигены.

5. Активированные Т-хелперы — главной функцией является усиление адаптивного иммунного ответа.

6. Лейкоциты — защита организма от инфекций, чужеродных белков и инородных тел, способных нанести ему вред.

7. Treg — контролировать силу и продолжительность иммунного ответа.

8. Активированные CTLs — непосредственно контактируют с поврежденными клетками и разрушают их.

Данные переменные являются предикторами. RECIST же в свою очередь является зависимой переменной. Группа пациентов была разбита на тестовую и генеральную выборки вслепую. При отсутствии данных по какому-либо показателю было принято решение не брать данный показатель при построении анализа. Количество таких отсутствующих показателей было незначительным относительно всех наблюдений, потому будем считать, что это не оказало суще-

ственного влияния на корреляционные матрицы. Экспериментальную группу составляли 43 пациента, в то время как контрольную группу 117 человек.

Как было уже ранее сказано, целью дискриминантного анализа является построение дискриминантной функции. Для ее построения необходимо провести промежуточные вычисления.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \\ x_{13} \end{pmatrix},$$

где x_1 — НКТ-клетки, x_2 — CD4/CD8, x_3 — лимфоциты, x_4 — лимфоциты%, x_5 — CD4/CD8 к Т-лимфоцитам, x_6 — CD4/CD8 к Т-лимфоцитам%, x_7 — Т-хелперы% от лимфоцитов, x_8 — Т-хелперы, x_9 — лейкоциты, x_{10} — Treg%, x_{11} — активированные Т-хелперы% от лимфоцитов, x_{12} — Treg, x_{13} — активированные CTLs.

$$\bar{\mathbf{x}}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \mathbf{x}_{ij}$$

где \mathbf{x}_{ij} — j-ый вектор i-го класса, $\bar{\mathbf{x}}_i$ — вектор выборочных средних i-го класса, N_i количество объектов в i-ом классе.

Вычислить матрицу ковариаций

$$\Sigma = \frac{1}{N} \left(\sum_{i=1}^k \sum_{j=1}^{N_i} (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)(\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)^T \right),$$

где $N = \sum_{i=1}^k N_i$ — количество всех объектов во всех классах,
 k — количество классов.

Линейная дискриминантная функция в общем виде выглядит следующим образом:

$$g_i(x) = \mathbf{x}^T \Sigma^{-1} \cdot \bar{\mathbf{x}}_i - 1/2 \cdot \bar{\mathbf{x}}_i^T \cdot \Sigma^{-1} \cdot \bar{\mathbf{x}}_i + \ln(\pi_k)$$

π_k — априорная вероятность класса (например, частота встречаемости в обучающей выборке).

Σ^{-1} — обратная матрица ковариаций.

$\bar{\mathbf{x}}_i$ — вектор выборочных средних i -го класса.

В итоге, по полученным данным, вычисляются «опорные точки» по которым строятся дискриминантные гиперплоскости:

$$g_1 = x^T \begin{vmatrix} 0.6209 \\ -0.0012 \\ -0.0041 \\ -0.0788 \\ 3.1481 \\ -0.0413 \\ -0.7575 \\ -0.008 \\ 0.002 \\ -0.0021 \\ 1.4521 \\ -0.0093 \\ 3.8633 \end{vmatrix} + 2.511 - 0.2293$$

и

$$g_2 = x^T \begin{pmatrix} 0.0437 \\ 0.002 \\ 0.0013 \\ -0.0239 \\ 3.3345 \\ 0.0003 \\ 0.6376 \\ -0.0091 \\ 0.0023 \\ -0.0009 \\ -2.6557 \\ 0.008 \\ -0.8728 \end{pmatrix} + 2.2698 - 0.591$$

Для программной реализации была выбрана среда обработки данных STATISTICA версии 10, разработанный компанией StatSoft, в настоящее время принадлежащей компании Dell, как наиболее подходящая для решения задач математической статистики [7]. А также пакет программ Microsoft Excel Версии 2016 года, для первоначальной обработки данных. Результаты анализов пациентов были представлены в виде таблиц. С образцом таблицы можно ознакомиться в приложении Г. Обработка анализов пациентов проводилась следующим образом:

- 1) В первую очередь нужно было убрать повторения. Для данного действия была использована встроенная функция удаления дубликатов;
- 2) Изменение вида таблицы, так как первоначальный вид не позволял произвести нужные расчеты;
- 3) Анализ данных и исключение выбросов;
- 4) Переход к нужному виду при помощи сводных таблиц, которые позволяют соотнести пациента с нужным анализом.

Таким образом был проведен первичный анализ данных и перенос таблицы в пакет программ STATISTICA [8]. Во вкладке "Анализ" находим пункт "Многомерный анализ" и в открывшемся поле выбираем графу "Общие модели дискриминантного анализа". В дан-

ном окне выбирается зависимая переменная и предикторы. После подтверждения появляется окно с результатами нашего исследования. Основными пунктами в проведенном анализе будут "Обратная матрица ковариации" и "Исходные коэф. канонической дискриминантной функции". Полученные значения переносятся в Microsoft Excel и производятся расчеты. На основе данных в таблицах была построена дискриминантная функция.

Дискриминантный анализ позволяет ответить на два существенных вопроса:

1. Надежно ли определены параметры модели (проверка гипотезы о значимости параметров модели);
2. Хорошо ли описывает модель имеющиеся экспериментальные данные (проверка гипотезы об адекватности модели).

Нарушение надежности модели может проявиться и в более скрытой форме, а именно через:

- наличие выбросов;
- наличие влиятельных наблюдений;
- нарушение основных предположений дискриминантного анализа.

Как было сказано ранее, группа пациентов была разбита на тестовую и генеральную выборки вслепую. Из полученных данных экспериментальной выборки можно заметить, что 72,1% указали на возможность правильных предсказаний распределения пациентов по трем группам, в то время как контрольной группа показала 68,4%, что является не удовлетворительным результатом.

Далее нужно построить дискриминантную функцию. Данная функция строится по формуле:

Можно заметить, что велика вероятность ошибки при использовании данного анализа. Сложней всего определить пациентов относящихся к 2 и 3 группам.

Итоги распределения по трем группам в экспериментальной и контрольной группах указаны в приложении А и Б.

Из поставленных цели и задач по результатам проведенного исследования, можно сделать следующие выводы:

Класс	Доля Правил.	1 p=,4419	2 p=,3488	3 p=,2093
1	84,21053	16,00000	2,00000	1,000000
2	60,00000	5,00000	9,00000	1,000000
3	66,66667	2,00000	1,00000	6,000000
Всего	72,09302	23,00000	12,00000	8,000000

Рис. 1.
Результаты экспериментальной группы пациентов

Класс	Доля Правил.	1 p=,5897	2 p=,2564	3 p=,1538
1	94,20290	65,00000	3,00000	1,000000
2	33,33333	18,00000	10,00000	2,000000
3	27,77778	10,00000	3,00000	5,000000
Всего	68,37607	93,00000	16,00000	8,000000

Рис. 2.
Результаты контрольной группы пациентов

1. В ходе исследования стало очевидно, что дискриминантный анализ не подходит для подобных исследований, так как сил требуется много, а результат себя не оправдывает;
2. Недостаточность анализов пациентов приводит к невозможности составления общего портрета заболевания;
3. Поведение отдельно взятого организма индивидуально и может искажать общую картину исследования. Кроме того, пациенты проходили разное лечение, что тоже могло иметь значительное влияние на результаты анализов.

Заключение

Система иммунитета играет важнейшую роль в сохранении генетической целостности организма. За последние годы стала ясна сложность ее структуры. Вместе с тем значение иммунитета при различных заболеваниях, в том числе онкологической патологии приобретает практическое значение.

Последние годы отмечается рост интереса к иммунотерапии злокачественных опухолей благодаря новым достижениям в этой области медицины. Основной принцип иммунотерапии онкологических заболеваний состоит в том, что, как и при любых других заболеваниях, активизировать собственные защитные ресурсы организма и направить их на борьбу с опухолью. Широкое применение иммунотерапия получила при лечении меланомы кожи.

Сложность системы иммунитета потребовала привлечение различных математических подходов – статистических, матмоделирования и пр. Распространенность этих исследований настолько широка, что в настоящее время предполагается введение термина «математическая иммунология».

При иммунном ответе организма на патогенное воздействие большое значение имеют показатели иммунного статуса. Представленная модель описывает реакцию пациента на новые способы лечения онкологических заболеваний.

Вместе с тем, не вполне ясны вопросы применения различных математических подходов. Это может быть связано, в том числе, с нерешенностью определенных аспектов математических методов. В данном исследовании продемонстрирована недостаточность дискриминантного анализа при изучении системы иммунитета у онкологических больных. Кратко дана оценка причины этого и возможность улучшения результатов.

Список литературы

- [1] Боровиков В.П. Популярное введение в программу STATISTICA. М.: Компьютер-Пресс, 1998. 267 с.
- [2] Многомерные статистические методы. Дубров А.М., Мхитарян В.С., Трошин Л.И. Москва. 2003г.
- [3] Аренс Х., Лейтер Ю. Многомерный дисперсионный анализ. М.: Финансы и статистика, 1985. 230 с
- [4] Афифи А., Эйзен С. Статистический анализ: Подход с использованием ЭВМ. М.: Мир, 1982. 488 с.
- [5] Кендалл М., Стьюарт А. Многомерный статистический анализ и временные ряды. М.: Наука, 1976. 736 с.
- [6] Айвазян С.А. [и др.]. Прикладная статистика: Классификация и снижение размерности: Справ. изд. / под ред. С.А. Айвазяна. М.: Финансы и статистика, 1989. 607 с.
- [7] Боровиков В.П., Боровиков И.П. STATISTICA – Статистический анализ и обработка данных в среде Windows. М.: Филинь, 1997. 608 с.
- [8] Ким Дж. Щ. Факторный, дискриминантный и кластерный анализ 22 / Дж. Щ. Ким, Ч.У. Мюллер, У.Р. Клекка и др.; под ред. И.С. Енюкова. – М.: Финансы и статистика, 1989. – 215 с.

Приложение

Приложение А

Пациент 1	1
Пациент 2	3
Пациент 3	2
Пациент 4	2
Пациент 5	3
Пациент 6	2
Пациент 7	3
Пациент 8	1
Пациент 9	1
Пациент 10	3
Пациент 11	2
Пациент 12	2
Пациент 13	1
Пациент 14	3
Пациент 15	1
Пациент 16	3
Пациент 17	2
Пациент 18	2
Пациент 19	3
Пациент 20	1
Пациент 21	2
Пациент 22	1
Пациент 23	1
Пациент 24	1
Пациент 25	2
Пациент 26	2
Пациент 27	1
Пациент 28	1
Пациент 29	1
Пациент 30	1
Пациент 31	3
Пациент 32	1
Пациент 33	1
Пациент 34	1
Пациент 35	2
Пациент 36	2
Пациент 37	2
Пациент 38	1
Пациент 39	2
Пациент 40	2
Пациент 41	3
Пациент 42	1
Пациент 43	1

Рис. 3.
Результаты исследования экспериментальной группы

Приложение В

Эффект	Корреляционная матрица эффектов Сигма-ограниченная параметризация Декомпозиция гипотезы			
	Уровень	1	2	3
Св.член	1	1,00000	1,00000	-1,00000
	2	1,00000	1,00000	-1,00000
	3	-1,00000	-1,00000	1,00000
Сумма по полю НКТ-клетки	1	1,00000	1,00000	-1,00000
	2	1,00000	1,00000	-1,00000
	3	-1,00000	-1,00000	1,00000
Сумма по полю CD4 CD8	1	1,00000	1,00000	-1,00000
	2	1,00000	1,00000	-1,00000
	3	-1,00000	-1,00000	1,00000
Сумма по полю Лимфоциты	1	1,00000	-1,00000	1,00000
	2	-1,00000	1,00000	-1,00000
	3	1,00000	-1,00000	1,00000
Сумма по полю Лимфоциты, %	1	1,00000	-1,00000	-1,00000
	2	-1,00000	1,00000	1,00000
	3	-1,00000	1,00000	1,00000
Сумма по полю Дв. Поп. CD4/CD8 Т-лимфоциты %	1	1,00000	1,00000	-1,00000
	2	1,00000	1,00000	-1,00000
	3	-1,00000	-1,00000	1,00000
Сумма по полю Дв. Поп. CD4/CD8 Т-лимфоциты	1	1,00000	-1,00000	1,00000
	2	-1,00000	1,00000	-1,00000
	3	1,00000	-1,00000	1,00000

Рис. 5.
Корреляционная матрица эффектов

Сумма по полю Активированные HLA-DR Т-хелперы, % от лимфоцитов	1	1,00000	1,00000	-1,00000
	2	1,00000	1,00000	-1,00000
	3	-1,00000	-1,00000	1,00000
Сумма по полю Активированные HLA_DR Т-хелперы	1	1,00000	1,00000	-1,00000
	2	1,00000	1,00000	-1,00000
	3	-1,00000	-1,00000	1,00000
Сумма по полю Лейкоциты	1	1,00000	-1,00000	-1,00000
	2	-1,00000	1,00000	1,00000
	3	-1,00000	1,00000	1,00000
Сумма по полю Treg, %	1	1,00000	-1,00000	1,00000
	2	-1,00000	1,00000	-1,00000
	3	1,00000	-1,00000	1,00000
Сумма по полю Активированные Т-хелперы, % (CD25 от лимф)	1	1,00000	-1,00000	-1,00000
	2	-1,00000	1,00000	1,00000
	3	-1,00000	1,00000	1,00000
Сумма по полю Treg	1	1,00000	-1,00000	1,00000
	2	-1,00000	1,00000	-1,00000
	3	1,00000	-1,00000	1,00000
Сумма по полю Активированные Т-хелперы (CD25+ от Th), %	1	1,00000	1,00000	-1,00000
	2	1,00000	1,00000	-1,00000
	3	-1,00000	-1,00000	1,00000
Сумма по полю HLA DR Активированные CTLs	1	1,00000	-1,00000	-1,00000
	2	-1,00000	1,00000	1,00000
	3	-1,00000	1,00000	1,00000

Рис. 6.
Корреляционная матрица эффектов

Таблица 1.
Приложение Г

ФИО	Тест	Текущий	RECIST
Пациент 1	НКТ-клетки	0,08	1
Пациент 1	Т-лимфоциты	1,13	1
Пациент 1	Цитотоксические лимфоциты	0,51	1
Пациент 1	HLA DR	0,06	1
Пациент 1	НКТ/цитотоксические Т-лимфоциты	0,15	1
Пациент 1	В-лимфоциты	0,14	1
Пациент 1	HLA DR	0,06	1
Пациент 1	Иммунорегуляторный индекс (CD4/CD8)	1,15	1
Пациент 1	В-лимфоциты	9,6	1
Пациент 1	В-лимфоциты	0,14	1
Пациент 1	HLA DR	0,06	1
Пациент 2	Лимфоциты	3,1	2
Пациент 2	Иммуноглобулин G	24	2
Пациент 2	Миграция	54	2
Пациент 2	IgA	6,4	2
Пациент 2	CD16	11	2
Пациент 2	IgA	6,4	2
Пациент 2	Миграция	63	2
Пациент 2	Моноциты	1,2	2
Пациент 2	Т-лимфоциты	2,39	2
Пациент 2	CD14	0,17	2
Пациент 2	Моноциты	1,2	2
Пациент 2	CD14	0,17	2
Пациент 2	CD16	0,34	2
Пациент 2	HLA DR	0,39	2
Пациент 2	В-лимфоциты	0,22	2
Пациент 2	CD14	0,17	2
Пациент 2	CD16	0,34	2
Пациент 2	Цитотоксические лимфоциты	0,9	2

ФИО	Тест	Текущий	RECIST
Пациент 2	В-лимфоциты	7	2
Пациент 2	Т-хелперы	1,27	2
Пациент 2	Иммунорегуляторный индекс (CD4/CD8)	1,4	2
Пациент 2	Т-хелперы	1,27	2
Пациент 3	Базофилы	0,03	0
Пациент 3	Цитотоксические лимфоциты	1,69	0
Пациент 3	HLA DR	0,63	0
Пациент 3	IgA	2,20	0
Пациент 3	Иммуноглобулин G	14,00	0
Пациент 3	Базофилы	0,03	0
Пациент 3	HLA DR	0,63	0
Пациент 3	Миграция	37,00	0
Пациент 3	Миграция	53,00	0
Пациент 3	Т-лимфоциты	3,25	0
Пациент 3	IgA	2,20	0
Пациент 3	Т-лимфоциты	3,25	0
Пациент 3	В-лимфоциты	0,82	0
Пациент 3	Цитотоксические лимфоциты	1,69	0
Пациент 3	В-лимфоциты	0,82	0
Пациент 3	CD14	0,53	0
Пациент 3	CD16	1,46	0
Пациент 3	Цитотоксические лимфоциты	1,69	0
Пациент 3	Иммуноглобулин G	14,00	0
Пациент 3	В-лимфоциты	0,82	0
Пациент 3	CD38	30,00	0
Пациент 3	Иммунорегуляторный индекс (CD4/CD8)	0,86	0
Пациент 3	CD71	0,00	0
Пациент 3	В-лимфоциты	0,82	0
Пациент 3	Цитотоксические лимфоциты	1,69	0

ФИО	Тест	Текущий	RECIST
Пациент 4	Т-лимфоциты	1,63	3
Пациент 4	HLA DR	0,49	3
Пациент 4	IgA	2,72	3
Пациент 4	Иммуноглобулин G	20,00	3
Пациент 4	Миграция	67,00	3
Пациент 4	плазмациты	0,00	3
Пациент 4	Т-лимфоциты	1,63	3
Пациент 4	CD16	0,28	3
Пациент 4	Цитотоксические лимфоциты	0,56	3
Пациент 4	Миграция	39,00	3
Пациент 4	В-лимфоциты	0,13	3
Пациент 4	CD16	0,28	3
Пациент 4	Цитотоксические лимфоциты	0,56	3
Пациент 4	В-лимфоциты	0,13	3
Пациент 4	CD16	0,28	3
Пациент 4	Иммунорегуляторный индекс (CD4/CD8)	2,80	3
Пациент 4	CD71	0,00	3
Пациент 4	CD14	0,01	3
Пациент 4	Цитотоксические лимфоциты	0,56	3

Приложение Д

0.036	0.000688	-0.000365	-0.002914	0.0563	0.001307	0.00473	-0.003335	-0.002166	-0.00258	-0.096083	0.000581	0.054617
0.001	0.000041	0.000022	-0.000003	-0.00046	-0.000059	0.000954	-0.000187	-0.00003	-0.000033	-0.004728	0.000061	-0.000677
0	0.000022	0.000116	-0.000175	-0.0005	-0.000085	0.001513	-0.000152	-0.00017	0.000039	-0.002534	0.000054	-0.003719
-0.003	-0.000003	-0.000175	0.021588	-0.16946	-0.001582	0.003665	-0.00016	0.000196	0.000416	-0.01862	-0.000304	-0.023767
0.056	-0.000464	-0.000498	-0.169456	13.86587	0.021743	0.539477	0.001539	-0.030033	0.010919	0.737189	0.007174	-0.808115
0.001	-0.000059	-0.000085	-0.001582	0.02174	0.006478	-0.010877	0.000635	-0.001175	-0.000363	0.026165	-0.000398	-0.061037
0.005	0.000954	0.001513	0.003565	0.53948	-0.010877	1.89641	-0.025905	0.004659	0.002448	-0.6622	0.002808	-0.789548
-0.003	-0.000187	-0.000152	-0.00016	0.00164	0.000635	-0.025905	0.002757	-0.00091	-0.000153	0.020945	-0.000133	-0.012117
-0.002	-0.00003	-0.00017	0.000196	-0.03003	-0.001175	0.004659	-0.00091	0.003316	0.000124	-0.016205	-0.000682	0.014036
-0.003	-0.000033	0.000039	0.000416	0.01092	-0.000383	0.002448	-0.000193	0.000124	0.001958	0.011589	-0.00045	-0.006861
-0.096	-0.004728	-0.002534	-0.01862	0.73719	0.026165	-0.6622	0.020945	-0.016205	0.011569	6.021314	-0.012057	-0.12651
0.001	0.000061	0.000054	-0.000304	0.00717	-0.000398	0.002808	-0.000133	-0.000582	-0.00045	-0.012057	0.000739	-0.009003
0.055	-0.000677	-0.003719	-0.023767	-0.80811	-0.061037	-0.789548	-0.012117	0.014036	-0.006861	-0.12651	-0.009003	4.519022

Рис. 7.
Обратная матрица ковариации

Приложение Е

0,573969	-0,144474
-0,088168	0,394065
-0,216031	-0,686024
-0,32866	0,327815
0,263323	0,262812
0,115533	-0,057209
-0,028681	0,028031
-0,452791	-0,067967
0,821071	1,219204
-0,056169	-0,146178
0,234492	-0,470982
-0,683777	-0,159179
0,891003	-0,156737

Рис. 8.
Средние выборочные для двух классов

Приложение Ж

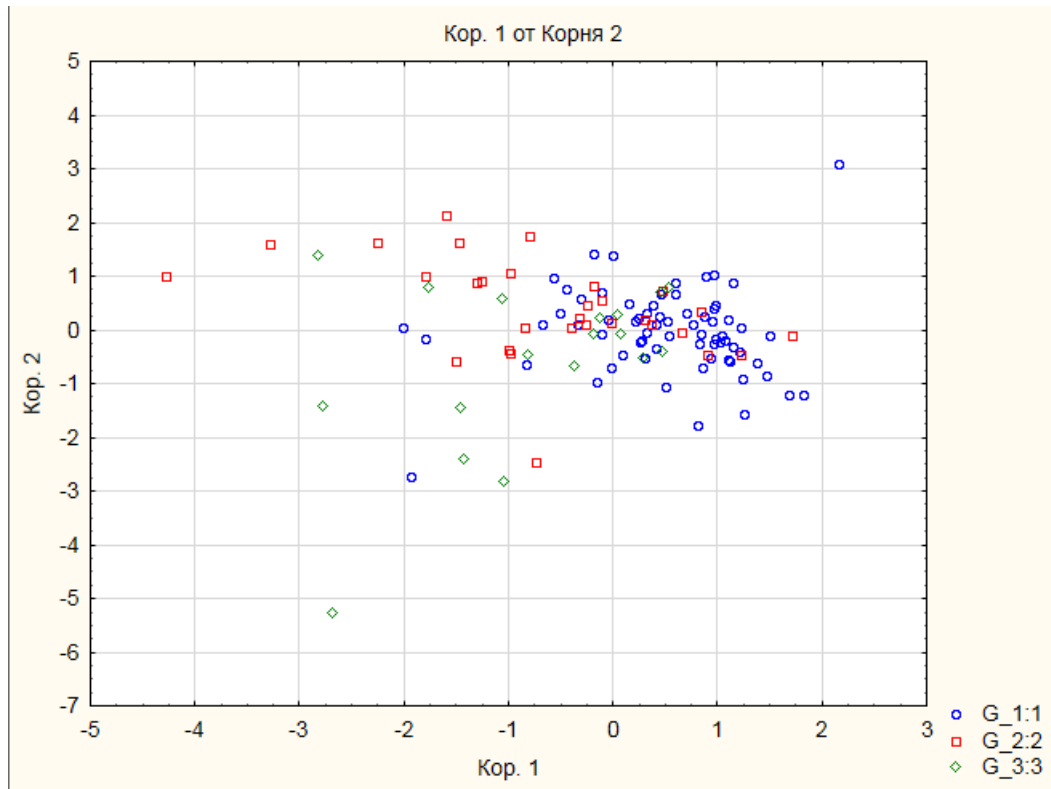


Рис. 9.
Диаграмма рассеяния для канонических значений