

УДК 61:007

МОДЕЛЬНЫЙ ПОДХОД К АНАЛИЗУ ОПЕРАбельНОСТИ ОНКОЛОГИЧЕСКИХ БОЛЬНЫХ

Т.И. БУЛДАКОВА, А.Е. ЗОЛОТЬКО, А.В. ЛАНЦБЕРГ, С.И. СУЯТИНОВ
E-mail: okb_sstu@sstu.ru, тел.: (8452) 79-86-09, г. Саратов

Резюме: Исследована задача принятия решения о возможности оперирования онкологических больных преклонного возраста. Определены критерии и предложено решение с помощью статистических моделей.

Ключевые слова: онкологические больные, преклонный возраст, функциональное состояние, дискриминантный анализ, прогнозирование.

THE MODEL APPROACH TO THE ANALYSIS OF AN OPERABILITY IN ONCOLOGICAL PATIENTS

T.I. BULDAKOVA, A.E. ZOLOTKO, A.V. LANTSBERG, S.I. SUYATINOV
E-mail: okb_sstu@sstu.ru, tel.: (8452) 79-86-09, city Saratov

Summary: The task of decision making about possibility of an operation in oncological patients with a declining age is researched. The criteria are defined and the solution with the help of statistical models is offered.

Key words: an oncologic ill, declining age, functional state, the discriminant analysis, forecasting.

Введение

Принятие решения об оперировании онкологических больных преклонного возраста всегда подразумевает дилемму: рискнуть с надеждой успешного исхода операции либо не рисковать жизнью больного, сохранив ему жизнь на период развития заболевания. Правильность решения основывается на оценке функционального состояния больного, которая ложится в основу прогноза исхода операции. Аналогичная проблема возникает и при принятии решения об оперировании пациентов с другими заболеваниями [1].

Поэтому, исходя из оценки функционального состояния пациента, необходимо решить, можно ли ему сделать радикальную операцию, предполагающую удаление пораженного органа, или общее состояние организма пациента позволяет провести лишь паллиативную операцию, направленную на облегчение состояния больного и на устранение угрожающих жизни симптомов при невозможности удаления пораженного очага.

Возможен также вариант, при котором функциональное состояние пациента не позволяет осуществить хирургическое вмешательство. В этом случае возможно лишь медикаментозное лечение. Следовательно, задача принятия решения об оперировании онкологических больных может быть представлена в следующем виде:

- 1) определить, позволяет ли общее функциональное состояние человека осуществить хирургическое вмешательство;
- 2) если это возможно, то позволяет ли оно удалить очаг поражения (радикальная операция) или только устранить угрожающие жизни симптомы (паллиативная операция), тем самым облегчив состояние больного.

Цель исследования - определение критериев принятия решения об оперировании онкологических больных преклонного возраста и разработка прогнозных моделей на их основе.

Задачи исследования:

- 1) выявить скрытые взаимосвязи между исходными параметрами и функциональным состоянием больного (диагнозом);
- 2) определить наиболее значимые для диагностики параметры;
- 3) разработать алгоритмы принятия решения об оперировании больных преклонного возраста.

Анализ исходных данных

В настоящее время в клинической практике лечения онкологических заболеваний для определения функционального состояния пациента используется набор из 39 признаков. Условно данные признаки можно разделить на следующие группы:

1. Общепатологические признаки: «Пол», «Возраст», «Температура при поступлении», «Сопутствующая патология», «Нарушение ритма сердца», «Наличие инфаркта».
2. Специальные онкологические признаки: «Длительность непроходимости», «Расположение опухоли», «Стадия TNM», «Наличие кишечной непроходимости», «Наличие метастазов в органы», «Наличие метастазов в лимфатические узлы», «Прорастание опухоли в соседние структуры», «Наличие асцита», «Гистологический тип опухоли».
3. Разновидности анализа крови:

3.1. Общий анализ крови: «Эритроциты», «Гемоглобин», «Гематокрит», «Лейкоциты», «Палочко-ядерные нейтрофилы», «Лимфоциты», «Сегментоядерные нейтрофилы», «Лейкоцитарный индекс» (признаки «Сегментоядерные нейтрофилы» и «Лимфоциты» используются только для вычисления признака «Лейкоцитарный индекс»).

3.2. Биохимический анализ крови: «Уровень билирубина», «Уровень мочевины», «Уровень калия», «Белок», «Уровень альбуминов», «Уровень глобулинов», «Глюкоза», «Уровень натрия».

3.3. Анализ крови коагулограмма: «Гепарин», «Ретракция», «Фибриноген».

4. Анализ мочи: «Белок в моче», «Сахар», «Ацетон», «Лейкоциты мочи», «Эритроциты мочи».

Часть исследуемых признаков представлена качественными характеристиками, часть - количественными, часть – как качественными, так и количественными. Такое представление исходных данных вызывает трудности, связанные с выбором шкал для их градуации.

Каждая группа признаков исследуется врачами на разных стадиях предоперационного обследования, однако на данный момент не выявлен точный алгоритм определения функционального состояния человека по всем 39 признакам, а также не существует единого мнения, стоит ли использовать все представленные признаки для решения поставленной задачи или достаточно использовать выборку признаков. Кроме того, нет ответа на вопрос, какие второстепенные признаки или комбинация второстепенных признаков оказывают наибольшее влияние на оценку функционального состояния пациента, и имеет ли смысл использовать их вообще.

Для решения поставленных задач была исследована выборка значений клинических параметров 281 пациента, прошедших лечение. Для каждого пациента из истории болезни были извлечены обозначенные выше признаки, а также диагноз, поставленный врачом, сведения о виде проведенной операции и состоянии пациента при выписке.

Выбор инструмента для анализа данных. Представляется необходимым построение модели, позволяющей выявить признаки, оказывающие наибольшее влияние на формирование оценки функционального состояния человека, и выбрать наиболее адекватный тип лечения онкологического больного. Если обозначить как $Z = \{z_1, z_2, z_3\}$ множество вариантов решения, где z_1 – радикальная операция, z_2 – паллиативная операция, z_3 - операция невозможна, то модель должна давать возможность выбора варианта лечения z_i . В основу такой модели должен быть положен инструмент, осуществляющий классификацию пациентов на 3 класса (радикальная операция, паллиативная операция, пациенты без операции) по заданным признакам и выявляющий наиболее значимые признаки [2]. Для реализации поставленной задачи был выбран дискриминантный анализ, позволяющий классифицировать объекты оптимальным способом (минимум математического ожидания потерь или минимум вероятности ложной классификации) на основе измеренных характеристик.

Предобработка исходных данных. Для проведения дискриминантного анализа необходимо выполнение следующего условия: измеренные характеристики должны быть представлены количественно. В связи с этим была проведена предобработка исходных данных, поскольку часть признаков была представлена качественно или имела как качественное, так и количественное представление. Поэтому с целью приведения качественных характеристик к количественному виду, необходимо было выбрать систему кодировки данных признаков. Следует отметить, что при проведении анализа данных несколько раз выполнялась перекодировка признаков «Наличие инфаркта», «Нарушение ритма», «Расположение опухоли» и «Прорастание опухоли в соседние структуры». Это связано с тем, что практически невозможно изначально выбрать кодировку, адекватно отражающую изменения признаков. Итеративный процесс перекодировки во время проведения анализа позволяет прийти к наилучшим результатам. В табл. 1 приведен пример кодирования.

Таблица 1

Пример кодирования исходных признаков

Номер кода	Качественные значения
Пол	
1	Мужской
2	Женский
Длительность непроходимости	
1	1 - 4 (Выражено в часах и приравнивается одному койкодню)
5-40	5 - 40 (Значения выражены в койкоднях, поэтому номер их кода соответствует количеству койкодней)
Сопутствующая патология	
1	Атеросклероз
2	Атеросклероз, нарушения работы системы пищеварения, мочевыводящей и дыхательной

	системы
3	Атеросклероз, гипертония, стенокардия
4	Атеросклероз, гипертония, ИБС
5	Случаи, требующие дополнительное операционное вмешательство.
6	Болезни, свидетельствующие об осложнениях на сосуды.
7	Тяжелые случаи, требующие особых условий.
Нарушение ритма	
0	Нет
1	Различные виды аритмий, а также синдром реполяризации желудочков и сочетания аритмий с реполяризацией
2	Сочетания нескольких видов аритмий, а также мерцательная аритмия
3	Нарушение внутрипредсердной и внутрижелудочковой проводимостей
4	Неполные блокады
5	Полные блокады, а также АВ-блокады
Наличие инфаркта	
0	Нет
1	Рубцы передней и задней стенок (а также рубцовые изменения)
2	Рубцы в передне-верхушечно-боковой, задне-перегородочно-верхушечной, а также передне-перегородочно-боковой областях
3	Инфаркт миокарда

В процессе предварительного изучения и подготовки признаков к анализу было выявлено, что в анамнезе многих больных отсутствуют некоторые группы биохимических признаков. Так, при изучении группы признаков «Анализ крови (коагулограмма)» было выявлено, что у 94 человек отсутствуют все признаки данной группы, у 14 человек отсутствуют 2 из 3 признаков, у 17 человек отсутствует какой-либо 1 признак. В группе признаков «Биохимический анализ крови» у 8 человек отсутствовали все признаки данной группы, у 37 человек отсутствовала подгруппа признаков «Альбумины, Глобулины, Калий, Натрий», у 93 человек отсутствовала подгруппа признаков «Калий, Натрий». В группе «Анализ мочи» признак «Эритроциты мочи» отсутствовал более чем у половины обследуемых, в связи этим он был удален из набора анализируемых признаков. Был исключен также признак «Наличие метастазов в органы», поскольку выявление данного признака осуществляется уже на этапе оперирования.

Данная ситуация показывает, что вышеназванные признаки не используются врачами для прогнозирования возможности проведения той или иной операции, а фиксируются лишь для описания состояния больного. Вероятно, это обусловлено либо отсутствием влияния, либо неизученностью влияния данных групп признаков на прогноз исхода операции.

Было сделано обоснованное предположение, что смертельный исход в результате проведенного лечения свидетельствует об ошибочности его назначения. В связи с этим общее количество пациентов было разделено на 2 группы. В первую группу вошли все пациенты со смертельным исходом в результате назначенного и проведенного лечения (таких пациентов оказалось 19 человек). На данной группе пациентов в дальнейшем была осуществлена проверка адекватности прогнозной модели. Во вторую группу вошли все пациенты, выписанные после лечения в удовлетворительном состоянии (таких пациентов оказалось 262 человека), то есть пациентам данной группы было назначено и проведено правильное лечение. Следовательно, сведения о пациентах данной группы служат основой для построения прогнозирующей модели. Далее в каждой из групп было выделено по 3 подгруппы пациентов: к первой подгруппе были отнесены пациенты, которым была сделана радикальная операция (117 пациентов с удовлетворительным состоянием и 7 - со смертельным исходом), ко второй - пациенты с паллиативной операцией (135 пациентов с удовлетворительным состоянием и 11 - со смертельным исходом), к третьей - пациенты, которым был противопоказан любой вид операции (10 пациентов с удовлетворительным состоянием и 1 - со смертельным исходом).

Результаты исследования

Для проведения дискриминантного анализа был выбран Модуль «Дискриминантный анализ» программного продукта Statistika 6.0. Для оценки достоверности и адекватности полученных результатов при проведении анализа контролировались следующие значения показателей:

1. Уровень значимости (p-level) – находится в убывающей зависимости от надежности результата. Его значение должно оставаться неизменным и равным 0.0000.

2. Лямбда Уилкса – свидетельствует о дискриминации, чем больше значение данного показателя, тем лучше дискриминация и меньше вероятность ошибки прогноза.

Для просмотра рассчитанных вероятностей отнесения данных к классификационной группе использовалась Классификационная матрица. Если значение хотя бы одной вероятности (из трех) низкое,

то можно говорить о том, что выбранная совокупность признаков оказывает влияние на формирование лишь двух классификационных групп (для которых значения вероятностей сравнительно высоки). Было принято, что пороговым значением для вероятности распознавания является 70%, то есть удовлетворительными считаются все значения больше 70%.

Для анализа была выбрана группа пациентов, выписанных в удовлетворительном состоянии (262 человека). Для определения механизма влияния совокупности признаков на оценку функционального состояния пациента было принято решение о проведении анализа для каждой подгруппы признаков. Целью является выбор в каждой подгруппе признаков, оказывающих наибольшее влияние на формирование оценки предоперационного состояния пациента, и формирование совокупности признаков, позволяющих построить модель прогнозирования. При этом при любом сочетании независимых признаков, зависящий признак остается неизменным – «Операция».

Исходя из постановки задачи, было принято решение о необходимости проведения трех отдельных анализов данных по следующей схеме. На первом этапе необходимо выявить, позволяет ли состояние пациента сделать ему операцию. А на втором этапе - выявить, какую операцию можно сделать. Так, если состояние пациента позволяет провести сложную радикальную операцию, то оно, следовательно, позволяет провести и более простую паллиативную операцию. Однако если состояние пациента позволяет провести только паллиативную операцию, то ему нельзя проводить радикальную операцию. Поэтому сначала был проведен дискриминантный анализ для 1-й и 3-й групп пациентов, а потом для 1-й и 2-й групп пациентов. Кроме того, был также проведен анализ для 2-й и 3-й групп пациентов, позволяющий ответить на вопрос: является ли состояние пациента настолько плохим, что не позволяет сделать операцию, направленную на облегчение состояния?

Предполагая, что состояние пациента позволяет сделать ему операцию, были проанализированы результаты исследования первых двух групп пациентов. При этом использовалась методика пошагового дискриминантного анализа с перебором и постепенным изъятием наименее значимых признаков.

Результаты дискриминантного анализа показали, что для определения степени сложности операции у пациента с онкологическим заболеванием достаточно определить 9 признаков, 2 из которых являются общепатологическими («Возраст» и «Сопутствующая патология»), 3 - специальными онкологическими («Наличие кишечной непроходимости», «Наличие метастазов в лимфоузлы», «Наличие асцита»), остальные - биохимическими («Билирубин», «Альбумины», «Калий», «Сахар») (табл. 2).

Таблица 2

Значимые параметры для определения степени сложности операции

Наименование признака	Лямбда Уилкса	Уровень значимости
Наличие метастазов в лимфоузлы	0,6361	p<0,0000
Наличие кишечной непроходимости		
Возраст		
Сопутствующая патология		
Наличие асцита		
Билирубин		
Альбумины		
Калий		
Сахар		

Вероятность правильного распознавания в этом случае составила 82,9% (85,3% - для группы 1 и 80,7% - для группы 2).

Исследование влияния признаков «Наличие инфаркта» и «Нарушение ритма» показало, что данные признаки не оказывают существенного влияния на формирование обеих групп:

Далее был проведен анализ для определения, является ли состояние пациента настолько плохим, что ему нельзя делать даже поддерживающую операцию. В результате было определено, что использовании признаков «Длительность непроходимости», «Наличие метастазов в лимфоузлы», «Наличие инфаркта» является оптимальным для определения пациентов, которым нельзя делать операцию вообще либо можно сделать только поддерживающую операцию. Вероятность правильного распознавания составила в среднем 97,2% (98,5% - для пациентов группы 2, 80,0% - для пациентов группы 3). Аналогичные результаты были получены, при замене третьего признака на «Нарушение ритма».

При исследовании пациентов группы 1 и группы 3 изначально для анализа были выбраны признаки, давшие наилучший результат: «Длительность непроходимости», «Наличие метастазов в лимфоузлы». Однако в дальнейшем оказалось, что в данном случае признак «Наличие метастазов в лимфоузлы» не является значимым. Добавление других признаков, например, «Наличие асцита», «Пол» и «Возраст», не

привело к улучшению результатов. Кроме того, добавление признаков, полученных в результате анализа крови (Гепарин, Эритроциты, Гемоглобин, Гематокрит, Лейкоциты), привело к существенному снижению показателей достоверности. Полученные оптимальные наборы параметров приведены в табл. 3.

Таблица 3

Значимые наборы параметров при исследовании пациентов групп 1 и 3

№№ п/п	Наименование признаков в наборах	Лямбда Уилкса	Уровень значимости	Вероятность правильного распознавания
1	Длительность непроходимости, наличие инфаркта	0,4931	p<0,0000	95,7 % (группа 1) 100,0 % (группа 3) 96,9 % (в среднем)
2	Длительность непроходимости, нарушение ритма	0,5232	p<0,0000	86,9 % (группа 1) 90,0 % (группа 3) 87,9 % (в среднем)
3	Длительность непроходимости, наличие инфаркта, нарушение ритма	0,4423	p<0,0000	91,3 % (группа 1) 100,0 % (группа 3) 93,9 % (в среднем)
4	Длительность непроходимости, наличие инфаркта, сопутствующая патология	0,4802	p<0,0001	95,6 % (группа 1) 100,0 % (группа 3) 96,9 % (в среднем)

Таким образом, было определено, что из 39 существующих признаков для первоначального выявления возможности какого-либо операционного вмешательства необходимо обязательное рассмотрение признака «Длительность непроходимости» из группы специальных онкологических признаков и одного из признаков группы общефизиологических признаков: «Наличие инфаркта» либо «Нарушение ритма» (оба этих признака оказали одинаковое влияние на дискриминацию групп пациентов). Исследование остальных признаков с целью дискриминации данных двух групп пациентов показало, что они не оказывают существенного влияния. Это свидетельствует о том, что при предварительном заключении о возможности проведения операции нет необходимости проведения анализа крови, мочи, а также исследования других общефизиологических и специальных онкологических признаков, требующих дополнительной затраты времени.

При проведении анализа были получены коэффициенты дискриминантных функций. Пример коэффициентов для п.2 табл. 3 представлен в табл. 4.

Таблица 4

Значения коэффициентов дискриминантных функций

Параметры	Группа 1	Группа 3
Длительность непроходимости	0,20338	0,58304
Нарушение ритма	0,43737	0,04963
b ₀ (Constant)	-1,50813	-7,57412

Полученные коэффициенты b_{ki} позволяют построить дискриминантные функции y_k [3], являющиеся основой модели, которая позволяет выбрать наиболее адекватный тип лечения онкологического больного:

$$y_k = b_{k0} + \sum_{i=1}^p b_{ki} \cdot x_i = b_{k0} + b_{k1} \cdot x_1 + b_{k2} \cdot x_2 + \dots + b_{kp} \cdot x_p,$$

где p – число переменных, k – номер группы, x_i – значение i – переменной. К примеру, при оценке состояния больного для принятия решения об его оперировании модель имеет вид:

$$Z = \begin{cases} z_1 & \text{при } y_1 > y_3, \\ z_3 & \text{в противном случае,} \end{cases}$$

$$y_1 = -1,50813 + 0,20338 \cdot x_1 + 0,43737 \cdot x_2,$$

$$y_3 = -7,57412 + 0,58304 \cdot x_1 + 0,04963 \cdot x_2.$$

Здесь x_1 – длительность непроходимости (в койкоднях), x_2 – код нарушения ритма. Аналогично могут быть получены модели по другим информативным признакам, например, для выбора типа операции.

Заключение

Проведенные исследования позволили выявить механизм принятия решения об оперировании онкологических больных преклонного возраста с учетом предварительно проведенного обследования. По-

лученные результаты позволяют определить наиболее значимые признаки для принятия решения об оперировании и выбора типа операции. Исключение из рассмотрения малозначимых признаков может существенно сократить время предоперационного обследования больного.

Исследования показали, что наиболее оптимальным является следующий алгоритм принятия решения об оперировании. На первом этапе необходимо выявить, позволяет ли общее функциональное состояние человека осуществить операционное вмешательство. В случае если сделано заключение о том, что операционное вмешательство невозможно, то анализ признаков прекращается. В противном случае необходимо проведение второго этапа, на котором определяется, какую операцию можно провести.

Было выявлено, что из используемых в настоящее время в клинической практике 39 признаков, значимыми являются лишь 15, остальные же не оказывают существенного влияния на принятие решения. Данное заключение может позволить в дальнейшем сократить время, затрачиваемое на принятие решения, в частности на проведение процедуры подготовки к оперированию. Следует отметить важность признаков «Нарушение ритма» и «Наличие инфаркта» при определении возможности проведения операционного вмешательства.

Литература

1. Шевченко Ю.Л., Щихвердиев Н.Н., Оточкин А.В. Прогнозирование в кардиохирургии. СПб: Питер Паблишинг, 1998. 208 с.

2. Лифшиц В.Б., Булдакова Т.И., Суятинов С.И., Колентьев С.В. Статистический и нейросетевой методы идентификации и прогнозирования в медицине // Информационные технологии. 2004. № 3. С. 60-63.

3. Медик В.А., Токмачев М.С. Математическая статистика в медицине. М.: Финансы и статистика, 2007. 800 с.

Работа выполнена при финансовой поддержке РФФИ (проект 07-07-12066).