

# **Математико-статистические модели прогнозирования эффективности оперативного лечения некоторых заболеваний**

**Е.И. Воронова, Ю.А. Костиков, Л.А. Муравей**

## **1. Введение**

В последние десятилетия во многих странах мира, в том числе и в России, происходит рост заболеваемости раком. Наиболее часто рак выявляется у лиц старше 50 лет с постепенным снижением уровня заболеваемости в группах населения в возрасте после 75 лет [1]. Но, несмотря на довольно четкое увеличение частоты рака в старших возрастных группах, все чаще это заболевание диагностируют в молодом возрасте, особенно его семейные и наследственные формы. Таким образом, проблема диагностики и лечения рака приобрела в настоящее время важное социальное значение.

Несмотря на улучшение методов диагностики среди впервые выявленных больных колоректальным раком преобладают запущенные формы. Применение различных математических приемов для описания опухолей дело не новое, первые попытки создания модели роста были предприняты еще в первой половине XIX века. Однако математические описания опухолей, начиная с первой модели Benjamin Gompertz (1825г.), направляются на оценку разрастания клеточных множеств (T.E Wheldon. "Mathematical Models in Cancer Research". A.Hilger, — Bristol and Philadelphia)[2], в том числе, с учетом различных начальных условий и характера внешнего вмешательства.

Современные методики исследования раковых болезней требуют автоматизированного подхода. Вследствие этого был предложен метод определения нового агрегированного показателя иммунитета, по которому можно определить вероятность развития рецидивов заболевания, тактику обследования, лечения на основании предложенной математической модели.

Моделирование и отображение процесса роста опухоли по данным клиники — важнейшая проблема математизации онкологии.

Целью нашего исследования явилась разработка научно-обоснованных подходов к повышению эффективности хирургического лечения злокачественных опухолей органов брюшной полости и малого таза на основе использования математико-статистических моделей прогнозирования переносимости оперативных вмешательств с применением внутрисполостной гипертермической химиотерапии (ВГХТ). С помощью статистических методов была проведена по изучению степени влияния ВГХТ на параметры иммунитета и оценке характера иммунологических изменений при различных видах оперативного вмешательства.

Методами бинарной логистической регрессии и дискриминантного анализа был рассчитан прогноз исхода операции — вероятность выживания пациента. На основе метода множественной регрессии были получены алгоритмы (методики) расчета продолжительности жизни после операции и безрецидивного периода. Исследования, основанные на факторном анализе, в частности, показали какие факторы оказывают существенное влияние на эти показатели, а какие — нет. Обработка данных производилась в программе SPSS.

Статья организована следующим образом: приводится описание соответствующего метода, затем его анализ и сравнение результата полученного метода с опытными данными.

## **2.Метод бинарной логистической регрессии.**

Вероятность наступления события для некоторого случая рассчитывается по формуле:

$$p = \frac{1}{1+e^{-z}},$$

где  $z = b_1 * x_1 + b_2 * x_2 + \dots + b_n * x_n + a$

$b_1$  — коэффициенты, расчет которых является задачей бинарной логистической регрессии,  $x_1$  - значения независимых переменных,  $a$ — некоторая константа.

Если для  $p$  получится значение меньше 0,5, то предполагается, что событие не наступит; в противном случае предполагается наступление события.[1,3]

Рассмотрим диагностический тест на предмет выживаемости больного.

Из большого количества переменных, были выбраны следующие :

Таблица № 1

Переменные бинарной логистической регрессии

Имя переменной	Расшифровка
live	Исход (0 = скончался, 1 = жив)
age	возраст
sex	Пол (1= женский, 2= мужской)
t_imm	Суммарный показатель значений Т-иммунитета
b_imm	Суммарный показатель значений В-иммунитета
resist	Неспецифическая резистентность
hystology	Гистология ( Высокодифференцированный рак – 1 Умереннодифференцированный рак – 2 Муцинозная опухоль(слизистая) -3 Низкодифференцированный рак-4 Недифференцированный рак - 5)

Наряду с переменной live (исход), имеются переменные, при первом взгляде на которые можно понять, что они с ней связаны.

Из-за вовлечения в анализ большого числа переменных был выбран пошаговый метод бинарной логистической регрессии[6,7].

Точность исполнения прогноза, которая достигается при использовании этих переменных, составляет 96,3 %.

Значения коэффициентов и константы для расчета вероятности выживаемости находятся в следующей таблице:

Таблица № 2

## Переменные в уравнении

Имя	B коэффициент регрессии	–Стандартная ошибка	df	Значимость.	Exp(B)
T_IMM	-0,111	0,140	1	0,427	0,895
B_IMM	-0,008	0,010	1	0,433	0,992
HYSTOLOG	-3,327	54,758	4	0,999	0,24
RESIST	0,004	0,018	1	0,844	1,004
AGE	-0,047	0,060	1	0,435	0,954
SEX	0,845	1,169	1	0,470	2,327
Constant	9,956	54,930	1	0,856	21074,939

Если мы рассмотрим случай с 80- летним пациентом, женского пола, с показателем Т- иммунитета = 5,91 и В – иммунитета =66,03 с высокодифференцированной опухолью и показателями резистентности= 72,3 то исходя из соотношения

$$z = 9,956 + 0,845 * 1 - 0,047 * 80 + 0,004 * 72,3 - 3,327 * 1 - 0,008 * 66,03 - 0,111 * 5,91 = 2,81895$$

получим вероятность выздоровления

$$p = \frac{1}{1 + e^{-2,81895}} = 0,943691.$$

Следовательно, вероятность выживания пациента = 0,943691

### 3.Дискриминантный анализ

С помощью дискриминантного анализа на основании некоторых признаков индивидум может быть причислен к одной из двух заранее известных групп.

Ядром дискриминантного анализа является построение так называемой дискриминантной функции:

$$d = b_1 * x_1 + b_2 * x_2 + \dots + b_n * x_n + a$$

Целью является определение таких коэффициентов, чтобы по значениям дискриминантной функции можно было с максимальной четкостью провести разделение по группам.[1,3]

Таблица № 3

Канонические коэффициенты дискриминантной функции :

	Функция
	1
AGE	0,040
SEX	-1,048
T_IMM	0,052
B_IMM	0,007
HYSTOLOG	0,560
RESIST	-0,003
(Константа)	-3,297

Распределение значений дискриминантной функции для двух групп :

«выжил» -  $d = 1.54 \pm 0.48$

и «умер» -  $d = 1 \pm 0.03$

Для случая, когда мужчине 72 года с показателем Т- иммунитета = 8,78 и В – иммунитета =85,28 с умереннодифференцированной опухолью и показателями резистентности= 75,5 то исходя из соотношения

$$d = -3,297 + 0,04 * 72 - 1,048 * 2 + 0,52 * 8,78 + 0,007 * 85,28 + 0,56 * 2 - 0,03 * 75,4 \\ = 1,050756$$

Опираясь на распределение значений дискриминантной функции, этого пациента можно отнести к группе выживших.

#### **4. Факторный анализ**

Факторный анализ – процедура, с помощью которой большое число переменных, относящихся к имеющимся наблюдениям, сводится к меньшему числу независимых, называемых факторами. При этом в один фактор объединяются переменные, сильно коррелирующие между собой [4].

Таким образом, нашей целью является нахождение таких комплексных факторов, которые как можно более полно объясняют наблюдаемые связи между переменными, имеющимися в наличии.

При анализе суммарных показателей было отобрано только три фактора. Первый фактор объясняет 25,355 % суммарной дисперсии, второй фактор – 20,85 % и третий фактор – 19,1 %.

После анализа повернутой матрицы компонент была оценена факторная нагрузка по каждому фактору. В итоге фактор 1 включает в себя такие показатели, как Т – иммунитет и резистентность, фактор 2 – гистологию и показатели В- иммунитета и фактор 3 – возраст и пол.

Большое значение имеют 1 и 2 факторы, поскольку в этом случае факторная нагрузка имеет наибольшее значение.

Таким образом, можно сделать вывод, что возраст и пол на продолжительность жизни и протекание болезни влияют меньше, чем другие показатели.

Рассмотрим еще одну модель факторного анализа – с включением размера и распространения опухоли. После анализа таблицы первичной статистики и повернутой матрицы компонент было выделено 7 факторов, приведенных ниже.

Таблица № 4

**Факторы и переменные, их характеризующие**

Фактор	Имена переменных	Расшифровка
1.	limf, t3, t4, t8, ae-roc, eac-roc	показатель значений Т- иммунитета: лейкоциты, лимфоциты, Т-3, Т-4,Т-8, АЕ-РОС, В-иммунитета: ЕАС-РОС
2.	lgm, lgg	Показатель значений В-иммунитета: lg M, lg G
3.	cik, lga, hystology	Показатель значений В-иммунитета: ЦИК, lg A Гистология (Высокодифференцированный рак – 1 Умеренно-дифференцированный рак – 2 Муцинозная опухоль(слизистая) -3 Низкодифференцированный рак-4 Недифференцированный рак - 5)
4.	fagoccis, inddiff,	Фагоц. число, инд. дифф.
5.	nst	Резистентность: НСТ
6.	livetim, leik	Срок жизни (в мес.), лейкоциты
7.	size	Размер опухоли – соотв. Т1-4 – 1-4

Первый фактор объясняет 28,871 % суммарной дисперсии, второй фактор 11,168 %, третий фактор 9,261%, четвертый – 8,493%, пятый – 7,279 %, шестой – 6,644 % и седьмой – 5,965 %.

Таким образом, можно сделать вывод, что использование суммарных показателей иммунитета не было ошибочно, и необходимо учитывать размер опухоли и ее распространение, т.к. это отдельный фактор.

**5.Метод множественной линейной регрессии**

В случае множественного регрессионного анализа речь идёт необходимо оценить коэффициенты уравнения

$$y = b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_n \cdot x_n + a,$$

где  $n$  -- количество независимых переменных, обозначенных как  $x_1$  и  $x_n$ , а -- некоторая константа.

Переменные, объявленные независимыми, могут сами коррелировать между собой; этот факт необходимо обязательно учитывать при определении коэффициентов уравнения регрессии для того, чтобы избежать ложных корреляций.)[3,5]

Уравнение регрессии, для прогнозирования значений безрецидивного периода выглядит следующим образом:

$$\text{brp} = -0,103 * \text{AGE} - 0,149 * \text{T\_IMM} + 0,01597 * \text{V\_IMM} + 2,179 * \text{HYSTOLOGY} + 0,06409 * \text{Resist} + 12,183$$

Для 45-летнего больного с показателями В-иммунитета 116,29 и Т-иммунитета 10,08 и резистентностью 69,51 с высокодифференцированной опухолью безрецидивный период будет равен

$$\text{brp} = -0,103 * 45 - 0,149 * 10,08 + 0,01597 * 116,29 + 2,179 * 1 + 0,06409 * 69,51 + 12,183 = 14,5 \text{ мес.}$$

Уравнение регрессии, для прогнозирования продолжительности жизни после операции выглядит следующим образом:

$$\text{LIVETIM} = 91,400 - 23,477 * \text{SIZE} + 9,580 * \text{LEIK} - 42,180 * \text{LIMF} + 48,051 * \text{T3} - 100,645 * \text{T4} - 29,208 * \text{T8} + 112,975 * \text{AEPOC} + 10,815 * \text{INDDIF} + 62,313 * \text{EACPOC} + 0,145 * \text{CIK} - 13,222 * \text{LGA} + 9,124 * \text{LGM} + 0,584 * \text{LGG} - 3,734 * \text{FAGOCCIS} + 15,760 * \text{NST} - 2,332 * \text{HYSTOLOG}$$

Для больного с 4 стадией высокодифференцированной опухоли и показателями лейкоцитов – 2,20, лимфоцитов – 0,83, Т-3 – 0,39, Т-4 – 0,25, Т-8 – 0,14, АЕ-РОС- 0, 23 ЦИК – 49, IgA – 0,70, IgM – 1,10, IgG – 15, фагоц.число -3,30, инд.дифф – 0,80, ЕАС-РОС - 116,29 продолжительность жизни будет равна

$$\text{LIVETIME} = 91,400 - 23,477 * 4 + 9,580 * 2,20 - 42,180 * 0,83 + 48,051 * 0,39 - 100,645 * 0,25 - 29,208 * 0,14 + 112,975 * 0,23 + 10,815 * 0,80 + 62,313 * 116,29 - 13,222 * 0,70 + 9,124 * 1,10 + 0,584 * 15 - 3,734 * 3,30 + 0,145 * 49 - 2,332 * 1 = 25,6.$$

Пациент прожил 27 месяцев.

## **6. Выводы**

С помощью методов математической статистики была изучена степень влияния ВГХТ на параметры иммунитета и получена оценка характера иммунологических изменений при различных видах оперативного вмешательства. В ходе исследования рассмотрены подходы к прогнозированию с использованием различных статистических методов[10]. Метод прогноза был проверен на статистических данных пациентов, проходивших лечение с ВГХТ и без ВГХТ. Функция прогнозирования зависит от нескольких параметров: ПИР, степени дифференцировки опухоли, полноты циторедукции, возраста пациента, применения ВГХТ. Полученные данные математического прогнозирования как для значений безрецидивного периода, так и для срока жизни коррелируются с опытными данными, т.е. вычисленные теоретические значения безрецидивного периода и срока жизни близки к достоверным .

Внедрение в лечебную практику статистических моделей прогнозирования позволит лечащим врачам объективно оценить показатели гомеостаза, переносимость планируемых оперативных вмешательств, более точно прогнозировать течение раковой болезни и вероятность развития рецидива заболевания, а также продолжительность и качество жизни пациентов[9]. Полученные результаты показывают целесообразность применения разработанных методик и лежащих в их основе статистических методов в повседневной клинической практике[11].

Использованы данные совместной обработки результатов операций в 3 Центральном военном клиническом госпитале имени А.А. Вишневого Министерства обороны РФ.

## Список литературы:

1. SPSS 19. Профессиональный статистический анализ данных: А. Наследов — Санкт-Петербург, Питер, 2011 г.- 400 с.
2. Corman M.L. Colon and Rectal Surgery.// - Philadelphia, 1992. – 233 p.
3. БюльАхим, Петер Цефель. SPSS 10.: Искусство обработки информации. Diasoft. — Москва, С-Петербург, Киев, 2002. — 601 с.80.
4. Харченко В.П. Высокие технологии в онкологии/В.П.Харченко, В.А.Паньшин.- Казань, 2000. -Т.2.- 42-44с.
5. Амосов, Н.М. Кибернетические методы и математическое моделирование в биологии и медицине / Н.М. Амосов // Диалектико-материалистический анализ основных методов исследования в биологии и медицине. Киев.- 2000.- С.80с.
6. С. М. Ермаков, Г. А. Михайлов. Курс статического моделирования. – Главная редакция физико-математической литературы издательства «Наука», М., 1976г.
7. Самарский А.А., Михайлов А.П. Математическое моделирование: Идеи, методы, примеры. – М.: Физматлит, 2008.
8. Двойрин В.В. Статистика злокачественных новообразований в России/В.В.Двойрин, Е. М.Аксель, А.А.Трапезников. М.: Медицина, 1995.- 167с.
9. Valentini V., Mantini G., Turriziani A., Balducci M., Trodella L. Research trends in the treatment of colorectal cancer. //Rays, 2000, № 25(3), P. 393-395
- 10.Виноградова Г.Л., Демчинова В.А., Метод прогнозирования запасов материальных ресурсов в многономенклатурном производстве, позволяющий учитывать классификацию материальных ресурсов с условием их взаимозаменяемости. Инженерный вестник Дона [Электронный журнал]. – 2010. – №4 URL: <http://ivdon.ru/magazine/archive/n3y2012/912>

11. Семенистая Е.С., Максимов А.В. О подходе к построению модели дистальных сосудов пригодной для оценки артериального давления. Инженерный вестник Дона [Электронный журнал]. – 2012. – №4(часть 2) URL: <http://ivdon.ru/magazine/archive/n4p2y2012/1469>