

10. Fukunishi I, Tsuruta T, Hirabayashi N, Asukai N. Association of alexithymic characteristics and posttraumatic stress responses following medical treatment for children with refractory hematological diseases. *Psychol Rep.* 2001 Dec;89(3):527-34.
11. Green B.L., Krupnick J.L., Rowland J.H. et al. Trauma history as a predictor of psychologic symptoms in women with breast cancer // *J. of Clinical Oncology.* – 2000. - Vol. 18 (5).
12. Kangas M, Henry J.L, Bryant R.A. Predictors of posttraumatic stress disorder following cancer. *Health Psychol.* 2005 Nov;24(6):579-85.
13. Kelly B, Raphael B, Smithers M, et al. Psychological responses to malignant melanoma: An investigation of traumatic stress reactions to life-threatening illness. *Hosp Psychiatry* 17:126-134, 1995.
14. Kessler R.C, Sonnega A, Bromet E, et al. Posttraumatic stress disorder in the National Comorbidity Survey. *Arch Gen Psychiatry.* 1995;52:1048-1060.
15. Koopman C, Gore-Felton C, Azimi N, et al. Acute stress reactions to recent life events among women and men living with HIV/AIDS. *Int J Psychiatry Med.* 2002;32(4):361-78.
16. Leopold K.A., Ahels T.A., Walch S. et al. Prevalence of mood disorders and utility of the PRIME-MD in patients undergoing radiation therapy // *Int. J. Radiat Oncol. Biol. Phys.* – 1998.-Vol. 42, N5. – P.1105-1112.
17. Martinez A, Israelski D, Walker C, Koopman C. Posttraumatic stress disorder in women attending human immunodeficiency virus outpatient clinics. *AIDS Patient Care STDS.* 2002 Jun; 16(6):283-91.
18. Maunsell E, Brisson J, Deschenes L. Psychological distress after initial treatment of breast cancer. *Cancer* 70:120-125, 1992.
19. Richardson J, Shelton D, Krailo M, et al. The effect of compliance with treatment on survival among patients with hematologic malignancies. *J Clin Oncol* 8:356-364, 1990.
20. dos Santos M.J., Pimentel P., Monteiro J.M. et al. Psychiatric disorders in hospitalized patients with hematologic neoplasms // *Acta Med. Port.* – 1991. – Vol. 4, N1. – P. 5-8.

УДК 616.62-006.6-037

© Ш.Х. Ганцев, А.А. Зимичев, В.Н. Маклаков, К.Ш. Ганцев, 2007

Ш.Х. Ганцев, А.А. Зимичев, В.Н. Маклаков, К.Ш. Ганцев  
**ПРИМЕНЕНИЕ МНОГОФАКТОРНОГО КОРРЕЛЯЦИОННО-  
РЕГРЕССИОННОГО АНАЛИЗА В ОПРЕДЕЛЕНИИ РИСКА РАЗВИТИЯ И В  
ПРОГНОЗИРОВАНИИ ИСХОДА РАКА МОЧЕВОГО ПУЗЫРЯ**

*ГОУ ВПО «Башкирский государственный медицинский университет Росздрава», г.Уфа  
ГОУ ВПО «Самарский государственный медицинский университет Росздрава», г.Самара  
ГОУ ВПО «Самарский государственный технический университет», г.Самара*

Корреляционно-регрессионный анализ позволяет выявить комплекс наиболее весомых независимых факторов, оказывающих влияние на возникновение и исход рака мочевого пузыря. Эти факторы могут быть положены в основу математических моделей риска развития и прогноза исхода рака мочевого пузыря. Разработана компьютерная программа для оценки прогноза жизни больных (свидетельство об официальной регистрации программ для ЭВМ №2005612134 от 19.08.2005) и компьютерная программа оценки риска развития рака мочевого пузыря. Применение программы позволяет разработать комплекс мероприятий по первичной и третичной профилактике рака мочевого пузыря.

**Ключевые слова:** рак мочевого пузыря, корреляционно-регрессионный анализ, прогноз, заболеваемость.

Sh.Kh. Gantsev, A.A. Zimichev, V.N. Maklakov, K.Sh. Gantsev  
**THE USE OF MULTIFACTORAL CORRELATION-REGRESSIVE  
ANALYSIS IN RISK ASSESSMENT OF THE DEVELOPMENT AND  
PROGNOSIS OF URINARY BLADDER CANCER OUTCOME**

Correlation-regressive analysis allows determining a complex of significant independent factors that effect the development and outcome of the urinary bladder cancer. Based on these factors, developmental risks and prognosis of outcome of the urinary bladder cancer can be modeled mathematically. A computer programme for prognosis of patients' life (official registration certificate N 2005612134 of 19.08.2005) and a programme for risk assessment of the development of the urinary bladder cancer have been introduced. The use of the programme contributes to carrying out a complex of measures for primary and tertiary prevention of the urinary bladder cancer.

**Key words:** urinary bladder cancer, correlation-regressive analysis, prognosis, morbidity.

Профилактика злокачественных новообразований является в настоящее время приоритетной с позиций современной онкологии. Для разработки комплекса мер по профилактике необходимо изучать основные факторы, приводящие к заболеванию. Наиболее перспективны в этом плане эпидемиологические факторы и в первую очередь экзогенные – профессия, вредные привычки, особенности питания и питьевого режима человека. На данный момент существующие методы лечения рака мочевого пузыря не всегда эффективны. Анатомо-физиологические особенности мочевого пузыря, биологические свойства опухолей, трудность определения стадии и степени злокачественности, склонность к рецидивированию, множественность поражения, возраст больного и сопутствующие заболевания – вот далеко не все параметры, которые влияют на исход лечения. Совокупность перечисленных факторов создает большие трудности при выборе адекватного способа лечения.

Целью данной работы является улучшение диагностики и результатов лечения на основе выявления комплекса факторов, способствующих развитию рака мочевого пузыря, и комплекса факторов, влияющих на прогноз злокачественного заболевания.

#### Материал и методы

Для оценки риска заболевания и прогнозирования результатов лечения рака мочевого пузыря использована выборка из 510 больных раком мочевого пузыря. В качестве входящих переменных ( $x_j$ ) использованы в обеих программах эпидемиологические информативные признаки: профессиональные особенности, курение, особенности питания, питьевой режим, пол и возраст пациента, генетическая детерминация опухолевых образований, наличие сопутствующих заболеваний мочеполовой системы, а в математической модели прогноза использованы клиникo-морфологические информативные признаки: своевременность начала лечения, ультразвуковая структура рака мочевого пузыря, размер опухоли, ее локализация, характер роста рака мочевого пузыря, количество опухолей, функциональное состояние почек, степень распространенности заболевания по TNM, объемы проведенного лечения и оперативного пособия, чувствительность рака мочевого пузыря к проводимой системной полихимиотерапии, гистологическая структура опухоли и ее гистопатологическая градация, степень инвазии рака в стенку мочевого пузыря, в лим-

фатические и венозные сосуды, рецидивный характер опухоли. В качестве выходящей переменной ( $y$ ) принят уровень состояния здоровья пациента (наличие заболевания или рецидив опухоли, летальный исход). Каждой из входящих и выходящих переменных присвоено то или иное дискретное числовое значение натурального ряда в соответствии со степенью проявления переменной для каждого пациента.

Определены коэффициенты линейной корреляции ( $R_{yx}$ ), характеризующие степень взаимосвязи между величинами ( $y, x$ ). Однако по абсолютному значению коэффициента корреляции проблематично судить о тесноте связей с количественной точки зрения. Поэтому в данном случае были привлечены процедуры проверки статистических гипотез [1, 4]. Процедура проверки значимости начинается с формулировки проверяемой гипотезы или нулевой гипотезы ( $H_0$ ). Например, если в рассматриваемом случае  $H_0 : R = 1$ . (Корреляция имеет место), тогда альтернативная гипотеза запишется как  $H_1 : R = 0$ . (Корреляция отсутствует). Оценка значимости коэффициента корреляции производится с помощью статистики F распределения. Если  $F_{набл} > F_{крит}$ , то нулевая гипотеза  $H_0$  принимается, а альтернативная  $H_1$  отвергается, при этом вероятность отвергнуть правильную гипотезу  $H_0$  равна уровню значимости  $\alpha$  (0,01). Значение статистики F наблюдаемого вычисляется по имеющимся данным, а значение F критического находится по таблицам F распределения с учетом степеней свободы, зависящих от объема выборки и уровня значимости  $\alpha$  (вероятность отвергнуть правильную гипотезу). Доверительный интервал для коэффициента корреляции строится с использованием преобразования Фишера [1].

Статистическая обработка проводилась на ЭВМ Intel Pentium 4 с использованием специально разработанной прогностической программы в среде Visual Basic 7.0 for Application (Свидетельство об официальной регистрации программ для ЭВМ №2005612134 от 19.08.2005), позволяющей построить математическую модель как по риску развития рака мочевого пузыря, так и по определению прогноза заболевания.

#### Результаты и обсуждение

В начале исследования вычислены парные коэффициенты корреляции для всех ин-

формационных признаков ( $x_j$ ) и выходящей переменной ( $Y$ ). При этом исключались факторы, для которых не получено достоверной корреляции с возникновением рака мочевого пузыря, и факторы, не влияющие на исход заболевания. Далее рассмотрен вопрос мультиколлинеарности. Явление мультиколлинеарности состоит в существовании линейной связи между объясняющими переменными. Наличие мультиколлинеарности вызывает технические трудности, связанные с уменьшением точности оценки тех или иных параметров или даже с невозможностью оценки вообще.

В результате из исследования были исключены факторы, не связанные с риском развития и исходом заболевания, а также связанные между собой. Всего было использовано 6 признаков ( $m = 6$ ) для оценки риска развития рака мочевого пузыря (профессиональные особенности, курение, особенности питания, питьевой режим, генетическая детерминация опухолевых образований, наличие инфравезикальных заболеваний мочеполовой системы) и 6 признаков ( $m = 6$ ) для прогноза исхода заболевания: курение сигарет, генетическая детерминация опухоли, глубина инвазии раковых клеток, рецидивная опухоль, локализация опухоли, характер лечения рака мочевого пузыря.

При исследовании влияния информативных факторов на развитие рака мочевого пузыря и выживаемость онкологических больных было использовано уравнение множественной линейной регрессии [2]. Но на практике в каждом конкретном случае следует предварительно убедиться, является ли зависимость между рассматриваемыми величинами  $x_j, j = \overline{1,6}$ ,  $Y$  линейной  $Y = ax + b$  [4].

Для всех независимых признаков  $x_j, j = \overline{1,6}$  и зависимой переменной ( $Y$ ) для уровня значимости  $\alpha = 0,01$  была проверена статистическая гипотеза о том, является ли зависимость между величинами ( $x_j, Y$ ) линейной. В случае, когда связь между величинами ( $x, Y$ ) не является линейной, можно воспользоваться квазилинейной зависимостью, так как известно, что при увеличении степени полинома точность аппроксимации увеличивается [4]:

$$y = a_0 + a_1x + a_2x^2 + \dots + a_kx^k = a_0 + \sum_{j=1}^k a_jx^j$$

Для всех перечисленных выше признаков, вплоть до  $k = 15$ , были вычислены коэффициенты соотношения по методу наименьших квадратов (МНК) [1,4], дисперсии

$$s_u^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$$

остатков, точность

$$L = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} * 100\% \quad [3], \text{ где } y_i -$$

исходные данные,  $\hat{y}_i$  - результаты расчетов.

Отметим, что дисперсию остатков ( $s_u^2$ ) можно рассматривать как относительную характеристику, тогда как величину  $L$  как абсолютную. Но увеличение порядка полинома ( $k$ ) практически не привело к уменьшению дисперсии остатков и величин ошибок. Поэтому можно воспользоваться уравнением множественной линейной регрессии.

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 = a_0 + \sum_{j=1}^6 a_jx_j$$

Для определения прогноза заболевания необходимо найти значения коэффициентов уравнения множественной линейной регрессии, для которых сумма квадратов отклонений опытных и теоретических значений зависимой переменной минимальна. По имеющимся данным, помимо МНК-оценок коэффициентов соотношения были найдены доверительные интервалы.

Для характеристики тесноты связи между зависимой величиной ( $Y$ ) и несколькими независимыми величинами ( $x_1, x_2, \dots, x_m$ ) вычислялся коэффициент множественной корреляции [1,4]. Для рассматриваемых данных коэффициент множественной корреляции оказался равным  $R_{y.12\dots m} = 0,82$  (F наблюдаемая = 32,0; F критическая = 1,49) для модели прогноза исхода заболевания и  $R_{y.12\dots m} = 0,70$  (F наблюдаемая = 135,82, что значительно превышает F критическую = 2,02) для оценки риска развития рака мочевого пузыря, что указывает на высокую тесноту

связи между величинами  $Y$  и  $x_1, x_2, \dots, x_m$  и в связи с этим может служить косвенным оправданием выбора функции в виде уравнения множественной линейной регрессии. Заметим, что для всех признаков коэффициент множественной корреляции оказался равным

$R_{y.12\dots m} = 0,83$  и практически не отличался от указанного выше, что косвенно подтверждает правильность исключения некоторых признаков из рассмотрения.

Изложенная методика может быть использована для оценки степени риска и прогнозирования состояния здоровья отдельного пациента. Для нахождения риска и прогноза, если коэффициенты  $a_0, a_1, \dots, a_m$  найдены по МНК, достаточно в уравнение линейной регрессии подставить отдельные значения признаков пациента в виде вектора  $x_j$  и вычислить  $\hat{y}$ . Для вычисленного риска и прогноза  $\hat{y}$  может быть построен доверительный интервал для уровня значимости  $\alpha=0,01$  с использованием неравенств  $\hat{y} - t_{f,\alpha} s_e \leq \tilde{y} \leq \hat{y} + t_{f,\alpha} s_e$ , где  $t_{f,\alpha}$  - квантиль  $t$ -распределения при заданном уровне значимости  $\alpha$  и числе степеней свободы  $f = n - m - 1$ ,  $s_e^2 = s_u^2 x'(X'X)^{-1} x + s_u^2$ , где  $X$  - матрица отдельных значений информативных признаков, символ "верхний штрих" означает операцию транспонирования, символ "-1" означает операцию вычисления обратной матрицы [1,4], тогда с вероятностью  $P = 0,99$  ( $1-\alpha$ ) можно утверждать, что истинное значение  $\tilde{y}$  при фиксированных значениях признаков  $x_j$  отдельного пациента находится в этом интервале.

С помощью специально разработанных программ рассчитаны величины риска развития и прогноза рака мочевого пузыря и доверительный интервал. Если значения информа-

тивных признаков состоят из значений натурального ряда, а результат в общем случае не будет принимать целые значения, то можно говорить лишь о попадании в интервал, границы которого равно отстоят от целочисленных значений. Риск развития рака мочевого пузыря считается минимальным, если значение ( $y$ ) менее 1,4; риск умеренный, когда число находится в интервале 1,5-2,4; высокий риск развития рака мочевого пузыря отмечается, если ( $y$ ) превышает 2,5. Прогноз считается хорошим (безрецидивная выживаемость 10 и более лет), если число находится в интервале 0-2,5; прогноз считается удовлетворительным (безрецидивная выживаемость от 5 до 10 лет), если число находится в интервале 2,5-3,5; прогноз считается сомнительным (выживаемость больного более 5 лет, но возник рецидив заболевания), если число находится в интервале 3,5-4,5; прогноз считается неудовлетворительным (летальный исход в течение 5 лет), если число превышает 4,5.

#### Заключение

С помощью корреляционно-регрессионного анализа создана математическая модель оценки риска и прогноза рака мочевого пузыря. Использование математической модели и разработанной на ее основе компьютерной программы позволяет рассчитать индивидуальный риск заболевания и прогноз для каждого пациента проспективно и разработать комплекс мер по первичной и третичной профилактике заболевания, учитывающий индивидуальные особенности пациента.

#### ЛИТЕРАТУРА

1. Закс Л. Статистическое оценивание.- М.: Статистика, 1976. - 598с.
2. Иванов О.А., Сухарев А.Е., Егоров С.Н. Влияние различных факторов на выживаемость онкологических больных // Российский онкологический журнал. - 1997. - № 5. - С.35-38.
3. Мацкевич И.П., Свирид Г.П., Булдык Г.М. Теория вероятностей и математическая статистика. - Минск: Вишэйшая школа, 1996. - 318 с.
4. Ферстер Э., Ренц Б. Методы корреляционного и регрессионного анализа. - М.: Финансы и статистика, 1983. - 302с.