

Інформаційні технології в медицині

УДК 519.237:616-006

А.И. Бых¹, Е.В. Высоцкая¹, А.П. Порван¹, А.С. Козюк¹, А.С. Моисеенко²

¹ Харьковський національний університет радіоелектроніки, Харків

² ГУ «Інститут загальної та неотложної хірургії НАМН України», Харків

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ОПРЕДЕЛЕНИЯ СТЕПЕНИ ТЯЖЕСТИ КОЛОРЕКТАЛЬНОГО РАКА

В современной клинической практике значение математического обеспечения медицинских исследований становится важнейшим фактором развития лечебно-диагностических технологий. В статье предложена математическая модель определения степени тяжести колоректального рака у человека, полученная в результате дискриминантного анализа клинических инструментальных и лабораторных показателей. Полученная математическая модель позволяет провести корректную дифференциацию степени тяжести и определить раннюю стадию развития заболеваемости по десяти информативным показателям метаболических маркеров обмена веществ в сыворотке крови пациентов.

Ключевые слова: дискриминантный анализ, колоректальный рак, территориальная карта.

Введение

Проблема изучения патогенеза, ранней диагностики и оптимизации лечения колоректального рака (КР) является на сегодняшний день чрезвычайно актуальной во всем мире. КР является широко распространенной патологией: ежегодная заболеваемость достигает 1 миллиона случаев, а ежегодная смертность превышает 500 тыс. человек. По прогнозам абсолютное число случаев КР в мире в следующие два десятилетия увеличится в результате роста населения в целом и его старения как в развитых, так и в развивающихся странах. КР занимает второе место по смертности от злокачественных новообразований [1]. В результате несвоевременной диагностики этого онкологического заболевания количество летальных исходов больных КР в Украине составляет более 40%. В развитых странах, где проводится обследование на наличие рака кишечника и предопухолевые состояния, этот показатель значительно ниже. Несмотря на это, отмечается рост нерезектабельных форм рака толстой кишки вследствие поздней диагностики заболевания, когда значительная часть больных приходит в клинику с осложнениями, что препятствует своевременному хирургическому вмешательству и проведению химио-лучевого лечения пациентов [2, 3]. Поэтому существует необходимость изучения патогенетических механизмов развития КР и обоснования использования маркерных показателей при его определении на ранней стадии развития.

Улучшение результатов диагностики и лечения больных с колоректальным раком невозможно без

многофакторной оценки тяжести состояния. Для этого целесообразно проведение анализа результатов клинических исследований с применением математических моделей и методов.

Анализ проблемы и постановка задачи исследования. Существуют различные способы определения степени тяжести рака кишечника по клинико-лабораторным показателям, которые включают в себя как традиционные инструментальные методы, так и методы, основанные на применении различных математических моделей.

Одним из таких методов является оценка распространенности КР по клинико-лабораторным данным с дополнительной диагностикой по Фоллю. При наличии асимметрии в 6-8 парах контрольно-измерительных точек, значения в которых превышают или равны 10 условным единицам, считают рак распространенным и ограничивают объем хирургического вмешательства до паллиативного [4].

Также известен метод прогнозирования безрецидивной выживаемости [5], который заключается в том, что у больных с метастазами КР в печени после ее резекции определяют максимальный размер, количество метастазов и их синхронность с первичной опухолью, наличие опухолевых клеток в лимфоузлах гепатодуоденальной связки и края резекции, степень дифференцирования метастазов. Далее определяют экспрессию маркеров СК 20, Ki 67 Max, причем для СК 20 определяют экспрессию в ядре, а для Ki 67 Max - индекс пролиферации 30% или > 30%. Затем в зависимости от степени дифференциации метастазов находят уровень значимости молекулярных факторов и соответствующие значения в

баллах, после чего полученные баллы суммируются.

К недостаткам рассмотренных методов можно отнести отсутствие возможности диагностики ранних форм КР и степени тяжести заболевания.

Существует ряд математических моделей и методов, характеризующих процесс определения степени тяжести онкопатологии, в том числе КР.

На основе законов распределения и значений вероятностей присутствия того или иного признака используются различные статистические методы – регрессионный, кластерный и дискриминантный анализы, которые показали высокую адекватность при решении различных медицинских задач определения информативных показателей и их классификации [6 – 8].

Рассмотрим некоторые из них. Известна регрессионная математическая модель на основе метода пропорционального риска Кокса, которая позволяет по 27 показателям результатов проведения инструментальных и лабораторных исследований определить прогрессирование метастатического КР после окклюзии правой ветви воротной вены. [9].

Другая известная модель, полученная с помощью метода множественной логистической регрессии, позволяет по результатам молекулярно-генетических скрининговых тестов онкомаркеров РЭА крови, СА, фекального Tumor M2-ПК и наличия в крови метилированной ДНК SEPT9 установить риск прогрессирования КР [10]. Основным недостатком данной модели является высокая стоимость проведения молекулярно-генетических скрининговых тестов.

Для выявления информативных показателей, качественной классификации исследуемых объектов и синтеза адекватной линейной математической модели в медицине успешно применяется метод дискриминантных функций. Он основан на составлении уравнения регрессии, для которого используют некоторую номинативную зависимую переменную, не являющуюся количественной, как в случае регрессионного анализа. Уравнение регрессии составляется на основе тех объектов, о которых известна групповая принадлежность, что позволяет максимально точно подобрать его коэффициенты. После того как уравнение регрессии получено, его используют для группировки интересующих объектов в целях прогнозирования.

Применение данного метода позволяет разработать математически обоснованные подходы к повышению эффективности хирургического лечения злокачественных опухолей, улучшению качества проведения диагностических мероприятий и т.д. Так, известна математическая модель, позволяющая с вероятностью 98.2% определить выживаемость пациента с резистентной формой колоректального рака в послеоперационный период по шести показа-

телям иммуно-ферментного анализа крови и гистологического анализа биоматериала [11]. Гистологическое исследование является для пациента травматичным методом, что является его недостатком.

К общему недостатку рассмотренных выше моделей и методов можно отнести невозможность проведения диагностики КР на ранних стадиях.

В связи с вышесказанным, целью данной работы является синтез математической модели определения степени тяжести КР, включая ранние стадии развития, которая позволит повысить эффективность определения данной онкопатологии, что будет способствовать своевременному и качественному лечению.

Основной материал

Для построения математической модели определения степени тяжести больных с КР были изучены клинические данные, а также данные определения метаболических маркерных показателей обмена веществ в сыворотке крови 290 пациентов. Из общего числа обследованных были сформированы три исследуемые группы.

Первую группу (n = 80) составили пациенты второй степенью тяжести КР. Вторую группу составили пациенты с третьей степенью тяжести КР (n = 150), а третью – пациенты с четвертой степенью тяжести, отягченной различными метастазами (n = 60).

В ходе построения математической модели определения степени тяжести больных с КР проводилось изучение взаимосвязи между качественным показателем выявленной степени тяжести, выступающим в роли зависимого (результатирующего) и подмножеством количественных признаков (аспарагиновая аминотрансфераза, аланиновая аминотрансфераза, билирубин, гаммаглутомилтрансфераза, щелочная фосфатаза, глюкоза, креатин фосфокиназа, лактатдегидрогеназа, мочевины, креатинин, общий белок, магний, фосфор, холестерин, альбумин, железо, возраст) и закодированных качественных признаков (пол, период деспансеризации).

Для отбора наиболее информативных признаков фильтрацию входных данных дискриминантного анализа проводили по принципу пошагового включения диагностически значимых переменных с критическим уровнем значимости для отбора предикторов равным 5%. Вся математическая обработка результатов проводилась с использованием пакета прикладных программ SPSS 19.0 [12].

Далее отобранное подмножество информативных признаков использовали для вычисления дискриминантных функций. В итоге в диагностическую модель определения КР вошли 10 наиболее значимых признака: аспарагиновая аминотрансфераза (X1), аланиновой аминотрансфераза (X2), билирубин (X3), гаммаглутомилтрансфераза (X4), глюкоза (X5), креатин фосфокиназа (X6), креатинин (X7),

фосфор (X8), холестерина (X9), альбумин (X10). Выделенные 10 признаков имели статистически значимые дискриминантные свойства при определении групп пациентов (значение лямбды Уилкса составило 0,157 при $p < 0.001$), что свидетельствует о хорошей дискриминации.

Класификационная способность модели, основанной на наиболее значимых для диагностики признаках, составила 93.4%.

Искомые канонические дискриминантные функции, согласно вычисленным в ходе дискриминантного анализа коэффициентам, были представлены следующими выражениями:

$$F1 = 0.001 * X1 + 0.01 * X2 + 0.002 * X3 + 0.047 * X4 - 0.349 * X5 + 0.005 * X6 - 0.009 * X7 + 0.494 * X8 - 0.156 * X9 - 0.029 * X10 + 1.768,$$

$$F2 = 0.004 * X1 + 0.017 * X2 - 0.004 * X3 - 0.027 * X4 + 0.311 * X5 + 0.004 * X6 - 0.004 * X7 - 0.535 * X8 - 0.091 * X9 + 0.002 * X10 - 0.972,$$

где F1 и F2 – первая и вторая канонические дискриминантные функции,

X1 – X10 – значения дискриминантных переменных для каждого наблюдения в исследуемых группах.

Корреляционные коэффициенты между рассчитанными значениями дискриминантных функций и показателями принадлежности к группам являются мерой удачного распределения на группы. Характеристики дискриминантных возможностей синтезированных функций приведены в табл. 1.

Таблица 1

Характеристика дискриминантных возможностей синтезированных функций

Функция	Собственные значения	% дисперсии	Общий %	Каноническая корреляция
1	2,294	71,0	71,0	0,835
2	0,938	29,0	100,0	0,696

Полученные собственные значения приведены в порядке убывания их величин. Величина собственного значения связана с дискриминантной возможностью функции: чем больше собственное значение, тем лучше различие. Таким образом, первая функция имеет наибольшие дискриминантные возможности, вторая – обеспечивает максимальную разницу после первой.

Фактические числовые значения, равные 2.294 и 0.938, указывают на то, что дискриминантные возможности первой дискриминантной функции в 2.45 раз больше другой.

Для того чтобы определить, имеет ли весомое влияние вторая функция, определялся канонический коэффициент корреляции функции. При оценке значений коэффициентов, которые составили 0.835

и 0.696, можно сделать вывод, что существует высокая положительная связь между степенью тяжести КР и значением первой дискриминантной функции, что подтверждается процентом дисперсии, равной 71%. Также была установлена хорошая связь между степенями тяжести КР и значением второй дискриминантной функции.

На основании значений двух дискриминантных функций была построена территориальная карта разделения пациентов на группы (рис. 1).

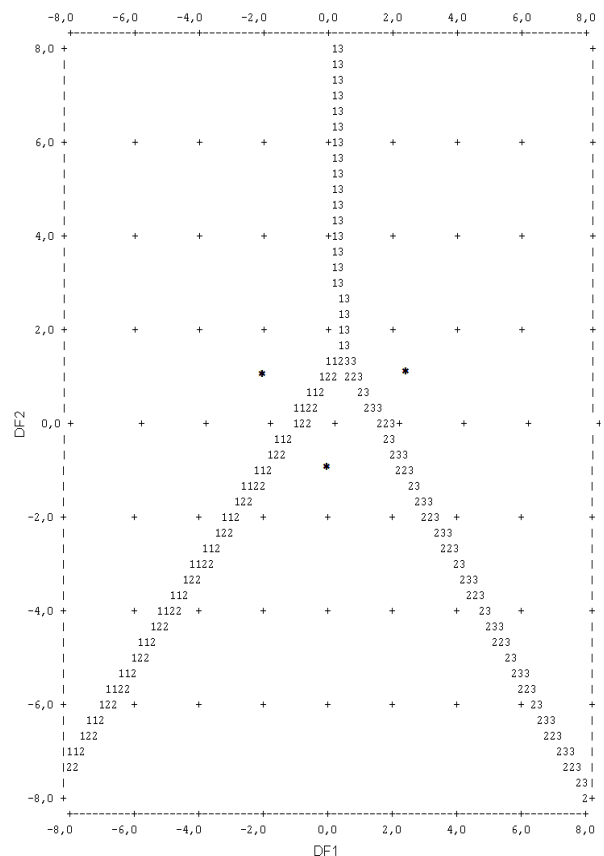


Рис. 1. Территориальная карта классификации объектов исследования:

- 1 – вторая степень тяжести,
- 2 – третья степень тяжести,
- 3 – четвертая степень тяжести,
- * – центр оид группы

На территориальной карте обозначены три области, определяющие принадлежность объекта (пациента) к той или иной группе.

При этом в пределах соответствующей области вероятность отнесения пациента к данной группе выше, чем для других, и одинакова на границах разделения.

Выводы

Таким образом, проведенный дискриминантный анализ наглядно продемонстрировал и подтвердил гипотезу о том, что все три группы различаются между собой.

Полученная математическая модель позволяет провести корректную дифференциацию степени тяжести больных с КР по 10 информативным показателям метаболических маркеров обмена веществ сыворотки крови пациентов, что дает возможность повысить эффективность диагностики данной онкопатологии и назначить своевременное качественное лечение.

Список литературы

1. Давыдов М.И. Заболеваемость злокачественными новообразованиями и смертность от них населения стран СНГ. [Электронный ресурс] / М.И. Давыдов, Е.М. Аксель. – Режим доступа к ресурсу: <http://www.ronc.ru/attachments/article/1708/%D0%A0%D0%A4%20%D0%B8%20%D0%A1%D0%9D%D0%93%20%202005.doc>.
2. Fontham E.T.H. Preventive Nutrition [Текст] / E.T.H. Fontham, Eds. A. Bendich, R.J. Deckelbaum. – Totowa, New Jersey: Humana Press. – 1997. – P. 33-55
3. Ferlay J., Bray F., Pisani P., Parkin D.M. Globocan 2002. Cancer Incidence, Mortality and Prevalence Worldwide. IARC CancerBase №5, version 2.0. IARC Press, Lyon, 2004.
4. Патент на изобретение Российской Федерации № 2336814. Диагностика рака желудка на ранней стадии. Ристимьяки Ари (Fi), Хярккенен Матти (Fi), Сиппонен Пентти (Fi). МПК А61В5/053 (2006.01). – Патентообладатель: БИОХИТ ОЙЙ (Fi). – Оpubл. 27.10.2008.
5. Патент на изобретение Российской Федерации № 2452373. Способ прогнозирования безрецидивной выживаемости по клинико-патологическим факторам у больных с метастазами колоректального рака в печени после ее резекции. Галанкина И.Е. (RU), Козлов М.В. (RU), Чжао А.В. (RU), Чузунов А.О. (RU), Коваленко Ю.А. (RU). МПК А 61В5/107, G01N 1/28. – Патентообладатель: Государственное учреждение здравоохранения Научно-исследовательский институт скорой помощи имени Н.В. Склифосовского Департамента здравоохранения г. Москвы. – Оpubл. 10.06.2012.
6. Порван А.П. Использование дискриминантного анализа для диагностики хронической сердечной недостаточности у подростков [Текст] / А.П. Порван, А.И. Бых,

Л.И. Рак, Е.В. Высоцкая // Вестник национального технического университета «ХПИ». – Х.: НТУ «ХПИ», 2010. – № 31. – С. 16-22.

7. Щукин Н.А. Использование дискриминантного анализа для диагностики доброкачественных и злокачественных опухолей [Текст] / Н.А. Щукин, Е.В. Высоцкая, А.П. Порван, С.Н. Пушкарь // Системы обработки информации. – Х.: ХУПС, 2011. – Вып. 2(92). – С. 234-238.
8. Высоцкая Е.В. Методика определения систолической дисфункции миокарда у подростков [Текст] / Е.В. Высоцкая, А.П. Порван, Л.И. Рак [та ін.] // Восточно-Европейский журнал передовых технологий. – 2012. – № 1/3(55). – С. 27-31.
9. Полищук Л.О. Прогрессирование метастатического колоректального рака после окклюзии правой ветви воротной вены [Текст] / Л.О. Полищук, О.Г. Скипенко, А.Г. Абдуллаев // Хирургия. – 2011. – № 10. – С. 4-12.
10. Омельченко В.П. Моделирование прогноза рака толстой кишки после радикального оперативного лечения по результатам скрининговых тестов [Электронный ресурс] / В.П. Омельченко, Д.В. Бурцев, О.И. Кум, А.А. Демидова // Фундаментальные исследования. – 2012. – №8 – Режим доступа к журналу: URL: http://www.rae.ru/fs/?section=content&op=show_article&article_id=9999589.
11. Воронова Е.И. Математико-статистические модели прогнозирования эффективности оперативного лечения некоторых заболеваний [Электронный ресурс] / Е.И. Воронова, Л.А. Муравей, Ю.А. Костилов // Электронный научный журнал «Инженерный вестник Дона», 2007–2013. – 2013. – №1. – Режим доступа к журналу: [www/URL: http://ivdon.ru/magazine/archive/n1y2013/1537](http://ivdon.ru/magazine/archive/n1y2013/1537).
12. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей [Текст] / А. Бююль, П. Цефель. – СПб.: ООО «ДиаСофтЮП», 2001. – 608 с.

Поступила в редколлегию 19.09.2014

Рецензент: д-р биол. наук, д-р мед. наук, проф. В.И. Жуков, Харьковский национальный медицинский университет, Харьков.

МАТЕМАТИЧНА МОДЕЛЬ ВИЗНАЧЕННЯ МІРИ ТЯЖКОСТІ КОЛОРЕКТАЛЬНОГО РАКА

А.І. Бих, А.С. Моїсеєнко, О.В. Висоцька, А.П. Порван, А.С. Козюк

У сучасній клінічній практиці значення математичного забезпечення медичних дослідів стає важливим фактором розвитку лікувально-діагностичних технологій. У статті запропонована математична модель визначення міри тяжкості колоректального рака у людини, що була отримана в результаті дискримінантного аналізу клінічних інструментальних і лабораторних показників. Отримана математична модель дозволяє провести неінвазивну коректну диференціацію міри тяжкості та визначити ранню стадію розвитку захворювання за десятьма інформативними показниками метаболічних маркерів обміну речовин у сироватці крові пацієнтів.

Ключові слова: дискримінантний аналіз, колоректальний рак, територіальна карта.

MATHEMATICAL MODEL FOR DETERMINING SEVERITY OF COLORECTAL CANCER

A.I. Bikh, A.S. Moiseenko, E.V. Visotskaya, A.P. Porvan, A.S. Kozuyk

In modern clinical practice, the value of the mathematical providing of medical researches becomes the major factor of development of curatively-diagnostic technologies. In the article, the mathematical model of determination of severity of colorectal cancer is offered for a man, got as a result of discriminant analysis of clinical instrumental and laboratory indexes. The got mathematical model allows conducting correct differentiation of severity and defining the early stage of development of morbidity on ten informing to the indexes of metabolic markers of metabolism in the serum of blood of patients.

Keywords: discriminant function analysis, colorectal cancer, the territorial map.