

УДК 618.19-007.42-089.844:51+004.94
DOI: 10.26435/UC.V014(41).789

М.И. Жуков, С.М. Тетюра, О.В. Стефкивская

ГОО ВПО «Донецкий национальный медицинский университет имени М. Горького», Донецк

ВОЗМОЖНОСТЬ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ХИРУРГИЧЕСКОЙ КОРРЕКЦИИ ВТОРИЧНОГО ПТОЗА МОЛОЧНЫХ ЖЕЛЕЗ

Для решения задачи прогнозирования результата хирургической коррекции вторичного птоза молочных желез при проведении маммопластики была предпринята попытка построения многофакторных математических моделей.

Целью работы было создать математическую модель исхода хирургической коррекции вторичного птоза молочных желез.

МАТЕРИАЛ И МЕТОДЫ

Исследования проводили посредством применения корреляционного, регрессионного и дисперсионного анализа. В процессе проведения хирургической коррекции участвовали данные 121 пациентки. В качестве качественных и количественных предикторов рассматривали следующие показатели:

- группа;
- возраст;
- степень подбора импланта;
- локализация импланта;
- физиологические причины птоза;
- характер технических ошибок при маммопластике;
- снижение общей массы тела;
- общий белок крови;
- эритроциты;
- лейкоциты крови;
- тромбоциты крови.

В качестве зависимой переменной, то есть критерия эффективности хирургической коррекции, использовали показатель «результат коррекции», который определяли как «хороший», «удовлетворительный» и «плохой».

Так как лейтмотивом исследования является сравнительный анализ основной группы и группы сравнения относительно эффективности лечения, в качестве одного из категориальных предикторов рассматривали переменную «группа», принимающую два значения – основная (1) и сравнения (2). В качестве второго предиктора принимали результативность хирургического лечения.

В качестве входных предикторов были выбраны 11 факторов, которые разделили на категориальные (присутствовали у пациента (признак равен 1) либо отсутствовали (признак равен 2)) и количественные (принимали определенные непрерывные числовые значения). Перед исследователями была поставлена задача прогнозирования переменной dR как «хорошего» результата (прогнозируемая переменная принимает значение «1») и «неудачного» результата (прогнозируемая переменная, в зависимости от степени выраженности неудачи, принимает значение «2» – «удовлетворительный» результат «3» – «плохой» результат). Математическая модель строилась на данных 121 истории болезни.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

При построении модели важнейшим фактором является правильный выбор ее структуры (линейная, параболическая полиномиальная и т.д.) и отбор наиболее значимых параметров модели [1, 2].

При выборе наиболее значимых входных переменных модели наиболее точным является путь перебора всех возможных типов моделей, построенных на выбранном наборе переменных, и отбор наиболее приемлемых по прогностическим способностям и наиболее простых по количеству учитываемых признаков [3, 4].

Для ускорения поиска приемлемого вида модели разработаны специальные подходы. Так, в методе пошагового включения первоначально строится модель, содержащая только свободный член. Затем постепенно один за другим добавляют остальные регрессоры. Процедура включения останавливается, когда добавление в модель новых факторных признаков уже существенно не уменьшает вариации отклика, что определяется по F-критерию Фишера. Метод пошагового исключения действует в обратном порядке.

На предварительном этапе выбора математической модели методом пошагового включения была построена многомерная линейная модель на всех 11-ти входных переменных, затем методом пошагового исключения были отобраны 4 переменных.

В случае построения линейной модели для выделения наиболее значимых переменных пригоден метод пошагового включения переменных, с помощью которого, как сказано выше, было отобрано 11 наиболее значимых признаков: группа (обозначим эту переменную X1); возраст (обозначим эту переменную X2); подбор импланта (обозначим эту переменную X3) и т.д.; и последний фактор – «лейкоциты, Г/л» (обозначим эту переменную X11). Выходную переменную обозначим через V1. Множественное значение коэффициента корреляции (R) составляет 0,9946, а скорректированное значение коэффициента детерминации (R2) – 0,9882. Значение F-критерия в первой модели составило 916,5 при $p=0,001$. Скорректированное значение показывает сокращение предсказательной мощности модели, которое незначительно отличается от грубого значения R2. Грубое значение показателя говорит нам, какая доля вариативности V1 может быть объяснена регрессионной моделью, построенной на данных нашей выборки. Скорректированное значение говорит о том, какую долю вариативности V1 объясняла бы эта модель, если бы она была построена на данных всей популяции, из которой была извлечена выборка.

Формула для расчета показателя «dR» в 1-й модели включает 11 входных предикторов. Абсолютные величины коэффициентов регрессии позволили расположить предикторы в ряд по убыванию их относительного вклада (β -коэфф.) в прогнозирование итоговой переменной регрессионного уравнения: «технические ошибки маммопластики» > «физиологические причины птоза» > «снижение массы тела» > «лейкоциты» > «эритроциты» > «общий белок крови»: $|0,600| > |0,359| > |0,134| > |0,090| > |0,057| > |0,052|$.

После расчета линейной модели были получены уравнение 1:

$$V1 = 5,76 + 0,01 \times X1 + 0,0001 \times X2 + 0,001 \times X3 - 0,02 \times X4 - 1,36 \times X5 - 0,01 \times X6 - 0,58 \times X7 + 0,04 \times X8 - 0,14 \times X9 + 0,001 \times X10 + 0,04 \times X11.$$

Учитывая, что статистическая значимость некоторых предикторов, таких как «группа», «возраст», «подбор импланта», «локализация импланта», «количество тромбоцитов», выше порогового уровня $p > 0,05$, данные предикторы оказывают незначительное влияние на зависимую переменную. Поэтому ими в уравнении можно

пренебречь. Таким образом, уравнение модели 1 будет представлено в следующем виде:

$$V1 = 5,76 - 1,36 \times X5 - 0,01 \times X6 - 0,58 \times X7 + 0,04 \times X8 - 0,14 \times X9 + 0,04 \times X11 + Z,$$

где Z(0,013) – поправочный коэффициент.

На следующем этапе моделирования методом пошагового исключения факторов для уравнения регрессии было отобрано 4 переменных: «технические ошибки маммопластики» (X1), «физиологические причины птоза» (X2), «снижение массы тела» (X3), «лейкоциты» (X4). Выходную переменную обозначим через V2.

Формула для расчета показателя «dR» во 2-й модели включает 4 входных предиктора. Абсолютные величины коэффициентов регрессии позволили расположить предикторы в ряд по убыванию их относительного вклада (β -коэфф.) в прогнозирование итоговой переменной регрессионного уравнения: технические ошибки маммопластики > физиологические причины птоза > снижение массы тела, кг > лейкоциты, Г/л: $|0,66| > |0,42| > |0,134| > |0,14| > |0,11|$.

После расчета линейной модели были получены уравнение 2:

$$V2 = 5,09 - 1,50 \times X1 - 0,68 \times X2 - 0,04 \times X3 + 0,05 \times X4,$$

где X1 – «технические ошибки маммопластики», X2 – «физиологические причины птоза», X3 – «снижение массы тела», X4 – «количество лейкоцитов». V2 – выходная переменная «результат коррекции».

Анализ качества модели указывает на достаточно высокие прогностические способности линейной модели на 4 признаках без ухудшения основных регрессионных показателей. При этом, исходя из уравнения 2, можно сделать вывод, что наличие даже только одного признака X1, X2, X3 в рамках предложенной модели приводит к прогнозу показателя dR, а значит, и к определенному виду результата коррекции. Для остальных выделенных признаков в уравнении 1 или 2 для прогнозирования показателя dR необходимо подкрепление его еще каким-либо признаком.

Расчет статистики по критерию Дарбин-Уотсон выявил, что коэффициент принимал значение 1,33 для 1-й модели и 1,32 для 2-й модели, при этом коэффициент сериальной корреляции был на уровне 0,162 в первом случае и 0,177 – во втором. Это подтверждает гипотезу, которая указывает на отсутствие автокорреляции остатков в моделях.

Таким образом, методом пошагового исключения из 11 факторов многофакторная линейная модель прогнозирования показателя «Результат коррекции» была оптимизирована до 4 предикторов без ухудшения итоговых статистик

модели, поэтому модель прогнозирования может быть описана уравнением 2.

Высокое значение коэффициента множественной корреляции прогнозируемых данных в модели ($R=0,994$ и $R^2=0,987$) свидетельствует об адекватности математической модели.

Линейная корреляционная связь между остатками и факторным признаком отсутствует. Линейная корреляционная связь между остатками и результирующим признаком отсутствует.

ВЫВОДЫ

1. Методом пошагового включения было отобрано 11 наиболее значимых факторов для предварительного прогнозирования показателя «результата хирургической коррекции».

2. Методом пошагового исключения из 11 факторов была оптимизирована многофакторная линейная модель прогнозирования показателя «результата коррекции» до 4 признаков без ухудшения итоговых статистик модели (модель может быть описана уравнением 2).

3. Модель обладает чувствительностью 99,9% и специфичностью 98,3%. Адекватность модели подтверждена на тестовом множестве.

М.И. Жуков, С.М. Тетюра, О.В. Стефкивская

ГОО ВПО «Донецкий национальный медицинский университет имени М. Горького», Донецк

ВОЗМОЖНОСТЬ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ХИРУРГИЧЕСКОЙ КОРРЕКЦИИ ВТОРИЧНОГО ПТОЗА МОЛОЧНЫХ ЖЕЛЕЗ

Для решения задачи прогнозирования результата хирургической коррекции вторичного птоза молочных желез при проведении маммопластики была предпринята попытка построения многофакторных математических моделей. Целью работы было создать математическую модель исхода хирургической коррекции вторичного птоза молочных желез. Исследования проводили посредством применения корреляционного, регрессионного и дисперсионного анализа. В процессе проведения хирургической коррекции в анализе использовали данные 121 пациентки. Методом пошагового включения было отобрано 11 наиболее значимых факторов для предваритель-

ного прогнозирования показателя «результат хирургической коррекции». Методом пошагового исключения из 11 факторов была оптимизирована многофакторная линейная модель прогнозирования показателя «результат коррекции» до 4 признаков без ухудшения итоговых статистик модели. Модель обладает чувствительностью 99,9% и специфичностью 98,3%. Адекватность модели подтверждена на тестовом множестве.

Ключевые слова: вторичный птоз молочных желез, хирургическая коррекция, математическое моделирование.

M.I. Zhukov, S.M. Tetyura, O.V. Stefkivskaya

SEI HPE «M. Gorky Donetsk National Medical University», Donetsk

POSSIBILITY OF MATHEMATICAL MODELING OF SURGICAL CORRECTION OF SECONDARY PTOSIS OF THE MAMMARY GLANDS

To solve the problem of predicting the result of surgical correction of secondary ptosis of the mammary glands during mammoplasty, an attempt was made to build multifactorial mathematical models. The aim of the work was to create a mathematical model of the outcome of surgical correction of secondary ptosis of the mammary glands. The studies were carried out using correlation, regression analysis and analysis of variance. During the surgical correction, data from 121 patients were used in the analysis. By the method of step-by-step inclusion, 11 most significant factors were selected for preliminary

prediction of the indicator “Outcome of surgical correction”. The method of step-by-step exclusion from 11 factors was used to optimize the multifactorial linear model for predicting the “Result of correction” indicator up to 4 features without deteriorating the final statistics of the model. The model has a sensitivity of 99.9% and a specificity of 98.3%. The adequacy of the model is confirmed on the test set.

Key words: secondary ptosis of the mammary glands, surgical correction, mathematical modeling.

ЛИТЕРАТУРА

1. Антомонов М.Ю. Математическая обработка и анализ медико-биологических данных. К.: Малий друк; 2006. 558.
2. Боровиков В.П., Боровиков И.П. Statistica. Статистический анализ и обработка данных в среде Windows. М.: Филин; 1998. 595.
3. Лях Ю.Е., Гурьянов В.Г. и др. Основы компьютерной биostatистики: анализ информации в биологии, медицине и фармации статистическим пакетом MedStat. Донецк: Папакица Е.К. 2006; 214.
4. Реброва О.Ю. Статистический анализ медицинских данных. Применение пакета прикладных программ STATISTICA. М.: МедиаСфера; 2002. 305.

REFERENCES

1. Antomonov M.Yu. Matematicheskaya obrabotka i analiz mediko-biologicheskikh dannykh. K.: Malii druk; 2006. 558 (in Russian).
2. Borovikov V.P., Borovikov I.P. Statistica. Statisticheskii analiz i obrabotka dannykh v srede Windows. M.: Filin; 1998. 595 (in Russian).
3. Lyakh Yu.E., Gur'yanov V.G. i dr. Osnovy komp'yuternoi biostatistiki: analiz informatsii v biologii, meditsine i farmatsii statisticheskim paketom MedStat. Donetsk: Papakitsa E.K. 2006; 214 (in Russian).
4. Rebrova O.Yu. Statisticheskii analiz meditsinskikh dannykh. Primenenie paketa prikladnykh programm STATISTICA. M.: MediaSfera; 2002. 305 (in Russian).