

УДК

Д. А. Чичеватов

МОДЕЛЬ ШКАЛЫ ПРОГНОЗИРОВАНИЯ БИНАРНЫХ ПЕРЕМЕННЫХ В МЕДИЦИНСКИХ ИССЛЕДОВАНИЯХ

Областной онкологический диспансер, Пенза

Одной из наиболее типичных задач многих медицинских исследований является вероятностное прогнозирование значения какой-либо бинарной переменной, которая по определению имеет два значения: «да» (в случае развития признака) и «нет» (при отсутствии признака). Примерами подобных переменных являются развитие какого-нибудь конкретного послеоперационного осложнения (да, нет), возникновение заболевания или его рецидива, наличие какого-либо симптома. Типичная задача исследования обычно состоит в том, чтобы выяснить, как несколько объясняющих независимых переменных (предикторов) влияют на значение исследуемой зависимой бинарной переменной.

Традиционным подходом в решении подобной задачи является выполнение регрессионного анализа. Поскольку зависимая переменная является бинарной, то наиболее подходящей становится бинарная логистическая регрессионная модель, которая удобно реализована, например, в компьютерном программном пакете SPSS [1]. Результатом регрессионного анализа является уравнение вида

$$p = 1 / (1 + e^{-B_0 - B_1 x_1 - \dots - B_n x_n})$$

где p — теоретическая вероятность значения зависимой переменной, B_0 — значение свободного члена уравнения, B_i — коэффициент регрессии, x_i — значение переменной-предиктора. Применение такого подхода оправдано в научных исследованиях, но чрезвычайно неудобно в практическом здравоохранении, поскольку предполагает выполнение подсчета по достаточно сложной формуле.

Другой подход состоит в использовании автоматизированных систем прогнозирования, построенных, например, на основе нейросетевого анализа [2]. В настоящее время распространены компьютерные программные пакеты, такие, например, как STA-TISTICA Neural Networks [3], которые позволяют построить и обучить нейронную сеть и использовать ее как средство прогнозирования. Задача исследователя в подобном случае заключается только во введении в программу значений переменных-предикторов. Недостатком подобного метода в практическом здравоохранении является обязательное наличие персонального компьютера и необходимость знания основ нейросетевого анализа, который не получил пока повсеместного распространения.

Оптимальным видится подход, состоящий в использовании шкалы баллов, на основании которой осуществляется прогнозирование. Примером может служить Мангеймский индекс перитонита [4], шкала APACHE II [5].

Недостаток данного подхода состоит в том, что известные на сегодняшний момент шкалы разработаны применительно к определенным нозологическим формам, они не являются универсальными.

© Д. А. Чичеватов, 2007

В то же время имеется объективная необходимость простой разработки подобной шкалы для практически любого интересующего исследователя признака, описываемого бинарной переменной.

В настоящем исследовании приводится модель, посредством которой возможно создать шкалу баллов для оценки значения любого признака, который может быть рассмотрен как бинарная переменная.

Материалы и методы исследования. Смысл предлагаемой модели раскрывается на конкретном примере. Имеется выборка из 394 пациентов, перенесших пневмонэктомию в Пензенском областном онкологическом диспансере в период с 1993 по 2005 г. Задача состоит в создании шкалы, позволяющей предсказывать развитие несостоятельности культи главного бронха с известной вероятностью. Исследуемая бинарная переменная имеет 2 значения: «да» (есть несостоятельность) и «нет» (отсутствие данного осложнения). Все статистические расчеты проводились с использованием программы SPSS 10.0.

1. Определяется список переменных, которые рассматриваются как потенциальные предикторы. В данном случае таких переменных было 28, в скобках приведены их возможные значения:

- 1) пол пациента (мужской или женский);
- 2) возраст больного (интервальная переменная);
- 3) стадия рака легкого (I, II, III, IV);
- 4) критерий T по международной системе стадирования опухолей TNM (1, 2, 3, 4);
- 5) наличие у больного раннего рака T1 или T2 (да, нет);
- 6) критерий N по международной системе стадирования опухолей TNM (0, 1, 2, 3);
- 7) наличие у больного регионарных метастазов рака, т. е. N>0 (да, нет);
- 8) критерий M по международной системе стадирования опухолей TNM (0, 1);
- 9) наличие осложнений опухоли (да, нет);
- 10) сторона операции (справа, слева);
- 11) наличие у больного хронического бронхита (ХОБЛ) (да, нет);
- 12) наличие у больного ишемической болезни сердца (ИБС) (да, нет);
- 13) проведение предоперационной лучевой терапии (ДГТ) (да, нет);
- 14) доза предоперационной лучевой терапии (0 Гр, 7,5 Гр, 20 Гр, 40 Гр);
- 15) проведение предоперационной химиотерапии (ПХТ) (да, нет);
- 16) индекс массы тела больного (интервальная переменная);
- 17) выполнение расширенной медиастинальной лимфодиссекции (да, нет);
- 18) поликомбинированная операция (да, нет);
- 19) резекция перикарда (да, нет);
- 20) резекция верхней полой вены (ВПВ) (да, нет);
- 21) резекция предсердия (да, нет);
- 22) резекция пищевода (да, нет);
- 23) резекция легочной артерии (да, нет);
- 24) резекция бифуркации трахеи (да, нет);
- 25) уровень лейкоцитов крови (интервальная переменная);
- 26) уровень лимфоцитов крови (интервальная переменная);
- 27) уровень лейкоцитарного индекса интоксикации (ЛИИ) (интервальная переменная);
- 28) уровень общего белка крови (интервальная переменная).

Предстояло определить, какие из перечисленных независимых переменных и в какой степени могут влиять на развитие несостоятельности швов культи бронха (зависимая

Таблица 1

Статистическая связь зависимой
переменной с потенциальными
предикторами

№ п/п	Предиктор	P
1	Пол	0,148
2	Возраст	0,946
3	Стадия рака	0,050
4	Критерий Т (по системе TNM)	0,013
5	Группы T1+2 и T3+4	0,024
6	Критерий N (по системе TNM)	0,059
7	Группы N=0 и N>0	0,007
8	Критерий M	0,572
9	Осложнения опухоли	0,455
10	Сторона операции	0,000
11	Наличие ХОБЛ	1,000
12	Наличие ИБС	0,659
13	Проведение ДГТ	0,038
14	_____ Доза ДГТ _____	0,019
15	Проведение ПХТ	0,595
16	Индекс массы тела	0,216
17	Объем лимфодиссекции	0,101
18	Поликомбинированная резекция 2-х и более соседних органов	0,024
19	Резекция перикарда	0,891
20	Резекция ВПВ	0,000
21	Резекция предсердия	0,427
22	Резекция пищевода	0,246
23	Резекция легочной артерии	0,494
24	Резекция бифуркации трахеи	0,000
25	Уровень лейкоцитов крови	0,349
26	Уровень лимфоцитов крови	0,058
27	ЛИИ	0,109
28	Уровень протеинемии	0,185

переменная). Естественно, что у всех 394 пациентов значения абсолютно всех переменных были известны.

2. Для редукции переменных-предикторов и повышения качества дальнейших исследований было проведено разделение 394 больных на 2 группы в зависимости от значения зависимой переменной, т. е. на больных, имевших несостоятельность швов (НШ) культы бронха, и на больных, не имевших ее. Эти две группы сравнивали по значениям каждого из 28 потенциальных предикторов. Для номинальных переменных использовался анализ таблиц сопряженности, для порядковых и интервальных — тест Kruskal-Wallis. Результат проведенного анализа представлен в табл. 1. При уровне значимости 0,05 достоверно были связаны с зависимой переменной только стадия рака, критерий Т, наличие раннего рака, наличие регионарных метастазов, сторона операции, лучевая терапия, поликомбинированность операции, резекция верхней полой вены, резекция бифуркации трахеи.

3. Предикторы, имевшие статистическую связь с зависимой переменной с уровнем значимости $p = 0,15$ и меньше, далее были включены в регрессионную модель. В качестве регрессионной модели была выбрана регрессия с оптимальным шкалированием (Regression with Optimal Scaling (CATREG)), которая реализована в статистической компьютерной программе SPSS версии 10.0. Поскольку данная регрессионная модель оперирует с категориальными переменными, все включенные интервальные и порядковые предикторы были категоризированы. Из нескольких предикторов, у которых заранее была известна высокая частная корреляция (например предикторы № 4 и 5, № 6 и 7) в регрессионную модель был отобран тот, который имел наименьший уровень значимости по табл. 1. Таким образом, с учетом категоризации в регрессионную модель были включены 10 из 28 потенциальных предикторов:

1) пол; 2) критерий

T>2; 3) критерий N>0; 4) сторона операции; 5) лучевая терапия 40 Гр и более; 6) наличие расширенной лимфодиссекции; 7) поликомбинированная операция; 8) резекция бифуркации трахеи; 9) количество лимфоцитов в мазке крови < 8; 10) лейкоцитарный индекс интоксикации по Кальф-Калифу >1,4. Зависимой бинарной переменной была несостоятельность швов культи бронха.

4. Регрессионная модель с оптимальным шкалированием выбрана в связи с тем, что, помимо стандартизированных коэффициентов регрессии, в результате анализа получают так называемые «коэффициенты важности» (importance), представленные в последнем столбце результирующей анализ таблицы (табл. 2). Совершенно очевидно,

Таблица 2

Результат регрессионного анализа с оптимальным шкалированием (фрагмент результирующей таблицы)

№ п/п	Предиктор	Standardized Coefficients		F	Correlations			Importance
		Beta	Std. Error		Zero-Order	Partial	Part	
1	Пол	9,823E-02	0,047	4,364	0,082	0,106	0,097	0,048
2	Критерий T>2	2,298E-02	0,064	0,130	0,132	0,018	0,017	0,018
3	Критерий N>0	8,015E-02	0,049	2,709	0,129	0,084	0,077	0,062
4	Сторона	0,239	0,051	22,085	0,300	0,233	0,219	0,427
5	40 Гр	0,111	0,047	5,438	0,144	0,118	0,109	0,095
6	Расширенная диссекция	-3,514E-02	0,050	0,497	0,084	-0,036	-0,03	-0,018
7	Поликомбинированная операция	3,679E-02	0,063	0,346	0,102	0,030	0,027	0,022
8	Резекция бифуркации трахеи	0,141	0,051	7,683	0,199	0,140	0,129	0,167
9	Лимфоцитопения <8	-6,955E-02	0,052	1,757	-0,135	-0,068	-0,06	0,056
10	ЛИИ >1,4	0,126	0,052	5,881	0,164	0,123	0,113	0,123

что абсолютные значения коэффициентов важности были пропорциональны коэффициентам регрессии, следовательно, они пропорциональны степени вклада каждого предиктора в объяснение значения зависимой переменной. Коэффициенты важности выбраны в качестве весовых значений для создания шкалы. Для каждого из 10 включенных в регрессионную модель предикторов был подсчитан его балл путем умножения абсолютного значения соответствующего коэффициента важности на 100 и округления до целых. Таким образом была создана шкала как таковая (табл. 3). Каждый пациент в зависимости от значения каждого из 10 предикторов и соответствующего предиктору балла набирал суммарный балл, который являлся числовым показателем вероятности того, что зависимая переменная (несостоятельность культи бронха) примет значение «да» или «нет».

5. Следующим этапом было определение порогового суммарного балла, после превышения которого зависимая переменная принимала бы одно из своих значений с вероятностью, интересующей исследователя. Такая вероятность является эмпирической, и ее величина зависит от практической важности исследуемой зависимой переменной. В данном конкретном примере несостоятельность культи бронха (зависимая переменная) представляет собой жизнеугрожающее послеоперационное осложнение.

Таблица 3

Шкала суммарной балльной оценки риска несостоятельности швов *

№ п/п	Фактор	Значение	Баллы
1	Пол пациента	Мужской	+5
2	Критерий T (TNM) _____	>2	+2
3	Критерий N (TNM) _____	>0	+6
4	Сторона операции	Правая	+43
5	Лучевая терапия	40 Гр _____	+10 _____
6	Медиастинальная лимфодиссекция	Расширенная	+2
7	Комбинированная операция	2 и более органов	+2
8	Резекция бифуркации трахеи	Есть	+17 _____
9	Лимфоцитопения	< 8 в мазке	+6
10	ЛИИ	>1,4	+12 _____

* Суммарный балл более 30 является показателем высокого риска развития несостоятельности швов.

Для вычисления порогового балла первоначально был проведен регрессионный анализ, в котором суммарный балл каждого больного выступал в качестве предиктора, а зависимая переменная оставалась прежней. По определению зависимой переменной применялась бинарная логистическая регрессия. Было получено уравнение

$$p = 1 / (1 + e^{3,698 - X \cdot 0,045x})$$

где p — теоретическая вероятность развития несостоятельности культи бронха (зависимая переменная), x — значение суммарного балла у конкретного больного. С помощью данного уравнения были вычислены теоретические значения вероятности развития несостоятельности швов для каждого пациента. Диаграмма рассеяния, отражающая данную зависимость, представлена на рис. 1.

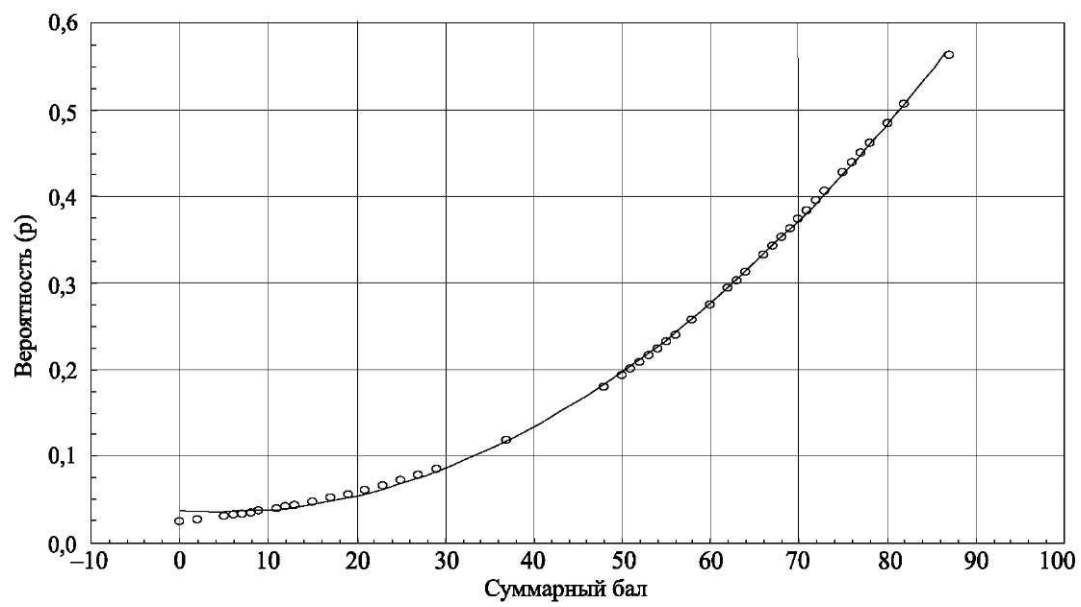
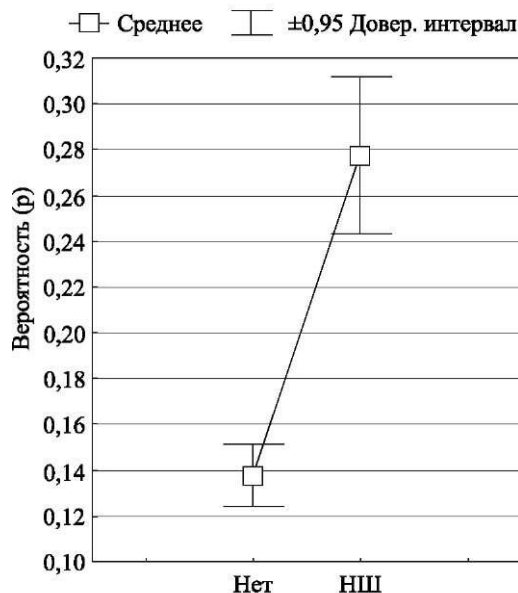


Рис. 1. Диаграмма рассеяния. Зависимость теоретической вероятности развития несостоятельности швов от суммарного балла.

Следующий момент — определение диапазона, в котором несостоятельность швов в выборке. Были подсчитаны средние значения вероятности в зависимой переменной «нет» и в группе больных переменной «да» (рис. 2). Оказалось, что практически не образовалась, если теоретическая находилась в диапазоне от 0,124 до 0,151. По нижней границе этого диапазона (0,124) соответствовал балл (30), при котором несостоятельность швов была принята как предварительное пороговое значение. Последним этапом построения шкалы явилась адекватности значения порогового балла. Была несостоятельности культи бронха у больных с Она оказалась равной 4,93 % (порядка 5 %). несостоятельности швов культи бронха у больных составила 27,75 % (порядка 30 %). С клинических порогового балла, поскольку ожидаемая частота % может быть рассмотрена в качестве



теоретических вероятностей, при практически не реализовалась. группе больных со значением со значением зависимой несостоятельности швов вероятность ее развития графику (см. рис.1) видно, что ствует интервалу от 30 до 40 минимальный ориентировочный фактически не образуется, и он значение суммарного балла. практическая проверка подсчитана фактическая частота суммарным баллом 30 и менее. Фактическая частота с суммарным баллом более 30 позиций—это удачное значение развития осложнения порядка 5 приемлемой, в то время как

ожидаемая частота осложнения порядка 30 % **Рис. 2. Диапазон теоретических вероятностей несостоятельности швов в группах больных с отсутствием («нет») и наличием («НШ») данного осложнения.** тяжести осложнения, т. е. суммарный балл 30 делит противоположные группы, что удобно для решения.

является явно чрезмерной с учетом пациентов на 2 диаметрально клинициста во время принятия

Таким образом, была создана шкала, которая на основании простой оценки 10 признаков у пациента позволяла объективно высказываться о развитии осложнения с четко градируемой и определенной долей вероятности.

В целом, общий алгоритм создания шкалы следующий: 1) определение списка потенциальных предикторов и их редукция; 2) категоризация отобранных предикторов и включение их в регрессионную модель с оптимальным шкалированием (Regression with Optimal Scaling (CATREG) пакета SPSS) с целью подсчета коэффициента важности (Importance), т. е. взвешивания предикторов; 3) определение балла для каждого предиктора на основании его коэффициента важности; 4) определение суммарного балла для каждого наблюдения (пациента) из выборки; 5) построение графика регрессионной зависимости теоретической вероятности исследуемой переменной от суммарного балла; 6) определение диапазона теоретических вероятностей, при котором зависимая переменная фактически принимала значение «нет» с последующим графическим определением диапазона баллов, соответствующих данному вероятностному интервалу; 7) выбор значения порогового балла из определенного выше диапазона;

8) проверка адекватности порогового балла путем подсчета фактической частоты значений зависимой переменной в группах пациентов с суммарным баллом меньше (или равным) и больше порогового.

Обсуждение. Во-первых, следует пояснить, почему выбрана регрессионная модель с оптимальным шкалированием (Regression with Optimal Scaling (CATREG)). Она позволяет взвесить предиктор так, что число, характеризующее его вес (коэффициент важности (Importance), см. табл. 2), заключено в интервале от 0 до 1, что удобно для присвоения конкретного пропорционального значения балла.

Поскольку шкала построена с использованием результатов регрессионного анализа, очевидно, что точность прогноза, полученного с помощью нее, не будет превышать точности регрессионного прогноза. Однако следует заметить, что ни одна из приведенных выше шкал [4, 5], не является высокоточной и совершенной с клинических позиций. Задача шкалирования состоит в оценке значения зависимой переменной с заранее заданной вероятностью. Предложенная модель шкалы как раз отвечает этому требованию, поскольку исследователь сам задает интересующий его порог вероятности, основанный на фактических частотах значения зависимой переменной в группах пациентов. Следует отметить, что достоинство такого подхода заключается в том, что шкала максимально адаптируется к конкретным условиям, конкретным пациентам, конкретному учреждению.

Смысл приведенных выше математических вычислений состоит лишь в том, чтобы, во-первых, трансформировать «неудобную» в практической медицине регрессионную модель в более подходящую упрощенную шкалу и, во-вторых, учесть при создании шкалы достоверный эмпирический опыт. Именно поэтому в качестве предикторов в рамках шкалы рассматривались не только факторы, значимо ($p < 0,05$) связанные с зависимой переменной, но и факторы, при оценке связи которых уровень значимости был выше, но близок к достоверному (до 0,15). Следует отметить, что коэффициенты значимости таких предикторов (а следовательно, и их вес), подсчитанные в рамках регрессионного анализа с оптимальным шкалированием, могли быть низкими.

Теоретическая адекватность порогового балла в приведенном выше примере проверялась бинарной логистической регрессией, в которой имелся единственный категоризированный предиктор — суммарный балл больного (<30 или >30), а зависимой переменной оставалась несостоятельность швов. Данная регрессионная модель осуществляла прогноз с точностью 84 % (коэффициент детерминации $R^2 = 0,178$). При независимом включении предикторов в регрессионную модель точность прогноза составила 84,7 % ($R^2 = 0,218$). Таким образом, суммарный балл, выступающий в качестве единственного интегрального категоризированного предиктора, давал несколько худшую модель. Тем не менее это было признано оправданным, поскольку шкала существенно удобнее в применении. Кроме того, приемлемость точности прогноза определяется субъективно в зависимости от клинической важности исследуемой переменной и вытекающих из нее последствий. В приведенном выше наблюдении суммарный балл был намеренно несколько снижен (до 30) с целью придания шкале тенденции к гиперпрогнозированию негативного результата (развитие несостоятельности швов).

В заключение можно сказать, что предлагаемая статистическая модель позволяет создавать клинически адаптированные шкалы для прогнозирования значений бинарных переменных в медицинских исследованиях.

Summary

Литература

1. *Наследов А. Д.* SPSS: Компьютерный анализ данных в психологии и социальных науках. СПб., 2005. 416 с.
2. *Боровиков В.* STATISTICA. Искусство анализа данных на компьютере: Для профессионалов. 2-е изд. СПб., 2003. 688 с.
3. *Нейронные сети. STATISTICA Neural Networks: Пер. с англ. М., 2001. 182 с.*
4. *Перитонит: Практич. руководство / Под ред. В.С. Савельева, Б.Р. Гельфанда, М.И. Филимонова. М., 2006. 208 с.*
5. *Абдоминальная хирургическая инфекция: клиника, диагностика, антимикробная терапия: Практич. руководство / Под ред. В.С. Савельева, Б.Р. Гельфанда. М., 2006. 168 с.*

Статья принята к печати 20 июня 2007 г.