

Построение прогнозной модели постинсультных пациентов

А.Б. Блинова

Факультет авионики, энергетики и
инфокоммуникаций
Уфимский государственный авиационный
технический университет
Уфа, Россия
e-mail: alla.blinova.95@mail.ru

К.Р. Шаймарданова

Факультет авионики, энергетики и
инфокоммуникаций
Уфимский государственный авиационный
технический университет
Уфа, Россия
e-mail: shaymardanova1994@mail.ru

Ю.О. Уразбахтина

Факультет авионики, энергетики и
инфокоммуникаций
Уфимский государственный авиационный
технический университет
Уфа, Россия
e-mail: urjuol@mail.ru

Л.Р. Ахмадеева

Башкирский государственный медицинский
университет
Уфа, Россия
e-mail: leila_ufa@mail.ru

Аннотация¹

В данной статье рассматривается построение модели прогнозирования реабилитационного потенциала постинсультных больных методом бинарной логистической регрессии. Оценивается качество построенной модели на основании коэффициентов R² Мак-Фаддена, Prob (LR statistic), log likelihood и теста Хосмера-Лемешоу и прогнозные свойства на основе графика исходных данных, смоделированных данных и остатков модели.

1. Введение (style -Title2)

Цереброваскулярные заболевания являются основной причиной заболеваемости, смертности и инвалидизации в России, занимая первое место среди причин первичной инвалидности взрослого населения (32 на 100 тыс. населения) [1,2]. Ежегодно в мире отмечается 16 млн впервые возникших случаев острого нарушения мозгового кровообращения (ОНМК) [3].

Степень и скорость восстановления после инсульта зависят от многих факторов: возраста пациента, тяжести поражения, выбора средств реабилитации, медикаментозного лечения. В связи с возникновением у большинства постинсультных пациентов неврологического дефицита, реабилитация включает в себя разнообразные мероприятия, направленные на восстановление или компенсацию нарушенных функций нервной системы, в основе

Труды Шестой всероссийской конференции
"Информационные технологии интеллектуальной
поддержки принятия решений", 28-31 мая, Уфа-
Ставрополь, Россия, 2018

которых лежат механизмы нейропластичности – способности нервной ткани к структурно-функциональной реорганизации после ее повреждения. [4,5].

В настоящее время нет критериев, точно устанавливающих реабилитационный прогноз конкретного пациента и эффективность реабилитационных мероприятий в стационаре. Таким образом, остро встал вопрос о создании прогнозной модели, предсказывающей результативность реабилитации постинсультных больных с заданным уровнем значимости на основе совокупности множества факторов.

2. Публикация (style -Title2)

Материалы конференции публикуются оргкомитетом, как в бумажном варианте, так и в виде электронного издания. (style Plane Text).

Это означает, что участники вместо бумажного варианта статьи, должны представить материал статьи в электронном виде (Winword 2003-XP). Файл следует называть по фамилии первого автора, с указанием даты отправки.

Цель исследования: построение прогнозной модели постинсультных пациентов методом бинарной логистической регрессии.

Материалы и методы: Для оценки реабилитационного потенциала пациентов, перенесших инсульт, и прогнозирования повторного инсульта на основе имеющихся данных была построена статистически значимая модель бинарной регрессии, оценены параметры и методом пошагового исключения, удалены все статистически не значимые переменные. Анализ проводился по 265

пациентам, проходивших реабилитационную программу в ГБУЗ РБ Городская клиническая больница №21 г. Уфы, за период с 1 июля 2017 г. по 31 декабря 2017 г. с подтвержденным клиническим диагнозом ОНМК (по ишемическому или геморрагическому типу). Выходной переменной является улучшение реабилитационного потенциала постинсультных пациентов на 2 балла и более по шкале NIHSS.

NIHSS расшифровывается как National Institutes of Health Stroke Scale (Шкала инсульта Национального института здоровья). Применяется для оценки неврологического статуса, локализации инсульта (в каротидном или вертебробазилярном бассейне), дифференциальной диагностики и результатов лечения [6].

Результаты и обсуждение:

Бинарная логистическая регрессия применяется для предсказания вероятности возникновения некоторого события по значениям множества признаков. Для этого вводится так называемая зависимая переменная y (в нашем случае это шкала NIHSS), принимающая лишь одно из двух значений — 0 (есть ухудшение, нет изменений или улучшение меньше, чем на 2 балла) и 1 (улучшение больше, чем на 2 балла), и множество независимых переменных (также называемых признаками, предикторами или регрессорами). На основе значений независимых переменных требуется вычислить вероятность принятия того или иного значения зависимой переменной [7].

Входными переменными являются:

BALANS_TERAPIYA - проведение лечения баланс-терапией;

BOBAT_TERAPIYA - проведение лечения Бобат-терапией (нейро-развивающим лечением, которое восстанавливает мышечный тонус и стимулирует развитие правильной моторики);

SAHARN_DIABET – наличие у пациента сахарного диабета II типа;

DAYS_NCO - количество койко-дней, проведенных пациентом в нейрососудистом отделении;

DAYS_OMP - количество койко-дней, проведенных пациентом в отделении медицинской реабилитации;

DIAGNOZ - диагноз пациент (геморрагический или ишемический инсульт);

MELKAYA_MOTORICA - проведении комплексов реабилитации мелкой моторики;

MMT-TERAPIYA - проведение мягкой мануальной терапии;

NIHSS_4 - значение по шкале NIHSS при поступлении 4 балла;

NIHSS_5 - значение по шкале NIHSS при поступлении 5 баллов;

NIHSS_6 - значение по шкале NIHSS при поступлении 6 баллов;

NIHSS_7 - значение по шкале NIHSS при поступлении 7 баллов;

NIHSS_8 - значение по шкале NIHSS при поступлении 8 баллов;

NIHSS_9 - значение по шкале NIHSS при поступлении 9 баллов;

NIHSS_10 - значение по шкале NIHSS при поступлении 10 баллов;

NIHSS_11 - значение по шкале NIHSS при поступлении 11 баллов;

NIHSS_12 - значение по шкале NIHSS при поступлении 12 баллов;

NIHSS_14 - значение по шкале NIHSS при поступлении 14 баллов;

NIHSS_15 - значение по шкале NIHSS при поступлении 15 баллов;

NIHSS_16 - значение по шкале NIHSS при поступлении 16 баллов;

NIHSS_17 - значение по шкале NIHSS при поступлении 17 баллов;

NIHSS_18 - значение по шкале NIHSS при поступлении 18 баллов;

ООКТ_TERAPIYA - проведение мягкой мануальной терапии;

PNF_TERAPIYA – проведение PNF-терапии (расшифровывается как, проприоцептивное нейромышечное протерение). Это одна из методик кинезиотерапии, т.е. лечение движением;

POL – пол пациента;

POZD_VOSST – поздний период восстановления пациента (3-6 месяцев);

PROTIVOPOKAZANITA_FTO – наличие противопоказаний к ФТО;

RANN_VOSST - поздний период восстановления пациента (6-12 месяцев);

VOZRAST – возраст пациента.

При проведении корреляционного анализа с целью исключения переменных, коррелирующих с другими переменными, выявлено, что данные переменные не коррелируют друг с другом (так как значения < 0.9).

Представленная зависимость является бинарной, следовательно, применяли бинарную логистическую регрессию. В результате была построена Логит-модель - это статистическая модель, используемая для предсказания вероятности возникновения

некоторого события путём подгонки данных к логистической кривой.

Variable	Coefficient	Std. Error	z-Statistic	Prob.
BOBAT_TERAPIYA	-1.039528	0.373019	-2.786796	0.0053
DAYS_OMP	0.102044	0.059016	1.729081	0.0838
NIHSS_6	2.315965	0.632056	3.664177	0.0002
NIHSS_7	3.290025	0.548639	5.996707	0.0000
NIHSS_8	2.901325	0.566990	5.117069	0.0000
NIHSS_9	3.925512	0.756912	5.186220	0.0000
NIHSS_10	2.749896	0.588540	4.672405	0.0000
NIHSS_11	2.843789	0.690946	4.115788	0.0000
NIHSS_12	2.777636	0.650139	4.272371	0.0000
NIHSS_14	4.174813	1.212308	3.443691	0.0006
NIHSS_15	1.935536	1.050446	1.842584	0.0654
NIHSS_16	2.526538	0.868419	2.909354	0.0036
NIHSS_17	3.221331	1.258865	2.558917	0.0105
NIHSS_18	2.226174	0.745294	2.986975	0.0028
POZD_VOSST C	-0.843348	0.496621	-1.698173	0.0895
	-3.000683	0.848600	-3.536040	0.0004

McFadden R-squared	0.271494	Mean dependent var	0.535849
S.D. dependent var	0.499657	S.E. of regression	0.419257
Akaike info criterion	1.126930	Sum squared resid	43.76824
Schwarz criterion	1.343065	Log likelihood	-133.3183
Hannan-Quinn criter.	1.213770	Deviance	266.6366
Restr. deviance	366.0046	Restr. log likelihood	-183.0023
LR statistic	99.36799	Avg. log likelihood	-0.503088
Prob(LR statistic)	0.000000		

Obs with Dep=0	123	Total obs	265
Obs with Dep=1	142		

Рис. 1 – Логистическое распределение.

Статистически значимые критерии и значения коэффициентов R² Мак-Фаддена, Prob (LR statistic) и log likelihood.

Оценка качества построенной модели проводится на основании коэффициентов R² Мак-Фаддена, Prob (LR statistic), log likelihood и теста Хосмера-Лемешоу (рис. 3).

1) Коэффициент детерминации McFadden R-squared равен 0,2715. Коэффициент детерминации меньше 0,6, т.е. колебания зависимой переменной не обусловлены колебаниями предикторов.

2) Критическая статистика для теста отношения правдоподобия Prob (LR statistic) равна 0. Это меньше уровня значимости 0.05, следовательно модель является значимой.

3) Критерий правдоподобия (log likelihood) равен -133.32. Это говорит о том, качество модели хорошее.

Критерий согласия Хосмера-Лемешоу, исследует расстояние между наблюдаемыми и ожидаемыми распределениями частот «живых» и «умерших» пациентов. Если уровень значимости является большим, то модель хорошо откалибрована и достаточно точно описывает реальные данные. Значение статистики Хосмера-Лемешоу не должно

быть меньше уровня значимости 0,05. Оптимальными считаются значения не меньше 0,5–0,6.

Quantile of Risk	Low		High		Actual	Dep=0	Expect	Actual	Dep=1	Expect	Total Obs	H-L Value
	Low	High	Low	High								
1	0.0021	0.1932	24	24.3917	2	1.60625	26	0.10171				
2	0.1932	0.1932	25	21.7846	2	5.21539	27	2.45695				
3	0.1932	0.2802	19	20.8821	7	5.11794	26	0.86173				
4	0.2802	0.5625	13	14.7646	14	12.2354	27	0.46541				
5	0.5625	0.6533	11	9.91697	15	16.0830	26	0.19121				
6	0.6533	0.7128	10	8.67798	17	18.3220	27	0.29679				
7	0.7128	0.7256	5	7.26003	21	18.7340	26	0.98080				
8	0.7256	0.8105	8	6.62004	19	20.3800	27	0.38109				
9	0.8105	0.8105	6	4.92667	20	21.0733	26	0.28850				
10	0.8105	0.8054	2	3.54976	25	23.4502	27	0.77901				
Total			123	122.761	142	142.219	265	6.80320				

H-L Statistic	6.8032	Prob. Chi-Sq(8)	0.5580
Andrews Statistic	19.8764	Prob. Chi-Sq(10)	0.0304

Рис. 2 – Результаты теста Хосмера-Лемешоу

Полученное значение (0,558) больше 0,05, следовательно, модель хорошо откалибрована и достаточно точно описывает реальные данные.

На рис. 3 представлен график исходных данных, смоделированных данных и остатков модели.

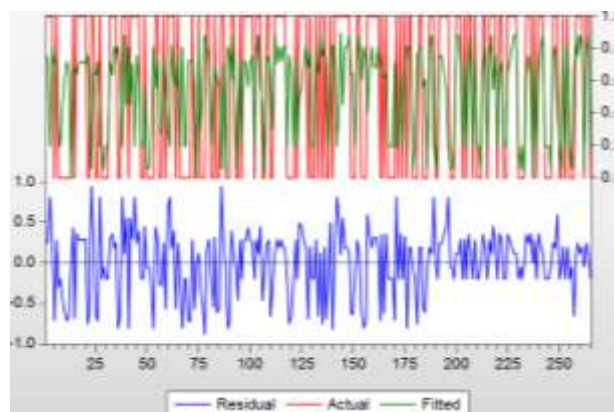


Рис. 3 – Фактические, расчетные значения и остатки модели

На графике видно, что фактические и расчетные значения плохо соответствуют друг другу, а остатки достаточно велики. То есть можно сделать вывод о плохих прогнозных качествах построенной модели.

Для улучшения качества полученной модели рекомендуется увеличить количество наблюдений. Это позволит установить связь между колебаниями зависимой переменной не обусловлены колебаниями предикторов и таким образом увеличить коэффициент детерминации McFadden R-squared до оптимального значения, равного 0,6.

Список используемых источников

1. Гусев Е.И., Скворцова В.И. Ишемия головного мозга. М.: Медицина, 2001.

2. Скворцова В.И. Реперфузионная терапия ишемического инсульта. *Consilium Medicum*. 2004; 6 (8): 610–614.
3. Wissel J., Olver J., Stibrant Sunnerhagen K. Navigating the poststroke continuum of care. *J. Stroke Cerebrovasc. Dis.* 2013; 22 (1): 1–8.
4. Дамулин И.В., Екушева Е.В. Клиническое значение феномена нейропластичности при ишемическом инсульте // *Анналы клинической экспериментальной неврологии*. 2016. (дата обращения: 10.04.2018)
5. Котов С.В. Новые технологии в диагностике и лечении больных в остром периоде инсульта. *Русский медицинский журнал*. 2014;22(10):712–6.
6. Неврологический портал. Шкала NIHSS [Электронный ресурс] – Режим доступа: <http://neurology.com.ua/2010/09/28/shkala-nihss.html>
7. Логистическая регрессия [Электронный ресурс] – Режим доступа: <https://dic.academic.ru/dic.nsf/ruwiki/1355147>