

С.В. Иванова¹, С.А. Кулева^{1,2}, А.П. Карицкий¹, Т.Ю. Семиглазова^{1,3}

Создание статистической модели для прогнозирования лимфомы Ходжкина у молодых взрослых

¹ФГБУ «НИИ онкологии им. Н.Н. Петрова» Минздрава России²ГБОУВПО «Санкт-Петербургский государственный педиатрический медицинский университет» Минздрава России³Северо-Западный государственный медицинский университет им. И.И. Мечникова, Санкт-Петербург

Целью настоящего исследования стало создание многомерной корреляционной матрицы детерминант, дающих информацию о степени влияния на выживаемость молодых взрослых с лимфомой Ходжкина каждой переменной, а также об эффекте взаимодействия этих переменных между собой.

Восемьдесят семь пациентов в возрасте от 19 до 29 лет (средний возраст 24 ± 4 года) включены в исследование. Сорок семь пациентов (54%) имели локализованные формы заболевания (I-II стадии заболевания), 40 больных (46%) – генерализованное заболевание (III–IV стадии заболевания). Экстранодальное поражение выявлено у 37 больных (42,5%). В 47 случаях (54%) имела место В-стадия злокачественного процесса. Большинство больных (76 или 87,4%) имели гистологическое заключение нодулярного склероза, у 2 больных (2,3%) выявлено лимфоидное преобладание и у 5 (5,7%) – смешанноклеточный вариант, в 4 случаях (4,6%) морфологический тип не диагностирован. Все пациенты получали от 2 до 11 циклов первичной полихимиотерапии (в среднем 4 цикла) по схемам ABVD (48 больных или 55,2%), BEACOPP (36 или 41,4%), MOPP (2 или 2,3%) и COPP (один больной). Пятидесяти семи больным (65,5%) проведено облучение в различных режимах в суммарных очаговых дозах 20–51 Гр.

Многокомпонентная корреляционная матрица, содержащая коэффициенты корреляции выживаемости и коэффициенты корреляции каждого из анализируемых факторов, а также коэффициенты, оценивающие степень тесноты связи между факторами, состояла лишь из 5 значимых детерминант (объем опухолевого поражения, IV стадия заболевания, Е-поражение, ускорение СОЭ, лейкоцитоз). Уравнение множественной регрессии представлено в следующем виде: $\bar{y} = 0,054097 - 0,231436x_1 [\text{объем}] + 0,118813x_2 [\text{IV стадия}] - 0,133359x_3 [\text{Е}] - 0,253969x_4 [\text{СОЭ}] - 0,147299x_5 [\text{Лей}]$; $R^2 = 0,11094850$. Часть коэффициентов в уравнении множественной регрессии оказалась статисти-

чески не значима, поэтому следующим этапом анализа стало построение корреляционной матрицы с целью отбора факторов для включения в уравнение, в котором содержались парные коэффициенты корреляции общей выживаемости и каждого из факторов, а также коэффициенты, оценивающие степень тесноты связи между ними. Оказалось, что на признак-результат влияет только объем опухолевого поражения (коэффициент корреляции 0,2570, $p = 0,026$). Новое уравнение представлено в следующем виде: $\bar{y} = 0,166667 - 0,227273x [\text{объем}]$; $R^2 = 0,0643674507$.

Полученное уравнение позволяет по заданным значениям фактора «объем опухолевого поражения» иметь теоретические значения результативного признака (выживаемости), подставляя в него фактические значения фактора.

Ключевые слова: лимфома Ходжкина, молодые взрослые, статистическая модель прогнозирования

Бимодальное распределение заболеваемости при лимфоме Ходжкина (ЛХ) делает необходимым выделение группы молодых взрослых в отдельную когорту с пиком заболеваемости 44 на млн в год (рис. 1) [3].

Очевидно, что выживаемость этой возрастной группы, равно как и прогностические факторы требует обязательного изучения. Тем более, что летальность 15–29-летних пациентов с ЛХ составляет 4,82 случая на 1 млн в год (рис. 2). Летальность больных в возрасте 20–24 лет в 2 раза выше, чем пациентов 15–19 лет, приближаясь к 5,23 на млн в год (у больных 15–19 лет эта цифра равна 2,31 на млн в год). Уровень летальности 25–29-летних достигает максимума (7,22 на 1 млн в год) [3].

Целью настоящего исследования стало создание многомерной корреляционной матрицы детерминант, дающих информацию о степени влияния на выживаемость молодых взрослых с лимфомой Ходжкина каждой переменной, а также об эффекте взаимодействия этих переменных между собой.

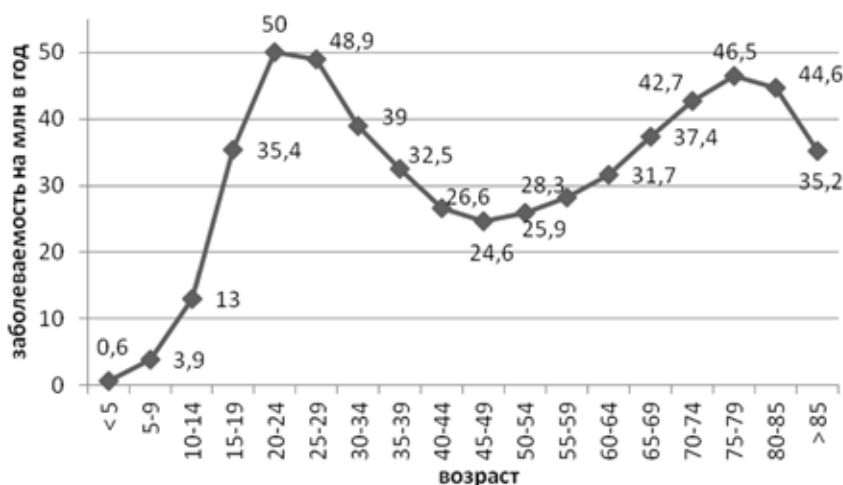


Рис. 1. Заболеваемость лимфомой Ходжкина по возрастам (бимодальная кривая МакМахона)

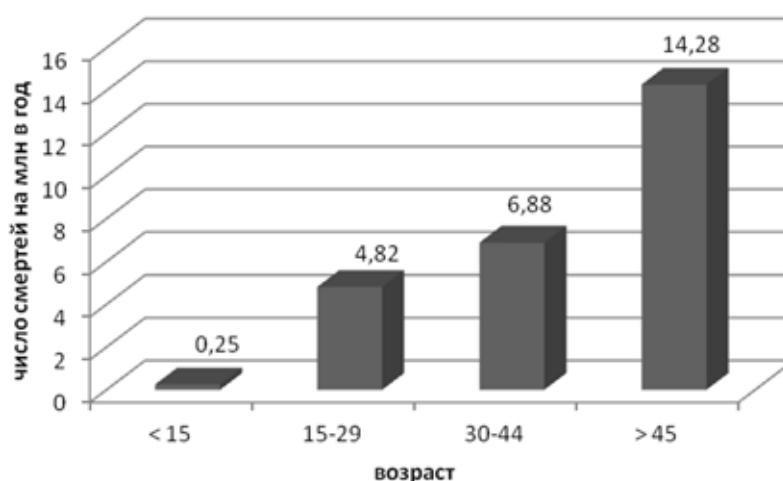


Рис. 2. Летальность от ЛХ

Материалы и методы

В исследование были включены сведения о пациентах в возрасте от 19 до 29 лет, лечившихся в НИИ онкологии им. Н.Н. Петрова в 2004-2011 гг. Число таких больных составило 87. Средний возраст был 24 ± 4 года. Мужчин – 34 (39,1%), женщин – 53 (60,9%), соотношение по полу 1 : 1,6 (табл. 1). Распределение по стадиям было практически одинаковым: у 47 пациентов (54%) диагностированы локализованные формы заболевания (I-II стадии заболевания), у 40 больных (46%) – генерализованное заболевание (III-IV стадии заболевания). Экстранодальное поражение выявлено у 37 больных (42,5%). Более половины больных (47 или 54%) имели В-стадию злокачественного процесса. У большинства (76 или 87,4%) гистологическое заключение звучало как нодулярный склероз, у 2 больных (2,3%) выявлено лимфоидное преобладание и у 5 (5,7%) – смешанно-клеточный вариант, в 4 случаях (4,6%) морфологический тип не диагностирован.

Лечебная программа состояла из 2-11 циклов первичной полихимиотерапии (в среднем 4 цикла) по схемам ABVD (48 больных или 55,2%), BEACOPP (36 или 41,4%), MOPP (2 или 2,3%) и COPP (один больной) (табл. 2).

Пятидесяти семи больным (65,5%) лекарственное лечение дополнено облучением в различных режимах. Лучевая терапия с методикой вовлеченных зон проведена 15 пациентам (17,2%), расширенных зон – 16 больным (18,4%), радикальную программу (субтотальное и тотальное облучение) получили 14 пациентов (16,1%), в остальных слу-

Таблица 1.
Инициальные характеристики больных с ЛХ (n = 87)

Характеристика	Количество больных	
	абс.	%
Пол:		
мужской	34	39,1
женский	53	60,9
Стадия:		
I	2	2,3
II	38	43,7
III	14	16,1
IV	33	37,9
Экстранодальное поражение		
есть	37	42,5
нет	38	43,7
нет данных	12	13,8
В-симптомы:		
есть (В)	47	54
нет (А)	40	46
Морфологический вариант:		
лимфоидное преобладание	2	2,3
нодулярный склероз	76	87,4
смешанно-клеточный	5	5,7
лимфоидное истощение	-	-
нет данных	4	4,6

Таблица 2.
Методы терапии в исследуемой группе пациентов

Методы терапии	Число больных	%
ABVD	48	55,2
BEACOPP	36	41,4
MOPP	2	2,3
COPP	1	1,1
ВСЕГО	87	100

чаях сведений о режиме лучевой терапии не было. Доза облучения варьировала от 20 до 51 Гр.

Для оценки одновременного влияния более чем одного фактора на результат был использован многомерный (или многофакторный) анализ, который давал информацию о степени влияния на исход каждой переменной, а также об эффекте взаимодействия этих переменных между собой. Количество переменных в модели было оптимизировано с помощью автоматизированных алгоритмов их подбора методом наилучшего подмножества, который подразумевал выбор путем подстановки такого набора переменных, которые наилучшим образом удовлетворяют условиям, определенным исследователем.

Для оценки силы связи переменных был использован коэффициент детерминации, который представляет собой квадрат коэффициента корреляции Пирсона (R^2). Очевидно, что чем больше коэффициент корреляции отклоняется от 1 или -1 (т.е. чем больше степень рассеяния точек от линии на рис.), тем меньше будет значение коэффициента детерминации и тем слабее будут переменные коррелировать между собой. Ввиду того, что в модель может входить несколько факторов риска, коэффициент рассчитывается с поправкой на их количество. При построении модели, помимо R и R^2 , были вычислены Adjusted R^2 – скорректированный коэффициент детерминации, F – расчетное значение критерия Фишера, Std. Error of estimate – стандартная ошибка уравнения.

Результаты и обсуждение

В многофакторный анализ были включены детерминанты, значимо влияющие на прогноз

заболевания, однако, некоторые из них обладали мультиколлинеарностью и были интеркоррелированы (например, объем опухоли и уровень фибриногена, объем опухоли и отсутствие эффективности лечения) и были исключены из окончательного анализа. Таким образом, многокомпонентная корреляционная матрица, содержащая коэффициенты корреляции выживаемости и коэффициенты корреляции каждого из анализируемых факторов, а также коэффициенты, оценивающие степень тесноты связи между факторами, состояла лишь из 5 значимых детерминант (табл. 3).

Т.о. уравнение множественной регрессии предстало в следующем виде:

$$\bar{y} = 0,054097 - 0,231436x_1[\text{объем}] + 0,118813x_2[\text{IV стадия}] - 0,133359x_3[\text{Е}] - 0,253969x_4[\text{СОЭ}] - 0,147299x_5[\text{Лей}]; R^2 = 0,11094850$$

где $\bar{y}$ – вероятность рецидива заболевания, $x_1[\text{объем}]$ – объем опухолевого поражения более $210 \text{ см}^2/\text{м}^3$, $x_2[\text{IV стадия}]$ – наличие у больного IV стадии, $x_3[\text{Е}]$ – наличие экстранодального поражения, $x_4[\text{СОЭ}]$ – СОЭ более 35 мм/ч , $x_5[\text{Лей}]$ – уровень лейкоцитов более $11,5 \times 10^9/\text{л}$

Из табл. 3 видно, что часть коэффициентов в уравнении множественной регрессии статистически не значима ($t_{\text{расч}} < t_{\text{табл}}$). Поэтому следующим этапом анализа стало построение корреляционной матрицы с целью обоснованного отбора факторов для включения в уравнение (табл. 4). В данной корреляционной матрице содержатся парные коэффициенты корреляции общей выживаемости и каждого из факторов,

Таблица 3.
Результаты расчета параметров уравнения линейной регрессии

	Beta	Std.Err. — of Beta	b	Std.Err. — of b	t	p-value
Intercept			0,054097	0,175505	0,30824	0,758832
Объем	0,272207	0,115266	-0,231436	0,098001	2,36156	0,021027
IV стадия	0,140867	0,155964	0,118813	0,131547	0,90320	0,369562
Е-поражение	-0,159135	0,154152	-0,133359	0,129183	-1,03232	0,305526
СОЭ	-0,151313	0,113855	-0,253969	0,191099	-1,32899	0,188228
Лей	-0,137085	0,120189	-0,147299	0,129144	-1,14058	0,257987

Таблица 4.
Корреляционная матрица

Параметры	IV стадия	Е-поражение	СОЭ	Лейк	Объем	ОВ
IV стадия	1,0000	,6366	,0215	,2178	,0895	,0308
	p= ---	p=,000	p=,854	p=,060	p=,445	p=,793
Е-поражение	,6366	1,0000	-,0428	-,0493	,1691	-,0102
	p=,000	p= ---	p=,715	p=,674	p=,147	p=,931
СОЭ	,0215	-,0428	1,0000	,0091	-,0072	-,1447
	p=,854	p=,715	p= ---	p=,938	p=,951	p=,216
Лей	,2178	-,0493	,0091	1,0000	,0148	-,0959
	p=,060	p=,674	p=,938	p= ---	p=,900	p=,413
Объем	,0895	,1691	-,0072	,0148	1,0000	-,2570
	p=,445	p=,147	p=,951	p=,900	p= ---	p=,026
ОВ	,0308	-,0102	-,1447	-,0959	-,2570	1,0000
	p=,793	p=,931	p=,216	p=,413	p=,026	p= ---

а также коэффициенты, оценивающие степень тесноты связи между ними. Как видно из матрицы, на признак-результат влияет только объем опухолевого поражения (коэффициент корреляции 0,2570, $p=0,026$).

В табл. 5 и 6 представлены основные показатели модели множественной регрессии.

Таблица 5.
Основные показатели модели множественной регрессии

	Value
Multiple R	0,253707412
Multiple R ²	0,0643674507
Adjusted R ²	0,053360009
F(1,85)	5,84763037
p	0,0177336093
Std.Err. of Estimate	0,425355047

Используя исходные данные табл. 5 и 6, строим новое уравнение.

$$\bar{y} = 0,166667 - 0,227273 \times [\text{объем}]; \\ R^2 = 0,0643674507$$

Таким образом, полученное уравнение позволяет по заданным значениям фактора «объем опухолевого поражения» иметь теоретические значения результативного признака, подставляя в него фактические значения фактора.

Лимфома Ходжкина представляет собой заболевание с мало выраженным опухолевым компонентом, составляющим не более 20%, и неопухолевой гетерогенной частью, представленной клетками воспаления и вспомогательными структурами (малыми лимфоцитами, эозинофилами, нейтрофилами, гистиоцитами, макрофагами, фибробластами и коллагеновыми волокнами) [9,17]. Данный факт необходимо учитывать при создании одномерных и многомерных моделей прогнозирования [1,2]. В связи с несовершенством моделей некоторые клинические рандомизированные исследования потерпели крах именно из-за неправильной стратификации пациентов на лечебные группы и проведения неадекватных объемов терапии.

Впервые предположение о значимом влиянии на прогноз объема опухолевого поражения высказали L. Specht и соавт. (1986, 1988, 1990, 1992). Ими было доказано, что влияние на прогноз даже приблизительных значений объема

опухолевого поражения значимо выше, чем влияние других клинических или лабораторных параметров, большинство которых так или иначе коррелировали с исследуемым фактором [10-16]. Авторами высказано предположение о неоднородности опухоли и существенном влиянии только неопластического компонента по сравнению с общим объемом, состоящим из опухолевого и реактивного компонентов, и о значимом влиянии межклеточных связей и факторов.

Спустя 20 лет к вопросу о влиянии относительного объема опухолевого поражения на показатели выживаемости вернулись P.G. Goobi и соавт. (2001), когда в диагностический алгоритм вошло обязательное использование КТ [4-7]. Вычисление относительного объема опухолевого поражения при лимфоме Ходжкина признано наиболее простым и значимым параметром, который можно использовать в повседневной клинической практике с целью оценки прогноза заболевания и стратификации пациентов на группы риска. Было доказано, что обозначенный параметр имеет даже большее предиктивное значение, чем такие известные прогностические модели и индексы, как международный прогностический индекс, индекс Мемориального онкологического центра им. Слоуна-Кеттеринга (Memorial Sloan-Kettering Cancer Center) и модели, взятые из международной базы данных по лимфоме Ходжкина [8]. Была очевидной тенденция падения этого влияния с возрастом больного, что объяснено снижением у взрослых и пожилых иммунологической реактивности организма и меньшим взаимодействием между неопластическими, воспалительными и стромальными клетками. Этим же объясняется преобладание у лиц старшего возраста морфологического варианта лимфоидного истощения и меньшего объема поражения, хотя абсолютная концентрация опухолевых клеток в некоторых случаях даже выше.

Таким образом, в исследованиях, проведенных за последние 10 лет, доказано, что интеграция в прогностические модели показателя общей суммы опухолевой массы (бремени, нагрузки, burden) заметно изменяет прогнозный вклад клинических и лабораторных параметров в регрессионных анализах, уменьшая при этом их предсказательное значение в десятки раз.

Построение нашей корреляционной матрицы с целью обоснованного отбора факторов,

Таблица 6.
Основные показатели модели множественной регрессии

	Beta	Std.Err. — of Beta	b	Std.Err. — of b	t(85)	p-value
Intercept			0,166667	0,057883	2,879347	0,005041
Объем	0,253707	0,104916	-0,227273	0,093985	2,418187	0,017734

в которой содержатся парные коэффициенты корреляции общей выживаемости и каждого из факторов, а также коэффициенты, оценивающие степень тесноты связи между ними, показало, что на признак-результат влияет только объем опухолевого поражения (коэффициент корреляции 0,2570, $p=0,026$).

ЛИТЕРАТУРА

1. Колыгин Б.А., Кулева С.А. Диагностика и лечение лимфомы Ходжкина. СПб.: Гиппократ, 2009. – 208 с.
2. Кулева С.А., Карицкий А.П. Прогностическое значение объема опухолевого поражения при лимфоме Ходжкина с современных позиций // Вестник РАМН. – 2014. – № 11 -12. – С. 67-71.
3. Bleyer W.A., O'Leary M., Barr R., Ries L.A.G. (eds) (2006) Cancer Epidemiology in Older Adolescents and Young Adults 15 to 29 Years of Age, Including SEER Incidence and Survival, 1975–2000. National Cancer Institute, NIH Pub. No. 06-5767. Bethesda, MD.
4. Gobbi P.G., Bassi E., Bergonzi E. et al. Tumor burden predicts treatment resistance in patients with early unfavorable of advanced-stage Hodgkin Lymphoma treated with ABVD and radiotherapy // Hematol. Oncol. – 2012. – Vol. 30. – P. 194-199.
5. Gobbi P.G., Bergonzi M., Bassi E. et al. Tumor burden in Hodgkin's lymphoma can be reliably estimated from a few staging parameters // Oncology reports. – 2012. – Vol. 28. – P. 815-820.
6. Gobbi P.G., Broglia C., Giulio G. et al. The clinical value of tumor burden at diagnosis in Hodgkin lymphoma // Cancer. – 2000. – Vol. 101. – P. 1824-1834.
7. Gobbi P.G., Ghirardelli M.L., Solcia M. et al. Image-aided estimate of tumor burden in Hodgkin's disease: evidence of its primary prognostic importance // J. Clin. Oncol. – 2001. – Vol. 19. – P. 1388–1394.
8. Hasenclever D., Diehl V. A prognostic score to predict tumor control in advanced Hodgkin's disease // N. Engl. J. Med. – 1998. – Vol. 339. – P. 1506–1514.
9. K p p e r s R., Schmitz R., Distler V. et al. Pathogenesis of Hodgkin's lymphoma // Eur. J. Haematol. – 2005. — Vol. 75 (Suppl. 66). – P. 26-33.
10. Specht L. Tumor burden as the main indicator of prognostic in Hodgkin's disease // Eur. J. Cancer. – 1992. – Vol. 28. – P. 1982-1985.
11. Specht L., Nissen N.I. Hodgkin's disease and stage // Eur. J. Haematol. – 1989. – Vol. 43. – P. 127-135.
12. Specht L., Nissen N.I. Prognostic factors in Hodgkin's disease stage III with special reference to tumour burden // Eur. J. Haematol. – 1988. – Vol. 41. – P. 80-87.
13. Specht L., Nissen N.I. Prognostic factors in Hodgkin's disease stage IV // Eur. J. Haematol. – 1988. – Vol. 41. – P. 359-367.
14. Specht L., Nissen N.I. Prognostic significance of tumour burden in Hodgkin's disease PS I and II // Scand. J. Haematol. – 1986. – Vol. 36. – P. 367-375.
15. Specht L., Nordentoft A.M., Cold S. et al. Tumor burden as the most important prognostic factor in early stage Hodgkin's disease. Relations to other prognostic factors and implications for choice of treatment // Cancer. 1988. – Vol. 61. – P. 1719-1727.
16. Specht L., Lauritzen A.F., Nordentoft A.M. et al. Tumor cell concentration and tumor burden in relation to histopathologic subtype and other prognostic factors in early stage Hodgkin's disease. The Danish National Hodgkin Study Group // Cancer. – 1990. – Vol. 65. – P. 2594–2601.
17. Steidl C., Connors J.M., Gascoyne R.D. Molecular pathogenesis of Hodgkin's lymphoma: increasing evidence of the importance of microenvironment // J. Clin. Oncol. – 2011. – Vol. 29. – P. 1812-1821.

Поступила в редакцию 20.12. 2015 г.

S.V.Ivanova¹, S.A.Kuleva^{1,2}, A.P.Karitsky¹, T.Yu. Semiglazova^{1,3}

Creation of a statistical model for prediction of Hodgkin's lymphoma in young adults

¹N.N.Petrov Research Institute of Oncology

²St. Petersburg State Pediatric Medical University

³I.I.Mechnikov North-West State Medical University
St. Petersburg

The aim of this study was to create a multi-dimensional correlation matrix of determinants giving information on the degree of influence on survival of young adults with Hodgkin's lymphoma each variable as well as the effect of the interaction of these variables with each other. 87 patients with Hodgkin's lymphoma at the age of 19 to 29 years (mean age 24 ± 4 years) were included in the study. Multiple matrix containing the coefficients of correlation of survival and correlation coefficients of 35 analyzed factors was of 5 significant determinants (volume of tumor lesion, stage IV of disease, E-damage, accelerated erythrocyte sedimentation rate, leukocytosis). Construction of the correlation matrix in order to select factors to be included in the equation, which contained the pair correlation coefficients of overall survival and each of factors, showed that only the volume of tumor lesion (correlation coefficient 0.2570, $p = 0.026$) influences on sign-result. Multiple regression equation is represented as follows $\bar{y}=0,166667-0,227273x[\text{volume}]$; $R^2=0,0643674507$. This equation allows for given values of the factor «volume of tumor lesion» to have theoretical values of resultant sign (survival), substituting the actual values of the factor in it.

Key words: Hodgkin's lymphoma, young adults, statistical model of prognosis