

УДК 614.2, 616.65-002

АНАЛИЗ СТРУКТУРЫ СОЦИАЛЬНО-ГИГИЕНИЧЕСКИХ ХАРАКТЕРИСТИК ПАЦИЕНТОВ С ДОБРОКАЧЕСТВЕННОЙ ГИПЕРПЛАЗИЕЙ ПРОСТАТЫ

С.Ю. Пехов

*Воронежская государственная медицинская академия им. Н.Н.Бурденко
Воронежский государственный технический университет*

Рассматриваются особенности реализации многомерных методов исследования структуры медицинских данных, а именно социально-гигиенических характеристик больных доброкачественной гиперплазией предстательной железы с целью разработки рациональной стратегии и тактики оказания специализированной медицинской помощи больным данной с патологией

Ключевые слова: доброкачественная гиперплазия предстательной железы, социально-гигиенические характеристики, методы многомерного статистического анализа

THE MULTIDIMENSIONAL ANALYSIS OF STRUCTURE OF THE CASE RATE THE GOOD-QUALITY HYPERPLASIA OF THE PROSTATE ON THE BASIS OF SOCIALLY-HYGIENIC CHARACTERISTICS OF PATIENTS

S.J. Pehov

The Voronezh state medical academy of N.N.Burdenko
The Voronezh state technical university

Features of realization of multidimensional methods of research of structure of copper-tsin'skih given, namely socially-hygienic characteristics of patients good-quality giperplazie-j a prostate for the purpose of working out of rational strategy and tactics of rendering of a spetsializiro-bathroom of medical aid by the patient given with a pathology are surveyed

Keywords: a good-quality hyperplasia of a prostate, socially-hygienic characteristics, methods of the multidimensional statistical analysis

Известно, что доброкачественная гиперплазия предстательной железы (ДГПЖ) является одним из самых распространенных заболеваний мужчин пожилого и старческого возраста. Пациенты с доброкачественной гиперплазией предстательной железы составляют 13,36% больных урологических стационаров, и этот показатель имеет тенденцию к увеличению [1]. С увеличением продолжительности жизни населения, страдающих гиперплазией простаты, становится все больше, а значит, вопросы, касающиеся оптимальной терапии данного заболевания, приобретают еще большую значимость. Более того, крупномасштабные популяционные исследования последних лет убедительно доказали, что гиперплазия простаты является заболеванием, имеющим прогрессирующий характер [2]. Появление и увеличение с возрастом степени выраженности симптомов заболевания, присоединение различных, порой тяжелых, осложнений болезни, крайне отри-

цательно сказываются на качестве жизни пациентов и делают её во многих случаях просто невыносимой.

В связи с актуальностью исследований по организации медицинской помощи больным ДГПЖ предлагается возможный вариант схемы исследования пациента с ДГПЖ ориентированная на методы системного анализа, обработку информации и принятие решений (рис.1).

В связи с необходимостью уточнения и структуризации социально-гигиенических и медико-биологических факторов риска в развитии доброкачественной гиперплазией предстательной железы и осложнений нами были исследованы две группы пациентов: основная (196 человек) с подтвержденным данными биопсии диагнозом ДГПЖ и контрольная (141 человек) - практически здоровых мужчин, не страдающих урологическими заболеваниями но у которых в течении жизни хотя бы раз выявлены незначительные нарушения мочеиспускания. На основе сформированной информационной базы была рассчитана частота появления различных ответов в анализируемых группах. Одной из задач анализа являлось изучения связи между распределением исследуемых показателей внутри сравниваемых групп, что приводило к построению соответствующих таблиц сопряженности. Для каждой из таких таблиц вычислялось значение критерия χ^2 – Пирсона для установления наличия статистически значимых различий между двумя качественными признаками. В качестве допустимой ошибки первого рода во всех выводах использовался порог 0,05.

Так, например, из рис.2 можно сделать вывод, что среди исследуемого контингента больных первой (основной) группы преобладают лица 66-74 лет, что составляет 42,34% в то время как в контрольной группе лица того же возраста составляют 9,6%. Такая разница (32,74%) еще раз подтверждает значимость возраста как фактора риска доброкачественной гиперплазии предстательной железы.

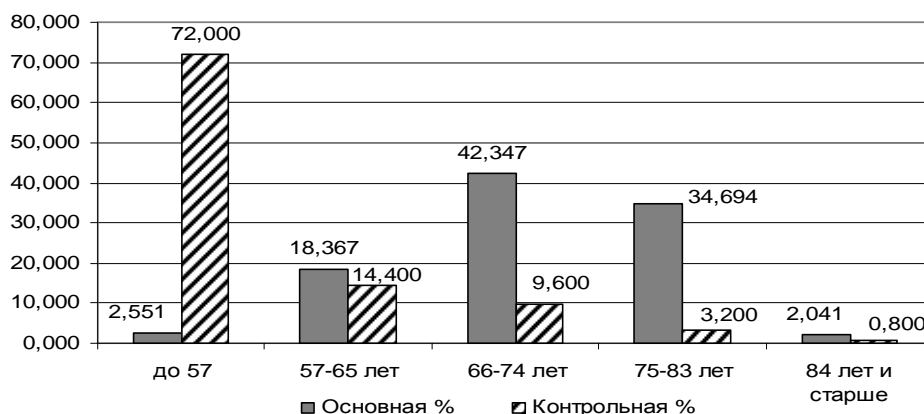


Рис. 2. Распределение числа исследуемых лиц по признаку «Возраст»

Далее для статистической обработки и моделирования все данные должны быть представлены в численном виде. Учитывая, что анкетные данные были представлены как в количественном, так и в качественном виде, мы сочли уместным преобразовать информацию, содержащую фиксированные смысловые (лингвистические) значения к численному виду [3]. Пример градаций численных значений для показателя «уровень образования» представлен в таблице 1.

Таблица 1.

Численные оценки значений показателя «уровень образования»

Лингвистическое значение показателя	Численная оценка показателя	Нормированная оценка
Высшее	8,500	1,0000
Неполное высшее	6,875	0,8088
Среднее специальное	5,000	0,5882
Среднее	2,750	0,3235
Неполное среднее	0	0

Взаимосвязь социально-гигиенических характеристик обследованных контингентов представлялась в виде коэффициентов парной и множественной корреляции. В таблице 2 представлены некоторые корреляционные связи проранжированные в зависимости от степени влияния на риск возникновения ДГПЖ и возможность развития онкологических осложнений.

Таблица 2.

Корреляционные связи

№	Наименование показателя	Риск ДГПЖ		Риск онкологии		№	Наименование показателя	Риск ДГПЖ		Риск онкологии	
		эфф.	ранг	ко-эфф.	ранг			ко-эфф.	ранг	ко-эфф.	ранг
1	Возраст	0,745	3	0,321	9	10	Характер сна	0,774	2	-0,019	15
2	Образование	0,319	10	0,038	14	11	Гимнастика	-0,188	12	-0,059	13
3	Сфера производства	0,691	6	-0,612	3	12	Курите ли Вы	0,103	14	0,445	7
4	Стаж работы	0,740	4	0,412	8	13	Проф. группа	0,690	7	0,561	5
5	Семейное положение	0,263	11	-0,087	12	14	Материальное положение	0,444	9	0,572	4
6	Жилищно-бытовые условия	0,560	8	0,285	10	15	Оценка здоровья	0,780	1	0,781	2
7	Прием пищи	0,018	15	0,089	11	15	Характер сна	0,774	2	-0,019	15
8	Состав пищи	-0,137	13	0,872	1						
9	Часы сна	-0,694	5	-0,473	6						

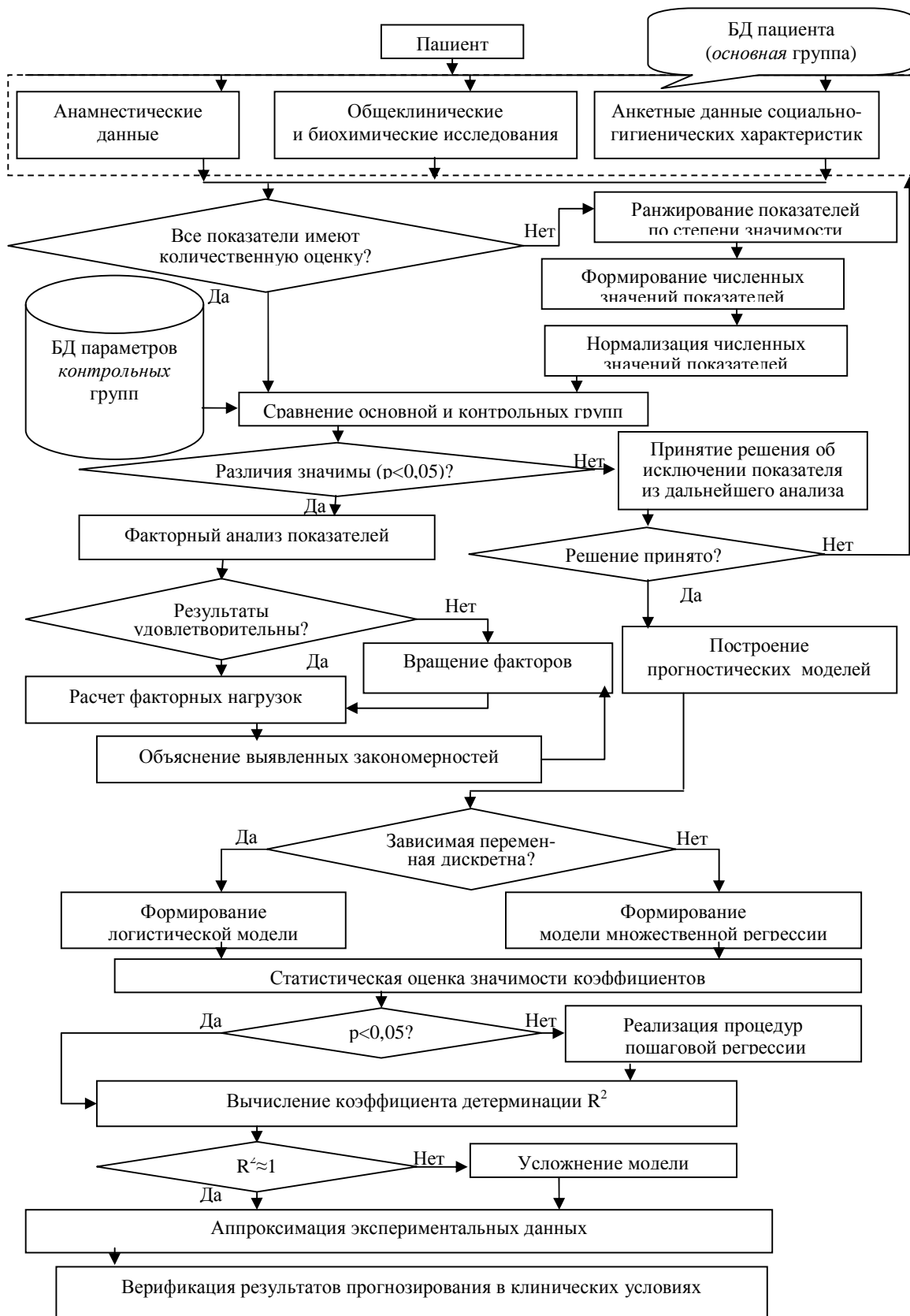


Рис. 1. Схема многомерного анализа структуры заболеваемости доброкачественной гиперплазией простаты

Для выявления скрытых факторов (критериев), характеризующих состояние больных проведен факторный анализ по данным анкетирования 196 человек у которых данными биопсии подтвержден диагноз ДГПЖ. Для установления количества факторов, которые необходимо выделить использовались критерии Кайзера или Кейтеля. На рис. 3 и 4 представлены фрагмент таблицы собственных значений и график «каменистой осыпи».

Extraction: Principal components				
Value	Eigenvalue	% Total variance	Cumulative Eigenvalue	Cumulative %
1	7,630697	14,67442	7,63070	14,67442
2	5,865963	11,28070	13,49666	25,95512
3	5,452980	10,48650	18,94964	36,44162
4	4,264894	8,20172	23,21453	44,64333
5	3,764419	7,23927	26,97895	51,88260
6	3,059113	5,88291	30,03807	57,76551
7	2,842248	5,46586	32,88031	63,23137
8	2,380710	4,57829	35,26102	67,80966
9	1,947463	3,74512	37,20849	71,55478
10	1,878374	3,61226	39,08686	75,16704
11	1,892635	3,25507	40,77950	78,42211
12	1,597821	3,07273	42,37732	81,49484
13	1,470049	2,82702	43,84737	84,32186
14	1,254400	2,41231	45,10177	86,73416
15	1,175106	2,25982	46,27687	88,99398
16	1,021262	1,96397	47,29813	90,95795

Рис.3. Собственные значения факторов

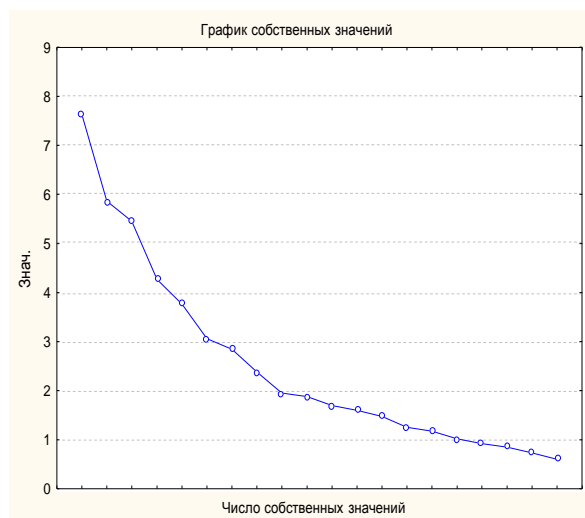


Рис.4. График каменистой осыпи

Во втором столбце таблицы приведены дисперсии выделенных факторов — собственные числа. В третьем столбце для каждого фактора приводится процент от общей дисперсии (в данном примере она равна 53). Как видно, первый фактор объясняет 14,7 % общей дисперсии, второй фактор — 11,2 % и т. д.

Далее программа (Statistica 8.0) выделила 4 независимых фактора (новые переменные), в которых доминирующую роль играют показатели, у которых большие факторные нагрузки. Так как общая дисперсия у *Фактора 1* выше, чем у остальных, *Фактор 1* более коррелирован с переменными, чем *Факторы 2* и *3*. Это следовало ожидать, потому что факторы выделяются последовательно и содержат все меньше и меньше общей дисперсии.

Для улучшения результатов факторного анализа и получения более интерпретируемых результатов можно применить повороты осей, которые достигаются вращением факторов [2]. Приведенный результат был получен методом вращения квартимакс нормализованных. На рис. 5. дано графическое представление факторных нагрузок.

Считается, что факторы выделены успешно, если группы переменных, составляющие факторы, расположены «плотно» в диаметрально про-

тивоположенных частях плоскости. Как видно из графика параметры в пространстве двух факторов сгруппированы по трем уровням в зависимости от значений коэффициентов корреляции.

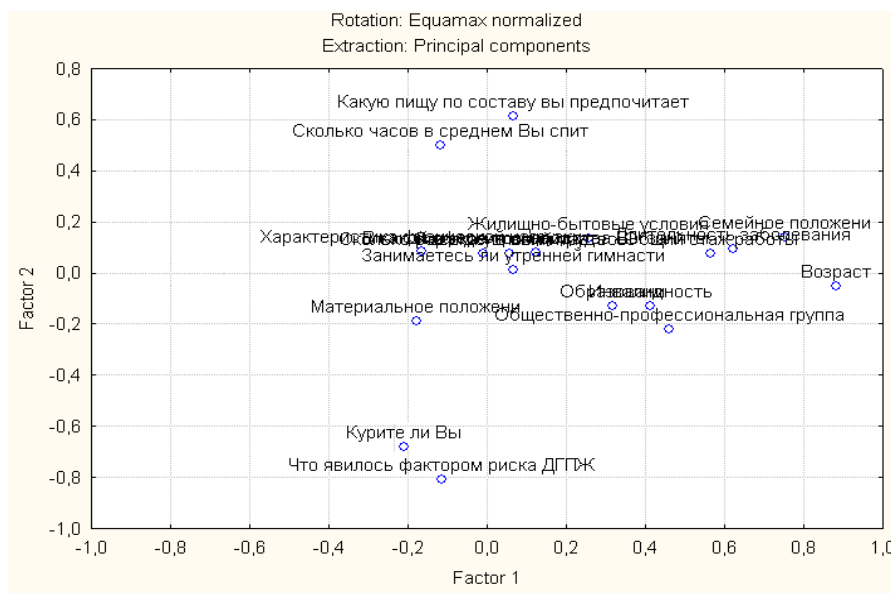


Рис. 5. Графическое представление факторных нагрузок

Например для *Фактора 1* линейная модель имеет вид:

$$\begin{aligned} \text{Фактор}_1 = & 0,254 * \text{Возраст} + 0,030 * \text{Образование} - 0,020 * \text{В какой сфере про-} \\ & \text{изводства Вы заняты} + 0,105 * \text{Общий стаж работы} + \\ & 0,224 * \text{Длительность заболевания} - 0,028 * \text{Что явилось фактором риска} \\ & \text{ДТПЖ} - 0,049 * \text{Характеристика физической нагрузки} - 0,057 * \text{Оценка усло-} \\ & \text{вий труда} + 0,225 * \text{Семейное положение} + 0,011 * \text{Жилищно-бытовые усло-} \\ & \text{вия} + 0,071 * \text{Сколько раз в день вы пьтаетесь} + 0,031 * \text{Какую пищу по со-} \\ & \text{ставу вы предпочитаете} - 0,213 * \text{Сколько часов в среднем Вы спите} - \\ & 0,026 * \text{Занимаетесь ли утренней гимнастикой} - 0,254 * \text{Курите ли Вы} + \\ & 0,076 * \text{Общественно-профессиональная группа} - 0,084 * \text{Материальное по-} \\ & \text{ложение} + 0,082 * \text{Инвалидность} \end{aligned}$$

Из таблицы (рис.3) видно, что для Факторов 1,2 показателям с сильными корреляциями в большинстве случаев соответствуют большие значения коэффициентов, а показателям со слабыми корреляциями — меньшие значения, близкие к 0.

Таким образом, применение факторного анализа к санитарно-гигиеническим показателям больных ДТПЖ позволило выделить факторы риска, группирующие совокупность социально-гигиенических характери-

стик пациентов и следовательно сократить число исследуемых параметров в дальнейшем при построении прогностических моделей.

На следующем этапе для прогнозирования вероятности развития доброкачественной гиперплазии предстательной железы применялись методы логистической регрессии. Для этого вводится зависимая переменная y , принимающая лишь одно из двух значений — 0 (событие не произошло) и 1 (событие произошло), и множество независимых переменных — вещественных x_1, x_2, \dots, x_n , на основе значений которых требуется вычислить вероятность принятия того или иного значения зависимой переменной [2].

В качестве выходной переменной фигурировала переменная Y_1 оценивающая риск развития ДГПЖ. В качестве информативных факторов при отсутствии мультиколлинеарности были отобраны следующие показатели: X_1 - оценка состояния здоровья; X_2 - материальное положение; X_3 - общественно-профессиональная группа; X_4 - жилищно-бытовые условия; X_5 - время сна; X_6 - возраст; X_7 - курение.

Уравнение модели имеет вид

$$\text{Риск ДГПЖ} = \exp(\eta) / (1 + \exp(\eta))$$

где

$$\eta = -0,51525 + 5,21468 * X_1 - 3,70954 * X_2 + 3,80166 * X_3 + 3,14887 * X_4 - 1,99568 * X_5 + 0,0996765 * X_6 + 1,99076 * X_7$$

Во второй модели в качестве зависимой была выбрана переменная Y_2 оценивающая риск развития онкологических осложнений. Независимые факторы следующие: X_1 - частота посещения уролога; X_2 - состав пищи; X_3 - жилищно-бытовые условия; X_4 - занятия утренней гимнастикой; X_5 - курение; X_6 - индекс I_PSS; X_7 - длительность заболевания

Уравнение модели имеет вид

$$\text{Онкологический риск} = \exp(\eta) / (1 + \exp(\eta))$$

где

$$\eta = -28,1985 + 28,7847 * X_1 + 7,48668 * X_2 - 2,78676 * X_3 - 46,7392 * X_4 + 1,793 * X_5 - 0,0462611 * X_6 - 0,087362 * X_7$$

В таблице 3 представлены результаты проверки точности моделей с использованием тестовых выборок ($N=50$).

Таблица 3.

Матрица распознавания по тестовой выборке

Риск развития ДППЖ	Правильное распознавание		Ошибочное распознавание		Риск развития онкологии	Правильное распознавание		Ошибочное распознавание	
	абс.	отн.	абс.	отн.		абс.	отн.	абс.	отн.
Высокий	24	48,0 %	1	2,0%	Высокий	21	42,0 %	4	8,0%
Низкий	21	42,0 %	4	8,0%	Низкий	17	34,0 %	8	16,0%
Итого	45	90,0 %	5	10,0 %	Итого	38	76,0 %	12	24,0%
Всего	50 человек				Всего	50 человек			

Как видно из таблицы построенная модель оценивающая риск развития ДППЖ точнее описывает результаты эксперимента – до 90% правильно распознанных объектов. В тоже время, относительно низкая точность модели прогнозирующей риск развития онкологических осложнений (76% правильно распознанных объектов) объясняется малым числом объектов включенных в тестовую выборку, среди которых были пациенты, у которых подтвержденные патоморфологическим исследованием злокачественные новообразования простаты были связаны с различными этиопатогенетическими причинами и слабо коррелировали с доброкачественной гиперплазией простаты в анамнезе.

Таким образом, исследование влияния характеристик медико-социальных факторов риска на заболеваемости доброкачественной гиперплазией предстательной железы на основе многомерных методов математической статистики является основой для принятия врачебных решений при проведении профилактических мероприятий.

Литература

1. Доброкачественная гиперплазия предстательной железы / Под ред. Акад. РАМН Н.А. Лопаткина. М., 1999. 216 с.
2. Pickard R, Emberton M, Neal D.E. The management of men with acute urinary retention. Br J Urol 1998;81:712–20.
3. Халафян А.А. Современные статистические методы медицинских исследований. М.: изд. ЛКИ, 2008, - 320 с.