

Типологизация заболеваемости городского населения Республики Башкортостан болезнями мочеполовой системы

Е.И.Нестеренко¹, М.А.Шарафутдинов², Е.М.Гареев²

¹Российский государственный медицинский университет, кафедра общественного здоровья и здравоохранения, Москва (зав. кафедрой – акад. РАМН, проф. Ю.П.Лисицын);

²Башкирский государственный медицинский университет, кафедра общественного здоровья и организации здравоохранения, Уфа (зав. кафедрой – проф. Н.Х.Шарафутдинова)

Изучение заболеваемости населения Республики Башкортостан болезнями мочеполовой системы провели на основе кластерного анализа. Представлены результаты исследования 20 городов (с прилегающими к ним районами) республики. В качестве «признаков» фигурировали сведения об обращениях взрослого населения по поводу 12 заболеваний мочеполовой системы, приведенные к стандартному виду «число обращений на 100 000 населения». С помощью кластерного анализа проведена классификация объектов – выделение «типологических групп» городов и районов соответственно сходству-различию «профилей описания» первичных обращений или «обращений в целом» по поводу рассматриваемой патологии. Исходные массивы данных предварительно были подвергнуты факторному анализу по способу «главных компонент», чтобы избежать включения в «профиль описания» коррелированных признаков. Выявлены территории с высоким уровнем заболеваемости городского населения Республики Башкортостан болезнями мочеполовой системы и определены значимые причины высокой или низкой заболеваемости данной территории страны.

Ключевые слова: взрослое население, заболеваемость, мочеполовая система, кластерный анализ, факторы риска

Typology of urogenital system morbidity among urban population in the Bashkortostan Republic

E.I.Nesterenko¹, M.A.Sharafutdinov², E.M.Gareev²

¹Russian State Medical University, Department of Public Health and Health Care, Moscow (Head of the Department – Acad. RAMS, Prof. Yu.P.Lisitsyn);

²Bashkortostan State Medical University, Department of Public Health and Health Care Organization, Ufa (Head of the Department – Prof. N.Kh.Sharafutdinova)

The level of prevalence and structure of urogenital system's morbidity of urban population constantly living in the Bashkortostan Republic are represented. The women's and men's chronic morbidity is analyzed depending on their age and residence in the Republic with the help of cluster statistics. The areas of danger of getting pathology of urogenital system are determined and the corresponding risk factors are identified. It allows the development of prophylactic medical and social programs.

Key words: adults, morbidity, urogenital system, cluster statistics, risk factors

В последние годы отмечается неуклонный рост заболеваемости населения болезнями мочеполовой системы (МПС). Причины этого – социально-экономические условия жизни, в том числе образ жизни человека, условия труда, быта, организация питания, неблагоприятная экологическая обстановка, неуклонный рост числа лиц пожилого и старческого возраста.

Материалы и методы

Изучение заболеваемости населения Республики Башкортостан болезнями мочеполовой системы провели на основе кластерного анализа. Исходным материалом для анализа послужили отчетные формы №12 учреждений здравоохранения и медико-информационно-аналитического центра Министерства здравоохранения РБ за 2001–2005 гг.

Исследованием охвачено 20 городов (с прилегающими к ним районами). В качестве «признаков» фигурировали сведения об обращениях взрослого населения по поводу 12 заболеваний МПС, приведенные к стандартному виду «число обращений на 100 000 населения». Список заболеваний МПС включал в себя: гломерулонефрит (ГЛН), почечную недостаточность (ПН), мочекаменную болезнь (МКБ), болезни

Для корреспонденции:

Нестеренко Елена Ивановна, доктор медицинских наук, профессор кафедры общественного здоровья и здравоохранения Российского государственного медицинского университета

Адрес: 117997, Москва, ул. Островитянова, 1

Телефон: (495) 434-0700

E-mail: nesei@yandex.ru

Статья поступила 24.12.2007 г., принята к печати 30.06.2008 г.

предстательной железы (БПЖ), мужское бесплодие (МБП), доброкачественную дисплазию молочной железы (ДДМ), сальпингит (СЛП), эндометриоз (ЭДМ), эрозию шейки матки (ЭШМ), расстройства менструаций (РМС), нарушения менопаузы (НМП) и женское бесплодие (ЖБП). Были составлены таблицы, отражающие число первичных обращений по поводу этих заболеваний в календарном году. Вариации числа обращений по всей примененной совокупности признаков (заболеваний МПС) представляли собой «профиль описания» объекта.

Задачей исследования являлась классификация объектов – выделение «типологических групп» городов и районов сообразно сходству-различию «профилей описания» первичных обращений или «обращений в целом» по поводу МПС. Для решения этой задачи был использован математико-статистический аппарат кластерного анализа [1–3]. Метод кластерного анализа позволяет обнаруживать (или создавать) компактные группировки (кластеры) внутри некоего рассматриваемого множества объектов. Предварительно осуществляется расчет взаимного положения объектов в m -мерном пространстве признаков, заданном размерностью их «профиля описания». Кластеры представляют собой «сгущения» объектов в этом пространстве и отличаются тем, что расстояние между входящими в них объектами существенно меньше, чем между самими кластерами.

Поскольку все признаки имели одну и ту же размерность и являлись количественными переменными, для построения матрицы расстояний был использован расчет «квадрата евклидова расстояния». Это позволяло более отчетливо различать объекты как по уровням числовых значений их «профилей описания» (по «сдвигу»), так и по различиям значений отдельных компонент профиля (по «форме» или «рисунку», «паттерну» профилей). Для выделения кластеров был использован метод Уорда (Ward), обеспечивающий получение

очень компактных и «отчетливых» кластеров примерно равных размеров.

Кроме того, чтобы избежать включения в «профиль описания» коррелированных признаков, исходные массивы данных предварительно были подвергнуты факторному анализу по способу «главных компонент» [4–6]. Факторный анализ основан на предположении, что корреляция между исследуемыми признаками порождена наличием «скрытых» (непосредственно не измеряемых) признаков-«факторов», опосредованным проявлением которых как раз и являются признаки, измеряемые исследователем. Математико-статистический аппарат факторного анализа позволяет без значительной потери содержания перейти от исходной размерности пространства признаков к существенно меньшей, включающей в себя ограниченный набор независимых друг от друга (ортогональных) факторов. Благодаря этому возникает возможность оперировать «значениями факторов» (factor scores), являющимися координатами положения объектов в факторном пространстве, причем крайне важно, что, несмотря на ту или иную степень коррелированности исходных признаков, корреляция между «значениями факторов» всегда равна нулю. Это позволяет эффективно использовать их как «профиль описания» объектов для проведения кластерного анализа, поскольку каждый фактор вносит свой независимый информационный вклад, а «значения факторов», независимо от числа и размерности абстрагированных ими исходных признаков, представляют собой нормированные отклонения объектов по факторным осям от начала координат.

Результаты исследования и их обсуждение

Рассмотрим сначала результаты обработки данных по городам. Начнем с анализа «первичных обращений». Обработка исходных данных методом главных компонент приве-

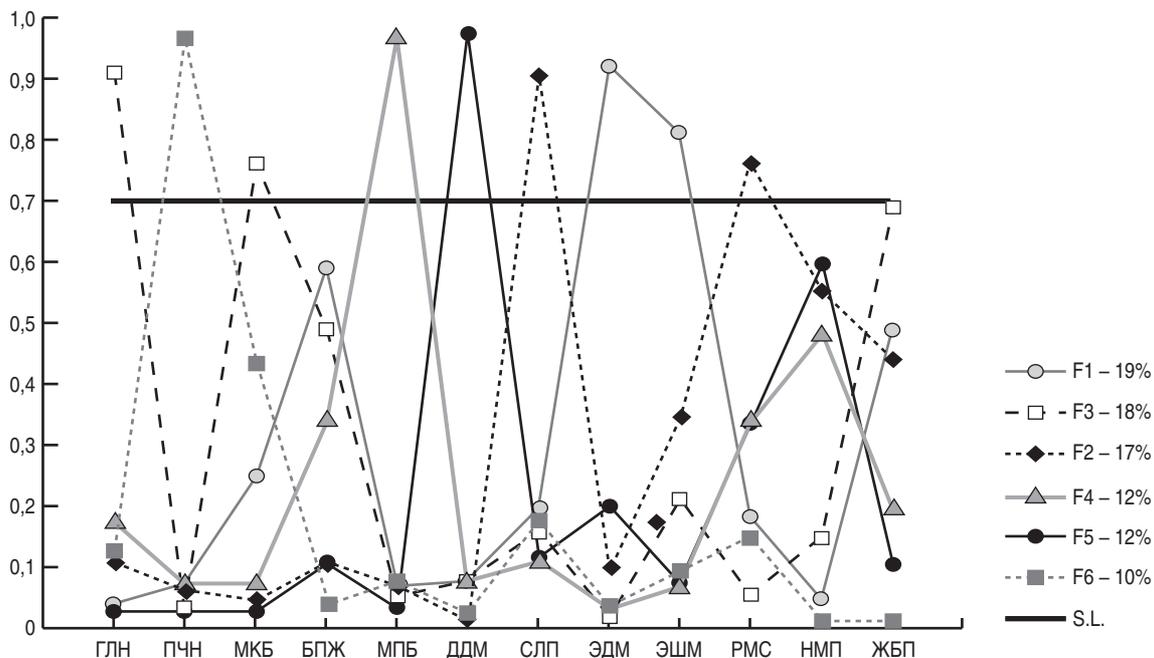


Рис. 1. Факторная структура «первичных обращений» по поводу заболеваний МПС по городам РБ. По оси абсцисс – «признаки» – заболевания МПС, включенные в «профиль описания»; по оси ординат – факторные нагрузки (корреляции признаков с осью независимых факторов); F1–F6 – «главные компоненты», «факторы» с указанием доли объясняемой ими дисперсии.

ла к выделению шести факторов, описывающих совместно 88% дисперсии корреляционной матрицы (рис. 1). Как видно на рисунке, лишь 3 фактора можно отнести к категории «общих», т.е. являющихся «источником вариации» двух и более переменных.

Первый фактор отражает, прежде всего, весьма тесную корреляцию ($r = 0,68$) между частотой первичных обращений по поводу ЭДМ и ЭШМ. Нагрузку на этот фактор, близкую к значимой, имеет и частота обращений по поводу БПЖ, имеющая относительно тесные ($r > 0,4$) корреляции с ЭДМ и ЭШМ. Вторая компонента отражает достаточно тесные связи ($r = 0,53$ и $0,61$) частоты обращений по поводу ГЛН, МКБ и ЖБП.

Третья компонента порождена, прежде всего, достаточно тесной корреляцией ($r = 0,55$) частоты обращений по поводу СЛП и РМС. Три прочих компоненты можно отнести к разряду «характерных», отражающих относительную самостоятельность вариаций частоты обращений по поводу МБП, ДДМ и ПН. Отмечена «факторная сложность» (наличие значимых или близких к значимым нагрузок более чем по двум факторным осям) переменных ЖБП, НМП и в какой-то мере БПЖ. Необходимо заметить, что эта переменная ЖБП из 11 корреляций с прочими имеет четыре корреляции более 0,6 (с ГЛН, МКБ, СЛП и ЭШМ) и две – порядка 0,5 (с БПЖ и ЭДМ). В этой связи обращает на себя внимание наличие корреляции ($r = 0,52$) частоты обращений по поводу женского бесплодия и заболеваний предстательной железы (БПЖ) при отсутствии существенных корреляций частоты ЖБП с частотой обращений по поводу иных заболеваний женской половой сферы – РМС и НМП ($r < 0,3$). Отметим и практическое отсутствие корреляции частоты обращений по поводу мужского и женского бесплодия ($r = -0,08$).

Классификационное «дерево», построенное описанным выше образом по значениям факторов, представлено на рис. 2. Узлы ветвления «древа» отражают расстояние, на котором произошло объединение объектов в один кластер (типологическую группу), а равно включение в кластер нового объекта или кластера более низкого уровня.

Исходя из особенностей диаграммы последовательного «слияния» объектов, мы сочли возможным выделить 6 кластеров (рассечь «дерево классификации» на 6 ветвей). Первые три кластера представлены «уникальными объектами», т.е. включали в себя по одному городу – это города Учалы, Благовещенск и Мелеуз. Четвертый кластер включал в себя пять городов: Бирск, Ишимбай, Белебей, Сибай и Баймак. Пятый был образован одним единственным городом Дюртюли. Зато шестой включал в себя сразу 11 городов: Салават, Октябрьский, Нефтекамск, Давлеканово, Стерлитамак, Уфа, Кумертау, Туймазы, Белорецк, Янаул и Агидель.

Поскольку «значения факторов» – полезные для понимания скрытой структуры явления, но абстрактные и, следовательно, лишенные наглядности характеристики, в ходе дальнейшего анализа мы вернулись к исходным признакам – частотам первичных обращений в каждом из выделенных кластеров. При этом если кластер включал в себя два и более городов, мы рассчитывали усредненный «профиль описания». Кроме того, для сравнения был рассчитан усредненный профиль частот первичных обращений для всех 20 городов. «Профили описания», отражающие межгрупповые особенности распределения частоты первичных обращений по поводу различных проявлений заболеваний МПС, представлены в таблице.

Необходимо отметить, что в среднем по всем городам по частоте первичных обращений (далее указаны среднее число обращений \pm Sd) лидируют такие заболевания, как эрозия шейки матки (586 ± 263) и сальпингит (544 ± 347). Следующий «уровень» образуют гломерулонефрит (339 ± 303) и расстройство менструаций (328 ± 200). Далее следуют болезни предстательной железы (252 ± 174), нарушения менопаузы (158 ± 114), доброкачественная дисплазия молочной железы (145 ± 191) и мочекаменная болезнь (107 ± 80). В следующей группе заболеваний частота обращений исчисляется уже лишь несколькими десятками или единицами: женское бесплодие (52 ± 31), эндометриоз (31 ± 43) и муж-

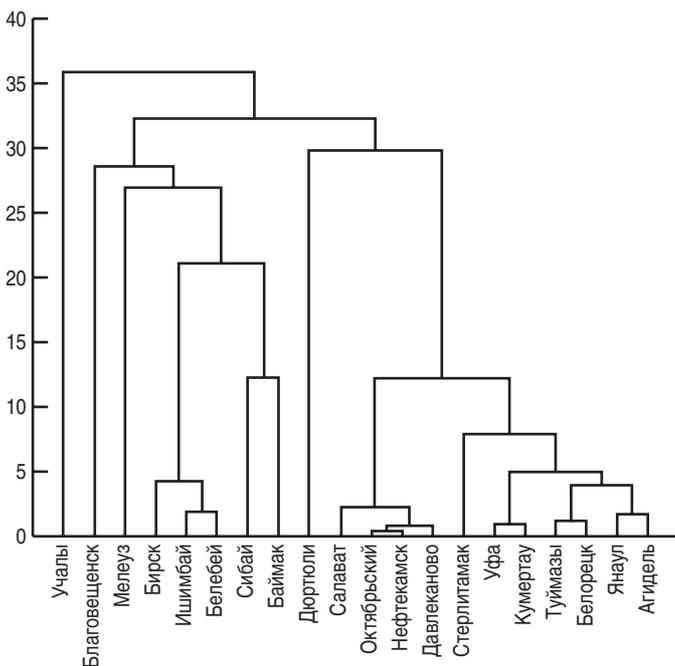


Рис. 2. «Дерево классификации» 20 городов РБ по структуре профилей первичных обращений по поводу заболеваний МПС. По оси абсцисс – города РБ; по оси ординат – квадрат евклидова расстояния между объектами (городами) в 6-мерном пространстве признаков.

Таблица. Профили частоты первичных обращений по поводу заболеваний МПС по шести кластерам городов РБ и в среднем по городам в целом

Болезни МПС	Кл. 1	Кл. 2	Кл. 3	Кл. 4	Кл. 5	Кл. 6	Все города
ЭШМ	794	229	475	693	1250	500	586
СЛП	437	368	1137	904	559	348	544
ГЛН	621	326	83	609	259	223	340
РМС	638	556	412	425	499	199	328
БПЖ	463	61	144	291	548	214	251
НМП	392	390	153	225	132	87	157
ДДМ	79	799	6	122	323	99	145
МКВ	126	41	77	153	110	92	106
ЖБП	41	22	46	89	98	36	52
ЭДМ	36	8	28	35	193	17	31
МБП	165	0	12	4	2	4	12
ПН	7	5	31	4	1	4	5

В таблице указана округленная до целых значений частота первичных обращений, рассчитанная на 100 000 взрослого населения. Заболевания МПС расположены согласно убыванию частоты обращений, усредненной по всем городам.

ское бесплодие (12 ± 36 случаев), а замыкает ранжированный по частоте ряд почечная недостаточность (5 ± 7). Таким образом, максимальные и минимальные частоты первичных обращений по поводу разных заболеваний МПС различаются практически в сто раз.

Из приведенных данных хорошо видно, что распределения заболеваний по частоте первичных обращений, как и сами эти частоты, могут резко различаться для некоторых городов или групп городов. В частности, следует отметить, что коэффициент вариации (процентное отношение Sd к среднему значению) частоты обращений по всем заболеваниям составляет от 45% (ЭШМ) до 300% (МБП), что свидетельствует о более или менее выраженной, но повсеместно высокой неоднородности городов по рассматриваемым признакам. Так, например, по городу Дюртюли (Кл. 5) отмечается наибольшая частота первичных обращений по поводу эрозии шейки матки – 1250 на 100 000 населения, что примерно вдвое больше, чем в среднем по всем городам, в 1,6 раза больше, чем в ближайшем по этому признаку г. Учалы (Кл. 1. – 793 случая) и в 5,4 раза больше, чем в г. Благовещенске (Кл. 2), имеющем самые низкие значения первичных обращений по поводу ЭШМ (229 случаев). Город Мелеуз резко выделяется частотой обращений по поводу сальпингита – 1136 случаев. Это также вдвое выше общего среднего уровня и примерно втрое больше, чем в наиболее представительном (11 городов) шестом кластере (348 случаев) или в г. Благовещенске (Кл. 2 – 367 случаев). Кроме того, Мелеуз имеет чрезвычайно высокий уровень первичных обращений по поводу почечной недостаточности (31 случай). Это в 5,6 раза выше, чем в среднем по всем городам, и в 30 раз больше, чем в г. Дюртюли, имеющем минимальное число обращений (1 случай). При этом в г. Мелеузе отмечен минимум обращений по таким заболеваниям, как ГЛН (83 случая) и ДДМ (6 случаев). По ГЛН это в 4 раза ниже среднего уровня и в 7,5 раза ниже, чем в г. Учалы, имеющем самое высокое среди всех городов число обращений по поводу ГЛН (621 случай). По ДДМ г. Мелеуз в 24 раза уступает среднему уровню и в 133 раза г. Благовещенску, «лидирующему» по ДДМ (799 случаев). Благовещенск интересен еще и в том отношении, что, кроме частоты первичных обращений по поводу ЭШМ, имеет минимальные частоты обращений по поводу:

- БПЖ – 61 случай, что в 4 раза ниже среднего показателя и в 9 раз меньше, чем у «лидера» г. Дюртюли – 548 случаев;
- МКБ – 41 случай, что в 2,6 раза ниже среднего уровня и в 3,7 раза ниже уровня 4-го кластера – 153 случая;
- ЖБП – 22 случая против 52 в среднем и 98 в г. Дюртюли;
- ЭДМ – 8 случаев против 31 в среднем и 193 в том же г. Дюртюли;
- МБП – 0 случаев против 12 в среднем по всем городам и 165 в г. Учалы.

Рассмотрим два кластера, включающих в себя по несколько городов. Четвертый кластер, в который вошли пять городов, интересен тем, что, являясь «рекордсменом» по числу первичных обращений по поводу МКБ (в среднем 153 ± 125 случаев), имеет по всему списку заболеваний (исключая лишь мужское бесплодие и почечную недостаточность) уровень числа обращений существенно или многократно более высокий, чем в среднем по всем городам. При

этом надо отметить, что 4-й кластер занимает второе место после отмеченных выше городов Мелеуз и Учалы по числу первичных обращений из-за СЛП (в среднем – 904 ± 316) и ГЛН (в среднем – 609 ± 468).

Особенностью шестого кластера, в котором объединены 11 городов РБ (т.е. более половины), включая столицу г. Уфу, является то, что здесь средний уровень первичных обращений по всему списку заболеваний заметно или даже значительно (в разы) ниже, чем в среднем по всем городам. Важно отметить также, что ни в одном случае этот уровень не является не только «рекордно высоким», но и «рекордно низким».

Применение критерия χ^2 показало, что по профилю распределения частот первичных обращений все кластеры достоверно отличаются от усредненного распределения для всех городов, но для «представительных» 6-го и 4-го кластеров эти отклонения от «общих средних уровней» оказались, естественно (они вносят основной вклад в этот средний уровень), наименьшими ($\chi^2 = 245$ и 608 соответственно; $p < 0,0001$) в сравнении с городами, образовавшими уникальные кластеры: Мелеуз ($\chi^2 = 1190$), Дюртюлей ($\chi^2 = 2326$), Учалы ($\chi^2 = 3136$) и Благовещенск ($\chi^2 = 3948$).

Таким образом, кластерный анализ позволил выявить в Республике Башкортостан территории с высоким уровнем заболеваемости городского населения болезнями мочеполовой системы, сгруппировать города по уровню заболеваемости и определить значимые причины высокой или низкой заболеваемости в кластерах. Различия в заболеваемости населения болезнями МПС могут быть обусловлены факторами риска их возникновения на тех или иных территориях, а также уровнем организации медицинской помощи и наличием соответствующих специалистов.

Литература

1. Жамбю М. Иерархический кластер-анализ и соответствия. – М.: Финансы и статистика, 1988. – 342 с.
2. Олдендерфер М.С., Блэшфилд Р.К. Кластерный анализ. – В кн: Факторный, дискриминантный и кластерный анализ. – М.: Финансы и статистика, 1989. – С.139–200.
3. Уиллиамс У.Т., Ланс Дж. Н. Методы иерархической классификации. Статистические методы для ЭВМ. – М.: Наука, 1986. – С.269–300.
4. Енюков И.С. Методы, алгоритмы, программы многомерного статистического анализа. – М.: Финансы и статистика, 1986. – 232 с.
5. Иберла К. Факторный анализ. – М.: Статистика, 1980. – 398 с.
6. Ким Дж.-О., Мьюллер Ч.У. Факторный анализ: статистические методы и практические вопросы. – В кн.: Факторный, дискриминантный и кластерный анализ. – М.: Финансы и статистика, 1989. – С.5–77.

Сведения об авторах:

Шарафутдинов Марат Амирович, кандидат медицинских наук, ассистент кафедры общественного здоровья и организации здравоохранения Башкирского государственного медицинского университета
Адрес: 450000, Уфа, ул. Ленина, 3
Телефон: (347) 272-4221
E-mail: nazira-h@rambler.ru

Гареев Евгений Мусинович, кандидат биологических наук, доцент кафедры общественного здоровья и организации здравоохранения Башкирского государственного медицинского университета
Адрес: 450000, Уфа, ул. Ленина, 3
Телефон: (347) 272-4221
E-mail: nazira-h@rambler.ru