



УДК 303.722.4+314.482 (477.75)

ИСПОЛЬЗОВАНИЕ КЛАСТЕРНОГО АНАЛИЗА ДЛЯ КЛАССИФИКАЦИИ РАЙОНОВ РЕСПУБЛИКИ КРЫМ ПО УРОВНЮ СМЕРТНОСТИ**USING CLUSTER ANALYSIS FOR THE CLASSIFICATION OF THE REPUBLIC OF CRIMEA TERRITORIES ON MORTALITY RATE****Т.Н. Голубова, Н.М. Овсянникова, З.Р. Махкамова
T.N. Golubova, N.M. Ovsyannikova, Z.R. Makhkamova**

*Медицинская академия имени С.И. Георгиевского
ФГАОУ ВО «КФУ имени В.И. Вернадского»
Россия, 295006, г. Симферополь, бульвар Ленина, 5/7
Medical Academy named after S.I. Georgievsky of Vernadsky CFU
Russia, 295006, Simferopol, Lenin Boulevard, 5/7*

E-mail: tn.golubova@yandex.ru

Ключевые слова: кластерный анализ, классификация, район, смертность, Республика Крым.
Key words: cluster analysis, classification, territory, mortality rate, Republic of Crimea.

Аннотация. В статье представлены результаты анализа территориальных особенностей общей смертности населения по основным причинам в Республике Крым по среднему значению за 2000-2012 гг. с использованием кластеризации методом k – средних. По 10 из 13, выбранных для анализа причин смертности, установлены статистически значимые различия, что позволило достоверно объединить территории Республики Крым в 4 кластера. Первый кластер объединил территории с наивысшим средним уровнем смертности по причинам болезней органов пищеварения, инфекционных и паразитарных болезней, болезней органов дыхания, нервной и мочеполовой системы. Статистически значимый максимальный средний уровень смертности от новообразований, эндокринных болезней, включая сахарный диабет, отмечен во 2-м кластере. Третий кластер характеризуется второй позицией по среднему уровню смертности в Крыму по классу болезней органов дыхания, самоубийствам, болезням нервной, эндокринной системы, включая сахарный диабет. Средние значения уровня смертности по причине болезней системы кровообращения и токсического действия алкоголя в этом кластере занимают третью позицию в республике. В четвертый кластер вошли территории с наивысшим средним уровнем смертности вследствие психических расстройств.

Resume. The paper presents the mortality rate analysis by main causes in the Republic of Crimea on the average value for the period of 2000-2012 by applying k-means clustering. Statistically significant difference was determined in the 10 causes of death out of 13 selected for analysis. It allowed to reliably combine the territory of the Republic of Crimea in 4 clusters. The first cluster includes the territories with the highest average level of mortality by cause of digestive diseases, infectious and parasitic diseases, respiratory, nervous and urinary systems diseases. Statistically significant maximum average mortality rate from neoplasms, endocrine diseases, including diabetes, marked in the 2nd cluster. The third cluster is characterized by the second position on the mortality rate in the Crimea in the class of respiratory diseases, suicides, diseases of the nervous, endocrine system, including diabetes. The average values of mortality rate due to cardiovascular diseases and toxic effects of alcohol in this cluster occupied the third place in the region. The territories with the highest average level of mortality due to mental disorders entered in the fourth cluster.

Введение

Различие в уровнях здоровья населения, как на международном, так и национальном уровнях является результатом не только бедности и ограничения доступности медицинской помощи, но и условий, в которых люди живут [Bellis et al., 2012]. Развитые страны обладают более богатым опытом борьбы с низким уровнем здоровья и больше стремятся к повышению продолжительности жизни, чем развивающиеся, которым тяжело справиться с различными факторами, такими как социальная исключенность, плохое обеспечение питанием, неразвитость транспорта и отсутствие политики в области занятости [Wilkinson, Marmot, 2003]. Применяя различные методы статистического анализа, можно выделить различные районы в зависимости от профиля здоровья, одним из которых необходимо внешнее вмешательство для устранения проблем, а другие могут справиться сами, перенимая определенные стратегии и активно взаимодействуя между собой. Стандартные методы кластерного анализа являются объективной методологией для комбинирования показателей здравоохранения с целью лучшей идентификации отдельных групп районов с похожими профилями здоровья населения и для установления связей между трендами каждого показателя, которые могут быть визуализированы по когортам [Friedman, Rubin, 1967; 4; Айвазян и др., 1985; Lisboa et al., 2008].

Цель

Цель исследования – оценка территориальных особенностей уровня и структуры общей смертности по основным причинам в Республике Крым по среднему значению показателей за период 2000-2012 гг.

Задачи исследования:

1. Оценка достоверности различий между территориями РК по основным причинам смертности для определения числа кластеров.
2. Определение районов, входящих в разные кластеры, с учетом различий по уровню смертности от основных причин.
3. Анализ структуры смертности по причинам в установленных кластерах.

Материалы и методы исследования

В качестве первичных данных использована официальная статистика уровня смертности по основным причинам и состояниям населения Республики Крым (РК) за период 2000-2012 гг. Для оценки уровня и структуры общей смертности в комплексе использована база данных по 13 основным причинам смертности (заболеваниям и состояниям) населения по всем городам и районам РК. Обозначение классов заболеваний и состояний, послужившими причинами смертности, использованных в кластерном анализе:

- S1–Некоторые инфекционные и паразитарные болезни A00-B99,
- S2–Новообразования C00-D48,
- S3–Болезни эндокринной системы, расстройства питания и нарушения обмена веществ E00-E90; включая S4–Сахарный диабет E10-14,
- S5–Психические расстройства и расстройства поведения F00-F99,
- S6–Болезни нервной системы G00-G99,
- S7–Болезни системы кровообращения I00-I99,
- S8–Болезни органов дыхания J00-J99,
- S9–Болезни органов пищеварения K00-K93,
- S10–Болезни мочеполовой системы N00-N99,
- S11–Травмы, отравления и некоторые другие последствия воздействия внешних причин S00-T98; включая S13–Токсическое действие алкоголя T51,
- S12–Внешние причины заболеваемости и смертности: преднамеренное самоповреждение (самоубийства) X60-X84.

Методика кластеризации

По данным уровня смертности по 13 основным причинам в городах и районах Крыма за период 2000-2012 гг. рассчитаны средние значения. После расчета средних значений уровня смертности за исследуемый период, проведена кластеризация методом k - средних. Для кластеризации использовалась программа Statistica (StatSoft Statistica 10.0 En). Кластеризация методом k-средних предполагает распределение N наблюдений по K кластерам (число K задается в начале процедуры) таким образом, чтобы каждое наблюдение попало в кластер с ближайшей средней.

Результаты и их обсуждение

Основываясь на проведенном ранее кластерном анализе заболеваемости населения РК, мы посчитали оптимальным, по наибольшему количеству статистически значимых различий между городами и районами по причинам смерти (10 из 13), произвести разделение на 4 кластера [Ефстафьева, Овсянникова 2010; Овсянникова и др., 2013]. Кластеризация методом k - средних по среднему уровню смертности населения РК за период 2000-2012 гг. статистически значимых различий данного показателя по причинам болезней системы кровообращения, травм и отравлений, и самоубийств не выявила. По всем остальным, выбранным для анализа причинам смертности, установлены статистически значимые различия, что позволило достоверно объединить административные территории РК в 4 кластера (см. таблицу).

В первый кластер вошли город Феодосия и Джанкойский район с наибольшими средними значениями уровней смертности по причине болезней системы кровообращения, второй позицией среди всех кластеров по причине смертности в результате новообразований ($p=0.004$) и травм и отравлений. В данном кластере достоверно выявлены, также, самые высокие в Крыму средние уровни смертности в результате болезней органов пищеварения, инфекционных и паразитарных болезней, болезней органов дыхания, болезней нервной и мочеполовой системы. Уровень смертности вследствие инфекционных и паразитарных болезней превышает средние значения в других кластерах в 2 раза (наивысший показатель отмечен в г. Феодосия – $71.1 \pm 23.5\%$), в результате бо-



лезней нервной системы – в 2-3 раза (максимум также в г. Феодосия – $33.3 \pm 20.3\%$). Анализ динамики смертности вследствие инфекционных и паразитарных болезней за период 2000–2012 гг. выявил два пика роста показателя, начиная с 2000 по 2006 гг. и с 2007 по 2009 гг., далее отмечен тренд снижения к 2011 г. и намечающийся рост с 2012 г. (рис.1). Данная картина свидетельствует о серьезной проблеме в регионе.

Средний уровень смертности в кластерах по причинам смерти
The Mean of the mortality rate in the clusters on the death causes

Знак класса	Причина смерти	Уровень смертности на 100 000 населения, $X \pm s$				Значимость различий, p
		1-й кластер	2-й кластер	3-й кластер	4-й кластер	
S1	Инфекционные болезни	65.7 ± 7.6	39.0 ± 11.6	33.1 ± 8.0	33.2 ± 9.6	0.001
S2	Новообразования	181.8 ± 71.4	228.3 ± 34.7	159.0 ± 27.0	169.2 ± 27.1	0.004
S3	Болезни эндокринной системы	0.9 ± 0.1	4.6 ± 0.9	2.7 ± 1.4	2.5 ± 0.6	0.007
S4	Сахарный диабет	0.6 ± 0.3	3.9 ± 0.8	2.3 ± 1.3	1.7 ± 1.1	0.02
S5	Психические расстройства	0.3 ± 0.2	0.5 ± 0.6	0.5 ± 0.5	1.7 ± 0.4	0.001
S6	Болезни нервной системы	22.8 ± 14.8	10.3 ± 1.9	10.9 ± 4.7	7.0 ± 1.6	0.01
S7	Болезни системы кровообращения	1155.7 ± 43.9	937.3 ± 96.0	958.4 ± 91.8	1025.5 ± 129.5	0.07*
S8	Болезни органов дыхания	62.1 ± 0.4	25.8 ± 2.9	47.0 ± 20.3	34.7 ± 9.7	0.05
S9	Болезни органов пищеварения	75.4 ± 2.1	62.1 ± 8.7	46.3 ± 4.8	54.1 ± 5.4	0.00001
S10	Болезни мочеполовой системы	13.0 ± 0.6	11.3 ± 4.9	7.0 ± 2.9	7.1 ± 2.8	0.04
S11	Травмы и отравления	134.9 ± 12.7	128.6 ± 10.9	119.8 ± 22.1	168.0 ± 63.3	0.11*
S12	Самоубийства	24.2 ± 1.5	21.1 ± 3.9	26.8 ± 3.9	27.1 ± 4.0	0.08*
S13	Токсическое действие алкоголя	14.0 ± 1.5	15.9 ± 3.2	15.1 ± 6.1	26.3 ± 9.7	0.03

Примечание: * различия статистически не значимы

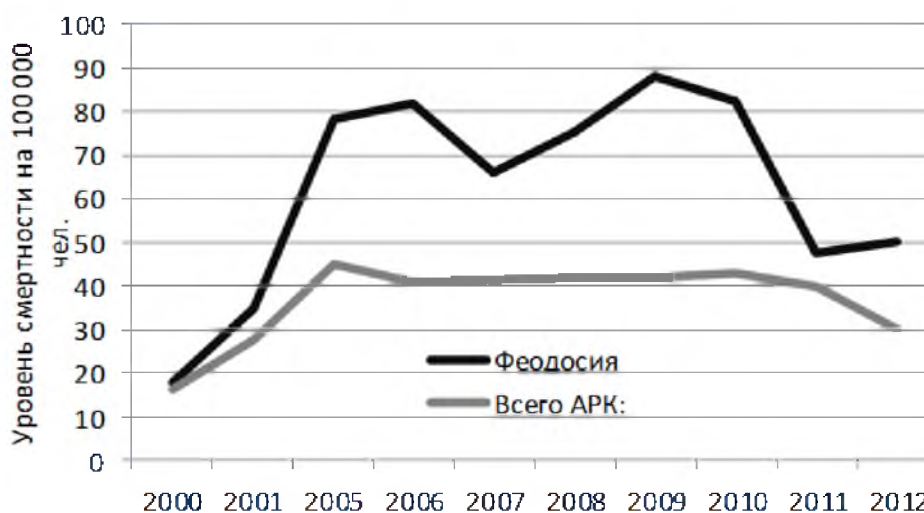


Рис. 1. Динамика уровня смертности по классу инфекционных и паразитарных болезней в РК и г. Феодосия за 2000–2012 гг.

Fig.1. Dynamics of death rate on a class of infectious and parasitic diseases in RK and Feodosiya for 2000–2012.

Уровень смертности в результате болезней нервной системы в г. Феодосии имел резкую тенденцию к росту с 2006 по 2009 гг. с последующим трендом снижения к 2012 г. (Рис. 2).

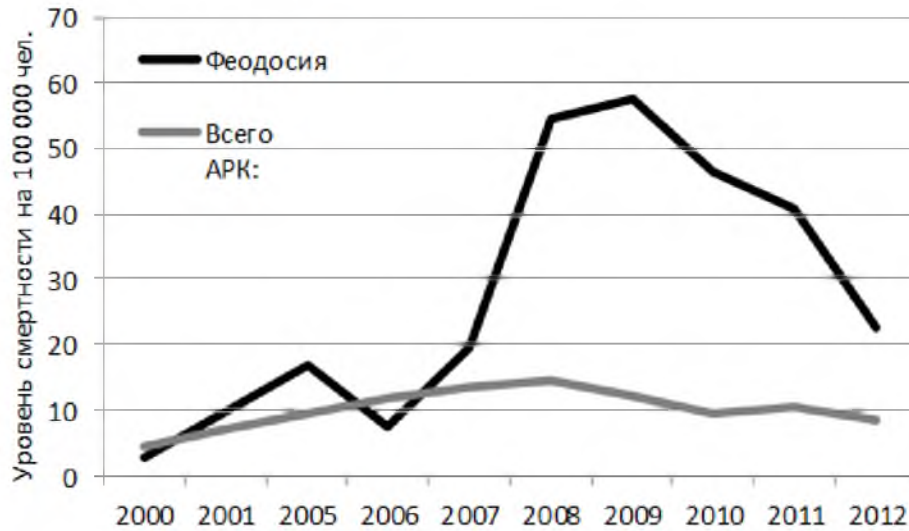


Рис. 2. Динамика уровня смертности по классу болезней нервной системы в РК и г. Феодосия за 2000-2012 гг.
 Fig.2. Dynamics of death rate on a class of diseases of nervous system in RK and Feodosiya for 2000-2012.

Второй кластер включил в себя города Симферополь, Керчь, Ялта и Евпатория с лидирующим среди кластеров средним значением уровня смертности в результате новообразований и третьей позицией по причине смертности в результате травм и отравлений. В этом кластере, кроме проанализированных причин, установлено максимальное среднее значение уровня смертности в результате болезней эндокринной системы ($p=0.007$), включая сахарный диабет ($p=0.02$); вторая позиция по уровню смертности как следствие инфекционной патологии ($p=0.001$), болезней органов пищеварения ($p=0.00001$), мочеполовой системы ($p=0.04$) и токсического действия алкоголя ($p=0.03$); третье место – в результате травм и отравлений, болезней нервной системы ($p=0.01$) и психических расстройств ($p=0.001$). Наивысший средний уровень смертности в этом кластере и Крыму в целом по причине новообразований отмечен в г. Керчь ($255.9 \pm 32.5\%$), в результате болезней эндокринной системы ($5.8 \pm 2.7\%$), включая сахарный диабет ($5.0 \pm 2.8\%$) – в г. Евпатория. При анализе динамики смертности от новообразований в г. Керчь видно, что этот показатель практически оставался неизменным с 2006 по 2011 гг. с тенденцией роста с 2012 г. (рис. 3).

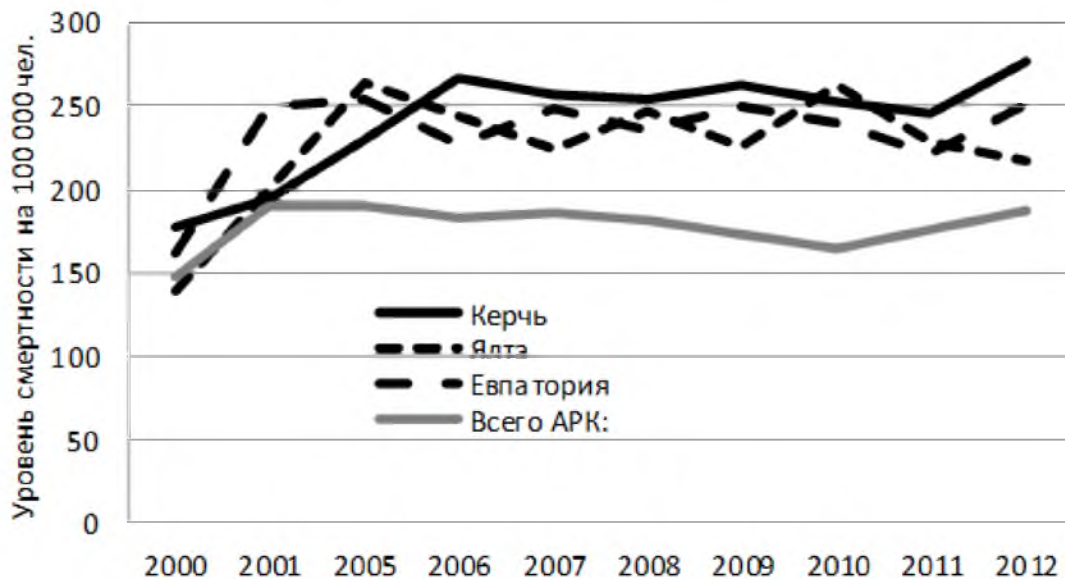


Рис. 3. Динамика уровня смертности по классу новообразований в РК и 2-ом кластере в 2000-2012 гг.
 Fig.3. Dynamics of death rate on a class of new growths in RK and the 2nd cluster in 2000-2012.



Третий кластер объединил город Судак и десять районов: Белогорский, Кировский, Красногвардейский, Ленинский, Нижнегорский, Первомайский, Раздольненский, Симферопольский, Советский и Черноморский. Этот кластер характеризуется второй позицией по среднему уровню смертности в Крыму по классу болезней органов дыхания, самоубийствам, болезням нервной, эндокринной системы, включая сахарный диабет. Средние значения уровня смертности по причине болезней системы кровообращения и токсического действия алкоголя в этом кластере занимают третью позицию в республике.

В четвертый кластер вошли города Алушта, Армянск, а также Бахчисарайский, Краснопереконский и Сакский районы. Данный кластер занимает первую позицию по уровню смертности в результате психических расстройств, вторую - по причине сердечно-сосудистых заболеваний, третью - по классу новообразований, болезней органов дыхания, пищеварения, эндокринной системы, включая сахарный диабет. Наивысший средний уровень смертности в этом кластере и Крыму в целом по классу психических расстройств установлен в Симферопольском районе ($1.7 \pm 1.3\%$) после г. Симферополь, который занимает по республике вторую позицию ($1.4 \pm 0.7\%$).

Структура смертности в результате наиболее часто регистрируемых причин внутри каждого кластера соответствует в целом показателям в регионе, но при этом отличается от Российской Федерации в целом. Большую долю традиционно во всех кластерах занимает смертность в результате сердечно - сосудистой патологии, на втором месте - новообразования, на третьем - травмы и отравления. Анализ различий между кластерами по структуре смертности по причинам показал, что 2-ой кластер отличается меньшей долей смертности по причине болезней системы кровообращения (63.0% против 66.0% и 67.0%) и большей по причине новообразований (15.0% против 10.0% и 11.0%). В 4-ом кластере выявлена наибольшая доля смертей в результате травм и отравлений (11.0% против 8.0% и 9.0%). При этом в Российской Федерации в целом, доля смертей в результате болезней системы кровообращения составила 54.9%, вследствие новообразований - 15.4%, в т. ч. злокачественных - 15.2%, травм и отравлений - 10.2% по данным 2012 года. Суммарная доля смертности по трем основным причинам (в сравнении с остальными причинами) по кластерам составила: в 1-ом - 84.0% (16.0%), во 2-ом и 3-м - 87.0% (13.0%) и в 4-ом - 88.0% (12.0%) соответственно.

Выводы

1. В результате кластеризации административных территорий Республики Крым по среднему уровню смертности за период 2000-2012 гг. статистически значимых различий по классу болезней системы кровообращения, травмам и отравлениям, и самоубийствам не установлено. По всем остальным, выбранным для анализа причинам смертности, установлены статистически значимые различия, что позволило объединить территории Крыма в 4 кластера.

2. В первый кластер вошли территории, объединенные наивысшими средними показателями смертности по классу болезней органов пищеварения, инфекционных и паразитарных болезней, болезней органов дыхания, нервной и мочеполовой системы. Внутри кластера в г. Феодосия выявлен наивысший в Крыму средний уровень смертности за период 2000-2012 гг. вследствие инфекционной патологии и болезней нервной системы. Статистически значимо максимальный средний уровень смертности за исследуемый период от новообразований, эндокринных болезней, включая сахарный диабет, отмечен во 2-м кластере. Внутри кластера наивысший уровень смертности от новообразований выявлен в г. Керчь с дальнейшей тенденцией роста с 2012 г. В г. Евпатория установлен максимальный по Крыму уровень смертности от эндокринной патологии, включая сахарный диабет. Третий кластер характеризуется второй позицией по среднему уровню смертности в Крыму по классу болезней органов дыхания, самоубийствам, болезням нервной, эндокринной системы, включая сахарный диабет. Средние значения уровня смертности по причине болезней системы кровообращения и токсического действия алкоголя в этом кластере занимают третью позицию в республике. В четвертый кластер вошли территории с достоверно наивысшим средним уровнем смертности вследствие психических расстройств.

3. Анализ структуры смертности по причинам показал, что 2-ой кластер отличается меньшей долей смертности по причине болезней системы кровообращения и большей - по причине новообразований; в 4-ом кластере выявлена наибольшая доля смертей в результате травм и отравлений.

Для установления причин территориальной неоднородности уровня смертности в Крыму, подтвержденной результатами проведенной кластеризации, необходимо исследовать связь между показателями здоровья населения в регионе и медико-организационными, социально-экономическими и эколого-гигиеническими факторами риска.

Список литературы References

- Айвазян С.А. и др. 1985. Прикладная статистика: Исследование зависимостей. Справочник. М., Финансы и статистика, 182.
- Ajvazjan S.A. i dr. 1985. Prikladnaja statistika: Issledovanie zavisimostej. Spravochnik. M., Finansy i statistika, 182 (in Russian).
- Ефстафьева Е.В., Овсянникова Н.М. 2010. Многомерный статистический анализ: первые шаги и апробации для решения медико-биологических проблем на территории Крыма. Ученые записки ТНУ. Т.23(62): 78-88.
- Efstafeva E.V., Ovsjannikova N.M. 2010. Mnogomernyj statisticheskiy analiz: pervye shagi i aprobacii dlja reshenija mediko-biologicheskikh problem na territorii Kryma. Uchenye zapiski TNU. T.23(62): 78-88 (in Russian).
- Овсянникова и др. 2013. Территориальные особенности заболеваемости в АР Крым по данным кластерного анализа. Научная дискуссия: вопросы медицины. Сборник статей по материалам XVI международной заочной научно-практической конференции. Москва. № 5 (12): 62-67.
- Ovsjannikova N.M. i dr. 2013. Territorial'nye osobennosti zaboлеваemosti v AR Krym po dannym klaster'nogo analiza. Nauchnaja diskussija: voprosy mediciny. Sbornik statej po materialam XVI mezhdunarodnoj zaochnoj nauchno-prakticheskoy konferencii. Moskva. № 5 (12): 62-67 (in Russian).
- Bellis M. A., et al. 2012. Using clustering techniques to identify localities with multiple health and social needs. Health & Place. 18(2): 138-143. doi:10.1016/j.healthplace.2011.08.003.
- Friedman, H.P., Rubin, J., 1967. On some invariant criteria for grouping data. Journal of the American Statistical Association 62: 1159—1178.
- Lisboa, P. et al. 2008. Cluster-based visualization with scatter matrices. Pattern Recognition 29: 1814—1823.
- Wilkinson, R.G., Marmot, M. 2003. Social determinants of health: the solid facts. World Health Organization, Denmark, 2nd, 31. www.euro.who.int/__data/assets/pdf_file/0005/98438/e81384.pdf