

© Рапаков Г. Г., Банщиков Г. Т., Горбунов В. А.,
Харахнин К. А., Ревелев И. М., 2019

Рапаков Георгий Германович

Кандидат технических наук, доцент,
Вологодский государственный университет
(Вологда, Россия)
E-mail: grapakov@yandex.ru

Rapakov Georgii Germanovich

PhD in Technical Sciences,
Associate professor,
Vologda State University
(Vologda, Russia)
E-mail: grapakov@yandex.ru

Банщиков Геннадий Трофимович

Доктор медицинских наук, врач-кардиолог,
Вологодская областная клиническая
больница (Вологда, Россия)
E-mail: vologdauzo@inbox.ru,
vol_obl_boll@mail.ru

Banshchikov Gennadii Trofimovich

Doctor of Medicine, M.D. Cardiologist,
Vologda region Clinical Hospital
(Vologda, Russia)
E-mail: vologdauzo@inbox.ru,
vol_obl_boll@mail.ru

Горбунов Вячеслав Алексеевич

Доктор физико-математических наук,
профессор,
Вологодский государственный университет
(Вологда, Россия)
E-mail: gorbunov1945@inbox.ru

Gorbunov Vyacheslav Alekseevich

Doctor of Physico-mathematical Sciences,
Professor,
Vologda State University
(Vologda, Russia)
E-mail: gorbunov1945@inbox.ru

Харахнин Константин Аркадьевич

Кандидат технических наук, доцент,
Череповецкий государственный
университет
(Череповец, Россия)
E-mail: harahninka@chsu.ru

Harahnin Konstantin Arkadievich

PhD in Technical Sciences,
Associate professor,
Cherepovets State University
(Cherepovets, Russia)
E-mail: harahninka@chsu.ru

Ревелев Игорь Михайлович

Врач, сердечно-сосудистый хирург,
БУЗ ВО Вологодская городская
больница № 2 (Вологда, Россия)
E-mail: postservice2016@yandex.ru

Revelev Igor' Mikhailovich

Physician, Cardiovascular Surgeon,
Vologda municipal hospital no 2
(Vologda, Russia)
E-mail: postservice2016@yandex.ru

**ПРИМЕНЕНИЕ КОМПЬЮТЕРНОГО
МОДЕЛИРОВАНИЯ ПРИ ОЦЕНКЕ
ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ
РЕГИОНАЛЬНОГО
ЗДРАВООХРАНЕНИЯ НА ОСНОВЕ
МЕТОДОВ СТАТИСТИЧЕСКОГО
АНАЛИЗА ДАННЫХ**

**COMPUTER SIMULATION
OF REGIONAL PUBLIC HEALTH
EFFICIENCY USING STATISTICAL
ANALYSIS METHODS**

Аннотация. Изучены возможности применения статистических алгоритмов и методов в социально-экономических исследованиях общественного здоровья региона. Обосновано их использование в сложном прикладном объекте исследования – системе здравоохранения Вологодской области. Представлены результаты корреляционного анализа и непараметрических мер взаимозависимости показателей эффективности регионального здравоохранения. На базе метода множественной регрессии выполнен анализ медико-демографического показателя здоровья населения (смертности) для обеспечения аналитической поддержки при принятии управленческих решений в ходе реализации государственной программы «Развитие здравоохранения Вологодской области» и совершенствования системы медико-демографического состояния территорий. С помощью результатов кластерного анализа и трехмерной визуализации данных, на основе карты линий уровня и контурного графика обнаружены аномалии и выявлен диспаритет для территорий Вологодского региона по ряду медико-демографических показателей мощности медицинских учреждений и финансовых показателей деятельности муниципального здравоохранения. Предложены мероприятия по преодолению выраженных территориальных диспропорций с целью улучшения демографической ситуации в регионе.

Ключевые слова: компьютерное моделирование, статистические методы, множественная регрессия, кластерный анализ, смертность населения, принятие решений

Abstract. The article presents the results of application of statistical algorithms and methods in social-economic studies of public health in the region. The use of these algorithms and methods to a complicated research object – health care system in the Vologda region – was justified. The comparative analysis of partial correlations as well as Kendall's tau and Spearman rank correlations of health care efficiency was conducted. The multiple regression method applied to population mortality analysis provides analytical support in the process of executive decision-making while implementing the state-run programme “Enhancement of health care system in the Vologda region” and improving the medico-demographic situation in the regional territories. Cluster analysis and data visualization, mapping programs and counter plots were used to reveal demographic anomalies and territorial disproportion in the system of medico-demographic parameters, capacity of health care facilities and financial performance of municipal public health. The authors suggest the system of measures to overcome this territorial disproportion and improve demographic situation in the region.

Keywords: computer modeling, statistical methods, multiple regression, cluster analysis, population mortality, decision making

Введение

В государственной программе «Развитие здравоохранения Вологодской области на 2014–2020 гг.», утвержденной постановлением Правительства Вологодской области от 28.10.2013 г. № 1112, отмечается, что демографическая ситуация в регионе характеризуется стойким процессом убыли населения. К задачам подпрограммы «Совершенствование системы территориального планирования Вологодской области» отнесены: оценка медико-демографического состояния территории и улучшение демографической ситуации в регионе [9]. В связи с этим актуальной является задача изучения возможностей компьютерного моделирования в социально-

экономических исследованиях здравоохранения Вологодской области. Практическая значимость работы определяется результатами применения методов статистического анализа для снижения социально-экономического бремени смертности населения. Цель работы состоит в изучении возможностей методов системного анализа для сложного прикладного объекта исследования – системы регионального здравоохранения, включая вопросы аналитической поддержки принятия управленческих решений, на основе статистических алгоритмов и методов обработки данных. Научная новизна работы заключается в сопоставлении методов статистических исследований и обосновании их использования в анализе общественного здоровья региона.

Основная часть

Вопросы применения методов статистического анализа при изучении общественного здоровья и здравоохранения нашли широкое отражение в работах [4], [15]. Практика использования современного программного обеспечения на основе статистических методов представлена в работах [2], [16], [17]. Информационной базой исследования являются показатели эффективности здравоохранения по территориям Вологодского региона [5]. К показателям мощности работы медицинских учреждений была отнесена обеспеченность врачами и средним медицинским персоналом на 10 тыс. человек населения; к медико-демографическим показателям – заболеваемость и болезненность на 100 тыс. человек, а также число умерших в расчете на 1 тыс. человек населения. Были использованы следующие финансовые показатели деятельности муниципального здравоохранения: расходы на здравоохранение по муниципальным образованиям Вологодской области из всех источников финансирования и среднемесячная зарплата в отрасли по районам.

При анализе общественного здоровья региона исследованы возможности применения и сопоставлены результаты следующих статистических методов: корреляционного анализа зависимостей, поиска связей на основе непараметрических методов, множественной регрессии и кластерного анализа. Для выявления сложных многомерных и нелинейных взаимосвязей использована трехмерная визуализация данных на основе карты линий уровня и контурного графика.

Результаты изучения источников и их анализ сопоставлялись с целями и методами, представленными в литературе. В работах [6]–[8] вопросы смертности, продолжительности жизни и общественного здоровья рассматриваются с позиций системного анализа. Широко используются математические методы. Так, в задаче количественной диагностики старения человека для определения биовозраста используется алгоритм множественной регрессии. Обсуждаются недостатки моделей и оценивается эффективность методик. В разработанном специализированном программном обеспечении учитывается нелинейный характер возрастной динамики. С помощью множественных корреляций исследуется системный характер старения. Управление целевым параметром – качеством жизни населения – предлагается рассматривать с позиций кибернетического подхода: на основе систем с обратными связями. Представлена методика расчета индекса качества жизни и обоснован выбор индикаторов на основе требований информативности и надежности. Решению задач прогнозирования содействуют смешанные подходы, основанные на привлечении традиционной статистики и методов DataMining [1].

При выявлении корреляционных зависимостей проверку гипотезы о виде распределения (нормальное по критерию χ^2) прошли следующие медико-демографические и финансовые показатели (переменные): заболеваемость на 100 тыс. человек; болезненность на 100 тыс. человек; число умерших в расчете на 1 тыс. человек; расходы на здравоохранение из всех источников финансирования в расчете на 1 человека (руб.); зарплата.

Коэффициент корреляции определяет уровень линейной зависимости между случайными величинами. При исследовании степени связи двух переменных (l и h) может возникнуть ситуация, при которой каждая из них связана с третьей переменной (d) или группой переменных. Необходимо иметь возможность рассчитать степень зависимости между l и h независимо от d . Для этого используются частные коэффициенты корреляции.

Частные коэффициенты корреляции первого порядка определяются по формуле

$$\rho_{lh-d} = \frac{\rho_{lh} - \rho_{ld}\rho_{hd}}{\sqrt{1-\rho_{ld}^2}\sqrt{1-\rho_{hd}^2}},$$

где l, h – две переменные из множества X_1, X_2, \dots, X_k ; d – подмножество оставшихся $k-2$ переменных; ρ_{lh-d} – степень линейной зависимости между l и h , если влияние d переменных исключено.

Парный коэффициент корреляции вычисляется по формуле

$$\rho_{12} = \frac{\text{cov}(X_1, X_2)}{\sigma_1\sigma_2},$$

где $\text{cov}(X_1, X_2)$ – ковариация, $\text{cov}(X_1, X_2) = M[(X_1 - m_1)(X_2 - m_2)]$; m_1, m_2 – математические ожидания случайных величин X_1, X_2 ; σ_1, σ_2 – их среднеквадратичные отклонения [2], [14].

В результате расчетов были выявлены зависимости для следующих пар переменных: сильная прямая (0,89) – между заболеваемостью и болезненностью (на 100 тыс. человек); средняя обратная (-0,46) – между числом умерших (в расчете на 1 тыс. человек) и расходами на здравоохранение из всех источников финансирования (сумма на 1 человека (руб.)); средняя обратная (-0,48) – между числом умерших (в расчете на 1 тыс. человек) и зарплатой.

Коэффициент ранговой корреляции Спирмена вычисляется по формуле

$$r_s = 1 - \frac{6 \sum (x_i^I - y_i^I)^2}{n(n^2 - 1)},$$

где n – объем выборки (x_i, y_i) , $i = 1, 2, \dots, n$, двух переменных (X, Y) ; (x'_i, y'_i) – пара рангов, соответствующая паре (x_i, y_i) .

Коэффициент ранговой корреляции Кендалла (τ) вычисляется по формуле

$$\tau = 1 - \frac{4k}{n(n-1)},$$

где k – число инверсий (нарушений порядка) в ряду рангов последовательности значений переменной y'_i , $i = 1, 2, \dots, n$ при условии, что ранги переменной x'_i упорядочены.

Расчитаны непараметрические меры взаимозависимости для следующих показателей деятельности муниципального здравоохранения:

1) «обеспеченность врачами (на 10 тыс. человек)» – «зарплата»: $r_s = 0,411$, $\tau = 0,296$;

2) «число умерших (в расчете на 1 тыс. человек)» – «расходы на здравоохранение из всех источников финансирования (сумма на 1 человека (руб.))»: $r_s = -0,39$, $\tau = -0,27$.

Коэффициенты ранговой корреляции значимы ($\alpha = 0,05$).

Исследуем зависимость показателя «число умерших (в расчете на 1 тыс. человек)» от показателей мощности медицинских учреждений – обеспеченность врачами и средним медицинским персоналом (на 10 тыс. человек), а также от показателей финансирования – расходы на здравоохранение из всех источников (сумма на 1 человека (руб.)) и зарплаты.

Матричная форма записи для уравнения множественной линейной регрессии имеет вид

$$Y = A \cdot \beta + \varepsilon,$$

где Y – вектор значений зависимой переменной, $Y = (y_1, y_2, \dots, y_n)^T$; A – регрессионная матрица ($n \times k$) (где n – количество случаев наблюдения зависимой переменной Y ; k – количество параметров (факторов)),

$$A = \begin{pmatrix} 1 & x_{11} & x_{21} & \dots & x_{(k-1)1} \\ 1 & x_{12} & x_{22} & \dots & x_{(k-1)2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} & \dots & x_{(k-1)n} \end{pmatrix};$$

β – вектор значений параметров, $\beta = (\beta_0, \beta_1, \dots, \beta_{k-1})^T$; ε – вектор ошибок наблюдений зависимой переменной, $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$.

Для нахождения оценок $\tilde{\beta}$ параметров модели по результатам наблюдений используется метод наименьших квадратов:

$$\tilde{\beta} = (A^T \cdot A)^{-1} A^T Y.$$

Таким образом, задача множественной линейной регрессии заключается в нахождении коэффициентов уравнения регрессии (ее детерминированной составляющей):

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_{k-1} x_{k-1},$$

где Y – выходная (зависимая) переменная модели или отклик; x_i – входные (независимые) переменные или факторы; β_i – коэффициенты (параметры) регрессии.

Независимость переменных x_i понимается как возможность установления их значений произвольно или независимо при расчете отклика Y . Сами же факторы x_i могут влиять друг на друга. Линейная зависимость столбцов регрессионной матрицы A называется мультиколлинеарностью и обуславливает вычислительные проблемы при выполнении регрессионного анализа. Кроме объясняющих факторов свой вклад в значение Y вносит случайная составляющая ε . Ее наличие обусловлено влиянием множества неучтенных факторов, а также ошибок наблюдений [3], [12].

Для эффективного решения этой задачи в рамках подхода множественной регрессии используется пошаговый метод отбора наиболее существенных факторов. Множественная регрессионная модель прогноза для показателей, отобранных на уровне значимости $\alpha = 0,05$, имеет следующий вид:

$$Y = \beta_0 + \beta_1 x_1 = 30,84 - 0,003x_1.$$

Для проверки гипотезы о незначимости регрессионной модели был использован дисперсионный анализ. Уравнение регрессии значимо ($p = 0,008961$). Коэффициент множественной корреляции R , описывающий степень линейной зависимости между Y и факторами (или стандартизованный регрессионный коэффициент X -beta), равен 0,484623. Для проверки согласия регрессионной модели и исходных данных рассчитаны остатки – разности между наблюдаемыми и предсказанными моделью значениями отклика Y , а также использован критерий Дарбина – Уотсона. Гипотеза H_0 о равенстве сериальных корреляций принимается на уровне $2\alpha = 0,1$. Дисперсии ошибок наблюдений постоянны. Для проверки гипотезы о нормальном распределении остатков был использован критерий χ^2 . Так как вычисленный уровень значимости $p = 0,60104 > \alpha = 0,05$ (заданного уровня значимости), то гипотеза о нормальном распределении принимается. Таким образом, регрессионная модель значима, адекватна результатам наблюдений и может быть использована при прогнозировании: при удвоении зарплаты c число умерших (в расчете на 1 тыс. человек) снижается с 25 до 19, т. е. на 24 % при 95 %-м доверительном интервале от 16,35 до 21,3.

Кластерный анализ используется для группировки объектов с близкими значениями признаков. Результат классификации должен быть интерпретирован с учетом особенностей предметной области. В качестве критерия оптимальности разбиения используется расстояние (метрика) между любой парой объектов рассматриваемой

совокупности. Основная сложность состоит в том, что выбор расстояния неоднозначен. Наиболее популярна евклидова метрика $\rho_E(X_i, X_j)$.

$$\rho_E(X_i, X_j) = \left(\sum_{k=1}^p (X_{ki} - X_{kj})^2 \right)^{1/2},$$

где X_{ki} – значение k -го признака i -го объекта.

Кластерный анализ был выполнен для медико-демографического показателя (число умерших (в расчете на 1 тыс. человек)) и показателя финансирования (среднемесячная зарплата в отрасли по районам) с помощью древовидной кластеризации (соединения), которая относится к группе иерархических методов. Предварительная визуализация данных выполнялась при помощи диаграммы рассеяния. В качестве метрики, определяющей расстояние между кластерами, использован метод Варда со статистическим расстоянием между классами ρ_s ; в качестве метрики, определяющей расстояние между объектами, – евклидова метрика.

$$\rho_s = \frac{n_l n_m}{n_l + n_m} (\bar{X} - \bar{Y})^T (\bar{X} - \bar{Y}),$$

где n_i – число элементов в i -м классе; \bar{X}, \bar{Y} – «центры тяжести» классов [2], [10]–[12].

Выполнен сравнительный анализ при разбиении исходной совокупности объектов – территорий Вологодской области – на 3, 4 и 5 кластеров с помощью метода K -средних, который относится к группе эталонных методов кластерного анализа. На основании экспертной оценки результатов принято разбиение региона на 5 кластеров, состоящих из 7, 1, 9, 2 и 9 территорий. При помощи геоинформационной системы построена тематическая карта, отражающая географическое распределение переменной – номера кластера (рис. 1). Тематическая карта позволяет наглядно выделить неблагоприятные территории, объединенные в 1, 3 и 5-й кластеры.

Исследование таблицы профилей кластеров позволило выделить их характерные признаки:

- наилучшие показатели демонстрируют кластеры № 2 (г. Череповец) и № 4 (г. Вологда и Нюксенский район);

- смертность для кластера № 2 составляет 68 % от средней по области. Отрыв от аутсайдера по данному показателю – кластера № 3 составляет 1,52 раза;

- среднемесячная зарплата для кластера № 2 в 1,54 раза превышает среднюю зарплату по области. Отрыв от аутсайдера по данному показателю – кластера № 1 составляет 1,89 раза;

- ближайшие показатели по отношению к кластеру № 2 демонстрирует кластер № 4. Смертность для кластера № 2 составляет 82 % от средних значений по области, а зарплата – в 1,2 раза выше данных значений;

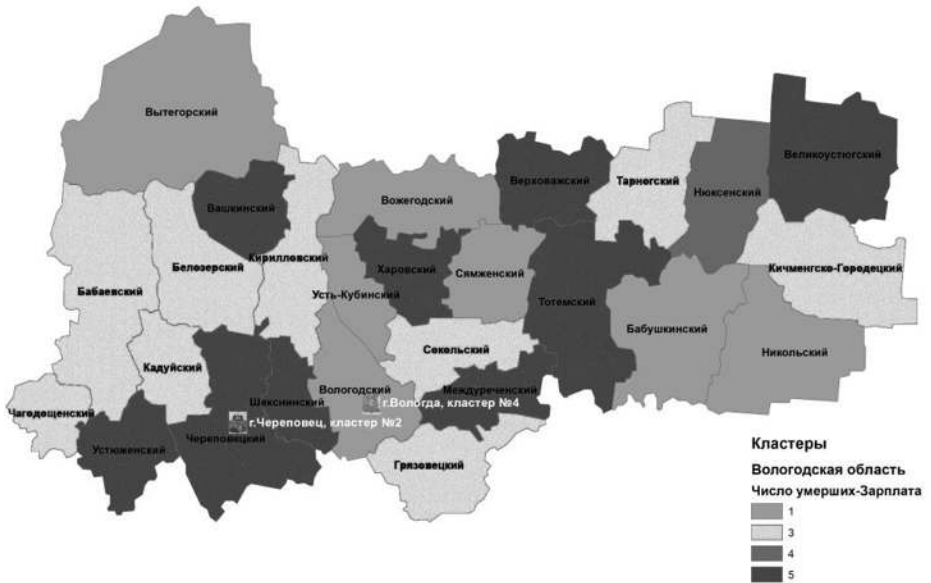


Рис. 1. Тематическая карта результатов кластеризации

- тем не менее, смертность для кластера № 4 в 1,2 раза выше, чем для кластера № 2, а зарплата составляет 78 %;
- число умерших в кластерах № 1 и 5 в 1,5 и в 1,48 раза выше, чем в кластере № 2, а зарплата составляет 53 и 69 % соответственно;
- региональный диспаритет для двух территорий, объединенных процедурой в кластер № 4, выражен меньше, чем в кластере № 2. Отрыв по смертности от кластеров № 1, 3, 5 составляет 1,25; 1,27 и 1,23 раза. Зарплата составляет 68, 81 и 89 % соответственно;
- выраженные территориальные диспропорции для кластеров № 1, 3, 5, объединяющих в себе 25 территорий, отсутствуют. Смертность для них практически не превышает среднюю по области: разница между значениями составляет 1,03; 1,04 и 1,01 раза соответственно. Самые плохие показатели зарплаты в кластерах № 1 и 3 (81 и 97 % от среднего значения соответственно). Для кластера № 5 оплата труда превышает среднее значение по области в 1,07 раза;
- различие показателей смертности в парах кластеров № 1 и 3, №1 и 5, №3 и 5 незначительно (99 %, 1,02 и 1,03 раза соответственно). Однако присутствует дифференциация по оплате труда (84, 63 и 91 % соответственно). Именно она обусловила разбиение территорий на кластеры № 1, 3 и 5.

Эффективный способ исследования данных – их визуализация, которая позволяет выявлять значимые сведения, особенно для сложных многомерных и нелинейных связей между переменными. Для качественного исследования зависимости медико-демографического показателя смертности от показателя мощности работы медицинских учреждений (обеспеченности врачами на 10 тыс. человек населения) и финансового показателя (среднемесячной зарплаты в отрасли) была выполнена трехмерная

визуализация. Использовались карта линий уровня с подгонкой методом Bicubic Spline Smoothing (рис. 2) и контурный график с подгонкой на основе метода Wafer Smoothing (рис. 3).

Визуализация способствует поиску рационального решения, которое максимизирует ожидаемую выгоду – требуемое снижение смертности при минимальных финансовых затратах. Карты линий уровня позволяют исследовать характер поверхности Z . На основе показателей графика можно сделать вывод о неоднородности исходных данных. Результаты сглаживания свидетельствуют о наличии области повышенной смертности в четвертом квадранте (см. рис. 2), что согласуется с результатами исследований [13] и является следствием недофинансирования регионального здравоохранения.

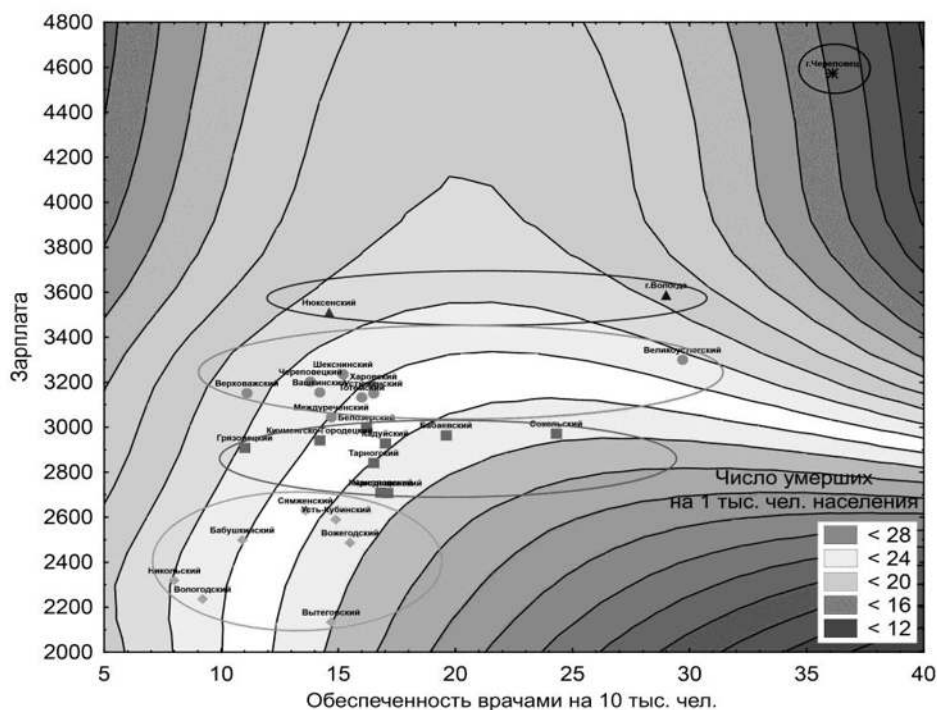


Рис. 2. Карта линий уровня с подгонкой методом Bicubic Spline Smoothing

Для обеспечения эффективности принимаемых управленческих решений по снижению смертности в качестве одного из сценариев предлагается отказ от простого наращивания мощности медицинских учреждений за счет увеличения количества врачей при сохранении текущей финансовой политики по отношению к зарплате. В силу существенной нелинейности, присущей объекту регулирования Z , это решение может оказаться контринтуитивным в максимальной степени для кластера № 1, в меньшей степени – для кластеров № 3 и 5. Для снижения смертности, при прочих равных условиях, необходимо сначала увеличить зарплату в кластерах № 1 и 3 как минимум до уровня показателей кластера № 5 (в 1,31 и 1,1 раза соответственно). Более предпочтительным является увеличение средней зарплаты в кластерах № 1, 3 и 5 до уровня кластера № 4 (в 1,47; 1,23 и 1,12 раза соответственно). После этого можно

решать кадровый вопрос с перспективой выхода на изолинию г. Череповца при большей обеспеченности врачами (~40 против 36,1 на 10 тыс. человек населения), но при меньшем уровне зарплат (в 1,29 раза).

Для наглядного и точного представления пространственной структуры данных был использован контурный график с подгонкой методом Wafer Smoothing (рис. 3).

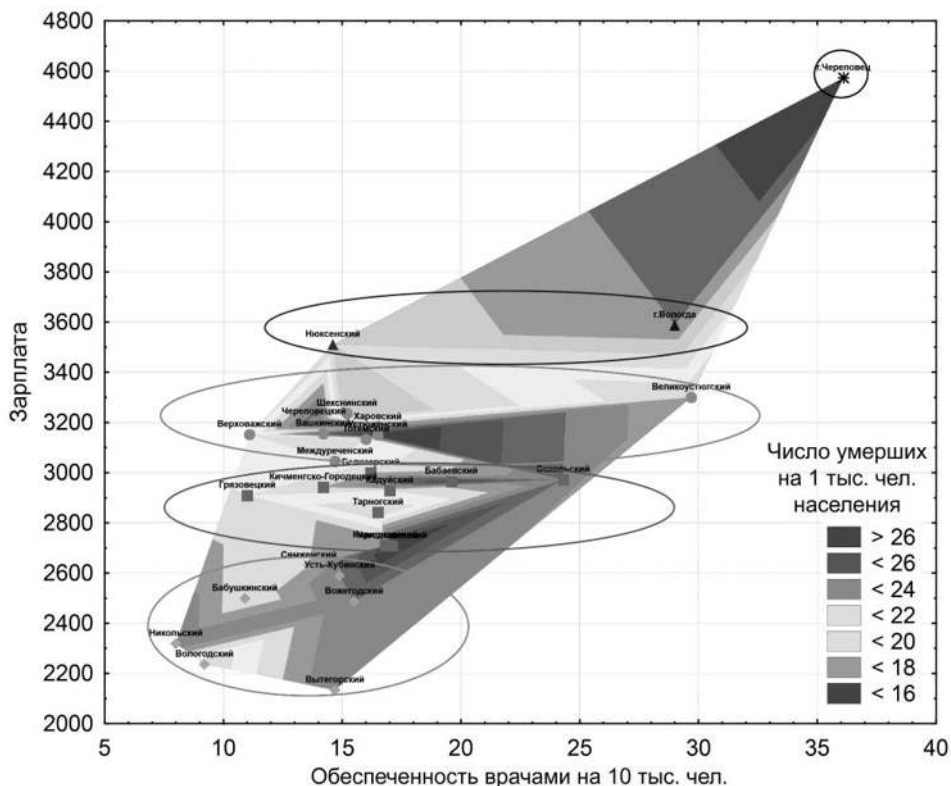


Рис. 3. Контурный график с подгонкой на основе метода Wafer Smoothing

Особенностью этого метода является организация первичных данных в триады (x, y, z) . Пара (x, y) определяет ориентацию координатной сетки, а z – значение для данной координаты. Если в результате преобразования одной паре (x, y) соответствует несколько значений z , то выполняется усреднение. Поверхность создается путем формирования триад между точками координатной сетки. Поскольку метод использует необработанные данные, форма трехмерной поверхности может быть исследована без влияния искажений, неизбежных в ходе аппроксимации.

Анализ графика, представленного на рис. 3, позволил эффективно выявить аномалии в сложноорганизованных данных. Обратим внимание на то, что в Вологодском районе показатель смертности в 1,19 раза меньше среднего значения аналогичного показателя в кластере № 1 и близок к показателю сравнительно благополучного по обеспеченности врачами и к зарплате кластера № 4. Также привлекает внимание низкая смертность в Шекснинском (кластер № 5) и Тарногском (кластер № 3) районах, демонстрирующих отрыв по числу умерших в 1,23 и 1,2 раза от средних значений соответственно. Превышение показателя смертности зафиксировано в Чагод-

шенском и Устюженском районах в 1,19 и 1,22 раза от средних значений соответственно. По показателям обеспеченности врачами и уровня зарплат Чагодощенский район близок к Кирилловскому, но по смертности уступает ему в 1,17 раза. Одной из возможных причин диспропорции является разница в расходах на здравоохранение в 1,22 раза (на 1 человека) из всех источников финансирования. Аналогичным образом Устюженский район по показателям обеспеченности врачами и уровня реального содержания заработной платы близок к Тотемскому и Харовскому районам и уступает им по показателям смертности в 1,35 и 1,17 раза. При этом разница в расходах на здравоохранение из всех источников финансирования составляет 1,18 и 1,24 раза соответственно. Для выявления причин аномалий необходимо расширять мерность пространства данных информационной базы исследования.

Выводы

Изучены возможности применения статистических алгоритмов в социально-экономических исследованиях здравоохранения области. Представлены результаты корреляционного анализа и непараметрических мер взаимозависимости показателей эффективности регионального здравоохранения. Полученные оценки параметров регрессионной модели позволили выявить зависимость смертности от размера среднемесячной заработной платы в отрасли по районам. При помощи кластерного анализа и трехмерной визуализации данных обнаружены аномалии и выявлен диспаритет для территорий региона по ряду медико-демографических показателей мощности медицинских учреждений и финансовых показателей деятельности муниципального здравоохранения. Дальнейшие перспективы связаны с уточнением модели сценарного прогнозирования для достижения уровня смертности, не превышающего заданный, при минимальных затратах на основе методов оптимизации.

Литература

1. Воронов А. М., Пахолков И. В., Ершов Е. В. Использование Data Mining для прогнозирования и оптимизации работы системы управления производства // Оптико-электронные приборы и устройства в системах распознавания образов, обработки изображений и символьной информации. Распознавание – 2017: сборник материалов XIII Международной научно-технической конференции. Курск: Юго-Западный государственный университет, 2017. С. 112–114.
2. Вуколов Э. А. Основы статистического анализа: практикум по статистическим методам и исследованию операций с использованием пакетов Statistica и Excel. М.: ФОРУМ, 2008. 464 с.
3. Вучков И., Бояджиева Л., Солаков Е. Прикладной линейный регрессионный анализ. М.: Финансы и статистика, 1987. 238 с.
4. Гланц С. Медико-биологическая статистика. М.: Практика, 1999. 459 с.
5. Ильин В. А. Эффективность здравоохранения региона. Вологда: Вологодский научно-координационный центр Центрального экономико-математического института Российской академии наук, 2006. 189 с.
6. Крутько В. Н., Смирнова Т. М. Анализ динамики и причин смертности россиян в конце XX века // Общественное здоровье и профилактика заболеваний. 2005. №3. С. 24–28.
7. Крутько В. Н., Смирнова Т. М. Анализ тенденций смертности и продолжительности жизни населения России в конце XX века. М.: Едиториал УРСС, 2002. 48 с.
8. Крутько В. Н., Смирнова Т. М. Человеческий капитал: проблема и ресурс инновационного развития России. М.: Цифровичок, 2012. 227 с.
9. Развитие здравоохранения Вологодской области на 2014–2020 годы: Государственная программа: постановление Правительства области от 28.10.2013 № 1112 // КонсультантПлюс: справочно-правовая система. URL: <https://www.consultant.ru>. (дата обращения: 01.02.2019).

10. Рапаков Г. Г., Банщиков Г. Т. Организация системы раннего выявления больных артериальной гипертензией и доступность антигипертензивных средств в Вологодской области: опыт использования кластерного анализа // Архив внутренней медицины. 2013. №4. С. 16–23.
11. Рапаков Г. Г. Статистические методы в социально-экономических исследованиях здравоохранения Вологодской области. Вологда: Вологодский государственный университет, 2015. 125 с.
12. Симчера В. М. Методы многомерного анализа статистических данных. М.: Финансы и статистика, 2008. 400 с.
13. Улумбекова Г. Э. Обоснование уровня государственного финансирования здравоохранения для улучшения здоровья населения Российской Федерации // Проблемы социальной гигиены, здравоохранения и истории медицины. 2013. № 4. С. 32–34.
14. Ферстер Э., Ренц Б. Методы корреляционного и регрессионного анализа. М.: Финансы и статистика, 1983. 302 с.
15. Флетчер Р., Флетчер С., Вагнер Э. Клиническая эпидемиология. Основы доказательной медицины. М.: Медиа Сфера, 1998. 352 с.
16. Allison Paul D. Multiple regression: a primer. London: Sage, 1999. 220 p.
17. Paulson D. Handbook of Regression and Modeling: Applications for the Clinical and Pharmaceutical Industries. Boca Raton: Chapman & Hall, 2007. 503 p.

References

1. Voronov A. M., Pakholkov I. V., Ershov E. V. Ispol'zovanie Data Mining dlia prognozirovaniia i optimizatsii raboty sistemy upravleniia proizvodstva [Using Data Mining to predict and optimize the operation of the production management system]. *Optiko-elektronnye pribory i ustroistva v sistemakh raspoznavaniia obrazov, obrabotki izobrazhenii i simvol'noi informatsii. Raspoznavanie – 2017: sbornik materialov XIII Mezhdunarodnoi nauchno-tekhnicheckoi konferentsii* [Optoelectronic devices and devices in image recognition, image processing and character information systems. Recognition – 2017: Proceedings of XIII International scientific and technical conference]. Kursk: Southwestern State University, 2017, pp. 112–114.
2. Vukolov E. A. *Osnovy statisticheskogo analiza: Praktikum po statisticheskim metodam i issledovaniiu operatsii s ispol'zovaniem paketov Statistica i Excel* [Fundamentals of statistical analysis. Workshop on statistical methods and investigation of operations using the Statistics package and Excel]. Moscow: FORUM, 2008. 464 p.
3. Vuchkov I., Boiadzhieva L., Solakov E. *Prikladnoi lineinyi regressiionnyi analiz* [Applied linear regression analysis]. Moscow: Finansy i statistika, 1987. 238 p.
4. Glants S. *Mediko-biologicheskaiia statistika* [Medico-biological statistics]. Moscow: Praktika, 1999. 459 p.
5. Il'in V. A. *Effektivnost' zdравookhraneniia regiona* [Efficiency of region health]. Vologda: Vologodskii nauchno-koordinatsionnyi tsentr Tsentral'nogo ekonomiko-matematicheskogo instituta Rossiiskoi akademii nauk, 2006. 189 p.
6. Kru't'ko V. N., Smirnova T. M. *Analiz dinamiki i prichin smertnosti rossiian v kontse 20-go veka* [Analysis of dynamics and causes of deaths of the Russians in the late 20th century]. *Obshchestvennoe zdorov'e i profilaktika zabolevanii* [Public health and disease prevention], 2005, no. 3, pp. 24–28.
7. Kru't'ko V. N., Smirnova T. M. *Analiz tendentsii smertnosti i prodolzhitel'nosti zhizni nasele-niia Rossii v kontse XX veka* [Analysis of mortality trends and life expectancy of the population of Russia at the end of the 20th century]. Moscow: Editorial URSS, 2002. 48 p.
8. Kru't'ko V. N., Smirnova T. M. *Chelovecheskii kapital: problema i resurs innovatsionnogo razvitiia Rossii* [Human capital: the problem and resource of Russia's innovative development]. Moscow: Tsifrovichok, 2012. 227 p.
9. *Razvitie zdравookhraneniia Vologodskoi oblasti na 2014 – 2020 gody: Gosudarstvennaia programma: postanovlenie Pravitel'stva oblasti ot 28.10.2013 № 1112* [Enhancement of health care system in the Vologda region in the period of 2014-2020: state-run programme: the Decree of Regional Government from 28.10.2013 № 1112]. Konsul'tantPlus: sprav.-pravovaia sistema; Kompani-

ia “KonsultantPlus” [ConsultantPlus: the system of legal reference information]. Available at: <https://www.consultant.ru> (accessed: 01.02.2019).

10. Rapakov G. G., Banshchikov G. T. Organizatsiia sistemy rannego vyivleniia bol'nykh arteriial'noi gipertenziei i dostupnost' antigipertenzivnykh sredstv v Vologodskoi oblasti: opyt ispol'zovaniia klasternogo analiza [Organization of early detection system for patients with arterial hypertension and availability of antihypertensive drugs in the Vologda region: experience of cluster analysis]. *Arkhiv" vnutrennei meditsiny* [The Russian Archives of Internal Medicine], 2013, no. 4, pp. 16 – 23.

11. Rapakov G. G. *Statisticheskie metody v sotsial'no-ekonomicheskikh issledovaniiax zdra-vookhraneniia Vologodskoi oblasti* [Statistical methods in socio-economic studies of the health care in the Vologda region]. Vologda: Vologda State University, 2015. 125 p.

12. Simchera V. M. *Metody mnogomernogo analiza statisticheskikh dannykh* [Methods of multi-dimensional analysis for statistical data]. Moscow: Finansy i statistika, 2008. 400 p.

13. Ulumbekova G. E. Obosnovanie urovnia gosudarstvennogo finansirovaniia zdavookhraneniia dlia uluchsheniia zdorov'ia naseleniia Rossiiskoi Federatsii [Justification of the level of public health care financing to enhance health of population in the Russian Federation]. *Problemy sotsial'noi gigieny, zdavookhraneniia i istorii meditsiny* [Problems of Social Hygiene, Health and History of Medicine], 2013, no. 4, pp. 32–34.

14. Ferster E., Rents B. *Metody korreliatsionnogo i regressionnogo analiza* [Methods of correlation and regression analysis]. Moscow: Finansy i statistika, 1983. 302 p.

15. Fletcher R., Fletcher S., Vagner E. *Klinicheskaia epidemiologiia. Osnovy dokazatel'noi med-itsiny* [Epidemiology. Fundamentals of Evidence-Based Medicine]. Moscow: Media Sfera, 1998. 352 p.

16. Allison Paul D. *Multiple regression: a primer*. London: Sage, 1999. 202 p.

17. Paulson D. *Handbook of Regression and Modeling: Applications for the Clinical and Pharmaceutical Industries*. BocaRaton: Chapman&Hall, 2007. 503 p.

Для цитирования: Рапакoв Г. Г., Банщикoв Г. Т., Горбунов В. А., Харахнин К. А., Ревелев И. М. Применение компьютерного моделирования при оценке показателей эффективности регионального здравоохранения на основе методов статистического анализа данных // Вестник Череповецкого государственного университета. 2019. № 2 (89). С. 56–68. DOI: 10.23859/1994-0637-2019-2-89-5

For citation: Rapakov G. G., Banshchikov G. T., Gorbunov V. A., Harahnin K. A., Rev-elev I. M. Computer simulation of regional public health efficiency using statistical analysis methods. *Bulletin of the Cherepovets State University*, 2019, no. 2 (89), pp. 56–68. DOI: 10.23859/1994-0637-2019-2-89-5