

УДК [616-006.04:519.237.5](613.32:614.76)

DOI: [10.26102/2310-6018/2023.40.1.022](https://doi.org/10.26102/2310-6018/2023.40.1.022)

Прогноз общей онкологической заболеваемости в регионах и муниципалитетах России на основе многофакторной модели

В.С. Степанов 

*Центральный экономико-математический институт РАН, Москва,
Российская Федерация
_stepanov@cemi.rssi.ru *

Резюме. В работе исследуется вопрос о связи общей заболеваемости онкологическими патологиями на территории региона с возрастным составом населения и набором эколого-гигиенических переменных. Показано наличие зависимости между распространенностью этой патологии на территории и лаговыми значениями переменных. Комплекс факторных объясняющих переменных содержит показатель загрязненности атмосферного воздуха в поселениях территории, качественную переменную по степени ее техногенного загрязнения, удельный объем сброса загрязненных сточных вод, уровень концентраций бенз(а)пирена и долю пожилого населения. На этой основе построена эконометрическая модель, имеющая высокое статистическое качество: коэффициент детерминации выше 95 %, ошибка аппроксимации ниже 3 %. Один из факторов был ординального типа, поэтому получилась модель регрессии с переменной структурой, содержащая фиктивные переменные. Параметры модели оценивались методом наименьших квадратов по панельным данным за 2017-18 гг. Эти данные включали наблюдения с контингентами больных раком в каждом регионе и факторные переменные. Наличие лагов у последних позволяет прогнозировать численность контингента лиц, страдающих опухолями любой локализации, с горизонтом планирования в 1 год. Такие прогнозы выполнялись как на уровне регионов, так и в разрезе муниципальных образований. На основе полученного уравнения можно принимать научно-обоснованные управленческие решения, направленные на профилактику онкологической заболеваемости населения на конкретной территории, а также строить имитационные модели агентного типа.

Ключевые слова: модель регрессии, загрязнение атмосферного воздуха, загрязнение сточными водами, бенз(а)пирен, техногенное загрязнение, курение, злокачественные новообразования, регион, муниципальное образование.

Для цитирования: Степанов В.С. Прогноз общей онкологической заболеваемости в регионах и муниципалитетах России на основе многофакторной модели. *Моделирование, оптимизация и информационные технологии*. 2023;11(1). URL: <https://moitvvt.ru/ru/journal/pdf?id=1286> DOI: 10.26102/2310-6018/2023.40.1.022

The forecast of cancer prevalence in the regions and municipalities of Russia based on a multivariate model

V.S. Stepanov 

*Central Economics and Mathematics Institute of the RAS, Moscow, Russian Federation
_stepanov@cemi.rssi.ru *

Abstract. The article considers the issue of the relationship between the cancer prevalence in some Russian regions and age composition along with a set of environmental and hygienic variables. The presence of the relationship between the prevalence of this pathology in the territory and the lag values of the variables are shown. The complex of factor explanatory variables contains an indicator of atmospheric air pollution in the settlements of the territory, a qualitative variable according to the degree of its technogenic pollution, the specific volume of polluted wastewater discharge, the concentration of

benzo(a)pyrene, and the age composition. On this basis, an econometric model has been built that has a high statistical quality: the coefficient of determination is above 95 %, the approximation error is less than 3%. One of the factors was of the ordinal type; therefore, the result was a linear regression model with a variable structure containing two dummy variables. The model parameters were estimated by the least squares method using panel data with Russian regions for 2017-18. These data included observations on the populations of people with oncologic pathologies in the regions and factor variables. The presence of lags in the latter makes it possible to predict the number of people suffering from tumors of any localization with a planning horizon of 1 year. Such forecasts were carried out both at the level of regions and in the context of municipalities. Based on the resulting equation, one can make evidence-based managerial decisions aimed at cancer prevention in specific regions of Russia as well as to build agent-type simulation models.

Keywords: regression model, atmospheric air pollution, sewage pollution, benz(a)pyrene, technogenic pollution of the territory, smoking, malignant neoplasm, region, municipality.

For citation: Stepanov V.S. The forecast of cancer prevalence in the regions and municipalities of Russia based on a multivariate model. *Modeling, Optimization and Information Technology*. 2023;11(1). URL: <https://moitvvt.ru/ru/journal/pdf?id=1286> DOI: 10.26102/2310-6018/2023.40.1.022 (In Russ.).

Введение

Развитие онкологической патологии у граждан является важной медико-социальной проблемой, так как в 2020 году новообразованиями в РФ было обусловлено около 13,8 % случаев смерти, вслед за болезнями системы кровообращения: ≈ 44 % [1]. Кроме прямых потерь от смертности, большую роль здесь играют высокие затраты, связанные с лечением патологии и ростом инвалидности, как и потери в экономическом развитии российских регионов. Из этого следует, что проблема борьбы со злокачественными новообразованиями (ЗНО) важна для граждан и властей России, как и многих других стран. Основной целью проекта «Борьба с онкологическими заболеваниями» в РФ является снижение смертности из-за новообразований, включая ЗНО; среди дополнительных показателей этого проекта есть «распространенность новообразований», заданная как 3108,5 случаев на 100 тыс. населения в 2024 г. В целях реализации проекта во всех субъектах РФ разработаны региональные программы.

В работе [2] была представлена стратегия борьбы против ЗНО. Ее первым звеном является первичная профилактика, нацеленная на снижение риска развития опухолей. Это звено включает: изучение вредоносных факторов окружающей среды и борьбу с ними, раннее выявление заболеваний, предшествующих канцерогенезу, их лечение, диспансеризацию лиц, попавших в группу риска [3]. Здесь отмечалось, что меры первичной профилактики развития ЗНО крайне редко обсуждаются в отечественной литературе. Следующим звеном является скрининг, а также создание раковых регистров и их анализ математическими методами посредством алгоритмов и программ [4]. Результатами этих действий являются улучшение эффективности диспансеризации и более высокая точность при диагностике. При создании алгоритмов принятия решений, лежащих в основе компьютерных программ, в подобных задачах часто используются [5]: логистическая и множественная модели регрессии [6-8], дискриминантный анализ [4], нейронные сети [9], деревья решений [7] и т. д.

Если говорить о причинах возникновения рака, то считается, что развитие ЗНО в организме обусловлено многими онкогенными факторами среды обитания: химическое загрязнение окружающей или производственной среды, загрязнение воздуха вредными веществами, загрязнение поверхностных вод и почвы, воздействие радиации и прочее. Важную роль играют и другие причины: возраст, качество питьевой воды, курение, характер питания, генетические факторы и т. д. [4, 6-8, 10, 11]. Комплексное воздействие разных факторов анализируется сравнительно редко; это сделано в [12, 13].

Распространенность ЗНО часто выше вблизи промышленных центров и на территориях, где проживает значительная доля лиц старшего возраста [8]. Похожая ситуация наблюдается в других странах; например, жители крупных городов немецкой Баварии чаще страдают от развития ЗНО, чем ее сельское население [14]. Качество атмосферного воздуха в достаточно крупных городах РФ нередко бывает хуже установленных нормативов; например, по опросам Росстата [15] около 37 % жителей считали важной проблемой «загрязненность окружающей среды», причем 48 % из недовольных проживает в мегаполисах, 43 % – в городах 250-500 тыс. населения. [16].

Значимость учета ряда факторов развития ЗНО: «возраст», «курение», «проживание или работа при воздействии химических канцерогенов», часто подтверждалась в исследованиях «случай-контроль» [11, 17, 18]. Например, в [18] рассмотрен контингент жителей Белгородской области и выделены значимые факторы риска. Для этого наблюдения над $N_1=320$ лицами основной группы, страдающими от ЗНО, сравнивались с контрольной группой ($N_2=336$ жителей без ЗНО). У каждого участника исследования фиксировались десятки показателей, также содержащих и факторы риска развития ЗНО. Анализ значимости каждого показателя выполнялся через сравнение его средних значений в этих группах на основе t -критерия Стьюдента. В результате высокие значения тестовой t -статистики оказались для таких факторов: а) «возраст пациента» (ему далее соответствует X_1); б) «интенсивность курения сигарет» (ИКС) (входит в X_4); в) «загазованность на рабочем месте» (отдаленно походит на наш фактор X_2); г) «токсико-химические вещества; там же» (в нашем случае это связано с X_3 , X_2 , X_4) [18]. Аналогично, в [11] было взято две группы трудоспособных горожан, с чуть большими объемами N_1 , N_2 , и для них сравнивались частоты факторов риска через применение точного критерия Фишера. Оказалось, что наличие работы или проживания «при воздействии химических канцерогенов» и «курение человека» являются факторами риска развития ЗНО, с уровнями значимости $p < 0,045$ [11].

По определению, контингент пациентов включает совокупность лиц, страдающих ЗНО с любой локализацией, зарегистрированных в одной из лечебно-профилактических организаций Минздрава РФ в регионе в году t , как и в предыдущие годы [1]. Этот контингент в расчете на 100 тыс. населения используется в качестве зависимой переменной Y в модели регрессии, предлагаемой ниже. Исходя из описанного выше и следуя [8], на роль переменных X_1, \dots, X_3 («факторов риска»), которые имеют значимые корреляционные связи с величиной Y на территории (регион, муниципальное образование), ниже выбраны несколько ее эколого-гигиенических характеристик, а также возрастной состав жителей. Дополнительно посредством X_4 учитывается постоянное воздействие на них бенз(а)пирена (БП): как из окружающего воздуха, так и из сигаретного дыма; эти факторы – сильные канцерогены [10, 19, 20]. Также переменные X_1, \dots, X_4 оказались информативными в модели канцерогенеза, построенной по 33-м поселениям Приморья в работе [21], где анализировались панельные данные с показателями по ЗНО в сопоставлении со средовыми факторами.

Целью настоящего исследования является построение многофакторной модели регрессии (с оценкой ее по точности), которая связывает общую заболеваемость различными формами рака на территории в году t с комплексом объясняющих переменных, взятых с лагами. Применение модели проиллюстрировано на муниципалитетах из ПФО, СФО, ДФО с Y за 2016-20 гг., как и на регионах Сибири.

Иначе говоря, в работе строится модель регрессии с переменной структурой, связывающая распространенность контингента лиц с этой патологией Y с долей жителей старшего возраста на территории (мужчин с 60 и женщин с 55 лет), а также – с комплексом факторов окружающей среды. Один из факторов имел качественную природу, поэтому уравнение модели содержит константу, зависящую от значений

фиктивных (dummy) переменных. На основе этого уравнения можно оценить, насколько можно снизить Y на территории, если уменьшить концентрации вредных веществ в воздухе, улучшить очистку сточных вод и / или снизить объемы их сброса, ослабить ИКС среди взрослого населения.

Материалы и методы

Объектами исследования были: а) данные Росстата, собранные на территориях (регионы или муниципалитеты) за 2014-20 гг.; б) данные Росгидромета по среднегодовым концентрациям БП в воздухе городских округов (г.о.); в) данные по объемам сброса загрязненных сточных вод (Росводресурсы); г) информация о техногенных загрязнениях (содержательные знания, карты загрязнений регионов). Обработка данных и статистический анализ выполнялись в MS-Excel. В таблицу с материалом обучения, на основе которой оценивались параметры модели, переменная Y входила за два года: 2017 и 2018 [1].

Во введении был обоснован выбор объясняющих факторных переменных. Их ежегодные значения легкодоступны: для X_1 , X_2 они опубликованы на <https://rosstat.gov.ru> в нескольких статистических таблицах, а переменные X_3 , X_4 очень легко вычисляются. В уравнение модели включались следующие факторы:

- X_1 – доля жителей старших возрастов или *Age* (Таблица 2), в % [22, 23];
- X_2 – характеристика состояния атмосферного воздуха в населенных пунктах – доля исследованных проб воздуха, где наблюдалось превышение ПДК вредных веществ (Q/ПДК, или *Air* в таблице), в %, [24];
- $X_3 = (V_{св}/S)^{0.5}$ – квадратный корень из удельного сброса загрязненных сточных вод в поверхностные водные объекты [25], где $V_{св}$ – объем сброса стоков в году t в регионе, [млн. м³], а S – площадь территории региона в [тыс. км²], взятая за вычетом площади, занимаемой его лесными землями, (имя *Water*) [22]. При этом в числителе в X_3 берется взвешенная сумма объемов воды, сброшенной без очистки, вместе со сбросами ненормативно очищенной воды (объем последних делился на число 4,5);
- X_4 – переменная с воздействием на среднего жителя бенз(а)пирена: как при вдыхании им атмосферного воздуха (с усреднением по основным населенным пунктам данной территории и за все месяцы года), так и из-за сигарет (имя *Benzap*);
- X_5 – качественная переменная с градациями: “А”, “В”, “С”, характеризующими уровень техногенного загрязнения территории (радиационного и др. [26]); здесь “А” – ситуация на территории сравнительно благополучна, “С” – наиболее острая и “В” – уровень загрязнения на территории – средний между уровнями “А” и “С”.

Бенз(а)пирен (БП) имеет формулу $C_{20}H_{12}$; он образуется при сжигании угля, нефти и ее продуктов, другой органики (пожары и пр.). На урбанизированных территориях он сильнее воздействует через воздух, почвенную пыль и пищевые цепи [27]. Это особенно важно там, где объекты теплоэнергетики работают на угле и / или там, где много печного отопления. В результате воздушной миграции выбросов в форме мелкодисперсного аэрозоля БП может возникнуть в значительной концентрации в 5-10 км от источника загрязнения. Затем он постепенно накапливается в организме и позднее может индуцировать в нем ряд форм ЗНО [19]; в условиях техногенных выбросов в виде газопылевой смеси основной его мишенью являются легкие [27]. Среднегодовая ПДК_{с.г.}, после ее усреднения за месяцы года, задана по БП, как 1 нг/ м³.

При расчете фактора X_4 с БП нами использовались данные за 2018 г. по объемам ежесуточного потребления сигарет в регионе, с градациями: «выкуривают до 5 шт./день», «тоже, до 10 шт.», «тоже, до 1 пачки» и «более одной». Они опубликованы для курящих ежедневно лиц возраста 15+ в [15]. Затем найдены их средние значения: 2½,

5, 10, 15 [шт./день] и, через распределение из [15], – количество, что в среднем ежедневно выкуривается всеми такими лицами. Затем оно соотносилось с общей численностью жителей этого возраста в 2018 г. и находилась интенсивность курения сигарет (ИКС): число сигарет Sig , что приходится на каждого ([22]). Например, ИКС среднестатистического жителя Подмосковья с возрастом 15+ получилась $Sig \approx 1\frac{1}{2}$ шт. Данных для других отсчетов по годам t не было и поэтому принято, что $Sig = const$. С учетом содержания БП в сигарете мы полагаем, что:

$$X_4 = Ln [(Q_{ср.г.} + 0.03 \cdot Sig) / ПДК_{с.г.}], \quad (1)$$

где $Q_{ср.г.}$ – среднегодовая концентрация БП в атмосферном воздухе территории всех основных населенных пунктов региона (или данного муниципалитета), установленная на постах наблюдения Росгидромета [28] (с усреднением по всем постам и месяцам года t). То есть фактор X_4 оценивает в логарифмической шкале воздействие БП на органы дыхания жителя: из воздуха и курения сигарет (в доле от $ПДК_{с.г.}$).

Информация о техногенном и, в частности, радиационном загрязнении территорий извлекалась в виде знаний, а также из карт данной местности. Например, после изучения карты загрязнений [29], категориям фактора X_5 в Пронском районе, а также в Михайловском (без г. Михайловск), был приписан код “С”; Рязскому (с г. Рязск) и Сасовскому районам (с г. Сасово) – код “В” (Таблица 1). Для Рязани также выбиралось “С”. С учетом [30], для гг. Бийск, Барнаул приписывалась “С”, а для г. Новоалтайск, где уровень загрязнений ^{137}Cs был ниже, выбирался код “В”.

Предлагаемая далее модель строилась в варианте “объединенная регрессия” по панельным данным, которые ниже названы “обучающими”. В состав этих данных вошли не все регионы России. Сначала был исключен ряд из них; регионы с бежевым цветом на Рисунке 1, входили в таблицу обучения лишь за 1 год, а с зеленым цветом – за оба года.

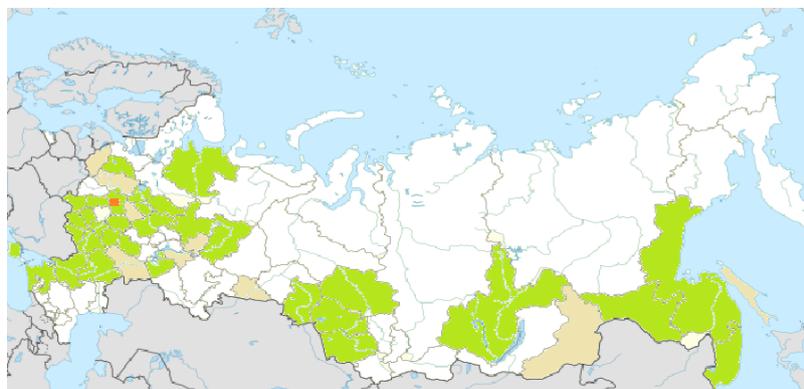


Рисунок 1 – Регионы материала обучения на карте РФ
Figure 1 –Regions of Russia that formed the learning data

Почти во всех исключенных регионах, помеченных белым цветом, общие коэффициенты смертности были относительно малы: ниже 26-го места в 2018 г. [22]; еще в таких регионах часто проживает высокая доля молодежи. Также оказалось, что около $\frac{3}{4}$ регионов из таблицы обучения содержат большую долю этноса “русские”. В таблицу обучения вошло $n=73$ строки с наблюдениями над регионами; около половины с Y за 2017 г., а прочие с Y за 2018 г. Регионы были взяты из ЦФО, реже – СЗФО, ПФО, ЮФО и редко – в УФО, СФО, ДФО. На следующем этапе под названием “экзамен”, оцененная при обучении модель применяется к новым данным.

Основными методами были корреляционный и регрессионный анализы, методы теории проверки гипотез. Качество модели регрессии оценивалось по критерию Фишера,

скорректированному коэффициенту детерминации R^2 (доля дисперсии Y , которая объясняется моделью) и ошибке аппроксимации. Дополнительно использовались тесты через критерий Стьюдента для проверки гипотез H_0 : “конкретный β – коэффициент в модели равен нулю” и тест через DW -статистику Дарбина и Уотсона для выявления отсутствия автокорреляции остатков [31]. Для корректного использования тестов применялись критерии согласия распределения остатков регрессии с нормальным законом Гаусса. Условие гомоскедастичности проверялось посредством анализа диаграмм рассеяния остатков.

Результаты

Ниже предлагается линейная модель регрессии:

$$Y = \beta_0 + \beta_1 \cdot X_1 + \dots + \beta_4 \cdot X_4 + \beta_5 \cdot X_5 + \varepsilon, \quad (2)$$

где Y – контингент лиц с ЗНО в году t в расчете на 100 тыс. жителей (или $Y^{(t)}$), а X_1, \dots, X_5 – вышеописанные факторы, наблюдаемые на территории в некоторые годы, что предшествуют году t ; β_0, β_1, \dots – неизвестные числовые параметры, оцениваемые по региональным наблюдениям, причем константа β_0 отражает влияние на Y прочих, не учтенных в (2) переменных. В выражении (2) предполагается, что качественная переменная X_5 в том или ином виде ранее была “оцифрована”, т. е. представлена числами. Величина ε определяется как случайная со средним ноль и дисперсией σ^2 , которая нам неизвестна. Переменные Y, X_1, \dots, X_4 имеют сверху индекс, привязанный к году t . Например, для Y он записывается в виде $Y^{(t)}$, но иногда, для простоты, он ниже не показывается. Оценки b_0, b_1, \dots для β -параметров, а также s – для параметра σ , находятся по методу наименьших квадратов (МНК) по таблице, содержащей панельные данные в среде *MS-Excel* [31]. На этапе “обучения” эти данные используются при построении модели, а на “экзамене” модель будет тестироваться на новых данных.

Вместо трех градаций фактора X_5 , входящего в (2), вводились фиктивные переменные d_1 и d_2 ; в результате чего его категория “А” кодируется через $d_1=d_2=0$, т. е. через (0, 0); аналогично, категория “В” кодируется через (1, 0), “С” – (0, 1) (Таблица 1).

Таблица 1 – Ненулевые значения фиктивных переменных для ряда областей и краев
Table 1 – Dummy variables for the regions and territories of the Russian Federation

Область, край	d	d	Область	d	d	Область	d	d	Область	d	d
	1	2		1	2		1	2		1	2
Амурская	1	0	Ивановская	1	0	Орловская	0	1	Рязанская	1	0
Алтайский кр.	0	1	Калужская	0	1	Новгородская	1	0	Тверская	0	1
Архангельская	1	0	Курганская	1	0	Новосибирская	0	1	Томская	1	0
Брянская	0	1	Курская	0	1	Нижегородская	0	1	Ярославская	1	0
Воронежская	1	0	Липецкая	1	0	Пензенская	1	0			
Забайкальский	1	0	Омская	0	1	Псковская	1	0			

Исходя из выше описанного, после подстановки в (2) оценок неизвестных нам β_j -коэффициентов и введения dummy-переменных вместо категорий X_5 , получим:

$$Y^{(t)} = b_0 + b_1 \cdot X_1^{(t-1)} + b_2 \cdot X_2^{(t-2)} + b_3 \cdot X_3^{(t-3)} + b_4 \cdot X_4^{(t-1)} + b_5 \cdot d_1 + b_6 \cdot d_2, \quad (3)$$

где $Y^{(t)}$ (или Y) – контингент больных со ЗНО к концу года t на территории на 10^5 жителей, для которой также наблюдается набор значений X_1, \dots, X_4 , взятых с лагами по t .

В Таблице 2 приводятся значения оценок b_0, \dots, b_6 для β -коэффициентов, которые находятся по методу наименьших квадратов (МНК), а также – значения всех t -статистик Стьюдента, используемых при проверке гипотез $H_0: \beta_j = 0$ о значимости каждого фактора (при $j = 1, \dots, 6$) и достигнутые уровни значимости (p -значение). Гипотеза о нормальности распределения остатков регрессии не отвергалась (см. ниже).

Таблица 2 – Результаты оценивания $\beta_0, \beta_1, \dots, \beta_6$, значения лагов и статистик

Table 2 – Estimation results of $\beta_0, \beta_1, \dots, \beta_6$, lag values and statistics

№	Факторная объясняющая переменная	Её имя	Лag T	Оценка (МНК)		t-статистика	Значение p
				4	5		
	1	2	3	4	5	6	7
1	Константа			b_0	188,90	(1,97)	0,053
2	Доля лиц старших возрастов, в [%]	Age	1 год	b_1	80,69	(21,4)	0,001*
3	Доля проб воздуха, где Q/ПДК > 1.0	Air	2 года	b_2	111,58	(10,9)	0,001*
4	Фактор X_3 по сбросам загрязнённой воды в поверхностные водные объекты территории	Water	3 года	b_3	82,89	(18)	0,001*
5	Фактор X_4 с БП для жителя территории	Benzap	1 год	b_4	23,99	(2,0)	0,049
6	Умеренный уровень загрязнений территории	d_1	0	b_5	179,63	(10,4)	0,001*
7	Высокий уровень загрязнений территории	d_2	0	b_6	381,96	(17)	0,001*

Примечание: p -значение, обозначенное через 0,001*, если получилась вероятность $p < 0,001$.

В нашем случае F -статистика Фишера получилась 252,7 с 66-ю степенями свободы и поэтому гипотеза $H_0: \beta_1 = \beta_2 = \dots = 0$ отвергается с вероятностью выше 0,999. Значение t -статистики из 6-го столбца определено через отношение значения b_j к средней квадратичной ошибке оценивания. Несмещенная оценка s параметра σ (или “стандартная ошибка”) равна 57; оценка R_{adj}^2 коэффициента детерминации R^2 , скорректированная нормирующим множителем, равна 95,4 % [31]. То есть модель (3) со значениями b -оценок из 5-го столбца Таблицы 2 объясняет высокую долю дисперсии $Y^{(t)}$. Применение модели к n наблюдениям материала обучения приводит к средней процентной ошибке 1,5 ($\pm 1,2$)%. Эта ошибка, или “ошибка аппроксимации”, находится так: $MAPE = (100/n) \cdot \sum_{i=1, \dots, n} |Y_{i,est} - Y_{i,f}| / Y_{i,f}$; здесь MAPE – сокращение для mean absolute percentage error. Под

знаком суммы стоит модуль относительной разности между оценкой $Y_{i,est}$ по (3) для i -й территории и фактическим значением $Y_{i,f}$ величины Y (оно берется из [1], когда (3) применяется в региональном разрезе). На этапе “экзамен”, на котором проверяется точность модели (3), MAPE оказалась 2,8 ($\pm 1,87$)%; здесь были взяты те же 38 регионов, но с $Y_{i,f}$ за 2019 г.

На Рисунке 2 показан прогноз $Y_{i,est}$ по (3) для областей и края СФО, Забайкалья.

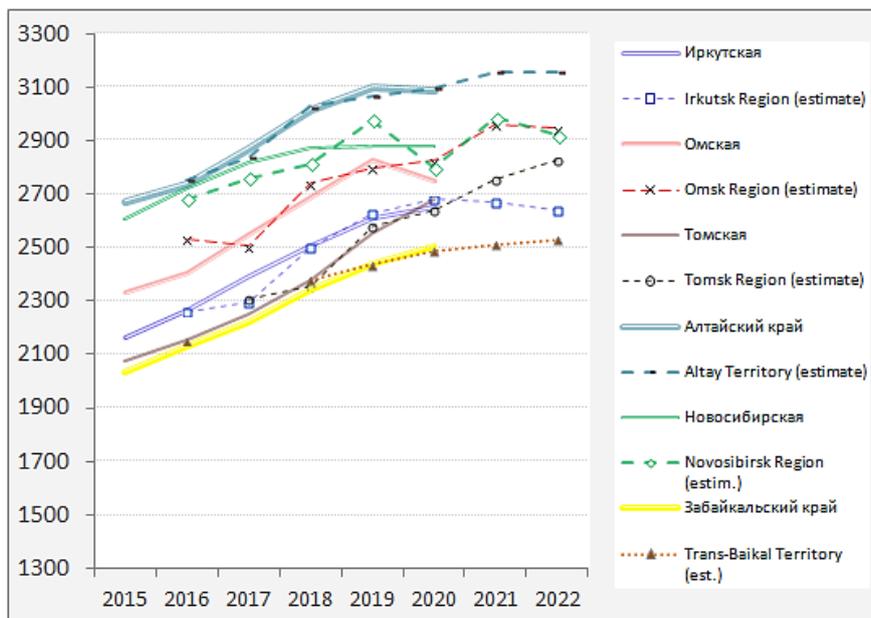


Рисунок 2 – Распространенность ЗНО в регионах
Figure 2 – Cancer prevalence in the regions of Russia

Точки на рисунке с $Y_{i,f}$ за 2017-18 гг. соответствуют строкам из таблицы обучения, а точки с $Y_{i,f}$ за 2019-20 гг. использовались лишь для “экзамена”. Чтобы не загромождать графики, точки $Y_{i,f}$ (т. е. фактические значения $Y^{(t)}$) не показаны; через них проходит непрерывная цветная линия. Для иллюстрации в Таблице 3 показано, как модель работает при прогнозе значения $Y^{(t)}$ (т. е. $Y_{i,est}$) в разрезе муниципальных образований из ЦФО, СФО, рассмотренных с 2017 по 2020 гг. На Рисунке 3 показаны аналогичные результаты для ряда городов: из ПФО, Алтайского и Красноярского края.

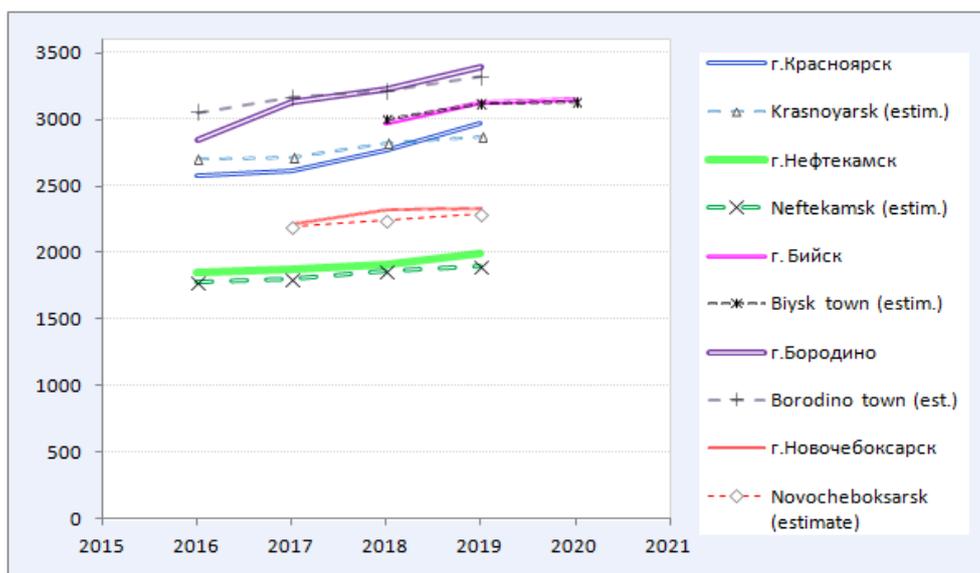


Рисунок 3 – Прогноз $Y^{(t)}$ в городах и $Y_{i,f}$ по факту
Figure 3 – Forecast of $Y^{(t)}$ in Russian cities and $Y_{i,f}$

Таблица 3 – Результаты прогноза для некоторых муниципалитетов из ЦФО и СФО
Table 3 – Forecast values of $Y^{(t)}$ for some municipal territories in Central and Northern Federal Districts of Russia

№	Муниципалитет	$(Y_{est}-Y_j)/Y_f$ [%]	Год	$X_1^{(t-1)}$ [%]	d_1	d_2	Q _{с.г./ПДК}	ИКС Sig	Примечания
1	2	3	4	5	6	7	8	9	10
1	Барнаул (г.о.)	- 4,8	2019	24,23	0	1	6,87	1,83	
2	Брянск (г.о.)	1,67	2017	27,53	то же		0,2	1,41	
3	Рязань (г.о.)	- 1,3	2019	27,34	--		0,56	1,29	
4	Пронский район	2,3	то же	33,24	--		0,4	то же	БП – оценка
5	Михайловский	1,67	--	33,95	--		0,1	--	то же
6	Рязский район	- 0,1	--	31,32	1	0	то же	--	--
7	Сасовский район	- 3,2	--	34,97	то же		--	--	--
8	Новоалтайск (г.о.)	- 0,19	--	24,77	--		0,25	1,83	--

Примечания:

- 1) Значения Y_f взяты из программ “Борьба с онкологическими заболеваниями” каждого региона;
- 2) Для всех строк данные по X_2 , X_3 оценены приближенно;
- 3) В Таблице 3 показан небольшой фрагмент прогноза $Y^{(t)}$ (из почти 150 муниципалитетов).

Фактические данные с $Y_{i,f}$ для создания Рисунка 3, Таблицы 3 взяты из региональных программ “Борьба с онкологическими заболеваниями...”, а для Алтайского или Красноярского края сделан расчет через опубликованные численности лиц с ЗНО и населения края из [22, 23].

Обсуждение

Сначала обсудим аспект, связанный с лагами. Обучающая таблица содержит в каждой строке значения $Y^{(t)}, X_1, \dots, X_4, d_1, d_2$ для конкретного региона, причем из (3) очевидно, что X_1, \dots, X_4 входят в нее с различными лагами τ относительно года t : от года до трех лет. Например, при анализе общей заболеваемости $Y^{(t)}$ в регионах за 2018 г., с ними здесь использовались: доля старших возрастов X_1 , а также X_4 – за 2017 г., фактор X_2 с загрязнением атмосферного воздуха – за 2016 г., фактор X_3 с удельным сбросом загрязненной воды – за 2015 г. Кроме того, в выражении (3) предполагалось, что d_1 и d_2 не зависят от отсчета t . При выборе значений лагов предполагается, что фактор X_j способствует развитию ЗНО с запаздыванием от момента фиксации X_j [8]. На этот аспект, применительно к развитию ЗНО, также обращалось внимание в [19].

Затем обратим внимание на ряд техногенных аспектов. В верхней части Рисунков 2, 3 чаще лежат графики регионов, чьи территории в той или иной степени связаны с авариями, испытаниями или радиационными загрязнениями еще в СССР. Эти графики подтверждают призывы к экологической безопасности жизни, сделанные в [33]. Кроме того, недавно в опытах на мышах в Великобритании установлено, что нежелательны даже малые дозы облучения, эквивалентные дозе трех компьютерных томографий [8]. Также отметим, что некоторые территории ряда областей, а также Алтайского и

Красноярского края, ранее пострадали из-за ряда техногенных событий [19, 29]. Кроме того, ряд районов Рязанской области, а также Забайкальского края существенно загрязняются промпредприятиями [32]. В Таблице 3 также показано, как модель (3) работает для регионов, пострадавших от выбросов техногенных радионуклидов. Остаточное радиоактивное загрязнение ряда территорий обусловлено наличием медленно распадающихся радионуклидов: цезия-137 (^{137}Cs) и стронция-90 (^{90}Sr). Здесь г.о. Брянск с его окрестностями был загрязнен ^{137}Cs и другими вредными выбросами после аварии на ЧАЭС [18]. Это же относится и к Рязанской области.

Наконец, в ряде городов были сильные загрязнения сточными водами [23]. К сожалению, информация по сбросам загрязненных вод на муниципальном уровне труднодоступна, поэтому значения X_3 нередко оценивались приближенно; впрочем, как $Q_{\text{ср.г.}}$ из (1) или X_2 – доле превышений ПДК в пробах воздуха, взятых в населенных пунктах муниципалитета. Так, для сравнительно крупных городов использовалась информация по динамике объемов сброса загрязненных вод за 2005-2011 годы [24] с последующей экстраполяцией, а для прочих территорий использовались данные [25]. С другой стороны, ряд районов имеют, по-видимому, довольно малые значения ряда X_j и поэтому для них в расчетах по (3) принято: $X_2=0,1$ и $X_3=0,5$, а также то, что $Q_{\text{ср.г.}}/ПДК_{\text{с.г.}}=0,1$.

По поводу курения можно заметить, что в 2013 г. Россия присоединилась к Рамочной антитабачной конвенции ВОЗ. Для исследования перспектив курения в РФ применялся один из вариантов модели *SimSmoke* [34]. Это позволяет спрогнозировать уменьшение распространенности курения и числа летальных случаев, связанных с ним, после реализации комплекса антитабачных мер. Кроме того, из данных системы *ENHIS* следует, что в нашей стране в 2006-10 гг. сравнительно слабо боролись с воздействием активного и пассивного дыма табака на детей возраста 13-15 лет (относительно ряда стран ЕС, СНГ, Азии) [35]. Также по многим показателям в *ENHIS* для РФ стоит “нет данных”. Тем не менее, на промежутке 2005-15 гг. среднегодовой спад в потреблении табака российскими мужчинами был неплохим [36]. Если взять Россию среди ее аналогов с близким социально-демографическим индексом, исключив развитые страны, то нас здесь опередили Венгрия, Индия, Румыния, Турция, Украина. А у россиянок этот спад был совсем малым, хотя и курят они реже, чем женщины из ряда стран ЕС.

Для читателя с более продвинутыми знаниями по математической статистике может быть интересно следующее. Для обоснования корректности использования МНК-оценок проверялись условия теоремы Гаусса-Маркова, которая доказывает, когда такие оценки имеют хорошие свойства [31]. Поэтому ниже описана проверка остатков модели на гомоскедастичность, приводятся результаты для тестов на согласие с нормальным законом, как и теста на автокорреляцию 1-го порядка. Из диаграммы рассеяния остатков модели, когда оценки $Y^{(t)}$ отложены по оси абсцисс, следует, что их разброс был одинаковым, точки лежали в горизонтальной полосе; отсюда заметного нарушения условия гомоскедастичности нет. Нормальность распределения остатков изучалась по двум критериям: χ^2 и Лиллиефорса. Верхняя 5%-ная точка распределения χ^2 с 6-ю степенями свободы равна 12,6 и значение 1,62 для χ^2 -статистики не попадает в критическую область; отсюда гипотеза H_0 : “остатки распределены нормально”, – не отвергается. То же относится и к другому критерию, т. к. наибольшее расхождение между эмпирической функцией распределения регрессионных остатков (с серией восходящих ступенек) и соответствующей нормальной функцией было малым (Рисунок 4).

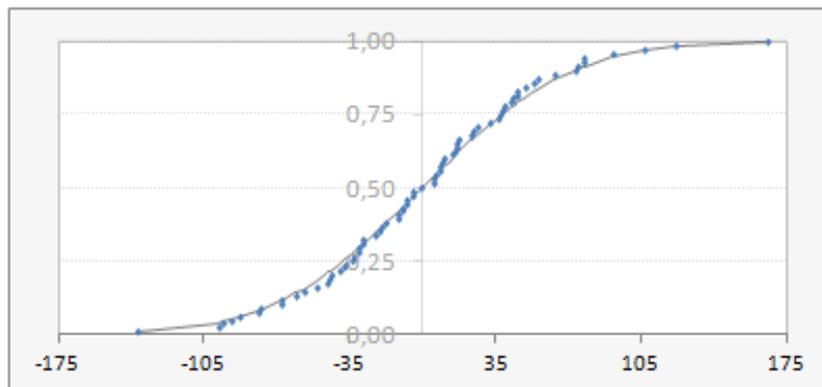


Рисунок 4 – Функция распределения остатков
Figure 4 – Distribution function of the residuals

Отсутствие автокорреляции проверялось DW -тестом; статистика $DW= 1,79$ и не попадает в критические интервалы при уровне значимости $\alpha = 5\%$. Отсюда гипотеза H_0 : “коэффициент автокорреляции равен нулю” не отвергается. Предпосылки теоремы Гаусса-Маркова выполнены; это подтверждает корректность выбора МНК-подхода.

Заклучение

Построена и оценена по точности многофакторная модель с переменной структурой, связывающая общую онкологическую заболеваемость населения с его возрастными свойствами и интенсивностью курения, а также с рядом факторов окружающей среды.

Подтверждена важность мер первичной профилактики, таких как снижение загрязнений атмосферного воздуха в населенных пунктах территории, улучшение степени очистки загрязненных сточных вод и/или снижение объема их сброса, снижение техногенного загрязнения территории и ослабление частоты курения сигарет.

На основе этой модели можно сделать краткосрочный прогноз контингента онкологических больных на ряде территорий (регионы, муниципалитеты), что было проиллюстрировано на примерах. Ее можно применять при разработке мероприятий по медико-демографической и экологической политике, проводимой на территориях. Модель применима и на уровне муниципалитетов, например, при прогнозе контингента лиц с ЗНО и оценке платежей за негативное вредное воздействие на природу. Кроме того, она может быть полезной при имитационном моделировании.

СПИСОК ИСТОЧНИКОВ

1. *Здравоохранение в России*. Приложение к сборнику (информация в разрезе субъектов Российской Федерации). М.: Росстат; 2017-2021. Доступно по: <https://rosstat.gov.ru/folder/210/document/13218> (дата обращения: 14.11.2022).
2. Каприн А.Д., Александрова А.М., Старинский В.В. Профилактика злокачественных новообразований в Российской Федерации как составная часть международной стратегии профилактики неинфекционных заболеваний. *Онкология. Журнал им. П.А. Герцена*. 2016;5(5):42–50. DOI: 10.17116/onkolog20165542-50.
3. Доможирова А.С. Комплексное медико-статистическое прогнозирование как основа перспективного планирования специализированной онкологической помощи в регионах. *Онкология. Журнал им. П.А. Герцена*. 2016;5(1):47–50. DOI: 10.17116/onkolog20165147-50.

4. Лазарев А.Ф., Петрова В.Д., Терехова С.А., Синкина Т.В. Многофакторный анализ при формировании групп высокого онкологического риска. *Бюллетень медицинской науки*. 2017;1(5):37–43. DOI: 10.31684/2541-8475.2017.1(5).37-43.
5. Чолоян С.Б., Екимов А.К., Байгазина Е.Н., Молодцов Н.С., Калинина Е.А., Поснов А.А. Использование методов моделирования для решения задач управления медицинской организацией. Обзор российских научных публикаций. *Менеджер здравоохранения*. 2021;7:11–24. DOI: 10.21045/1811-0185-2021-6-11-24.
6. Фельдблюм И.В., Алыева М.Х., Радионова М.В. Комплексное влияние медико-социальных и средовых факторов риска на вероятность развития колоректального рака. *Тихоокеанский медицинский журнал*. 2018;3(73):24–28. DOI: 10.17238/PmJ1609-1175.2018.3.24-28.
7. Борщук Е.Л., Боев В.М., Бархатова Л.А., Карпенко И.Л., Кудусова Л.Х. Прогностическая модель онкологической заболеваемости населения в условиях воздействия химических канцерогенов среды обитания. *Здоровье населения и среда обитания*. 2017;1(286):13–16. DOI 10.35627/2219-5238/2017-286-1-13-16.
8. Степанов В.С. Связь онкологической заболеваемости с возрастом населения, проживающего при неблагоприятных факторах окружающей среды. *Врач и информационные технологии*. 2021;3:38–49. DOI: 10.25881/18110193_2021_3_38.
9. Голубков А.В., Гаврилова М.П. Применение искусственных нейронных сетей в профилактической и клинической медицине (научный обзор). *Профилактическая и клиническая медицина*. 2020;4(77):30–39. DOI: 10.47843/2074-9120_2020_4_30.
10. *List of Classifications by cancer sites with sufficient or limited evidence in humans*, IARC Monographs of the Identification of Carcinogenic Hazard to Humans. 1–131. Доступно по: <https://monographs.iarc.fr/agents-classified-by-the-iarc> (дата обращения: 14.10.2022).
11. Ефимова Н.В., Мыльникова И.В. Оценка вклада онкогенных факторов в риск развития злокачественных новообразований у городского населения трудоспособного возраста. *Анализ риска здоровью*. 2021;3:99–107. DOI: 10.21668/health.risk/2021.3.09.
12. Корсаков А.В., Домахина А.С., Трошин В.П., Гегерь Э.В. Заболеваемость детского и взрослого населения Брянской области в зависимости от уровней радиационного, химического и сочетанного загрязнения: экологическое исследование. *Экология человека*. 2020;7:4–14. DOI: 10.33396/1728-0869-2020-7-4-14.
13. Новиков С.М., Шашина Т.А., Додина Н.С., Кислицин В.А., Воробьева Л.М., Горяев Д.В., Тихонова И.В., Куркатов С.В. Сравнительная оценка канцерогенных рисков здоровью населения при многосредовом воздействии химических веществ. *Гигиена и санитария*. 2015;94(2):88–92.
14. Radespiel-Troger M., Geiss K., Twardella D., Maier W., Meyer M. Cancer incidence in urban, rural, and densely populated districts close to core cities in Bavaria, Germany. *International Archives of Occupational and Environmental Health*. 2018;91:155–174. DOI: 10.1007/s00420-017-1266-3.
15. *Комплексное наблюдение условий жизни населения*. М.: Росстат; 2018. Доступно по: https://gks.ru/free_doc/new_site/KOUZ18/index.html (дата обращения: 14.11.2022).
16. Иванова А.Е. Прогноз смертности в России, исходя из контроля за основными социальными детерминантами. *Социальные аспекты здоровья населения*. 2020;66(6):6. DOI: 10.21045/2071-5021-2020-66-6-6.
17. Сыч Г.В., Косолапов В.П., Летникова Л.И., Шинкарева Е.В., Сыч А.В. Прогнозирование развития новообразований как одно из профилактических

- мероприятий онкологической заболеваемости. *Системный анализ и управление в биомедицинских системах*. 2018;17(3):802–807.
18. Чопоров О.Н., Агарков А.И., Клименко Г.Я., Шуршуков Ю.Ю. Прогнозирование развития онкологической заболеваемости по индивидуальным медико-социальным факторам риска. *Моделирование, оптимизация и информационные технологии*. 2013;3(3).
 19. Потылицина Е.Н., Тасейко О.В., Сугак Е.В. Оценка влияния загрязнения воздуха предприятиями машиностроения на здоровье населения. *Вестник Сибирского государственного аэрокосмического университета им. академика М.Ф. Решетнева*. 2015;16(4):958–969.
 20. Левшин В.Ф., Слепченко Н.И. Табакокурение и риск развития злокачественных опухолей и других хронических неинфекционных заболеваний. *Онкология. Журнал им. П.А. Герцена*. 2020;9(3):41–47. DOI 10.17116/onkolog2020903141.
 21. Веремчук Л.В., Кику П.Ф., Жерновой М.В. Системное моделирование экологической зависимости распространения онкологических заболеваний в Приморском крае. *Бюллетень физиологии и патологии дыхания*. 2011;41:48–53. Доступно по: <https://cfpd.elpub.ru/jour/article/view/411/389> (дата обращения: 01.12.2022).
 22. *Регионы России. Социально-экономические показатели: стат. сб.* М.: Росстат, 2016–2021. Доступно по: <https://rosstat.gov.ru/folder/210/document/13204> (дата обращения: 01.12.2022).
 23. *База данных показателей муниципальных образований*. Доступно по: <https://rosstat.gov.ru/storage/mediabank/munst.html> (дата обращения: 14.11.2022).
 24. *Охрана окружающей среды в России 2020*. Приложение (информация в разрезе субъектов Российской Федерации). М.: Росстат, 2012–2022. Доступно по: https://gks.ru/bgd/regl/b20_54/Main.htm (дата обращения: 1.12.2022).
 25. *Водные ресурсы и водное хозяйство РФ в 2017 г.*: Стат. сб. М.: НИА-Природа; 2018. 230 с.
 26. Александров Ю.А. *Основы радиационной экологии*. Йошкар-Ола: Издательство Марийского госуниверситета; 2007. 268 с.
 27. Галиулин Р.В., Галиулина Р.А., Кочуров Б.И. Техногенное загрязнение окружающей среды канцерогенными веществами. *Теоретическая и прикладная экология*. 2015;2:42–46.
 28. *Справка по результатам мониторинга бенз[а]пирена в воздухе городов РФ в 2015-19 гг.* Обнинск: НПО «Гайфун», 2016-2020. Доступно по: <https://www.rpatyphoon.ru/products/pollution-media.php> (дата обращения: 14.11.2022)
 29. Радиоактивное загрязнение территорий субъектов РФ цезием-137 (для Рязанской области). *ФИАЦ*. Доступно по: http://www.feerc.ru/Ru/pr1_13.htm (дата обращения: 14.11.2022).
 30. Korovina O.Y., Somin V.A. Environmental state of surface waters and bottom sediments in water bodies of the Southwestern part of the Altai Territory. *NNC RK Bulletin*. 2022;1:43–45. DOI: [10.52676/1729-7885-2022-1-43-45](https://doi.org/10.52676/1729-7885-2022-1-43-45).
 31. Карлберг К. *Регрессионный анализ в Microsoft Excel*. Пер. с англ. М.; 2017. 400 с.
 32. Черная В.В., Воронин Р.М., Сучков И.А., Горнов В.А., Дерягина Л.Е. Мониторинг загрязнения атмосферного воздуха и поверхностных вод в Рязанской области. *Экология человека*. 2021;8:28–35. DOI: [10.33396/1728-0869-2021-8-28-35](https://doi.org/10.33396/1728-0869-2021-8-28-35).
 33. Яковлев Р.М., Обухова И.А. Перспективы атомной энергетики в обеспечении энергетической и экологической безопасности России. *Биосфера*. 2021;3(3):120–136. DOI: [10.24855/biosfera.v13i3.591](https://doi.org/10.24855/biosfera.v13i3.591).

34. Levy D.T., Fouad H., Levy J., Dragomir A.D., Awa F.El. Application of the Abridged SimSmoke model to four Eastern Mediterranean countries. *Tobacco Control*. 2016;25(4):413–421. DOI: 10.1136/tobaccocontrol-2015-052334.
35. *EnvironmentHealth (ENHIS)* – Европейская система по окружающей среде и здоровью. Доступно по: <https://gateway.euro.who.int/ru/datasets/enhis/> (дата обращения: 11.09.2022).
36. GBD 2015 Tobacco Collaborators. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: a systematic analysis from the Global Burden of Disease Study 2015. *The Lancet*. 2017;389(10082):1885–1906. DOI: 10.1016/S0140-6736(17)30819-X.

REFERENCES

1. *Healthcare in Russia*. Appendix to the collection (information in the context of the subjects of the Russian Federation). M: Rosstat; 2007–2021. Available from: <https://rosstat.gov.ru/folder/210/document/13218> (accessed on 14.11.2022). (In Russ.).
2. Kaprin A.D., Aleksandrova L.M., Starinsky V.V. Malignancy prophylaxis in the Russian Federation as part of global strategy for the prevention of non-communicable diseases. *Onkologiya. Zhurnal im. P.A. = P.A. Herzen Journal of Oncology*. 2016;5(5):42–50. DOI: 10.17116/onkolog20165542-50. (In Russ.).
3. Domozhirova A.S. Complex health statistical prediction as a basis for the long-term planning of specialized cancer care in the areas of the Chelyabinsk Region. *Onkologiya. Zhurnal im. P.A. = P.A. Herzen Journal of Oncology*. 2016;5(1):47–50. DOI: 10.17116/onkolog20165147-50. (In Russ.).
4. Lazarev A.F., Petrova V.D., Terekhova S.A., Sinkina T.V. Multivariate statistical analysis when forming groups of high oncological risk. *Byulleten' meditsinskoj nauki = Bulletin of Medical Science*. 2017;1(5):37–43. DOI: 10.31684/2541-8475.2017.1(5).37-43. (In Russ.).
5. Choloyan S.B., Ekimov A.K., Baigazina E.N., Molodtsov N.S., Kalinina E.A., Posnov A.A. The use of modeling methods for solving problems of medical organization management. Review of Russian scientific publications. *Menedzher Zdravoohranenia = Public Health Manager*. 2021;7:11–24. DOI: 10.21045/1811-0185-2021-7-11-24. (In Russ.).
6. Feldblyum I.V., Alyeva M.H., Radionova M.V. Complex impact of medico-social and environmental risk factors on probability of colorectal cancer development. *Tikhookeanskiy meditsinskiy zhurnal = Pacific Medical Journal*. 2018;3(73):24–28. (In Russ.). DOI: 10.17238/PmJ1609-1175.2018.3.24-28.
7. Borshchuk E.L., Boev V.M., Barkhatova L.A., Karpenko I.L., Kudusova L.Kh. Predictive model of cancer morbidity of population under the influence of chemical carcinogen environment. *Zdorov'ye naseleniya i sreda obitaniya = Public Health and Life Environment*. 2017;1(286):13–16. DOI: 10.35627/2219-5238/2017-286-1-13-16. (In Russ.).
8. Stepanov V.S. The relationship of cancer prevalence with age of the population living under adverse environmental factors. *Vrach i informatsionnyye tekhnologii = Medical Doctor and Information Technology*. 2021;3:38–49. DOI: 10.25881/18110193_2021_3_38. (In Russ.).
9. Golubkov A.V., Gavrilova M.P. Application of artificial neural networks in preventive and clinical medicine (review). *Profilakticheskaya i klinicheskaya meditsina = Preventive and clinical medicine*. 2020;4(77):30–39. (In Russ.). DOI: 10.47843/2074-9120_2020_4_30.

10. *List of Classifications by cancer sites with sufficient or limited evidence in humans*, IARC Monographs of the Identification of Carcinogenic Hazard to Humans. 1–131. Available from: <https://monographs.iarc.fr/agents-classified-by-the-iarc> (accessed on 14.10.2022).
11. Efimova N.V., Myl'nikova I.V. Assessment of the contribution made by oncogenic factors to the risk of malignant neoplasms development for the urban population of working age. *Analiz riska zdorov'yu = Human Risk Analysis*. 2021;3:99–107. DOI: 10.21668/health.risk/2021.3.09.eng.
12. Korsakov A.V., Domahina A.S., Troshin V.P., Geger E.V. Child and adult morbidity in the Bryansk Region by the level of radiatoactive, chemical and combined contamination: an ecological study. *Ekologiya cheloveka = HumanEcology*. 2020;7:4–14. DOI 10.33396/1728-0869-2020-7-4-14. (In Russ.).
13. Novikov S.M., Shashina T.A., Dodina N.S., Kislitsyn V.A., Vorobiova L.M., Goryaev D.V., Tikhonova I.V., Kurkatov S.V. Comparative assessment of the multimedia cancer health risks caused by contamination of the Krasnoyarsk Krai Regions' environment. *Gigiena i Sanitariya = Hygiene and Sanitation, Russian Journal*. 2015;94(2):88–92 (In Russ.).
14. Radespiel-Troger M., Geiss K., Twardella D., Maier W., Meyer M. Cancer incidence in urban, rural, and densely populated districts close to core cities in Bavaria, Germany. *International Archives of Occupational and Environmental Health*. 2018;91:155–174. DOI: 10.1007/s00420-017-1266-3.
15. *Comprehensive monitoring of living conditions*. M.: Rosstat, 2018. (In Russ.). Available from: https://gks.ru/free_doc/new_site/KOUZ18/index.html (accessed on 14.11.2022).
16. Ivanova A.E. Mortality forecast in Russia based on monitoring major social determinants. *Social'nye aspekty zdorov'a naselenia = Social aspects of population health [serial online]* 2020;66(6):6. DOI: 10.21045/2071-5021-2020-66-6-6. (In Russ.).
17. Sych G.V., Kosolapov V.P., Letnikova L.I., Shinkareva E.V., Sych A.V. Forecasting the development of tumors as one of the preventive measures of cancer incidence. *Sistemnyy analiz i upravleniye v biomeditsinskikh sistemakh = System Analysis and Management in Biomedical Systems*. 2018;17(3):802–807. (In Russ.).
18. Choporov O.N., Agarkov A.I., Klimenko G.J., Shurshukov Y.Y. Prediction oncological morbidity individual health and social risk factors. *Modelirovaniye, optimizatsiya i informatsionnyye tekhnologii = Modeling, Optimization and Information Technology*. 2013;3(3). (In Russ.).
19. Potylitsyna E.N., Taseiko O.V., Sugak E.V. Assessment of influence of air pollution produced by engineering enterprises on human health. *Vestnik Sibirskogo gosudarstvennogo aerokosmicheskogo universiteta im. akademika M.F. Reshetneva = The Journal "Vestnik of SibGAU"*. 2015;16(4):958–969. (In Russ.).
20. Levshin V.F., Slepchenko N.I. Tobacco smoking and risk of developing malignant tumors and other chronic noncommunicable diseases. *Onkologiya. Zhurnal im. P.A. Gertsena = P.A. Herzen Journal of Oncology*. 2020;9(3):41–47. (In Russ.). DOI: 10.17116/onkolog2020903141.
21. Veremchuk L.V., Kiku P.F., Zhernovoi M.V. System modeling of ecological dependence in distribution of oncologic diseases within the Primorye Territories. *Byulleten' fiziologii i patologii dykhaniya = Bulletin Physiology and Pathology of Respiration*. 2011;41:48–53. Available from: <https://cfpd.elpub.ru/jour/article/view/411/389> (accessed on 01.12.2022). (In Russ.).
22. *Regions of Russia. Socioeconomic indicators digest*. M.: Rosstat; 2016–2021. Available from: <https://rosstat.gov.ru/folder/210/document/13204> (accessed on 01.12.2022).
23. *Database of indicators of municipalities*. (In Russ.). Available from: <https://rosstat.gov.ru/storage/mediabank/munst.html> (accessed on 14.11.2022).

24. *Environmental protection in Russia. Appendix with information in the context of Russia's Regions.* М.: Rosstat, 2012–2022. Available from: https://gks.ru/bgd/regl/b20_54/Main.htm (accessed on 1.12.2022). (In Russ.).
25. *Water resources and water management in Russia in 2017.* М.: Nature; 2018. 230 p. (In Russ.).
26. Aleksandrov Y.A. *Fundamentals of radiation ecology.* Yoshkar-Ola: Izdatelstvo Mariyskogo gosuniversiteta; 2007. 268 p. (In Russ.).
27. Galiulin R.V., Galiulina R.A., Kochurov B.I. Technogenesis contamination of environment by cancerogenic substances. *Teoreticheskaya i prikladnaya ekologiya = Theoretical and Applied Ecology.* 2015;2:42–46. (In Russ.).
28. *Information on the results of monitoring benz[a]pyrene in the air of the Russian cities in 2015-2019.* Obninsk: “Typhoon”. Available from: <https://www.rpatyphoon.ru/products/pollution-media.php> (accessed on 14.11.2022). (In Russ.).
29. Radioactive contamination of the territories of the constituent entities of the Russian Federation with cesium-137 (for the Ryazan Region). *Federal Environmental Emergency Response Centre (FEERC).* Available from: http://www.feerc.ru/Ru/pr1_13.htm (accessed on 14.11.2022). (In Russ.).
30. Korovina O.Y., Somin V.A. Environmental state of surface waters and bottom sediments in water bodies of the Southwestern part of the Altai Territory. *NNC RK Bulletin.* 2022;1:43–45. DOI: [10.52676/1729-7885-2022-1-43-45](https://doi.org/10.52676/1729-7885-2022-1-43-45).
31. Carlberg C. *Regression Analysis Microsoft Excel.* USA; 2016.
32. Chernaya V.V., Voronin R.M., Suchkov I.A., Gornov V.A., Deryagina L.E. Monitoring of air and surface water pollution in the Ryazan Region. *Ekologiya cheloveka = Human Ecology.* 2021;8:28–5. (In Russ.). DOI: [10.33396/1728-0869-2021-8-28-35](https://doi.org/10.33396/1728-0869-2021-8-28-35).
33. Yakovlev R.M., Obukhova I.A. Prospects for nuclear energetic in securing the energy and environmental safety of Russia. *Biosfera.* 2021;3(3):120–136. DOI: [10.24855/biosfera.v13i3.591](https://doi.org/10.24855/biosfera.v13i3.591). (In Russ.).
34. Levy D.T., Fouad H., Levy J., Dragomir A.D., Awa F.El. Application of the Abridged SimSmoke model to four Eastern Mediterranean countries. *Tobacco Control.* 2016;25(4):413–421. DOI: [10.1136/tobaccocontrol-2015-052334](https://doi.org/10.1136/tobaccocontrol-2015-052334).
35. *Environment and Health Information System.* Available from: <https://gateway.euro.who.int/ru/datasets/enhis/> (accessed on 11.09.2022).
36. GBD 2015 Tobacco Collaborators. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: a systematic analysis from the Global Burden of Disease Study 2015. *The Lancet.* 2017;389(10082):1885–1906. DOI: [10.1016/S0140-6736\(17\)30819-X](https://doi.org/10.1016/S0140-6736(17)30819-X).

ИНФОРМАЦИЯ ОБ АВТОРЕ / INFORMATION ABOUT THE AUTHOR

Степанов Владимир Сергеевич, кандидат физико-математических наук, старший научный сотрудник Центрального экономико-математического института РАН, Москва, Российская Федерация
e-mail: vladstep0355@gmail.com,
stepanov@cemi.rssi.ru
ORCID: [0000-0002-4478-376X](https://orcid.org/0000-0002-4478-376X)

Vladimir Sergeevich Stepanov, Candidate of Physical and Mathematical Sciences, Senior Researcher at Central Economics and Mathematics Institute of the RAS, Moscow, Russian Federation.

*Статья поступила в редакцию 05.12.2022; одобрена после рецензирования 01.02.2023;
принята к публикации 10.03.2023.*

*The article was submitted 05.12.2022; approved after reviewing 01.02.2023;
accepted for publication 10.03.2023.*