

DOI: <https://doi.org/10.17816/humeco611099>

Моделирование и прогнозирование возрастных коэффициентов смертности на основе модели Ли–Картера

Е.Л. Борщук¹, Д.Н. Бегун¹, И.П. Болодурина^{1, 2}, Л.И. Меньшикова³,
С.В. Колесник², А.Н. Дуйсембаева¹

¹ Оренбургский государственный медицинский университет, Оренбург, Россия;

² Оренбургский государственный университет, Оренбург, Россия;

³ Северный государственный медицинский университет, Архангельск, Россия

АННОТАЦИЯ

Обоснование. Проблема высокой смертности населения остаётся одной из наиболее значимых проблем здравоохранения России. Одной из приоритетных задач государственной политики является сокращение смертности трудоспособного населения и увеличение продолжительности жизни. Прогнозирование уровня смертности населения является инструментом разумного распределения разных видов ресурсов.

Цель. Моделирование и прогноз показателей смертности населения Оренбургской области на основе модели Ли–Картера.

Материал и методы. Изучены половозрастные показатели смертности и численности всего населения Оренбургской области за 1991–2020 гг. Для моделирования и прогнозирования смертности населения применен метод Ли–Картера. На основе вычисленных параметров построена модель случайного блуждания с дрейфом и произведена оценка точности.

Результаты. Для Оренбургской области моделью Ли–Картера описана смертность мужского населения. В результате моделирования смертности мужского населения Оренбургской области получена точность подгонки 87%, достаточная для оценки результатов долгосрочного прогнозирования, построены прогнозы смертности до 2035 г.

Заключение. Анализ полученных результатов показал, что влияние пандемии на смертность населения носит краткосрочный характер, в ближайшие годы смертность мужского населения Оренбургской области продолжит снижаться.

Ключевые слова: смертность; численность населения; половозрастные коэффициенты смертности; моделирование; прогнозирование.

Как цитировать:

Борщук Е.Л., Бегун Д.Н., Болодурина И.П., Меньшикова Л.И., Колесник С.В., Дуйсембаева А.Н. Моделирование и прогнозирование возрастных коэффициентов смертности на основе модели Ли–Картера // Экология человека. 2024. Т. 31, № 1. С. 59–73. DOI: <https://doi.org/10.17816/humeco611099>

DOI: <https://doi.org/10.17816/humeco611099>

Modeling and prediction of age-specific mortality rates using the Lee–Carter model

Evgeny L. Borshchuk¹, Dmitry N. Begun¹, Irina P. Bolodurina^{1, 2}, Larisa I. Menshikova³, Svetlana V. Kolesnik², Aislu N. Duisembayeva¹

¹ Orenburg State Medical University, Orenburg, Russia;

² Orenburg State University, Orenburg, Russia;

³ Northern State Medical University, Arkhangelsk, Russia

ABSTRACT

BACKGROUND: High mortality remains one of the most significant health concerns in Russia. One of the priorities of the state policy is to reduce mortality rates among the working-age population and increase life expectancy. Predicting population mortality rates serves as a valuable tool for effectively allocating the available resources.

AIM: To perform mathematical modeling and prediction of mortality rates of the population of the Orenburg region using the Lee–Carter model.

MATERIAL AND METHODS: The age- and sex-specific mortality rates and the population size of the Orenburg region for the period 1991–2020 was used as a study base. The Lee–Carter method was applied to model and predict population mortality. By deriving key parameters, a random walk model with drift was developed, and an accuracy assessment was performed.

RESULTS: The Lee–Carter model has been utilized to analyze the mortality rates of the male population in the Orenburg region. Through this modeling process, an accuracy rate of 87% was achieved, providing a reliable basis for long-term prediction. Mortality forecasts have been generated up to the year 2035, allowing for a comprehensive evaluation of future trends in the region.

CONCLUSION: The analysis of the results indicates that the pandemic's impact on population mortality is expected to be short-term. In the upcoming years, the mortality rate of the male population in the Orenburg region is projected to continue decreasing.

Keywords: mortality; population; age and sex mortality rates; modeling; forecasting.

To cite this article:

Borshchuk EL, Begun DN, Bolodurina IP, Menshikova LI, Kolesnik SV, Duisembayeva AN. Modeling and prediction of age-specific mortality rates using the Lee–Carter model. *Ekologiya cheloveka (Human Ecology)*. 2024;31(1):59–73. DOI: <https://doi.org/10.17816/humeco611099>

Received: 20.10.2023

Accepted: 09.07.2024

Published online: 06.08.2024

DOI: <https://doi.org/10.17816/humeco611099>

基于 Lee-Carter 模型的特定年龄死亡率建模和预测

Evgeny L. Borshchuk¹, Dmitry N. Begun¹, Irina P. Bolodurina^{1,2}, Larisa I. Menshikova³,
Svetlana V. Kolesnik², Aislu N. Duisembayeva¹

¹ Orenburg State Medical University, Orenburg, Russia;

² Orenburg State University, Orenburg, Russia;

³ Northern State Medical University, Arkhangelsk, Russia

简评

论证。 人口死亡率高问题仍然是俄罗斯最重要的公共卫生问题之一。国家政策的优先任务之一是降低劳动年龄人口的死亡率，延长预期寿命。预测人口死亡率是合理分配各类资源的工具。

目标。 基于Lee-Carter模型的奥伦堡地区死亡率建模和预测。

材料与方法。 研究了1991–2020年奥伦堡地区的年龄和性别死亡率以及总人口。使用Lee-Carter方法对人口死亡率进行建模和预测。根据计算得到的参数，构建了带漂移的随机游走模型，并进行了精度评定。

结果。 对于奥伦堡地区，Lee-Carter模型描述了男性人口的死亡率。通过对奥伦堡地区男性人口死亡率的建模，获得了87%的拟合精度，足以评估长期预测的结果，并构建了到2035年的死亡率预测。

结论。 对所得结果的分析表明，大流行病对人口死亡率的影响是短期的，在未来几年中，奥伦堡州男性人口的死亡率将继续下降。

关键词： 死亡率；人口；性别-年龄死亡率；建模；预测。

引用本文：

Borshchuk EL, Begun DN, Bolodurina IP, Menshikova LI, Kolesnik SV, Duisembayeva AN. 基于 Lee-Carter 模型的特定年龄死亡率建模和预测. *Ekologiya cheloveka (Human Ecology)*. 2024;31(1):61–76. DOI: <https://doi.org/10.17816/humeco611099>

收到: 20.10.2023

接受: 09.07.2024

发布日期: 06.08.2024

ОБОСНОВАНИЕ

Высокий уровень смертности является основной демографической проблемой [1]. Сегодня наиболее остро стоит вопрос смертности населения трудоспособного возраста как экономически активной части общества [2], в результате преждевременной смертности которого экономика страны несёт колоссальные потери [3, 4]. При этом следует отметить, что большинство учёных, занимающихся данной тематикой, отмечают потери именно трудоспособной рабочей силы страны [5, 6].

Снижение уровня смертности населения — одна из главенствующих целей демографической политики России [7]. Достичь этого можно путём сокращения заболеваний системы кровообращения, новообразований, внешних причин, которые занимают лидирующие места в причинах смерти во всем мире и в Российской Федерации [8, 9].

Одним из основных показателей здоровья населения любого государства является ожидаемая продолжительность жизни [10] — интегральный показатель состояния здоровья населения и уровня социально-экономического развития, который активно используется как индикатор качества человеческих ресурсов.

Как известно, для расчёта данного показателя строятся таблицы смертности [11]. Если известны повозрастные коэффициенты смертности, то появляется возможность рассчитать вероятность дожития до определенного возраста — непосредственно её и экстраполируют на будущее, определяя среднюю ожидаемую продолжительность жизни для различных возрастных когорт.

Интерес к моделированию и прогнозированию показателя смертности населения обусловлен необходимостью её снижения для достижения более высоких показателей качества жизни и благосостояния [12], сохранения доли экономически активного населения и многими другими причинами, в том числе потребностью формирования механизмов снижения смертности [13]. Обострило проблему получения точных прогнозов смертности существенное старение населения в экономически развитых странах [14].

Таким образом, прогнозирование смертности в настоящее время является актуальной задачей. Для реализации мероприятий демографической политики в решении проблемы необходим мультидисциплинарный подход [15], позволяющий осуществлять аналитическую оценку и прогнозирование эффективности принимаемых решений для обоснования наиболее оптимальных мер как для краткосрочной, так и для более продолжительной перспективы [16]. Статистические модели позволяют разрабатывать рекомендации по рациональному перераспределению финансовых и других ресурсов для повышения эффективности и результативности оказания медицинской помощи [17]. Для эффективного решения задачи поэтапной трансформации системы оказания первичной медико-санитарной помощи были использованы методы математической

статистики [18], что позволило разработать рекомендации по управлению изменениями в процессе перехода к новой модели медицинской организации [19]. Результаты моделирования показателей деятельности онкологической службы показали возможности прогнозирования числа случаев смерти от злокачественных новообразований, что позволяет более эффективно и рационально распределять бюджет на диагностику и лечение [20].

Наиболее интересной для прогнозирования возрастных коэффициентов смертности является модель, разработанная демографами Р.Д. Ли и Л.Р. Картером в 1992 г. для прогнозирования смертности в США [21]. Особенностью данной модели является учёт возрастной структуры населения. Преимущества модели Ли–Картера по сравнению с другими прогнозными моделями возрастной смертности состоят в её относительно простой конструкции и устойчивости. Модель включает минимум параметров, каждый из которых можно демографически интерпретировать. Для их оценки требуется сравнительно небольшое количество информации.

Цель исследования. Моделирование и прогноз показателей смертности населения Оренбургской области на основе модели Ли–Картера.

МАТЕРИАЛ И МЕТОДЫ

Данные численности населения по полу и возрасту взяты из базы территориального органа Федеральной службы государственной статистики по Оренбургской области. Из ежегодных сборников медицинского информационно-аналитического центра Оренбургской области выкопировали абсолютные данные о количестве смертей в 1991–2020 гг. по полу и по возрастам. На основе собранных данных рассчитали повозрастные коэффициенты смертности. К полученным значениям применяли метод Ли–Картера. Рассчитали параметры модели методом сингулярного разложения, в том числе индекс смертности населения Оренбургской области, представляющий собой временной ряд. Все вычисления и расчёты, в том числе проверку статистических гипотез, проводили с помощью библиотек языка программирования Python.

Для моделирования и прогнозирования индекса смертности применяли интегрированные модели авторегрессии — скользящего среднего (ARIMA-модели). ARIMA определяется тремя параметрами (p, d, q):

- p — порядок авторегрессии (AR), позволяющий включить предыдущие значения временного ряда в модель;
- d — порядок интегрирования (I), отражающий количество шагов, требуемых для приведения ряда к стационарному виду;
- q — порядок скользящего среднего (MA), позволяющий учесть погрешность модели как линейную комбинацию наблюдавшихся ранее значений ошибок.

Проверку временного ряда на стационарность осуществляли на основе анализа автокорреляционной и частной автокорреляционной функции, а также на основе расширенного теста Дики–Фуллера.

Подбор параметров ARIMA-модели производили на основе анализа автокорреляционной и частной корреляционной функции, качество модели оценивали с помощью информационных критериев Акаике, Шварца, Ханнана–Куина.

Адекватность выбранной модели оценивали на основе коэффициента детерминации. Коэффициент детерминации — доля дисперсии зависимой переменной, объясняемая рассматриваемой моделью АРИМА:

$$R^2 = 1 - \frac{\sigma^2}{\sigma_y^2}$$

Чем ближе эта доля к единице, тем лучше качество модели.

Точность прогнозирования оценивали на основе MAPE — средней абсолютной процентной ошибки:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i^{real} - y_i^{model}|}{y_i^{real}}$$

В качестве исходных данных для построения модели использовали возрастные коэффициенты смертности за 1991–2020 гг. отдельно для мужского и женского населения Оренбургской области. Возрастные коэффициенты смертности представлены для 5-летних возрастных групп (x): 0–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84, 85–89, 90–94, 95–99, 100+ лет, их рассчитывали по формуле:

$$m_{x,t} = \frac{D_{x,t}}{\bar{S}_{x,t}}$$

где $\bar{S}_{x,t}$ — среднегодовая численность мужчин/женщин в возрастной группе x в году t ($t = (1991, 2020)$).

$D_{x,t}$ — абсолютное число умерших мужчин/женщин в возрастной группе x в году t .

Для дальнейшего моделирования из вычисленных коэффициентов смертности составляли матрицу M , где каждая строка отражает определённую возрастную группу, а каждый столбец представляет годы исследования. Размер полученной матрицы составил 21×30 .

К полученной матрице повозрастных коэффициентов смертности применяли метод Ли–Картера. Согласно методу, матрица подгоняется с помощью простой модели для описания изменений общей смертности как функции одного временного параметра k_t :

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} \quad \text{или} \quad m_{x,t} = e^{a_x + b_x k_t + \varepsilon_{x,t}}$$

где $m_{x,t}$ — наблюдаемая повозрастная смертность в возрасте x в течение времени t , коэффициенты a_x , b_x , k_t — параметры модели, $\varepsilon_{x,t}$ — ошибки, имеющие нормальное распределение с нулевой средней и постоянной

дисперсией. Интерпретация параметров довольно проста: a_x — средние значения $\ln(m_{x,t})$ в возрасте x в течение времени t , b_x представляет собой характер изменения смертности по возрасту, а индекс k_t — тенденцию изменений смертности во времени.

Эта модель явно недоопределена, то есть b_x и k_t можно выбрать множеством способов. Так, один из этих двух элементов можно умножить на константу, а другой разделить на ту же константу без изменения прогнозируемых значений, заданных моделью. Мы использовали нормализующее ограничение на b_x :

$$\sum_x b_x = 1.$$

Для полной идентификации также необходимо потребовать

$$\sum_t k_t = 0.$$

Параметры не могут быть найдены методами регрессии, так как в модели отсутствуют объясняющие переменные, в правой части уравнения у нас есть только оцениваемые параметры. Для решения этой проблемы расчёт параметров модели осуществляли при помощи метода сингулярного разложения (SVD).

Сингулярным разложением матрицы M порядка $m \times n$ является разложение следующего вида:

$$M = U \Lambda V^T,$$

где Λ — матрица размера $m \times n$ с неотрицательными элементами, у которой элементы, лежащие на главной диагонали, — это сингулярные числа (все элементы не на главной диагонали являются нулевыми), а матрицы U (порядка $m \times m$) и V (порядка $n \times n$) — это две унитарные матрицы, состоящие из левых и правых сингулярных векторов соответственно, V^T — это транспонированная матрица к V .

Метод SVD применяется для уменьшения размерности во многих практических задачах, где требуется приближать матрицу некоторой другой матрицей M_k с заранее заданным рангом k :

$$M = U_k \Lambda_k V_k^T,$$

где матрицы U_k , Λ_k , V_k получаются из соответствующих матриц в сингулярном разложении матрицы M обрезанием до ровно k первых столбцов. Таким образом, приближая матрицу M матрицей меньшего ранга, мы выполняем своего рода сжатие информации, содержащейся в M : матрица M размера $m \times n$ заменяется меньшими матрицами размеров $m \times k$ и $k \times n$ и диагональной матрицей с k элементами. При этом сжатие происходит с потерями — в приближении сохраняется лишь наиболее существенная часть матрицы M . Применительно к нашей задаче метод сингулярного разложения записывается в виде:

$$SVD(\tilde{M}_{x,t}) = U_{x,1} \lambda_1 V_{1,t}^T,$$

где $\tilde{M}_{x,t} = \ln(m_{x,t}) - a_x$, $b_x = U_{x,1}$, $k_t = V_{1,t}^T$.

Таким образом, мы приближаем матрицу смертности матрицей, представленной в виде произведения первого

левого (первый столбец матрицы U) и первого правого (первая строка матрицы V^T) сингулярных векторов, а в качестве λ выбираем первое (наибольшее) сингулярное число λ_1 .

Расчёт сингулярного разложения производили с помощью функции `linalg.svd()`, входящей в состав библиотеки `scipy` языка программирования Python.

Для оценки точности подгонки вычисляли объяснённую дисперсию. Это отношение дисперсии различий между фактическими и подобранными показателями к дисперсии фактических показателей:

$$\sigma = \left(1 - \frac{\sigma(m_{model} - m_{real})}{\sigma(m_{real})}\right) \times 100\%.$$

Для данных по Оренбургской области точность модели представлена в табл. 1.

Анализ остатков построенной ARIMA-модели осуществлялся на основе критерия Льюинга–Бокса на уровне значимости $p=0,05$.

В результате предварительного анализа исходных данных построили графики зависимостей коэффициентов смертности мужского и женского населения от времени и от возраста. На рис. 1 изображена динамика общих коэффициентов смертности для мужского и женского населения за 1991–2020 гг. Анализ приведённых зависимостей позволил сделать вывод о том, что смертность населения региона имеет нестабильную динамику, обусловленную различными факторами. К примеру, распад СССР, экономический кризис и ухудшение показателей уровня и качества жизни привели к резкому росту смертности в 1993–1994 гг. С 2005 до 2020 г. в Оренбургской области наблюдалась тенденция к снижению смертности населения, причём снижение смертности мужского населения имело более выраженный характер. Темп снижения среди мужчин с 2005 по 2019 г. составил 22,0%, а среди женщин — 10,5%. Скачок смертности в 2020 г., отражённый на рис. 1, связан с начавшейся пандемией COVID-19, темп прироста среди мужчин в 2020 г., по сравнению с 2019 г., составил 19,5%, среди женщин — 28,5%.

РЕЗУЛЬТАТЫ

На рис. 2, 3 приведены зависимости коэффициентов смертности мужского и женского населения по различным возрастным группам для 2005, 2019 и 2020 гг. Анализ приведённых зависимостей показал, что смертность как мужского, так и женского населения снижалась до 2019 г. для всех возрастных групп. В 2020 г. смертность мужского населения превысила смертность 2019 г. для возрастных групп старше 25 лет, а женского населения — для возрастных групп старше 15 лет.

Расчёт параметров модели осуществляли при помощи метода сингулярного разложения.

Результаты расчёта параметров модели представлены в табл. 2, 3.

Рассматривая соответствие для каждой возрастной группы отдельно, наблюдаем самую низкую точность для мужчин в возрастной группе 95–99 лет (10,27%), для женщин — 75–79 лет (10,46%). В целом точность модели составила 70,69% для мужчин и 51,91% для женщин.

Дальнейшее исследование было сфокусировано на мужской части населения, так как точность подгонки для женского населения оказалась недостаточной для дальнейшего прогнозирования.

Результаты расчёта параметров модели для мужского населения представлены графически на рис. 4–6.

Анализ зависимости на рис. 6 показал, что индекс смертности не подчиняется линейной зависимости, поэтому точность модели оказалась недостаточной. На рис. 6 показано уменьшение индекса смертности k_t , начиная с 2001 г. Предположительно точность подгонки можно

Таблица 1. Точность построенной модели для возрастных коэффициентов смертности Оренбургской области с 1991 по 2020 г.

Table 1. Accuracy of the constructed model for age-specific mortality rates in the Orenburg region from 1991 to 2020

Возрастные группы Age groups	Объяснённая дисперсия σ , % The explained variance σ , %	
	Мужчины Males	Женщины Females
0–4	89,82	58,28
5–9	73,06	57,75
10–14	78,44	69,68
15–19	84,61	92,31
20–24	87,91	81,94
25–29	61,80	63,27
30–34	66,74	34,18
35–39	66,13	52,59
40–44	72,51	65,03
45–49	79,07	63,40
50–54	68,17	51,99
55–59	73,14	51,34
60–64	62,03	62,62
65–69	53,93	36,88
70–74	65,73	51,14
75–79	54,32	10,46
80–84	27,06	29,89
85–89	30,71	53,39
90–94	35,16	22,58
95–99	10,27	18,40
100+	33,13	23,13

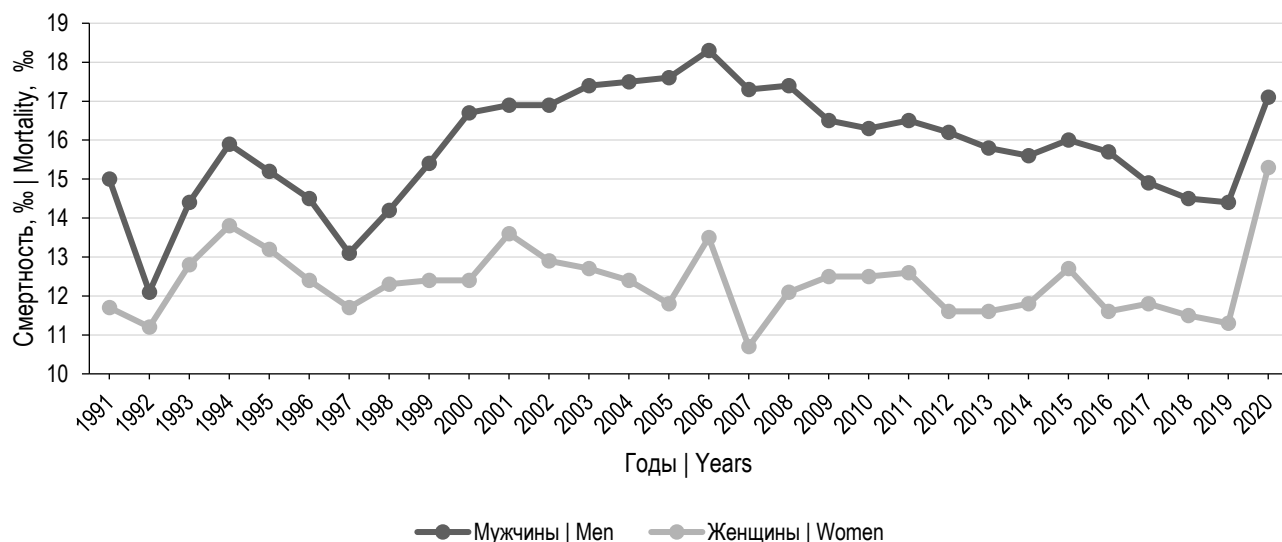


Рис. 1. Общие коэффициенты смертности мужского и женского населения Оренбургской области на 1000 населения.
Fig. 1. Mortality rates of the male and female population of the Orenburg region per 1000 population.

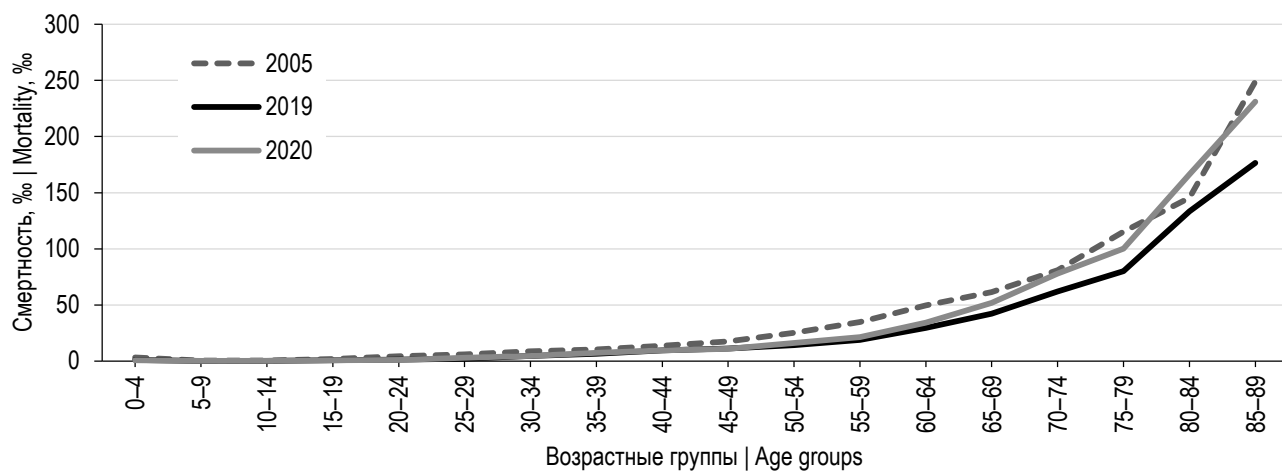


Рис. 2. Возрастные показатели смертности мужского населения Оренбургской области на 1000 населения.
Fig. 2. Age-specific mortality rates among males in the Orenburg region per 1000 population.

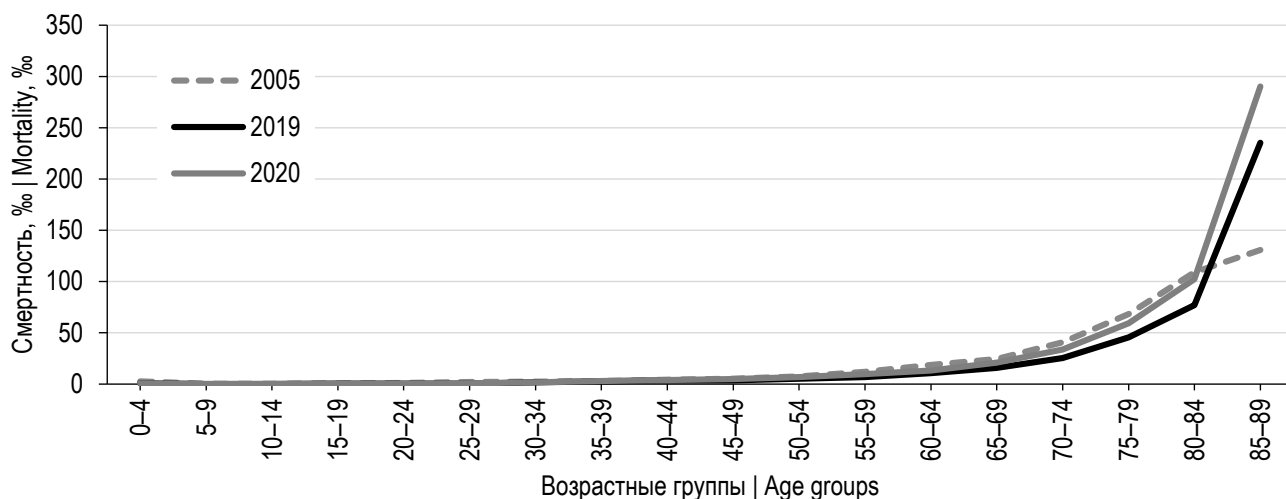


Рис. 3. Возрастные показатели смертности женского населения Оренбургской области на 1000 населения.
Fig. 3. Age-specific mortality rates among females in the Orenburg region per 1000 population.

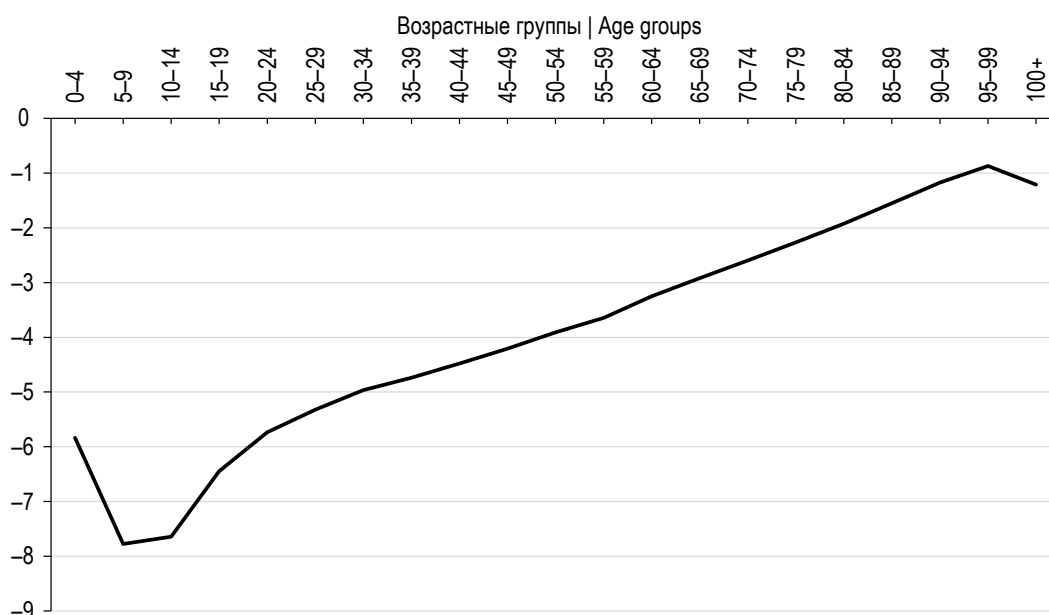


Рис. 4. Значения параметра a_x для базисного периода 1991–2020 гг.

Fig. 4. Values of the a_x parameter for the baseline period 1991–2020.

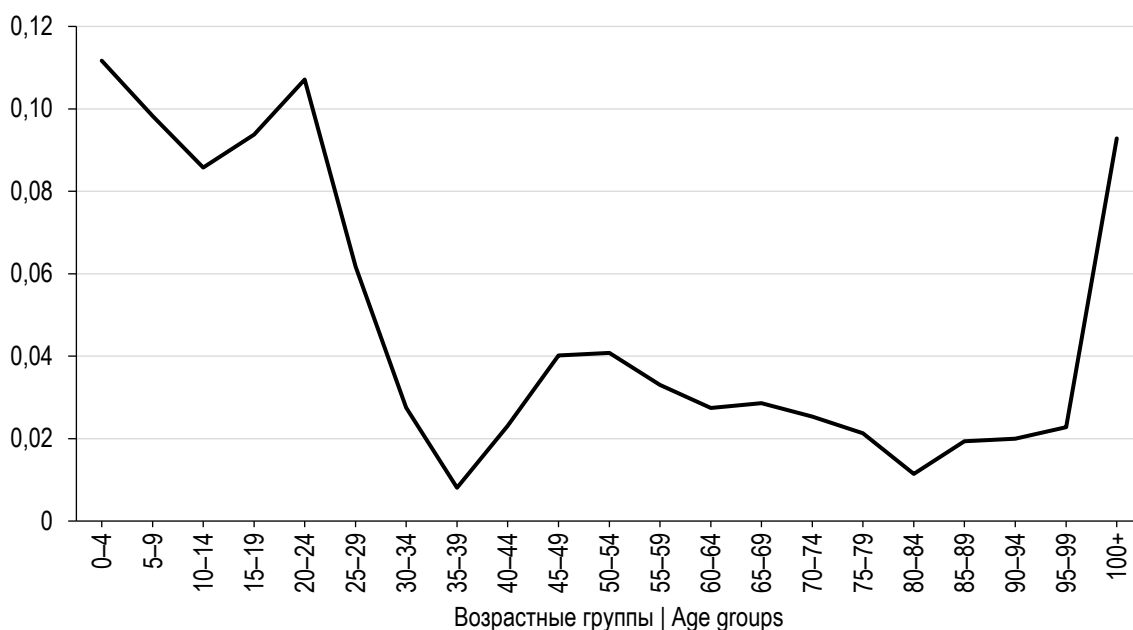
Таблица 2. Параметры a_x , b_x для мужского и женского населения Оренбургской области для базисного периода 1991–2020 гг.

Table 2. Parameters a_x , b_x for male and female population of the Orenburg region for the baseline period of 1991–2020

Возрастные группы Age groups	Мужское население Male population		Женское население Female population	
	a_x	b_x	a_x	b_x
0-4	-5,83	0,11	-6,10	0,16
5-9	-7,77	0,10	-8,19	0,18
10-14	-7,64	0,09	-8,26	0,25
15-19	-6,45	0,09	-7,32	0,32
20-24	-5,73	0,11	-7,00	0,25
25-29	-5,32	0,06	-6,62	0,06
30-34	-4,97	0,03	-6,28	-0,06
35-39	-4,74	0,01	-5,98	-0,03
40-44	-4,48	0,02	-5,73	0,04
45-49	-4,21	0,04	-5,40	0,09
50-54	-3,91	0,04	-5,05	0,03
55-59	-3,64	0,03	-4,63	-0,05
60-64	-3,25	0,03	-4,20	-0,05
65-69	-2,92	0,03	-3,84	0,09
70-74	-2,60	0,03	-3,30	0,16
75-79	-2,27	0,02	-2,85	0,01
80-84	-1,92	0,01	-2,28	-0,05
85-89	-1,55	0,02	-1,77	-0,14
90-94	-1,17	0,02	-1,32	-0,10
95-99	-0,87	0,02	-0,85	-0,06
100+	-1,21	0,09	-0,87	-0,09

Таблица 3. Параметр k_t для мужского и женского населения Оренбургской области**Table 3.** Parameter k_t for male and female population of the Orenburg region

Год Year	1991	1992	1993	1994	1995	1996
Мужчины Males	4,31	2,54	3,34	4,31	3,73	2,35
Женщины Females	2,04	1,99	2,41	2,30	2,52	2,75
Год Year	1997	1998	1999	2000	2001	2002
Мужчины Males	2,47	3,75	3,22	2,00	4,49	3,14
Женщины Females	1,41	1,91	1,26	1,57	2,85	1,07
Год Year	2003	2004	2005	2006	2007	2008
Мужчины Males	2,86	2,05	2,39	0,20	0,52	1,61
Женщины Females	0,89	0,41	0,43	0,34	0,61	-0,08
Год Year	2009	2010	2011	2012	2013	2014
Мужчины Males	-0,24	-0,74	-1,99	-2,29	-1,47	-3,06
Женщины Females	-0,72	-1,54	-0,50	-1,53	-1,93	-1,11
Год Year	2015	2016	2017	2018	2019	2020
Мужчины Males	-4,34	-6,72	-5,73	-6,52	-8,25	-7,94
Женщины Females	-2,68	-3,08	-2,12	-4,24	-3,43	-3,80

**Рис. 5.** Значения параметра b_x для базисного периода 1991–2020 гг.**Fig. 5.** Values of the b_x parameter for the baseline period 1991–2020.

увеличить, сократив период наблюдений. Также большой вклад в точность подгонки внесло разбиение на возрастные группы.

Для увеличения точности подгонки последовательно отбросили возрастные группы и моменты наблюдения, которые внесли наиболее большой вклад в необъяснённую дисперсию модели.

Ниже приведены ошибки модели, определённые как сумма квадратов отклонений реальных и модельных данных отдельно для каждой возрастной

группы и каждого года наблюдения. Они представлены в табл. 4, 5. Для увеличения точности подгонки удалим из периода наблюдений возрастные группы и года, соответствующие максимальной ошибке модели. В результате серии вычислительных экспериментов в качестве базисного периода выбран интервал с 1999 по 2020 г. Из возрастных групп исключили группы 0–4, 5–9, 10–14, 95–99, 100+, которые имеют слишком маленькую численность умерших, что приводит к статистической неточности и ошибкам в моделировании и прогнозировании.

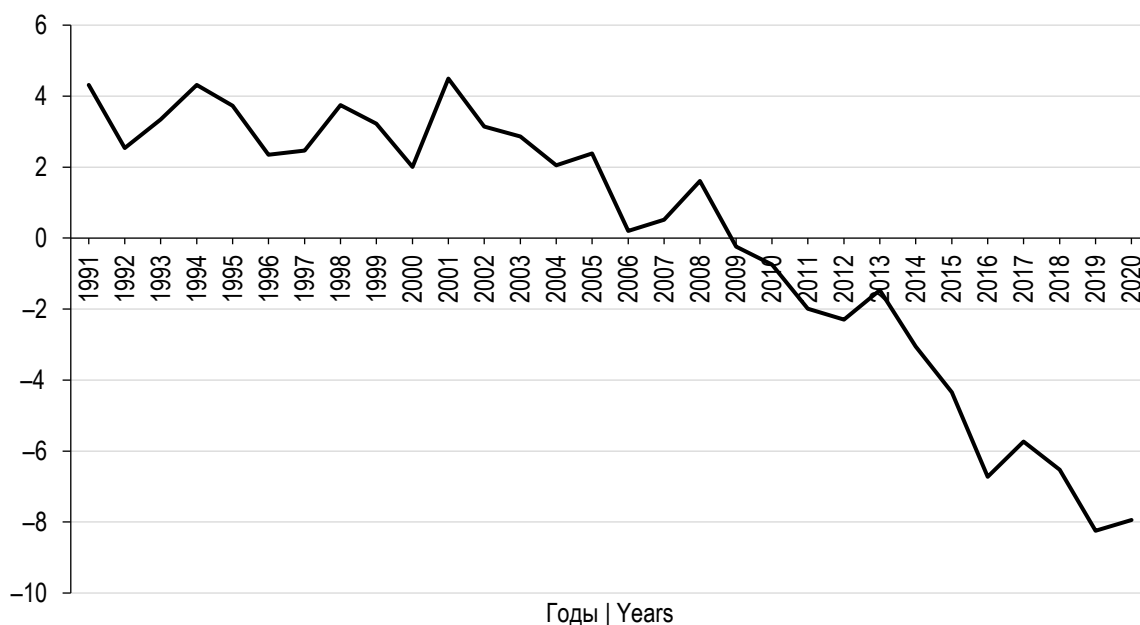


Рис. 6. Значения параметра k_t для базисного периода 1991–2020 гг.

Fig. 6. Values of the k_t parameter for the base period 1991–2020.

Таблица 4. Ошибки модели для каждой возрастной группы

Table 4. Model errors for each age group

Возрастные группы Age groups	Сумма квадратов остатков по возрастам Sum of squared residuals by age
0–4	0,64
5–9	1,61
10–14	0,91
15–19	0,72
20–24	0,71
25–29	1,07
30–34	0,94
35–39	0,45
40–44	0,22
45–49	0,19
50–54	0,35
55–59	1,63
60–64	0,21
65–69	0,32
70–74	0,15
75–79	0,17
80–84	0,16
85–89	0,38
90–94	0,33
95–99	2,05
100+	7,85

Точность подгонки построенной модели на выбранном периоде составила 87,21%.

Следующим этапом исследования является получение прогнозов показателей смертности. Прогнозирование, как правило, является основной целью моделирования показателей смертности. Главным преимуществом модели Ли–Картера является её простота для прогнозирования будущих значений показателей смертности и ожидаемой продолжительности жизни, поскольку значения коэффициентов a_x и b_x постоянны и не зависят от времени. Для этого необходимо построить экстраполяцию индекса смертности k_t . Прогнозы возрастных коэффициентов смертности выводятся из прогнозов k_t .

На практике для моделирования и прогнозирования коэффициента k_t используют интегрированные модели авторегрессии — скользящего среднего ARIMA (p, d, q), модель АРПСС, модель Бокса–Дженкинса.

Перед применением ARIMA-модели необходимо провести предварительный анализ временного ряда, определить присутствие трендов, цикличности, сезонных колебаний и прочих особенностей, которые могут повлиять на результаты моделирования. По результатам анализа сделан вывод о нестационарности исходного ряда, что подтверждается расширенным тестом Дики–Фуллера (ADF=1,699, $p=0,998$). Однако после применения операции взятия первой разности ряд стал стационарным, что также отражено в расширенном тесте Дики–Фуллера (ADF=-6,908817, $p < 0,0001$). Таким образом, определено, что параметр модели d должен быть равен 1, то есть динамика временного ряда может быть описана моделью ARIMA (p, l, q).

Для подбора параметров модели p и q произвели визуальный анализ корреляционной и частной корреляционной

Таблица 5. Ошибки модели для каждого наблюдаемого года**Table 5.** Model errors for each observed year

Годы Years	Сумма квадратов остатков по годам Sum of squared residuals by year
1991	2,20
1992	2,38
1993	0,38
1994	0,30
1995	0,31
1996	0,90
1997	1,11
1998	0,86
1999	0,33
2000	2,45
2001	0,43
2002	0,24
2003	0,39
2004	0,39
2005	0,47
2006	1,05
2007	0,40
2008	1,13
2009	0,26
2010	0,46
2011	0,57
2012	0,42
2013	0,76
2014	0,41
2015	0,24
2016	0,63
2017	0,30
2018	0,29
2019	0,35
2020	0,67

функции. Коэффициенты корреляционной и частной автокорреляционной функции близки к нулю, p -уровень Q -статистики для первых трёх коэффициентов автокорреляционной функции: 0,26; 0,51; 0,64. Для более точного определения p и q выполнили перебор по сетке, качество модели оценили с помощью информационных критериев Акаике, Шварца и Ханнана–Куина. Наилучшей моделью с минимальными значениями критериев оказалась

модель случайного блуждания ARIMA с параметрами $p=0$, $d=1$, $q=0$. В результате получена следующая модель случайного блуждания с дрейфом для описания индекса смертности:

$$k_t = k_{t-1} - 0,338 + \varepsilon_t,$$

где k_t — индекс смертности в году t , ε_t — случайные возмущения.

Параметр дрейфа $\theta = -0,338$ оказался значимым, его знак определяет убывание индекса смертности. Анализ остатков на основании критерия Льюнга–Бокса на уровне значимости 0,05 показал отсутствие автокорреляции ($p > 0,2$). Анализ автокорреляционной и частной автокорреляционной функций и тест на нормальность остатков подтвердили их соответствие процессу белого шума, то есть в остатках отсутствуют зависимости, модель пригодна для прогнозирования. Применение модели случайного блуждания с дрейфом для индекса смертности оказалось эффективным и показало высокий уровень достоверности, так как коэффициент детерминации $R^2 = 0,951$.

С использованием ARIMA-модели получили прогнозы индекса смертности (рис. 7).

Для дальнейшей оценки качества модели прогнозирования произвели сравнение прогнозируемых коэффициентов смертности с реальными данными. Элементы таблиц смертности найдены по формулам:

$$m_{x,t} = e^{b_x k_t + a_x},$$

где $t = 1999, 2020$, $x = 15-19, 20-24, \dots, 90-94$, a_x, b_x — постоянные во времени параметры модели Ли–Картера, $k_t = k_{t-1} - 0,338$. Средняя абсолютная процентная ошибка на базисном периоде ($t = 1999, 2020$): MAPE = 6,1%.

На основе полученных прогнозов индекса смертности k_t составили таблицы смертности мужского населения на 2022–2023 г.:

$$m_{x,t} = e^{b_x k_t + a_x},$$

где $t = 2022, 2023$, $x = 15-19, 20-24, \dots, 90-94$.

На прогнозируемом периоде также рассчитали метрику MAPE: MAPE(2022) = 5,6%, MAPE(2023) = 4,9%.

Таким образом, предложенная модель случайного блуждания с дрейфом пригодна для среднесрочного прогнозирования смертности мужского населения.

ОБСУЖДЕНИЕ

Пандемия COVID-19 значительно повлияла на жизнь людей по всему миру. К сожалению, многие страны столкнулись с увеличением числа смертей и болезней. Несмотря на эти трудности, стоит отметить положительные моменты. Прогнозы индекса смертности показывают, что в ближайшие годы смертность среди мужского населения в Оренбургской области продолжит падать, несмотря на неблагоприятное влияние пандемии. Это свидетельствует о росте уровня медицинского обслуживания в регионе, а также о высоком уровне медицинской подготовки и профессионализме медицинских работников.

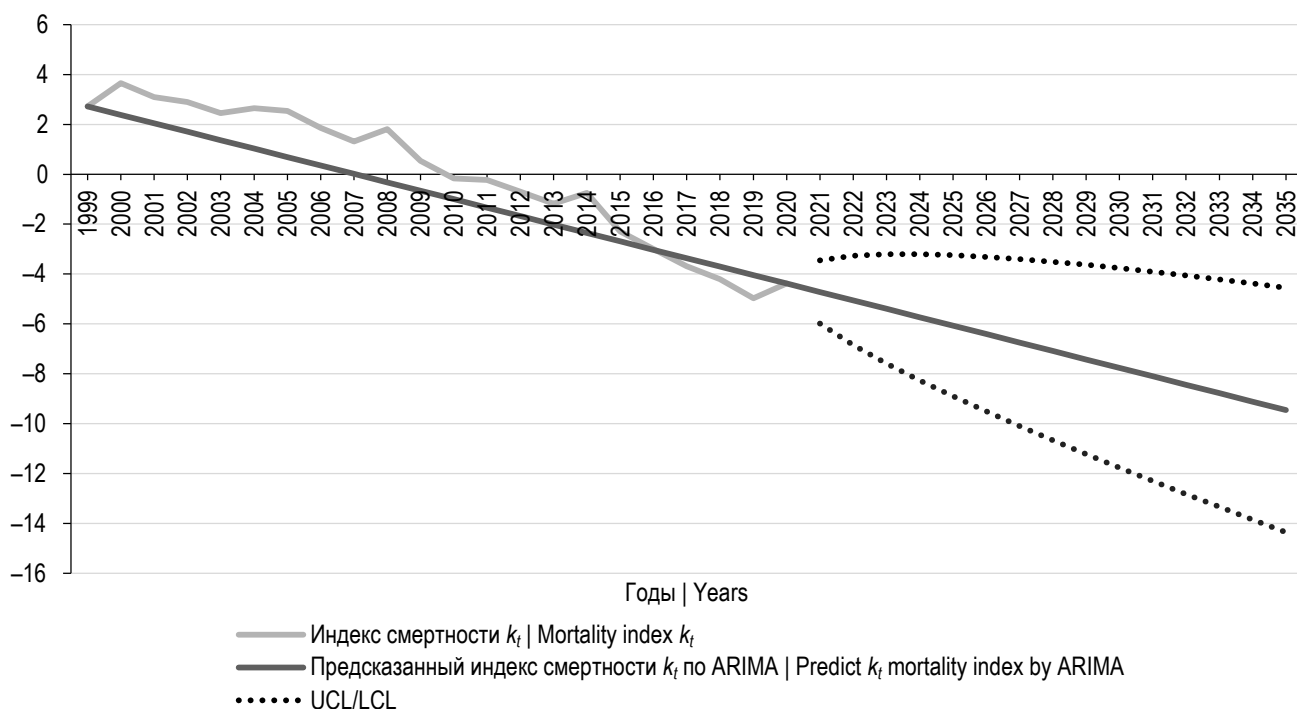


Рис. 7. Прогноз индекса смертности до 2035 г.
Fig. 7. Forecast of the mortality index until 2035.

С использованием прогнозов индекса смертности можно составить таблицы смертности на несколько ближайших лет и сделать вывод о том, какие группы населения могут испытывать определённые проблемы со здоровьем. Также можно провести анализ и прогнозирование для определённых возрастных групп, что даст возможность проводить более целенаправленную работу по предоставлению определённой категории населения качественной медицинской помощи.

В статье J. Cerda-Hernández и соавт. проиллюстрирована эффективность применения модели Ли–Катера при моделировании показателей смертности в Перу. В качестве панели данных были использованы показатели смертности населения за 14-летний период, составлен прогноз смертности и ожидаемая продолжительность жизни изучаемого населения, однако доверительные границы захватывали большой интервал значений, что, возможно, связано с вариабельностью прогноза [21].

В статье M. Belliard и соавт. описаны результаты прогноза ожидаемой продолжительности жизни к 2050 г. с помощью модели Ли–Катера. Модель имела высокий прогностический потенциал, но при увеличении зоны анализа прогностический уровень отличался от фактического [22].

В ходе исследования были выявлены серьёзные ограничения. При увеличении исследуемого диапазона данных прогностическая ценность модели снижалась. В связи с этим некоторые возрастные группы были исключены. Кроме того, особенностью полученной модели явилось то, что при высоком темпе прироста показателя

смертности в определённой возрастной группе также резко снижалась прогностическая ценность модели. Возможно, это связано с влиянием неизвестных факторов, которые оказывают воздействие на уровень смертности и на прогноз модели.

Кроме того, прогнозирование индекса смертности и составление на его основе таблиц смертности необходимо для определения средней ожидаемой продолжительности жизни в Оренбургской области. Будучи одним из наиболее важных показателей уровня жизни населения, средняя ожидаемая продолжительность жизни является важным критерием, для которого область может наметить различные социально-экономические цели и задачи. Таким образом, можно сделать вывод, что анализ показателей смертности имеет большое значение для определения приоритетов в работе медицинских учреждений, а также даёт возможность получить много полезной информации о здоровье и жизненном уровне населения области.

По результатам исследования может быть отмечено, что применение модели Ли–Катера для оценки данных по смертности не всегда является универсальным решением и может быть недостаточно точным. Так, модель Ли–Катера смогла описать только смертность мужского населения в Оренбургской области, но при этом описательная способность модели стала недостаточной для полного анализа данных. Возможно, проблема заключается в том, что базисный период, использованный для построения модели, был слишком коротким, поскольку метод изначально разработан и протестирован на данных о смертности населения США в течение столетия.

ЗАКЛЮЧЕНИЕ

Несмотря на вышеуказанные сложности, точность подгонки при моделировании смертности мужского населения Оренбургской области составила около 87%, что позволяет использовать этот результат для долгосрочного прогнозирования.

Наиболее подходящей ARIMA-моделью, описывающей полученный временной ряд — индекс смертности, оказалась модель случайного блуждания с дрейфом, что согласуется с результатами других авторов.

Исследование показало, что пандемия не оказывает долгосрочного влияния на смертность, поскольку, согласно полученным прогнозам, смертность мужского населения в Оренбургской области продолжит снижаться в ближайшие годы.

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Вклад авторов: Е.Л. Борщук — существенный вклад в концепцию и дизайн исследования, утверждение окончательного варианта статьи; Д.Н. Бегун — существенный вклад в концепцию исследования, утверждение окончательного варианта статьи; И.П. Болодурин — разработка концепции и дизайна исследования, подготовка окончательного варианта статьи; Л.И. Меньшикова — существенный вклад в концепцию исследования, окончательное утверждение присланной в редакцию рукописи; С.В. Колесник — обработка и интерпретация данных, моделирование, подготовка первичного варианта статьи;

СПИСОК ЛИТЕРАТУРЫ

1. Землянова Е.В. Отношение населения и экспертного сообщества к мерам по снижению смертности населения России (обзор литературы) // Социальные аспекты здоровья населения. 2013. № 3. С. 3. EDN: QINGER
2. Болотова Е.В., Самородская И.В., Дудникова А.В. Структура смертности и потерянных лет потенциальной жизни от болезней, ассоциированных с органами дыхания, населения экономически активного возраста (15–72 лет) Российской Федерации в 2019 г. // Врач. 2021. Т. 32, № 11. С. 5–10. EDN: EHMGS doi: 10.29296/25877305-2021-11-01
3. Козлова О.А., Макарова М.Н., Тухтарова Е.Х., Беленкова Т.В. Условия труда как фактор влияния на показатели смертности населения в трудоспособном возрасте // Фундаментальные исследования. 2015. № 7-1. С. 161–165. EDN: UDXTXB
4. Низамов И.Г., Садыкова Т.И. Здоровье экономически активного населения как база социально-экономического развития России и ее регионов // Общественное здоровье и здравоохранение. 2012. № 4. С. 28–33. EDN: PMSMIT
5. Родионова Л.А., Копнова Е.Д. Статистические подходы к анализу и моделированию сезонности в демографических данных // Демографическое обозрение. 2019. Т. 6, № 2. С. 104–141. EDN: FJVXKS
6. Денисенко М.Б., Варшавская Е.Я. Продолжительность трудовой жизни в России // Экономический журнал высшей школы экономики. 2017. Т. 21, № 4. С. 592–622. EDN: YLBIWS
7. Иванова А.Е., Кондракова Э.В. Обоснование прогноза продолжительности жизни населения в регионах России до 2025 г. // Социальные аспекты здоровья населения. 2008. № 1. С. 8. EDN: JWWSNB
8. Лазарев А.В., Калининская А.А., Васильева Т.П. Организационные резервы сбережения здоровья населения от болезней системы кровообращения // Проблемы социальной гигиены, здравоохранения и истории медицины. 2020. Т. 28, № 5. С. 762–765. EDN: DGCWUU doi: 10.32687/0869-866X-2020-28-5-762-765
9. Темешова Н.В. Реализация в 2016–2020 годах концепции демографической политики Российской Федерации на период до 2025 года. В кн.: Государственное и муниципальное управление в Камчатском крае: проблемы и пути решения. Петропавловск-Камчатский, 2018. С. 70–78. EDN: YTTQTJ
10. Соян Ш.Ч. Продолжительность жизни как индикатор качества жизни населения // Природные ресурсы, среда и общество. 2021. № 3. С. 42–46. EDN: KXIIUS doi: 10.24412/2658-4441-2021-3-42-46
11. Лакман И.А., Аскарлов Р.А., Прудников В.Б., и др. Прогнозирование смертности по причинам в Республике Башкортостан на основе модели Ли–Картера // Проблемы прогнозирования. 2021. № 5. С. 124–138. EDN: ISEHRS doi: 10.47711/0868-6351-188-124-138

А.Н. Дуйсембаева — сбор статистических данных, подготовка первичного варианта статьи. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией).

Финансирование. Авторы заявляют об отсутствии внешнего финансирования при проведении исследования.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

ADDITIONAL INFORMATION

Authors' contribution: E.L. Borshchuk — a significant contribution to the concept and design of the study, approval of the final version of the article; D.N. Begun — a significant contribution to the concept of the study, approval of the final version of the article; I.P. Bolodurina — development of the concept and design of the study, preparation of the final version of the article; L.I. Menshikova — a significant contribution to the concept of the study, final approval of the manuscript; S.V. Kolesnik — data processing and interpretation, modeling, preparation of the first draft; A.N. Duisembayeva — collection of statistical data, preparation of the first draft. All authors confirm that their authorship meets the international ICMJE criteria (all authors have made a significant contribution to the development of the concept, research and preparation of the article, read and approved the final version before publication).

Funding sources. No external funding.

Competing interests. The authors declare no conflicts of interest.

12. Booth H. Demographic forecasting: 1980 to 2005 in review // *International Journal of Forecasting*. 2006. Vol. 22, N 3. P. 547–581. doi: 10.1016/j.ijforecast.2006.04.001
13. Cancho-Candela R., Llano J.M.A., Ardura-Fernández J. Decline and loss of birth seasonality in Spain: analysis of 33 421 731 births over 60 years // *Journal of Epidemiology & Community Health*. 2007. Vol. 61, N 8. P. 713–718. doi: 10.1136/jech.2006.050211
14. Lee R.D., Carter L.R. Modeling and forecasting U.S. mortality // *Journal of the American Statistical Association*. 1992. Vol. 87, N 419. P. 659–671. doi: 10.1080/01621459.1992.10475265
15. Haberman S., Russolillo M. Lee–Carter mortality forecasting: application to the Italian population. 2005. Actuarial Research Paper No. 167.
16. Хубаев Г.Н. Регрессионные модели для прогнозирования продолжительности жизни населения административно-территориальных образований: построение и оценка качества // *Бюллетень науки и практики*. 2018. Т. 4, № 9. С. 206–217. EDN: XZBZZR doi: 10.5281/zenodo.1418761
17. Супранюк С.Б., Топрак С. Прогноз смертности от сердечно-сосудистых, цереброваскулярных заболеваний жителей трудоспособного возраста региона. В кн.: *Актуальные вопросы научного знания*. Курган, 2020. С. 255–261. EDN: DHZBVD
18. Эделева А.Н., Стародубов В.И., Федоткин М.А., и др. Математическое моделирование в управлении эффективностью работы медицинских организаций, оказывающих помощь лицам пожилого возраста в стационарных условиях // *Социальные аспекты здоровья населения*. 2018. № 2. С. 2. EDN: XUFBWH
19. Buleshov M.A., Alipbekova S.N., Tuktibayeva S.A. Modern medical and social aspects of morbidity and mortality rates in the working age population // *Вестник Казахского национального медицинского университета*. 2021. № 1. С. 285–288. EDN: TJDVBJ
20. Зайцева Н.В., Кирьянов Д.А., Камалдинов М.Р., и др. Анализ рисков потерь здоровья и комплексная оценка эффективности целевых мер территориальных систем здравоохранения по снижению смертности населения от сердечно-сосудистых и онкологических заболеваний // *Здравоохранение Российской Федерации*. 2021. Т. 65, № 4. С. 302–309. EDN: DVUUDG doi: 10.47470/0044-197X-2021-65-4-302-309
21. Cerda-Hernandez J., Sikov A. Lee–Carter method for forecasting mortality for Peruvian Population. 2018. Papers 1811.09622, arXiv.org.
22. Belliard M., Williams I. Proyección estocástica de la mortalidad. Una aplicación de Lee–Carter en la Argentina // *Revista Latinoamericana de Población*. 2013. Vol. 7, N 13. P. 129–148.

REFERENCES

1. Zemlyanova E.V. Attitude of population and experts to the measures for reduction of mortality in Russia (literature review). *Social Aspects of Population Health*. 2013;(3):3. EDN: QINGER
2. Bolotova E.V., Samorodskaya I.V., Dudnikova A.V. The structure of mortality and potential years life lost from respiratory diseases of the population of economically active age (15–72 years) of the Russian Federation in 2019. *Vrach (The Doctor)*. 2021;32(11):5–10. EDN: EHMGS doi: 10.29296/25877305-2021-11-01
3. Kozlova O.A., Makarova M.N., Tukhtarova E.H., Belenkova T.V. The labour conditions as a factor of the working age population mortality. *Fundamental Research*. 2015;(7-1):161–165. EDN: UDXTXB
4. Nizamov I.G., Sadykova T.I. Health of economically active population as a basis of social-and-economical development of Russia and its regions. *Public Health and Health Care*. 2012;(4):28–33. EDN: PMSMIT
5. Rodionova L.A., Kopnova E.D. Statistical approaches to analysis and modeling of seasonality in demographic data. *Demographic Review*. 2019;6(2):104–141. EDN: FJVXKS
6. Denisenko M.B., Varshavskaya E.Ya. Working life expectancy in Russia. *Higher School of Economics Economic Journal*. 2017;21(4):592–622. EDN: YLBIWS
7. Ivanova A.E., Kondrakova E.V. Grounds for life expectancy projections in Russian regions up to 2025. *Social Aspects of Population Health*. 2008;(1):8. EDN: JVWSNB
8. Lazarev A.V., Kalininskaya A.A., Vasilieva T.P. Organizational reserves saving the population's health from diseases of the circulatory system. *Problems of Social Hygiene, Public Health and History of Medicine*. 2020;28(S):762–765. EDN: DGCWUU doi: 10.32687/0869-866X-2020-28-s1-762-765
9. Temeshova N.V. Implementation of the concept of demographic policy of the Russian Federation for the period till 2025 in 2016–2020. In: *State and municipal administration in the Kamchatka territory: problems and solutions*. Petropavlovsk-Kamchatsky; 2018. P. 70–78. EDN: YTTQTJ
10. Soyán ShCh. Life expectancy as an indicator of the quality of life of the population. *Natural Resources, Environment and Society*. 2021;(3):42–46. EDN: KXIIUS doi: 10.24412/2658-4441-2021-3-42-46
11. Lakman I.A., Askarov R.A., Prudnikov V.B., et al. Predicting mortality by causes in the Republic of Bashkortostan using the Lee–Carter model. *Problemy Prognozirovaniya*. 2021;(5):124–138. EDN: ISEHRS doi: 10.47711/0868-6351-188-124-138
12. Booth H. Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 2006;22(3):547–581. doi: 10.1016/j.ijforecast.2006.04.001
13. Cancho-Candela R., Llano J.M.A., Ardura-Fernández J. Decline and loss of birth seasonality in Spain: analysis of 33,421,731 births over 60 years. *Journal of Epidemiology & Community Health*. 2007;61(8):713–718. doi: 10.1136/jech.2006.050211
14. Lee R.D., Carter L.R. Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*. 1992;87(419):659–671. doi: 10.1080/01621459.1992.10475265
15. Haberman S., Russolillo M. Lee–Carter mortality forecasting: application to the Italian population. 2005. Actuarial Research Paper No. 167.
16. Khubaev G. Regression models for forecasting life period of population of administrative-territorial education: construction and evaluation of quality. *Bulletin of Science and Practice*. 2018;4(9):206–217. EDN: XZBZZR doi: 10.5281/zenodo.1418761
17. Supraniuk S.B., Toprak S. Forecast of mortality from cardiovascular and cerebrovascular diseases of working-age residents of the region. In: *Current Issues of Scientific Knowledge*. Kurgan; 2020, P. 255–261. EDN: DHZBVD

18. Edeleva AN, Starodubov VI, Fedotkin MA, et al. Mathematical modeling in performance management of medical organizations providing inpatient care to the elderly. *Social Aspects of Population Health*. 2018;(2):2. EDN: XUFBWH
19. Buleshov MA, Alipbekova SN, Tuktibayeva SA. Modern medical and social aspects of morbidity and mortality rates in the working age population. *Bulletin of the Kazakh National Medical University*. 2021;(1):285–288. EDN: TJDVBJ
20. Zaitseva NV, Kiryanov DA, Kamaltdinov MR, et al. Health risks analysis and complex procedure for estimating the effectiveness of targeted activities performed within regional public healthcare systems and aimed at reducing mortality among the population caused by cardiovascular diseases and oncological diseases. *Health Care of the Russian Federation*. 2021;65(4):302–309. EDN: DVUUDG doi: 10.47470/0044-197X-2021-65-4-302-309
21. Cerda-Hernandez J, Sikov A. Lee-Carter method for forecasting mortality for Peruvian Population. 2018. Papers 1811.09622, arXiv.org.
22. Belliard M, Williams I. Proyección estocástica de la mortalidad. Una aplicación de Lee-Carter en la Argentina. *Revista Latinoamericana de Población*. 2013;7(13):129–148.

ОБ АВТОРАХ

Борщук Евгений Леонидович, д-р мед. наук, профессор;
ORCID: 0000-0002-3617-5908;
eLibrary SPIN: 9276-2040;
e-mail: be@orgma.ru

Бегун Дмитрий Николаевич, д-р мед. наук, профессор;
ORCID: 0000-0002-8920-6675;
eLibrary SPIN: 8443-4400;
e-mail: doctorbegun@yandex.ru

Болодурина Ирина Павловна, д-р техн. наук, профессор;
ORCID: 0000-0003-0096-2587;
eLibrary SPIN: 4848-0669;
e-mail: ipbolodurina@yandex.ru

Меньшикова Лариса Ивановна, д-р мед. наук, профессор;
ORCID: 0000-0002-3034-9014;
eLibrary SPIN: 9700-6736;
e-mail: mmm@gmail.com

Колесник Светлана Валентиновна;
ORCID: 0009-0009-3008-0308;
eLibrary SPIN: 7548-3688;
e-mail: svkolesnik_osu@mail.ru

***Дуйсембаева Айслу Нагашыбаевна**;
адрес: Россия, 460000, Оренбург, пр. Парковый, 7;
ORCID: 0000-0001-5762-4277;
eLibrary SPIN: 7164-7107;
e-mail: k.kro1@yandex.ru

AUTHORS' INFO

Evgeny L. Borshchuk, MD, Dr. Sci. (Medicine), Professor;
ORCID: 0000-0002-3617-5908;
eLibrary SPIN: 9276-2040;
e-mail: be@orgma.ru

Dmitry N. Begun, MD, Dr. Sci. (Medicine), Professor;
ORCID: 0000-0002-8920-6675;
eLibrary SPIN: 8443-4400;
e-mail: doctorbegun@yandex.ru

Irina P. Bolodurina, Dr. Sci. (Engineering), Professor;
ORCID: 0000-0003-0096-2587;
eLibrary SPIN: 4848-0669;
e-mail: ipbolodurina@yandex.ru

Larisa I. Menshikova, MD, Dr. Sci. (Medicine), Professor;
ORCID: 0000-0002-3034-9014;
eLibrary SPIN: 9700-6736;
e-mail: mmm@gmail.com

Svetlana V. Kolesnik;
ORCID: 0009-0009-3008-0308;
eLibrary SPIN: 7548-3688;
e-mail: svkolesnik_osu@mail.ru

***Aislu N. Duisembayeva**;
address: 7 Park Ave., Orenburg, 460000, Russia;
ORCID: 0000-0001-5762-4277;
eLibrary SPIN: 7164-7107;
e-mail: k.kro1@yandex.ru

*Автор, ответственный за переписку / Corresponding author