

УДК 004.378

Большие данные в сфере здравоохранения: влияние социально-экономических и других факторов на здоровье населения

Бисултанова Тамила Аслановна

Магистрант,
Институт математики, физики и информационных технологий,
Чеченский государственный университет им. А.А. Кадырова,
364093, Российская Федерация, Грозный, ул. Асланбека Шерипова, 32;
e-mail: mail@chesu.ru

Матыгов Мовсар Мусаевич

Ассистент,
Чеченский государственный университет им. А.А. Кадырова,
364093, Российская Федерация, Грозный, ул. Асланбека Шерипова, 32;
e-mail: Matygov.Movsar@gmail.com

Ибрагимова Марьям Адлановна

Магистрант,
Институт математики, физики и информационных технологий,
Чеченский государственный университет им. А.А. Кадырова,
364093, Российская Федерация, Грозный, ул. Асланбека Шерипова, 32;
e-mail: mail@chesu.ru

Аннотация

В данной статье рассматривается значимость использования больших данных в сфере здравоохранения с учетом факторов, влияющих на здоровье населения. Большие данные могут существенно улучшить качество медицинских услуг, предоставляя более точные и своевременные анализы здоровья. Кроме того, использование больших данных способствует выявлению тенденций в заболеваниях, что позволяет разрабатывать более эффективные стратегии профилактики и лечения. Растущее количество данных о показателях, влияющие на здоровье, и возможность связывать их с данными, полученными от отдельных индивидов, открывают потенциал для большого будущего в области здравоохранения. Расширение использования больших данных позволит лучше понимать факторы здоровья, добиться получения более сложных данных, которые с течением времени будут более точно отражать человека с социально-экономической, биологической, психологической точек зрения. Однако необходимо помнить о проблемах, связанных с конфиденциальностью данных. Надо приложить все усилия для защиты конфиденциальности, этического обращения с данными, а также о предоставлении равного доступа к медицинским услугам и ресурсам для всех членов общества.

Для цитирования в научных исследованиях

Бисултанова Т.А., Матыгов М.М., Ибрагимова М.А. Большие данные в сфере здравоохранения: влияние социально-экономических и других факторов на здоровье населения // Экономика: вчера, сегодня, завтра. 2024. Том 14. № 3А. С. 59-66.

Ключевые слова

Анализ больших данных, большие данные, социально-экономические факторы, факторы, здравоохранение.

Введение

Современная медицина находится на пороге новой эпохи. Большие данные изменили то, как люди управляют, анализируют и используют информацию в различных областях. Одной из наиболее заметных сфер, где анализ данных вносит большие изменения, является здравоохранение. Объем этих данных могут служить основой для планирования и оказания помощи в медицинских услугах. Использование анализа необходимо для экономии времени, денег и энергии. Целью анализа больших данных в сфере здравоохранения является извлечение ценной информации из большого объема данных для улучшения этой области, а также для принятия решений и предсказания дальнейших событий.

Большие данные в здравоохранении

Данные играют важную роль в современном обществе. Они проникли во все отрасли по всему миру, помогая разрабатывать новые технологии и инновации. Это особенно верно в среде бизнеса, где инициативы, основанные на данных, стали скорее требованием, чем выбором. То же самое происходит в отрасли здравоохранения.

Термин «Большие данные» употребляется для описания больших объемов информации, получаемых в результате внедрения технологий, которые применяются в здравоохранительных услугах. Количество информации будет представлять из себя состояния здоровья населения или конкретного человека. Под этими «данными» подразумевают медицинскую историю пациентов, результаты анализов, рентген и т.д.

Принцип больших данных основан на идее о том, что чем больше информации, тем больше понимания можно извлечь из этой информации и делать прогнозы относительно будущих событий [Эрл, 2018].

Исследования долгое время недостаточно обслуживались из-за бесконечно трудоемких вычислений требующие интеграции огромных объемов данных. Однако благодаря решениям для обработки больших данных теперь медицинские исследователи могут интегрировать огромные объемы информации автоматически, чтобы быстрее получать аналитические данные. Аналитика больших данных используется для снижения затрат на лечение, предотвращения заболеваний, прогнозирования вспышек эпидемий и улучшения качества жизни в целом [Амлаев и др., 2019].

Но не стоит забывать, что большой объем данных содержит чувствительную информацию о пациентах, включая медицинскую историю, генетическую информацию и личные данные. Необходимо обеспечить согласие пациентов на использование их персональных данных и соблюдать принципы конфиденциальности и анонимности при обработке [Погосян, 2019].

Сбор, хранение и обработка таких данных требуют строгих мер безопасности и защиты личной информации больного. Нарушение безопасности может привести к утечке.

Факторы, влияющие на здоровье населения

Растущее количество информации о показателях, влияющих на здоровье населения, и возможность связывания их с данными, полученные от отдельных лиц, показывают большой потенциал для улучшения здоровья граждан [Решетников, 2014]. Физическое благополучие человека и общества в целом зависит от множества взаимосвязанных факторов. Образ жизни человека и состояние здоровья определяется многими факторами, которые могут подвергаться различным изменениям. На здоровье населения влияют такие как социально-экономические, психологические, биологические. Они оказывают влияние в течение всей жизни человека.

На здоровье население могут влиять негативные социальные и экономические показатели. Люди, принадлежащие к низшим социальным классам, как минимум в два раза подвержены риску тяжелых заболеваний и ранней смертности по сравнению с представителями более высоких социальных классов [там же]. Большие расходы на хорошее медицинское обслуживание и препараты также могут привести к очень серьезным последствиям для них. Абсолютная бедность делает население крайне уязвимым к ряду заболеваний. Многие из них вызваны плохой гигиеной, а также нехваткой денег на лечение, загрязнением воздуха и т.д.

Все вышеперечисленные факторы являются отправной точкой для профилактики и укрепления здоровья населения. В связи с этим важно иметь представление, что они из себя олицетворяют.

Применение анализа данных

Сделаем небольшой анализ данных за счет информации из Росстата. Импортируем библиотеки для работы с информацией, построения будущих графиков и визуализации данных. Создаем два разных списка с данными и объединяем их по регионам как показано на рисунке 1.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

#Создание DataFrame для уровня безработицы
unemployment_data = {
    "Регион": ["Дагестан", "Ингушетия", "Кабардино-Балкария", "Карачаево-Черкесская",
              "Северная Осетия-Алания", "Чечня", "Ставропольский край"],
    "2020": [201.5, 77.7, 66.9, 30.1, 44.0, 120.8, 84.6],
    "2021": [206.4, 82.3, 52.8, 26.2, 42.1, 94.9, 72.2],
    "2022": [175.0, 81.3, 46.8, 21.6, 38.9, 72.0, 62.6]
}

unemployment_df = pd.DataFrame(unemployment_data)

#Создание DataFrame для численности населения с доходами ниже границы бедности/прожиточного минимума
poverty_data = {
    "Регион": ["Дагестан", "Ингушетия", "Кабардино-Балкария", "Карачаево-Черкесская",
              "Северная Осетия-Алания", "Чечня", "Ставропольский край"],
    "2020": [14.8, 29.4, 25.0, 23.4, 13.9, 20.4, 14.3],
    "2021": [14.9, 28.6, 18.8, 22.3, 13.4, 19.8, 13.5],
    "2022": [12.7, 26.8, 14.8, 21.5, 12.3, 19.1, 11.2]
}

poverty_df = pd.DataFrame(poverty_data)

#Объединение данных
merged_df = pd.merge(unemployment_df, poverty_df, on="Регион", suffixes=("_безработица", "_бедность"))
```

Рисунок 1 - Создание DataFrame

Здесь создаем графики прибегая к помощи библиотеки seaborn (рис.2.). В первом отображается уровень безработицы по регионам с 2020 по 2023, а во втором – численность населения с доходами ниже границы бедности. Оба графика находятся в одном общем окне (рис. 3).

```
#График уровня безработицы
plt.subplot(2, 1, 1)
sns.lineplot(data=merged_df.set_index("Регион")[[ "2020_безработицы",
| "2021_безработицы", "2022_безработицы"]], markers=True)
plt.title("Уровень безработицы по регионам (2020-2022)")
plt.ylabel("Уровень безработицы")
#График численности населения с доходами ниже границы бедности/прожиточного минимума
plt.subplot(2, 1, 2)
sns.lineplot(data=merged_df.set_index("Регион")[[ "2020_бедность",
| "2021_бедность", "2022_бедность"]], markers=True)
plt.title("Численность населения с доходами ниже границы бедности/прожиточного минимума (2020-2022)")
plt.xlabel("Год")
plt.ylabel("Уровень бедности")

plt.tight_layout()
plt.show()
#Вычисление корреляции между уровнем безработицы и уровнем бедности
correlation = merged_df[[ "2020_безработицы", "2021_безработицы", "2022_безработицы", "2020_бедность",
| "2021_бедность", "2022_бедность"]].corr()
```

Рисунок 2 - Создание Графика

Результат:

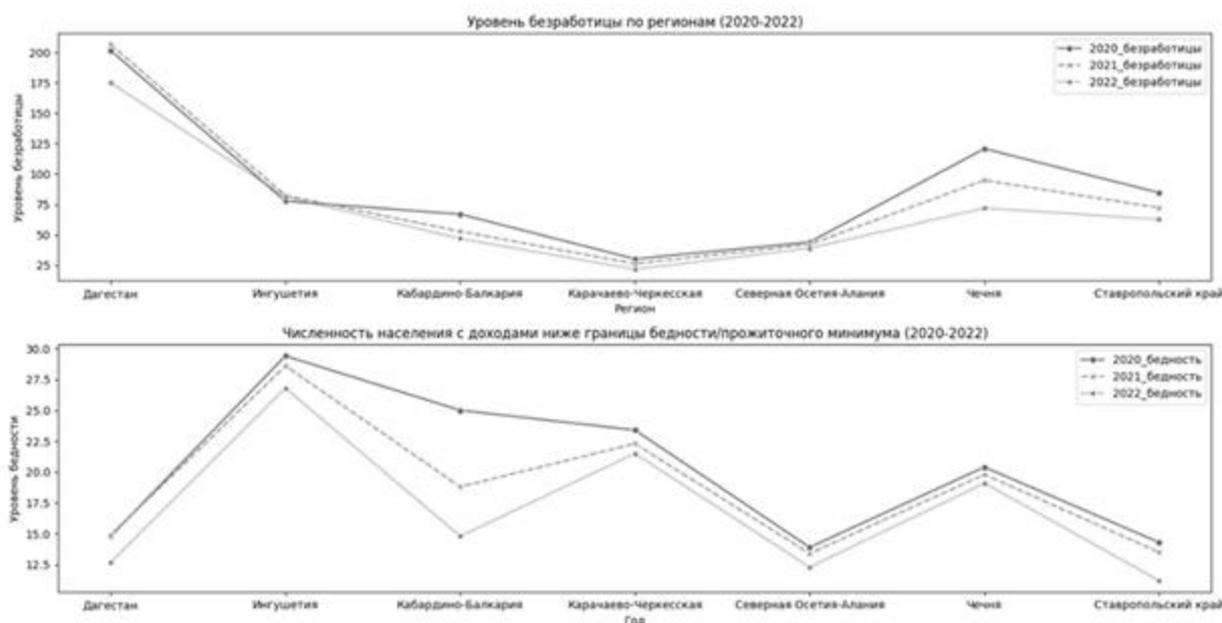


Рисунок 3 - Уровень безработицы и бедности СКФО

Этот блок кода выбирает столбцы из объединенного списка за три года. После «corr()» вычисляет корреляцию между всеми выбранными парами столбцов в «correlation» (рис.4.). Далее создаем тепловую карту (рис. 5).

Подводя итоги, можно сказать, что уровень безработицы и бедности Северо-Кавказского федерального округа с 2020 по 2022 г. снизился. Наибольший показатель безработицы характерен для республик Чечни и Дагестана, а наименьший – Ставропольского края и Карачаево-Черкесской республики. Наиболее высокий уровень бедности отмечается в

республиках Ингушетия и Дагестан, наименьший – Ставропольском крае. Из графика 2 видно, что есть положительная корреляция, что указывает на то, что уровень бедности и безработицы взаимосвязаны с друг другом.

```
#Вычисление корреляции между уровнем безработицы и уровнем бедности
correlation = merged_df[["2020_безработицы", "2021_безработицы", "2022_безработицы", "2020_бедность",
| | | | | | | | "2021_бедность", "2022_бедность"]].corr()

#Визуализация корреляции
plt.figure(figsize=(10, 8))
sns.heatmap(correlation, annot=True, cmap="coolwarm", fmt=".2f", linewidths=0.5)
plt.title("Корреляция между уровнем безработицы и уровнем бедности")
plt.tight_layout()
plt.show()
```

Рисунок 3 - Корреляция

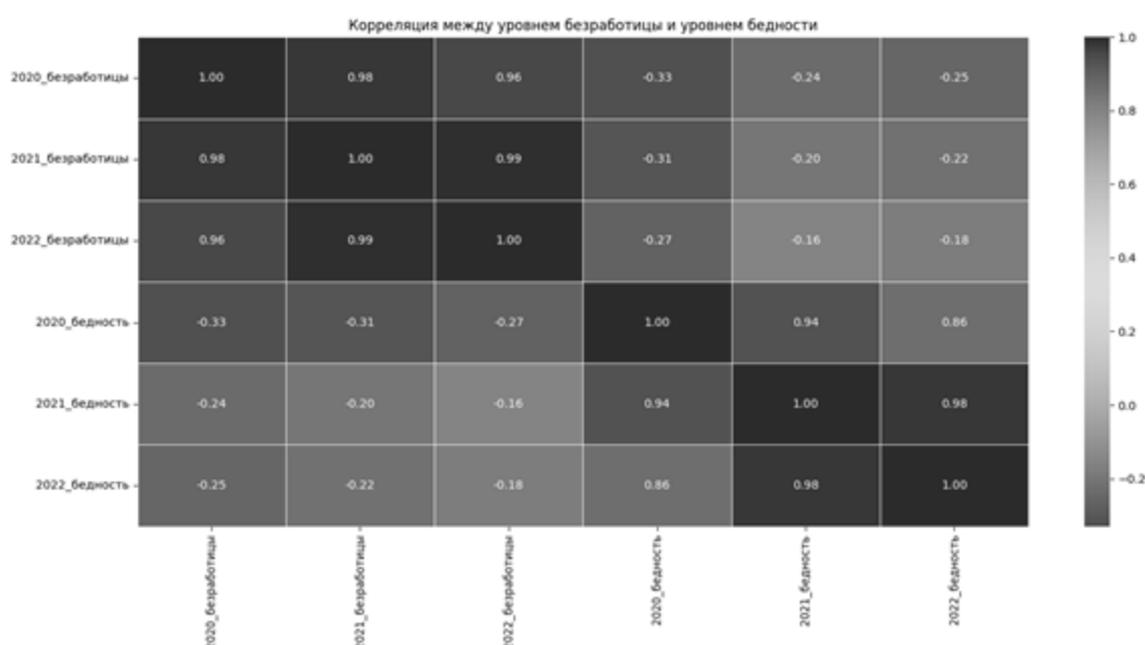


Рисунок 5 - Корреляция между уровнем безработицы и бедности

Данный анализ проводился как пример практической реализации больших данных в области здравоохранения с учетом социально-экономических факторов.

Заключение

Здоровье населения является самым важным фактором благополучия для каждой страны. Плохие показатели состояния могут оказывать пагубное влияние на само государство.

Большие данные играют немаловажную роль в современной медицине, предоставляя мощные инструменты для сбора, анализа и прогнозирования заболеваний. Объем этой огромной информации в здравоохранении отмечает необходимость учитывания социально-экономических и других факторов при анализе здоровья населения. Влияние таких факторов, как уровень дохода, образование, доступность медицинских услуг, стиль жизни и окружающая среда, бесспорно, влияет на здоровье и благополучие человечества и подчеркивают

необходимость их детального изучения. Растущее количество данных о показателях, влияющие на здоровье, и возможность связывать их с данными, полученными от отдельных индивидов, открывают потенциал для большого будущего в области здравоохранения. Расширение использования больших данных позволит лучше понимать факторы здоровья, добиться получения более сложных данных, которые с течением времени будут более точно отражать человека с социально-экономической, биологической, психологической точек зрения. Однако необходимо помнить о проблемах, связанных с конфиденциальностью данных. Надо приложить все усилия для защиты конфиденциальности, этического обращения с данными, а также о предоставлении равного доступа к медицинским услугам и ресурсам для всех членов общества.

Библиография

1. Алаудинов Б.Р. Что такое метавселенная и NFT // Тенденции развития науки и образования. 2022. № 86-1. С. 17-19.
2. Алаудинов Б.Р. Что такое метавселенная и какое у нее будущее? // Тенденции развития науки и образования. 2022. № 92-10. С. 26-28.
3. Амлаев К.Р. и др. Общественное здоровье и здравоохранение. М.: ГЭОТАР-Медиа, 2019. 560 с.
4. Джабраилов З.А. VR/AR-технологии // Мир в эпоху глобализации экономики и правовой сферы: роль биотехнологий и цифровых технологий. М.: КОНВЕРТ, 2021. С. 257-258.
5. Погосян И.А. Преимущества применения технологии Big Data в медицине. Образование и наука в современных реалиях. Чебоксары, 2019. С. 255-257.
6. Решетников А.В. Социология медицины: руководство. М.: ГЭОТАР-Медиа, 2014. 864 с.
7. Эрл Т., Хаттак В., Булер П. Основы Big Data. Концепции, алгоритмы и технологии. Днепр: Баланс Бизнес Букс, 2018. 320 с.
8. Rawat R., Yadav R. Big data: Big data analysis, issues and challenges and technologies // IOP Conference Series: Materials Science and Engineering. – IOP Publishing, 2021. – Т. 1022. – №. 1. – С. 012014.
9. Tao D., Yang P., Feng H. Utilization of text mining as a big data analysis tool for food science and nutrition // Comprehensive reviews in food science and food safety. – 2020. – Т. 19. – №. 2. – С. 875-894.
10. Lv Z., Singh A. K. Big data analysis of internet of things system // ACM Transactions on Internet Technology. – 2021. – Т. 21. – №. 2. – С. 1-15.

Big data in healthcare: influence of socio-economic and other factors on population health

Tamila A. Bisultanova

Master's Student,
Institute of Mathematics, Physics and Information Technologies,
Chechen State University,
364049, 32, Sheripova str., Grozny, Russian Federation;
e-mail: mail@chesu.ru

Movsar M. Matygov

Assistant,
Chechen State University,
364049, 32, Sheripova str., Grozny, Russian Federation;
e-mail: Matygov.Movsar@gmail.com

Mar'yam A. Ibragimova

Master's Student,
Institute of Mathematics, Physics and Information Technologies,
Chechen State University,
364049, 32, Sheripova str., Grozny, Russian Federation;
e-mail: mail@chesu.ru

Abstract

This article discusses the significance of using big data in healthcare, considering factors affecting public health. Big data plays an important role in modern medicine, providing powerful tools for collecting, analyzing and predicting diseases. The volume of this wealth of information in health care highlights the need to consider socioeconomic and other factors when analyzing population health. The influence of factors such as income level, education, access to health care services, lifestyle and environment undoubtedly influence the health and well-being of humanity and highlight the need for detailed study. Big data can significantly improve the quality of healthcare services by providing more accurate and timely health analyses. In addition, the use of big data helps identify trends in diseases, allowing for the development of more effective prevention and treatment strategies. The growing amount of data on health outcomes and the ability to link it with data from individuals holds the potential for a great future in health care. Expanding the use of big data will allow us to better understand health factors and obtain more complex data that, over time, will more accurately reflect a person from a socioeconomic, biological, and psychological point of view. However, you must be aware of data privacy concerns. Every effort must be made to protect privacy, ethical data management, and provide equal access to health services and resources for all members of society.

For citation

Bisultanova T.A., Matygov M.M., Ibragimova M.A. (2024) Bol'shie dannye v sfere zdavookhraneniya: vliyanie sotsial'no-ekonomiceskikh i drugikh faktorov na zdorov'e naseleniya [Big data in healthcare: influence of socio-economic and other factors on population health]. *Ekonomika: vchera, segodnya, zavtra* [Economics: Yesterday, Today and Tomorrow], 14 (3A), pp. 59-66.

Keywords

Big data analysis, big data, socio-economic factors, factors, healthcare.

References

1. Alaudinov B.R. (2022) Chto takoe metavselennaya i kakoe u nee budushchee? [What is the metaverse and what is its future?]. *Tendentsii razvitiya nauki i obrazovaniya* [Trends in the development of science and education], 92-10, pp. 26-28.
2. Alaudinov B.R. (2022) Chto takoe metavselennaya i NFT [What is the metaverse and NFT]. *Tendentsii razvitiya nauki i obrazovaniya* [Trends in the development of science and education], 86-1, pp. 17-19.
3. Amlaev K.R. et al. (2019) *Obshchestvennoe zdorov'e i zdavookhranenie* [Public health and healthcare]. Moscow: GEOTAR-Media Publ.
4. Dzhabrailov Z.A. (2021) VR/Ar-tekhnologii [VR/Ar technologies]. In: *Mir v epokhu globalizatsii ekonomiki i pravovoi sfery: rol' biotekhnologii i tsifrovyykh tekhnologii* [The world in the era of globalization of the economy and the legal sphere: the role of biotechnologies and digital technologies]. Moscow: KONVERT Publ.

5. Erl T., Khattak W., Buhler P. (2018) *Osnovy Big Data. Kontseptsii, algoritmy i tekhnologii* [Big Data Fundamentals. Concepts, Drivers & Techniques]. Dnepr: Balans Biznes Buks Publ.
6. Pogosyan I.A. (2019) *Preimushchestva primeneniya tekhnologii Big Data v meditsine. Obrazovanie i nauka v sovremennykh realiyakh* [Advantages of using Big Data technology in medicine. Education and science in modern realities]. Cheboksary.
7. Reshetnikov A.V. (2014) *Sotsiologiya meditsiny: rukovodstvo* [Sociology of medicine: a guide]. Moscow: GEOTAR-Media Publ.
8. Rawat, R., & Yadav, R. (2021). Big data: Big data analysis, issues and challenges and technologies. In IOP Conference Series: Materials Science and Engineering (Vol. 1022, No. 1, p. 012014). IOP Publishing.
9. Tao, D., Yang, P., & Feng, H. (2020). Utilization of text mining as a big data analysis tool for food science and nutrition. *Comprehensive reviews in food science and food safety*, 19(2), 875-894.
10. Lv, Z., & Singh, A. K. (2021). Big data analysis of internet of things system. *ACM Transactions on Internet Technology*, 21(2), 1-15.