



Разработка методов управления потребностью в медицинской помощи на региональном уровне с использованием демографических, социальных и экономических детерминант

**(Отчет за 2021-2022 гг.
Продление заявки на 2022-2023 гг.)**

**Лаборатория № 81
«Управление общественным здоровьем»**



Руководитель МНШ –
д.м.н. Черкасов С.Н., г.н.с., 81 лаб.

Публикации за 2019-2021 гг.

- WOS, Scopus – 14
- ВАК – 26

ПОДГОТОВЛЕНО УЧЕНИКОВ ЗА 2019-2021 ГГ.

- Докторов наук – 1
- Кандидатов наук – 4

Состав МНШ

- **Федяева А.В.**
- **Олейникова В.С.**
- **Широкий А.А.**
- **Мартин (Шевченко) М.А.**
- **Исянова А.М. (с 1.04.2022)**
- **Лобанов А.В. (с 1.04.2022)**

Публикации участников МНШ за отчетный период

Статус журнала	Заявлено на отчетный период	Опубликовано за отчетный период
WoS	1	2
Scopus	3	4
RSCI , ВАК, РИНЦ	6	9
ИТОГО	10	15



Федяева Анна Владимировна, 1989 г.р.

- Старший научный сотрудник лаб.81
- Кандидат медицинских наук
- Выпускница РГМУ им. Н.И. Пирогова
- Окончила клиническую ординатуру по специальности «Общественное здоровье и здравоохранение»
- Эксперт ВОЗ по семейству международных классификаций

Публикации Федяевой А.В. за отчетный период

1. Subjective and objective characteristics of physical activity in older adults / Sergey Cherkasov, Marina Shapovalova, Yuriy Boyko, Anna Fedyaeva, Alexander Shirokiy, Oleg Polozkov, Fedor Orlov // Archiv euromedica.–2021.-vol.11.-num.5.-p.62-65. DOI: 10.35630/2199-885X/2021/11/5.17 **(WOS)**
2. The influence of social determinants on the need for medical care in women / Sergey Cherkasov, Marina Shapovalova, Yuriy Boyko, Anna Fedyaeva, Dmitry Meshkov, Alexander Shiroki, Oleg Polozkov, Fedor Orlov // Archiv euromedica.–2021.-vol.11.-num.4.-p.31-33. DOI: 10.35630/2199-885X/2021/11/4.7 **(WOS)**
3. Организационные технологии мониторинга здоровья населения в российской федерации С.Н. Черкасов, А.В. Федяева, Д.О. Мешков, П.Н. Золотарев, И.Н. Мороз Судебная медицина 2022г. **(Scopus)**
4. Анализ состояния здоровья лиц, занимающихся спортом в возрасте 40 лет и старше / Хабриев Р. У., Черкасов С. Н., Григорьев Г. Ю., Федяева А. В. // Проблемы социальной гигиены, здравоохранения и истории медицины. 2021. Т. 29. № 2. С. 227-231. DOI: 10.32687/0869-866X-2021-29-2-227-231 **(Scopus)**
5. Возрастные и гендерные особенности питания лиц, занимающихся спортом в возрасте 40 лет и старше / Хабриев Р.У., Черкасов С.Н., Григорьев Г.Ю., Федяева А.В. // Спортивная медицина: наука и практика. 2021. Т. 11. № 1. С. 53-58. DOI: 10.47529/2223-2524.2021.1.7 **(RSCI)**
6. Влияние образования на распространенность ответственного отношения к своему здоровью населения старших возрастных групп / Черкасов С.Н., Авсadjанишвили В.Н., Камаев О.Ю., Федяева А.В. // Бюллетень Национального научно-исследовательского института общественного здоровья имени Н.А. Семашко. 2021. № 4. С. 64-71. DOI: 10.25742/NRIPH.2021.04.008 **(BAK)**
7. Attitude to informing the doctor about the treatment process in different age and sex groups of patients / Gorbunov A.L., Cherkasov S.N., Fedyaeva A.V., Moroz I.N., Martirosov A.V., Suhov A.A. // Bulletin Biomedicine and sociology. 2022. Т. 7. № 1. С. 66-75. **(BAK)**
8. Влияние уровня образования на степень комплаентности населения старших возрастных групп / Черкасов С.Н., Полозков О.И., Федяева А.В., Авсadjанишвили В.Н. // Медико-фармацевтический журнал Пульс. 2021. Т. 23. № 7. С. 55-61. DOI: 10.26787/nydha-2686-6838-2021-23-7-55-61. **(BAK)**
9. Роль признания ответственности за свое здоровье как социальной детерминанты здоровья в старших возрастных группах / Черкасов С.Н., Федяева А.В., Авсadjанишвили В.Н., Горбунов А.Л. // Медико-фармацевтический журнал Пульс. 2021. Т. 23. № 11. С. 12-17. **(BAK)**
10. Анализ гендерных особенностей возрастной динамики потребности населения старших возрастных групп в медицинской помощи, оказываемой в стационарных условиях, по поводу болезней ротовой полости, слюнных желез и челюстей Мартиросов А.В., Черкасов С.Н., Федяева А.В. Медико-фармацевтический журнал Пульс. 2022. Т. 24. № 1. С. 43-49. **(BAK)**
11. Влияние уровня образования на модели поведения, связанные с обращаемостью за медицинской помощью населения старших возрастных групп / Черкасов С.Н., Полозков О.И., Федяева А.В., Камаев Ю.О. // Медико-фармацевтический журнал Пульс. 2021. Т. 23. № 7. С. 31-37. DOI: 10.26787/nydha-2686-6838-2021-23-7-31-37. **(BAK)**
12. Базовая модель для сравнительной оценки результативности методов прогнозирования потребности в медицинской помощи на примере сердечно-сосудистых заболеваний / Олейникова В.С., Федяева А.В., Шевченко М.А., Земцов Д.В., Черкасов С.Н., Мешков Д.О. // Труды XVII Всероссийской школы-конференции молодых ученых «Управление большими системами. С.113-123. DOI: 10.25728/ubs.2021.010 **(BAK)**
13. Анализ возрастной динамики потребности в амбулаторной медицинской помощи по поводу состояний, определяющих базовую величину потребности у женщин при болезнях системы кровообращения Олейникова В.С., Черкасов С.Н., Федяева А.В. Наука молодых (Eruditio Juvenium). 2021. Т. 9. № 4. С. 543-552. **(BAK)**

В результате выполнения НИР «Изучение и мониторинг здоровья населения Российской Федерации» (закончена в 2019 году) Национального НИИ общественного здоровья имени Н.А.Семашко (руководитель темы академик Р.У.Хабриев) доказано, что:

1. Существующие алгоритмы анализа и прогнозирования не могут в полной мере дать комплексную оценку состояния здоровья населения и факторов его формирующих, а также определить потребность в объемах медицинской помощи
2. Отсутствуют методы управления потребностью в медицинской помощи

Факторы, влияющие на потребность в медицинской помощи



Перспективный подход к планированию



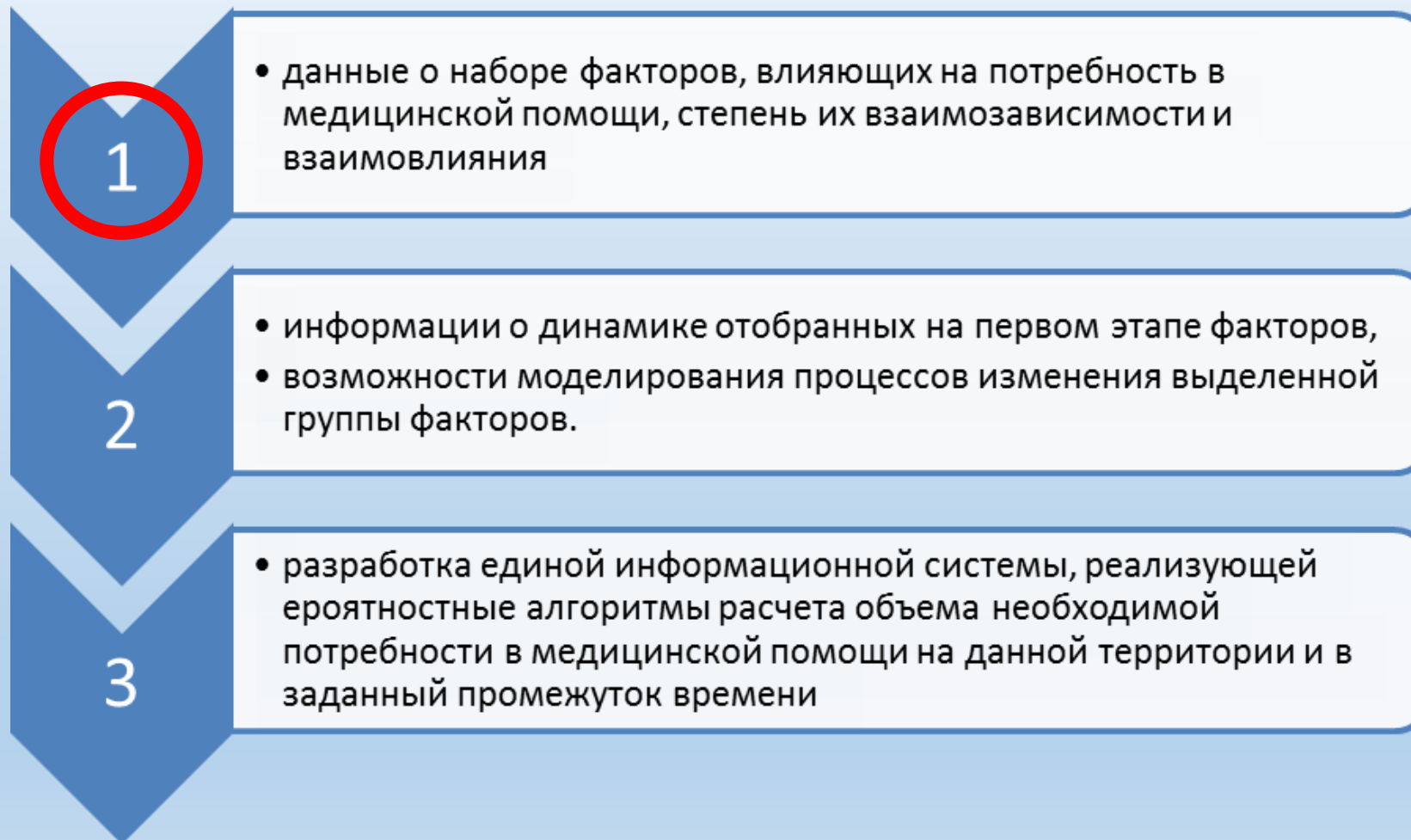
Цель проекта

Определение значимости демографических показателей и социально-экономических детерминант в процессах формирования потребности населения в медицинской помощи

Задачи проекта

1. Изучение зависимости обращаемости за медицинской помощью, оказываемой на амбулаторном и стационарном этапах, от демографических показателей у населения.
2. Исследование влияния социально-экономических детерминант здоровья и медицинской активности на потребность в медицинской помощи.
3. Прогноз изменений заболеваемости на среднесрочную перспективу как основа для моделирования потребности в медицинской помощи (2 год проекта).

Последовательность этапов разработки методики планирования объёмов медицинской помощи



Задача №1

Класс IX МКБ-10 2014-2016 (Болезни сердечно-сосудистой системы)

Рубрики, составляющие блок	Название блока
I00 – I02	Острая ревматическая лихорадка
I05-I09	Хронические ревматические болезни сердца
I10-I15	Гипертензивные болезни [Болезни, характеризующиеся повышенным артериальным давлением]
I20-I25	Ишемические болезни сердца
I26-I28	Легочное сердце и нарушения легочного кровообращения
I30-I52	Другие болезни сердца
I60-I69	Цереброваскулярные болезни [Болезни сосудов головного мозга]
I70-I79	Болезни артерий, артериол и капилляров
I80-I89	Болезни вен, лимфатических сосудов и лимфатических узлов, не классифицированные в других рубриках
I95-I99	Другие и неуточненные болезни системы кровообращения

Показатели, характеризующие потребность (на 1000 населения)

■ Частотные

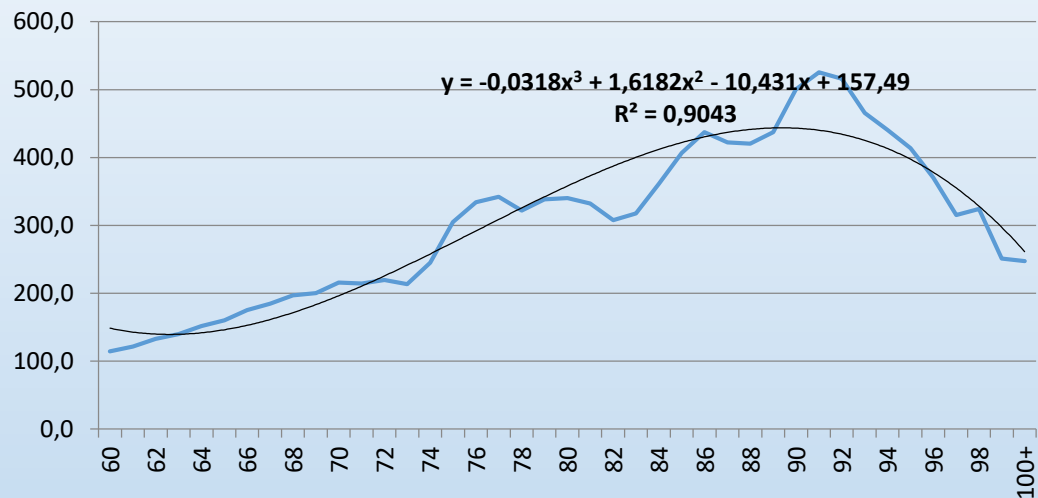
- Число посещений
- Число обращений
- Число госпитализаций (стационар круглосуточного пребывания)
- Число госпитализаций (стационар дневного пребывания)

■ Объёмные

- Длительность лечения в амбулаторных условиях
- Длительность лечения в условиях стационара круглосуточного пребывания
- Длительность лечения в условиях стационара дневного пребывания

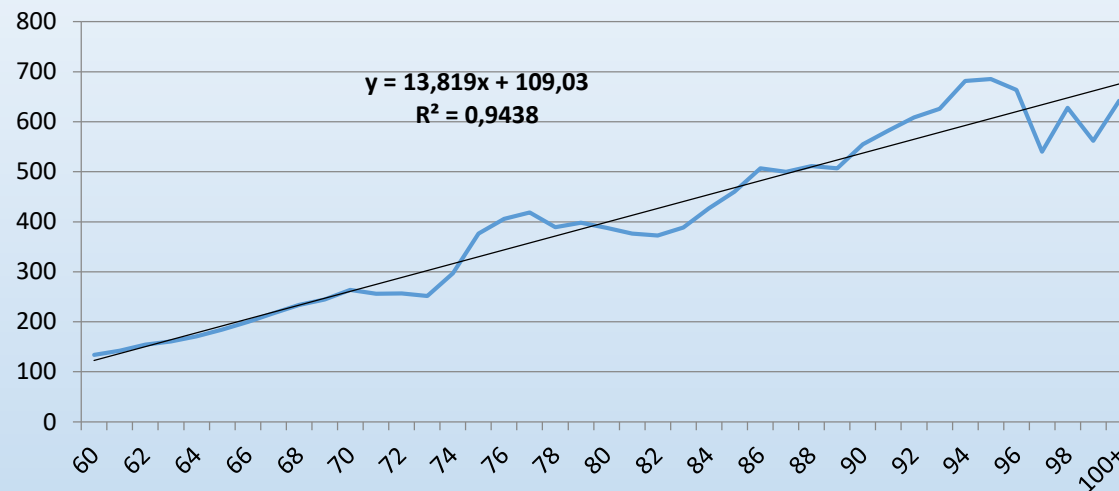
Отдельные результаты

мужчины



Интенсивность посещений по поводу причин, включенных в блок I60-I69 (Цереброваскулярные болезни [Болезни сосудов головного мозга]) у мужчин в различных возрастных группах (число посещений медицинских организаций, оказывающих медицинскую помощь в амбулаторных условиях в год на 1000 населения соответствующего возраста и пола)

женщины



Интенсивность посещений по поводу причин, включенных в блок I60-I69 (Цереброваскулярные болезни [Болезни сосудов головного мозга]) у женщин в различных возрастных группах (число посещений медицинских организаций, оказывающих медицинскую помощь в амбулаторных условиях в год на 1000 населения соответствующего возраста и пола)

Доля объясненных и случайных колебаний результирующей величины (потребность в медицинской помощи у мужчин) (на 1000 населения мужского пола)

Показатель	Доля объясненных колебаний объемов медицинской помощи (%)	Доля случайных колебаний объемов медицинской помощи (%)
Посещений	82	18
Обращений	83	17
Число госпитализаций в стационар круглосуточного пребывания	82	18
Число госпитализаций в стационар дневного пребывания	82	18
Длительность лечения в амбулаторных условиях	80	20
Длительность лечения в условиях круглосуточного стационара	84	16
Длительность лечения в условиях дневного стационара	85	15

Доля объясненных и случайных колебаний результирующей величины (потребность в медицинской помощи у женщин) (на 1000 населения женского пола)

Показатель	Доля объясненных колебаний объемов медицинской помощи (%)	Доля случайных колебаний объемов медицинской помощи (%)
Посещений	72	28
Обращений	71	29
Число госпитализаций в стационар круглосуточного пребывания	93	7
Число госпитализаций в стационар дневного пребывания	88	12
Длительность лечения в амбулаторных условиях	74	26
Длительность лечения в условиях круглосуточного стационара	94	6
Длительность лечения в условиях дневного стационара	86	14

Вывод

Полученные данные станут основой для создания практического инструмента планирования потребности в медицинской помощи на региональном уровне (2 год проекта).

Ожидаемые в конце реализации проекта научные результаты (2022-2023 гг.)

Федяева А.В.

- Построение модели потребности в медицинской помощи, оказываемой в амбулаторных и стационарных условиях у населения старших возрастных групп с онкологическими заболеваниями на территории субъекта Российской Федерации (технический прототип).

Шевченко М.А.

- Прогноз изменений заболеваемости на среднесрочную перспективу с использованием метода машинного обучения.

Исянова А.М.

- Алгоритмизация и практическая реализация методов моделирования потребности в медицинской помощи в виде компьютерных программ поддержки принятия управленческих решений на территориальном уровне.

Лобанов А.В.

- Идентификация мультиморбидности как фактора, определяющего обращаемость за медицинской помощью. Построение модели мультиморбидных состояний на примере сахарного диабета. Прогнозирование уровня потребности в медицинской помощи с учетом мультиморбидности.

Олейникова Валерия Сергеевна, 1988 г.р.

- Младший научный сотрудник лаб.81
- Выпускница РГМУ им. Н.И. Пирогова
- Окончила клиническую ординатуру по специальности «Общественное здоровье и здравоохранение»
- Эксперт ВОЗ по семейству международных классификаций

Публикации за отчетный период

1. Базовая модель для сравнительной оценки результативности методов прогнозирования потребности в медицинской помощи на примере сердечно-сосудистых заболеваний / Олейникова В.С., Федяева А.В., Шевченко М.А., Земцов Д.В., Черкасов С.Н., Мешков Д.О. // Труды XVII Всероссийской школы-конференции молодых ученых «Управление большими системами. С.113-123. DOI: 10.25728/ubs.2021.010 (ВАК)
2. Анализ возрастной динамики потребности в амбулаторной медицинской помощи по поводу состояний, определяющих базовую величину потребности у женщин при болезнях системы кровообращения Олейникова В.С., Черкасов С.Н., Федяева А.В. Наука молодых (Eruditio Juvenium). 2021. Т. 9. № 4. С. 543-552. (ВАК)
3. **Подготовлена к защите кандидатская диссертация (документы поданы в совет)**

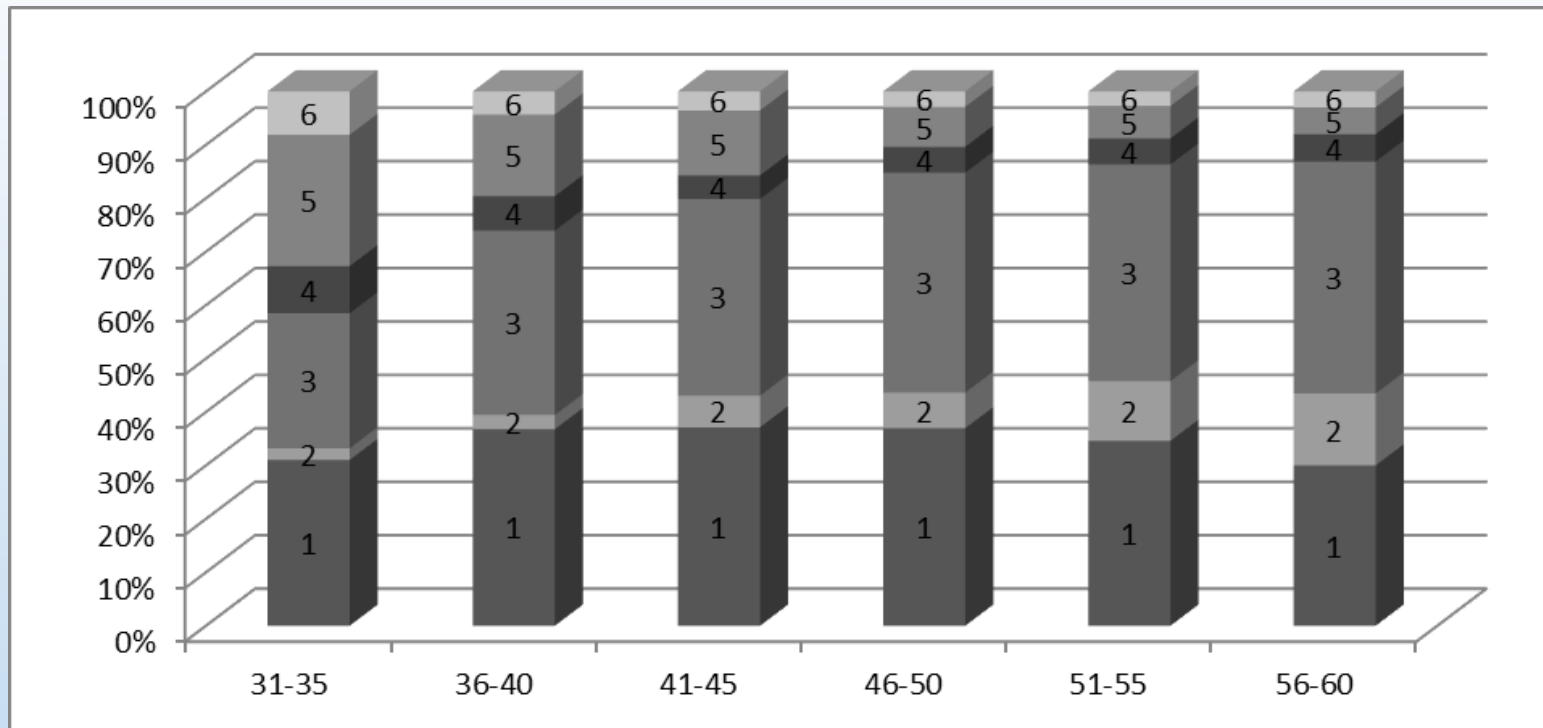
Задача № 2

Анализ возрастной динамики интенсивности потребления медицинской помощи, оказываемой в стационарных условиях, по причинам, включенным в Класс IX МКБ-10 («Болезни системы кровообращения») у населения трудоспособного возраста

- **Объект исследования:** медицинская помощь, оказываемая в стационарных условиях пациентам с болезнями системы кровообращения в возрасте 31-60 лет.

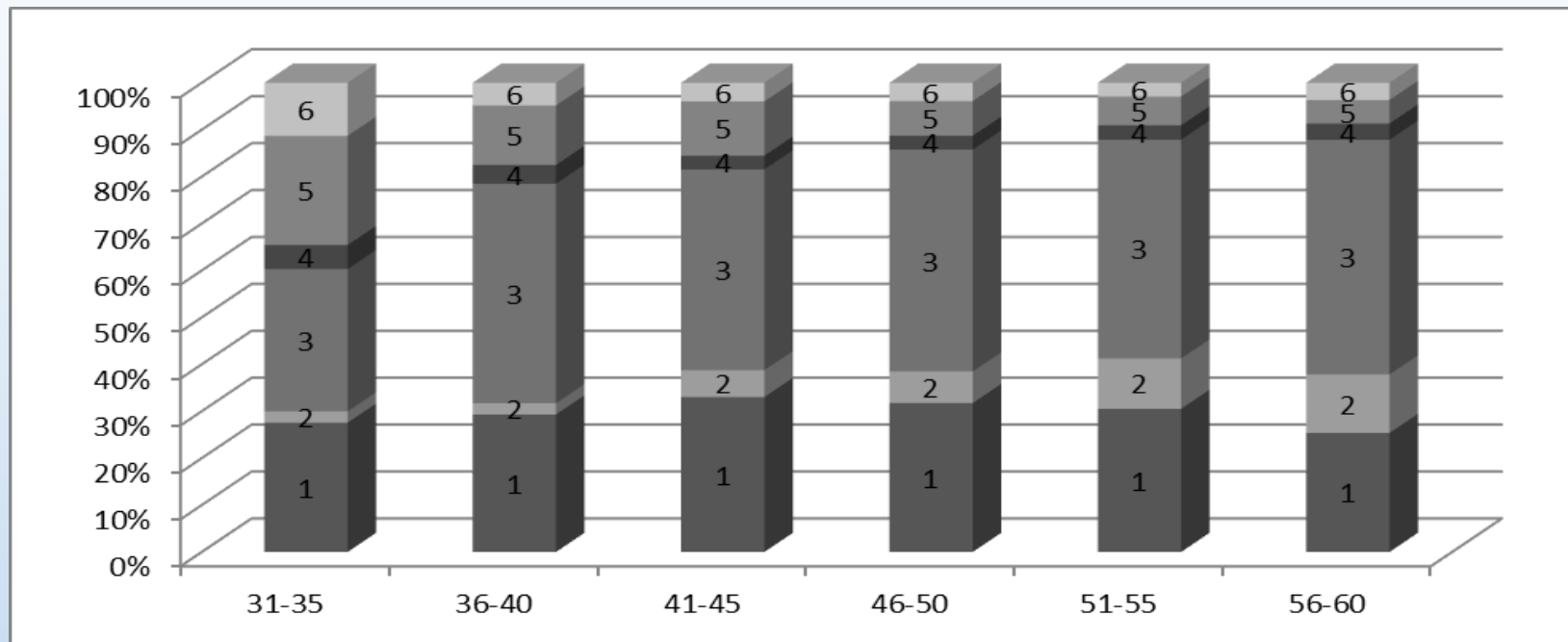
Частота обращений за стационарной медицинской помощью женщин и мужчин разных возрастных групп по причинам, включенным в Класс IX МКБ-10 («Болезни системы кровообращения») (на 1000 женщин и мужчин соответствующего возраста, обратившихся за стационарной медицинской помощью)

Возрастной интервал	31-35	36-40	41-45	46-50	51-55	56-60
Женщины	16,57	48,04	107,22	184,38	283,64	381,06
Мужчины	69,2	110,4	175,4	286,6	401,9	476,1
Отношение величины потребности (муж/жен)	4,17	2,29	1,63	1,55	1,41	1,24



Примечание – 1 – I10–I15 (Гипертензивные болезни [Болезни, характеризующиеся повышенным артериальным давлением]); 2 – I20–I25 (Ишемические болезни сердца); 3 – I60–I69 (Цереброваскулярные болезни [болезни сосудов головного мозга]); 4 – I44–I49 (Нарушение проводимости блок «Другие болезни сердца»); 5 – I80–I89 (Болезни вен, лимфатических сосудов и лимфатических узлов); 6 – остальные блоки, включенные в девятый Класс МКБ-10

Структура поводов обращений женщин за стационарной медицинской помощью по причинам из Класса IX МКБ-10 («Болезни системы кровообращения») (% от всех поводов обращений за стационарной медицинской помощью)



Примечание – 1 – I10–I15 (Гипертензивные болезни [Болезни, характеризующиеся повышенным артериальным давлением]); 2 – I20–I25 (Ишемические болезни сердца); 3 – I60–I69 (Цереброваскулярные болезни [болезни сосудов головного мозга]); 4 – I44–I49 (Нарушение проводимости блок «Другие болезни сердца»); 5 – I80–I89 (Болезни вен, лимфатических сосудов и лимфатических узлов); 6 – остальные блоки, включенные в девятый Класс МКБ-10

Удельный вес востребованных объемов МПОСУ по причинам из Класса IX МКБ-10 («Болезни системы кровообращения») (% от общего востребованного объема стационарной медицинской помощи)

Результаты анализа возрастной динамики интенсивности потребления МПОСУ у мужчин

Моделируемое состояние	Формула	$\sum \varepsilon^2$	$D_{\text{ост}}$	$\sigma_{\text{ост}}$	D_y	Доля влияния исследуемого признака (возраста)	Доля случайных колебаний результирующей величины
Блок I10–I15	$Y = -19,73 x^2 + 277,94 x - 158,02$	9501	1900	43,6	73244	0,986	0,014
Блок I60–I69	$Y = 93,005 x^{1,672}$	24534	4906	70,0	513569	0,995	0,005
Блок I20–I25	$Y = 68,204 x^{1,68}$	26668	5333	73,0	306717	0,991	0,009
Блок I80–I89	$Y = 0,5x^2 + 3,73 x + 6,92$	6,38	1,28	1,13	3,32	0,785	0,215

Примечание – X – номер возрастного интервала



**Мартин (Шевченко) Михаил Алексеевич,
1995 г.р.**

Инженер лаб.81
Магистрант МГТУ им. Н.Э. Баумана

Публикации

1. Базовая модель для сравнительной оценки результативности методов прогнозирования потребности в медицинской помощи на примере сердечно-сосудистых заболеваний / Олейникова В.С., Федяева А.В., Шевченко М.А., Земцов Д.В., Черкасов С.Н., Мешков Д.О. // Труды XVII Всероссийской школы-конференции молодых ученых «Управление большими системами. С.113-123. DOI: 10.25728/ubs.2021.010 **(ВАК)**
2. Математические подходы к отбору показателей потребности в медицинской помощи / Черкасов С.Н., Федяева А.В., Исянова А.М., Шевченко М.А., Лобанов А.В. // Вестник медицинского стоматологического института (принята к печати, **ВАК**)

Задача № 3

Исследование влияния социально-экономических детерминант здоровья и медицинской активности на потребность в медицинской помощи

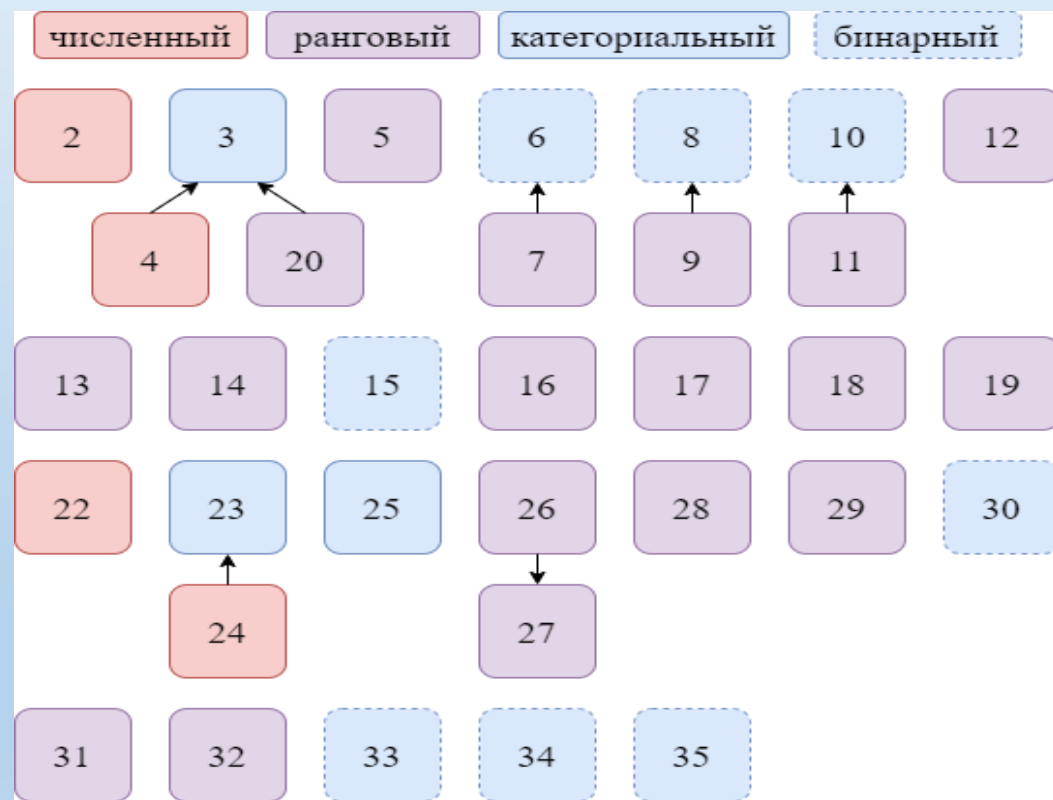
1 год: Поиск социально-экономических детерминант здоровья и медицинской активности на потребность в медицинской помощи с использованием метода машинного обучения.

Оценка значимости признаков на решение модели машинного обучения

- 1. Набор данных**
- 2. Модель машинного обучения**
- 3. Метод оценки значимости признаков**

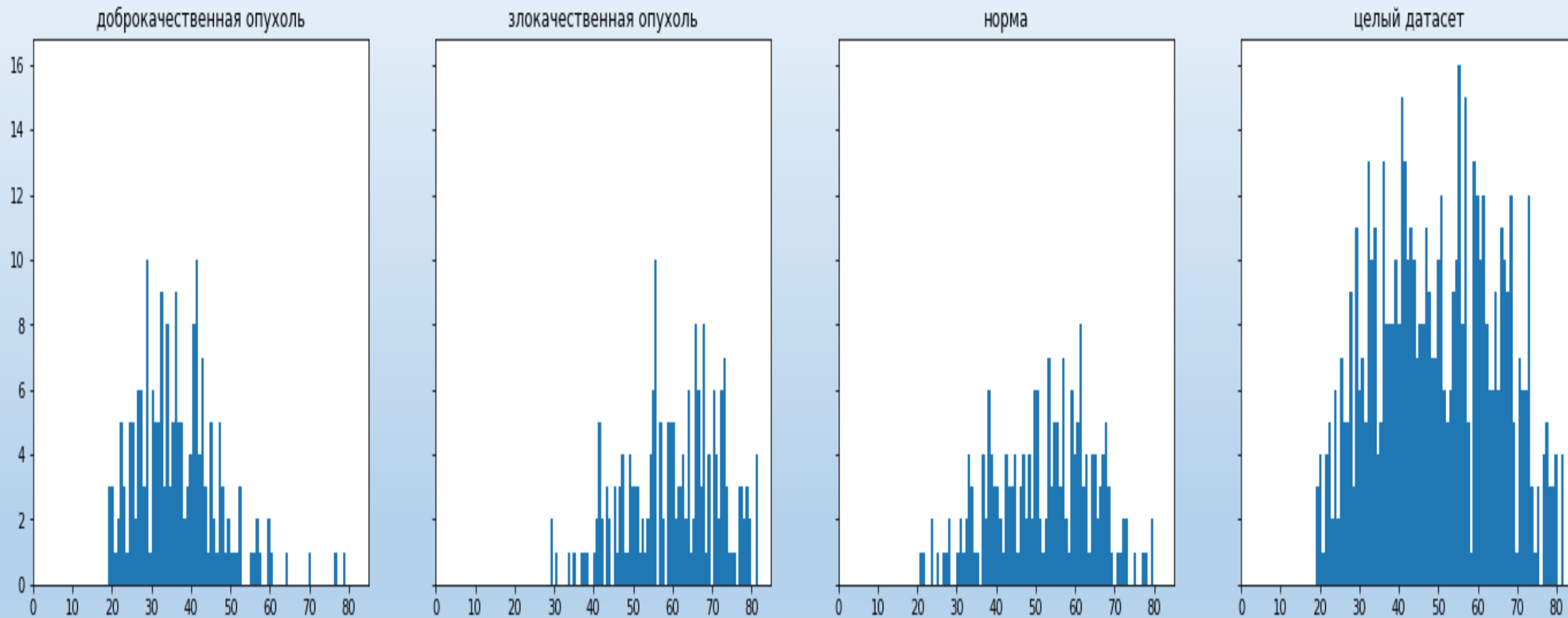
Набор данных

Признаки в анкете-опроснике по злокачественным новообразованиям молочной железы



Набор данных

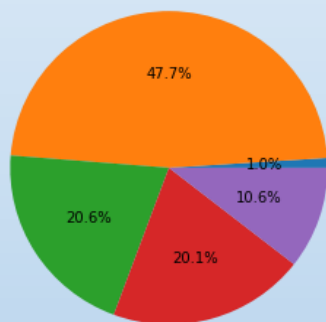
2. Возраст



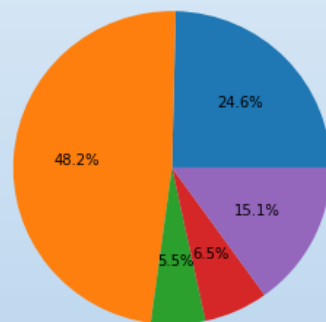
Набор данных

3. Семейное положение

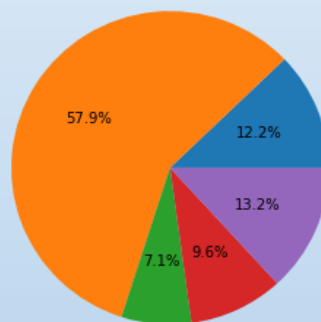
доброкачественная опухоль



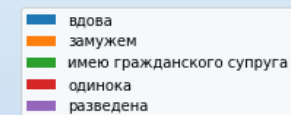
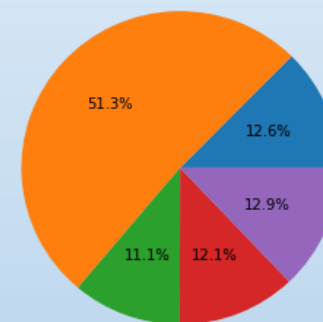
злокачественная опухоль



норма

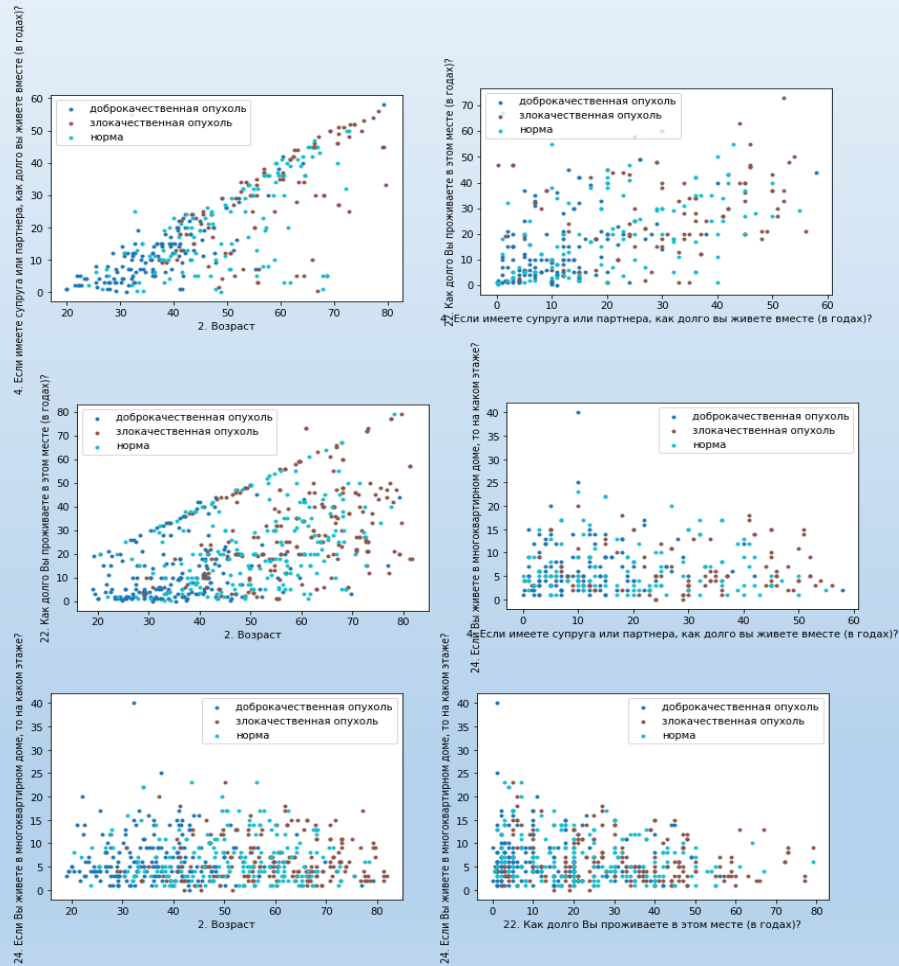


целый датасет

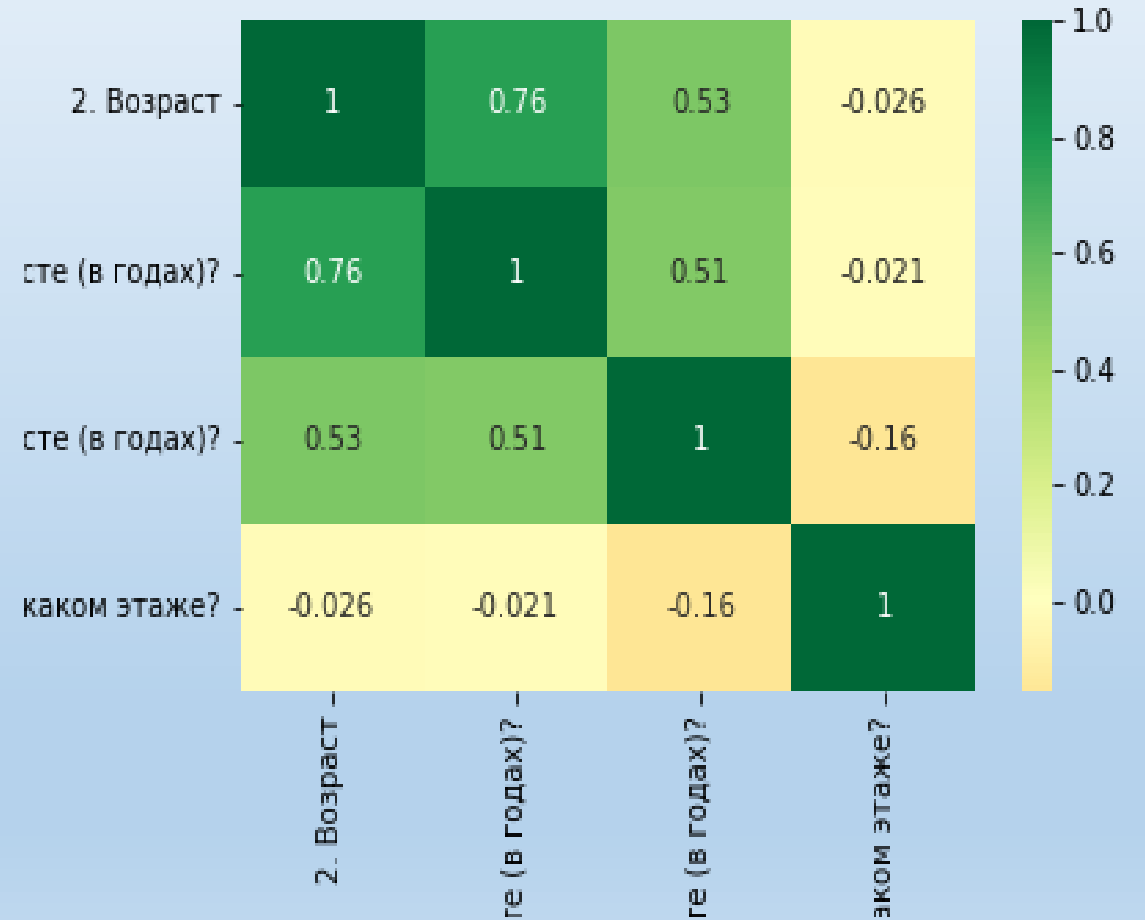


Набор данных

Диаграммы рассеяния



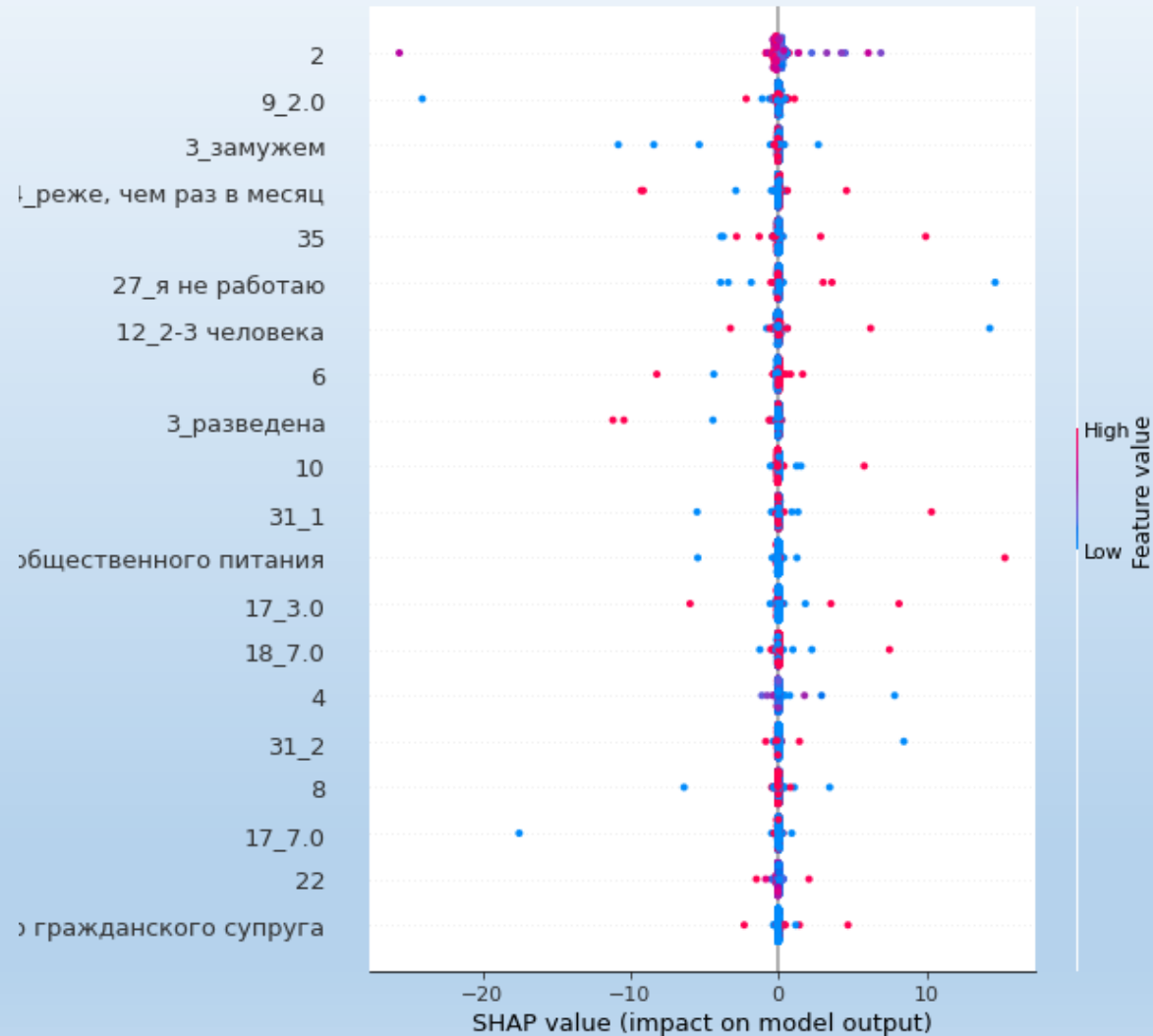
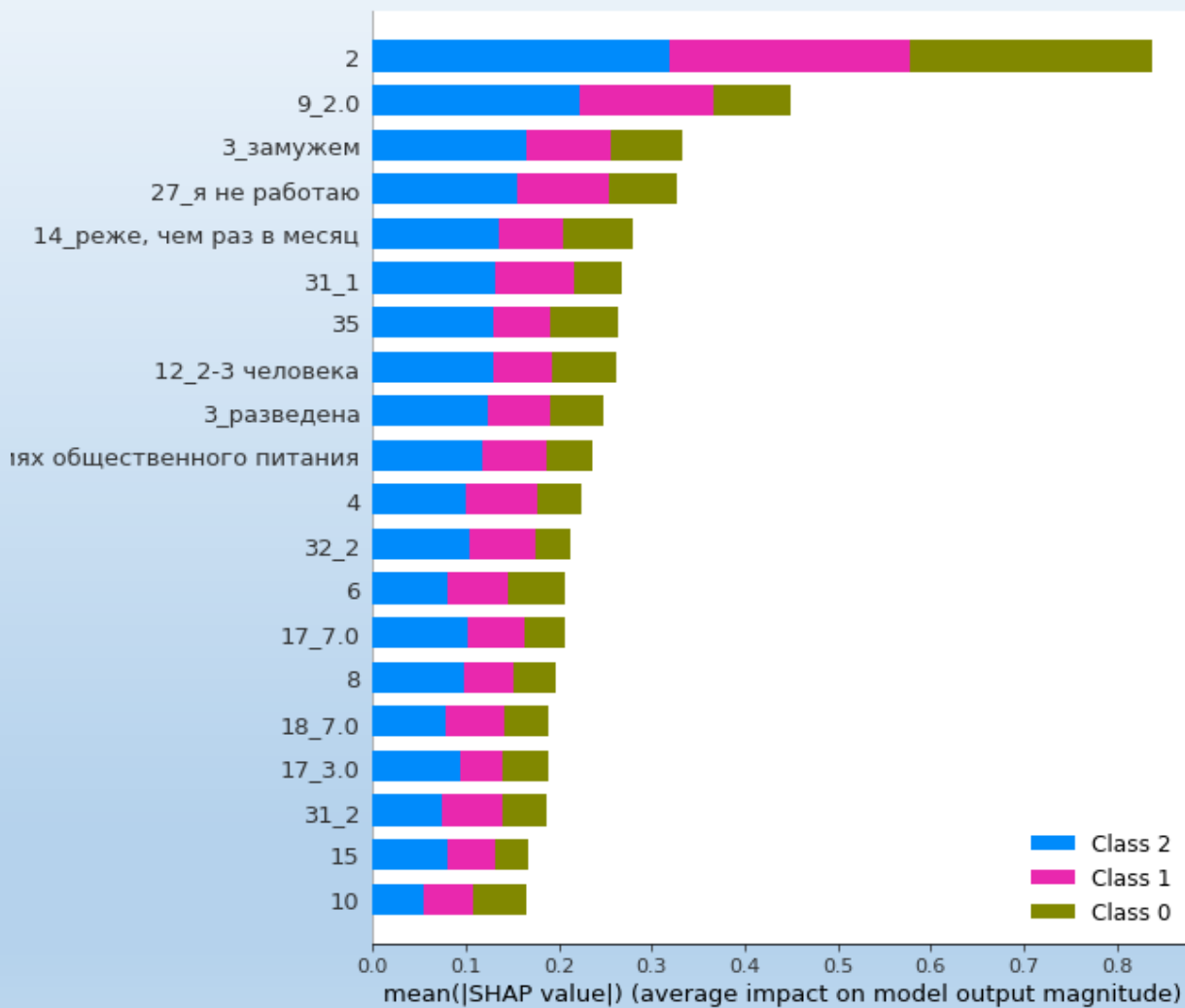
Корреляционная матрица



Модель машинного обучения

Модель	Метрика			
	accuracy	precision	recall	F1-score
scikit-learn DecisionTreeClassifier Label Encoding категориальных и ранговых признаков	0.594 ±0.03	0.578 ±0.034	0.588 ±0.026	0.58 ±0.03
scikit-learn DecisionTreeClassifier Label Encoding ранговых, One-Hot Encoding категориальных признаков	0.602 ±0.009	0.58 ±0.022	0.594 ±0.012	0.58 ±0.021
scikit-learn DecisionTreeClassifier One-Hot Encoding категориальных и ранговых признаков	0.586 ±0.04	0.572 ±0.041	0.58 ±0.036	0.568 ±0.038
scikit-learn DecisionTreeClassifier Label Encoding ранговых, Target Encoding категориальных признаков	0.582 ±0.021	0.564 ±0.028	0.574 ±0.02	0.56 ±0.028
scikit-learn DecisionTreeClassifier Target Encoding категориальных и ранговых признаков	0.59 ±0.038	0.57 ±0.044	0.584 ±0.034	0.566 ±0.038
Собственное дерево решений	0.576 ±0.031	0.455 ±0.056	0.57 ±0.027	0.498 ±0.24
Нейронная сеть	0.622±0.02	0.612±0.02	0.62±0.02	0.614±0.02

Метод оценки значимости признаков





**Широкий Александр Александрович,
1987 г.р.**

Старший научный сотрудник лаб.79

Публикации:

Всего 8, из них WoS — 3 (1 — Q1), SCOPUS — 4, RSCI — 1.

Конференции:

- Четырнадцатая международная конференция «Управление развитием крупномасштабных систем» (MLSD'2021)
- XVII Всероссийская школа-конференция молодых ученых «Управление большими системами» (УБС-2021)
- XXIX Международная конференция «Проблемы управления безопасностью сложных систем» (ПУБСС-2021)

Задача № 4

Идентификация модели управления системой регионального здравоохранения

1-й год проекта: выявление причинно-следственных влияний показателей качества региональной системы здравоохранения

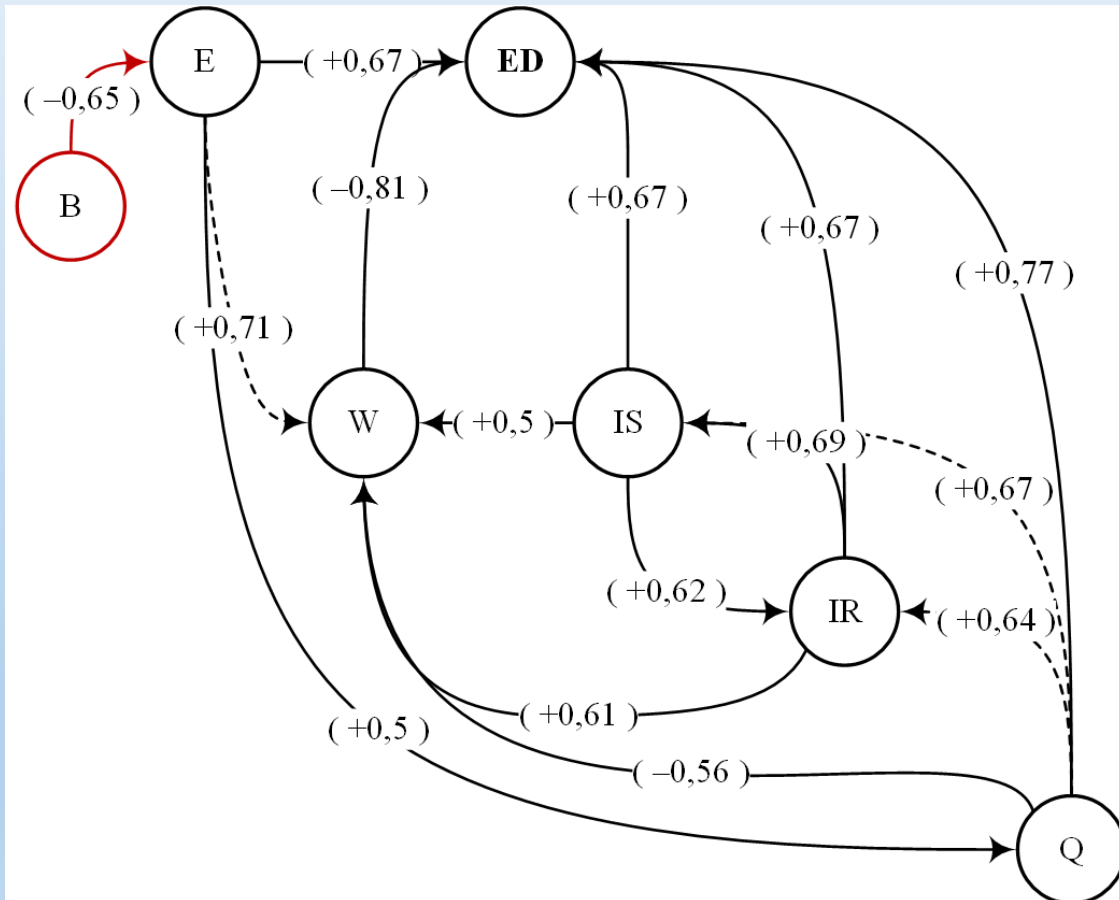
Две нозологии: онкологические заболевания и заболевания сердечно-сосудистой системы

Опрос группы экспертов в области организации здравоохранения

Вопросы:

- Перечислите ключевые количественные показатели, которые следует применять для оценки состояния системы здравоохранения в части организации медицинской помощи при онкологических/кардиоваскулярных заболеваниях.
- Оцените, пожалуйста, их важность по шкале от 0 до 3.
- Оцените, как изменение одного показателя может повлиять на изменение других показателей.

Модель причинно-следственных влияний показателей качества региональной системы здравоохранения в онкологии



(ED) Доля заболеваний на ранних стадиях среди впервые поставленных диагнозов

(E) обеспеченность диагностическим оборудованием

(W) средняя продолжительность ожидания приёма онколога в поликлинике по записи

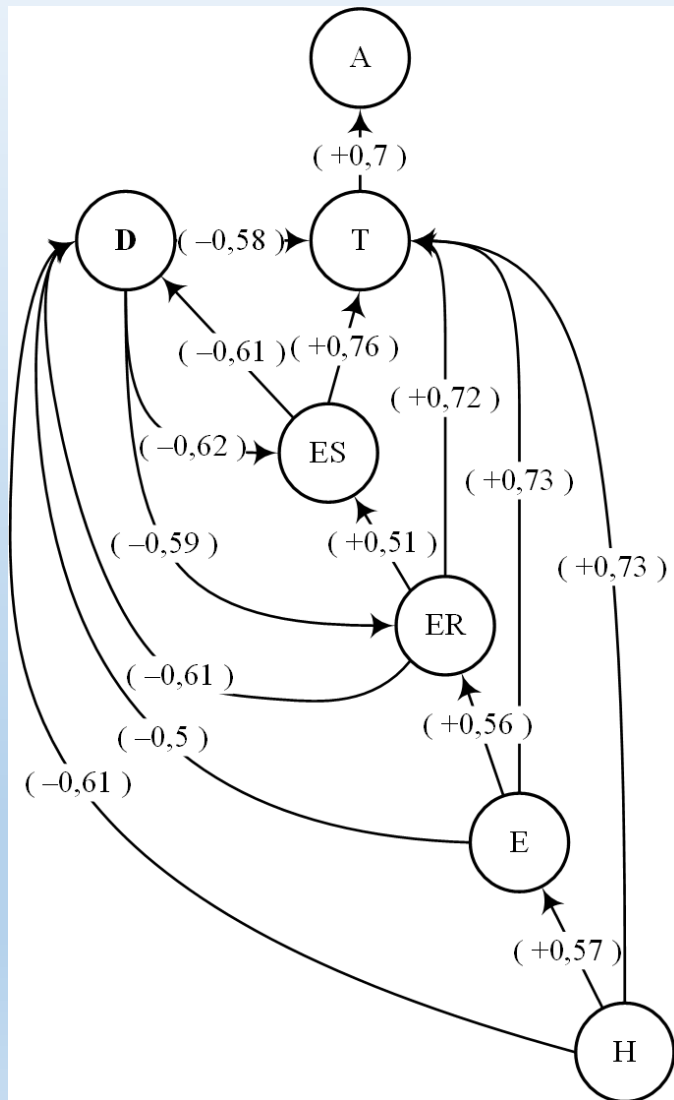
(IS) информированность населения о ранних симптомах заболевания

(IR) информированность населения о факторах риска

(Q) медианный уровень квалификации профильных специалистов

(B) кол-во коек в онкологических отделениях больниц в расчёте на 10 000 чел. населения

Модель причинно-следственных влияний показателей качества региональной системы здравоохранения в области заболеваний ССС



(D) повозрастные показатели смертности по причине сердечно-сосудистых заболеваний

(ER) доступность скорой медицинской помощи

(ES) уровень доступности специализированной медицинской помощи

(E) наличие оборудования

(A) медицинская активность пациентов

(H) доступность высокотехнологичной медицинской помощи

(T) доверие пациентов системе здравоохранения



Исянова Алина Маратовна, 2001 г.р.

Техник лаб.81 с 01.04.2022 года
Бакалавр МГТУ им. Н.Э. Баумана

Публикация

Математические подходы к отбору показателей потребности в медицинской помощи / Черкасов С.Н., Федяева А.В., Исянова А.М., Шевченко М.А., Лобанов А.В.
// Вестник медицинского стоматологического института
(принята к печати, **ВАК**)

Задача № 5

Алгоритмизация и практическая реализация методов моделирования потребности в медицинской помощи в виде компьютерных программ поддержки принятия управленческих решений на территориальном уровне.

Методы решения:

- Для разработки математических подходов к моделированию потребности в медицинской помощи на основании теории нелинейных уравнений и вероятностных подходов будут использоваться возможности нейронных сетей.
- Вероятностный подход планируется реализовать при помощи построения байесовского классификатора.
- Поставленная задача решается с использованием средств и инструментов языка программирования Python.



Лобанов Алексей Владимирович, 1996 г.р.

Инженер лаб.81 с 01.04.2022 года

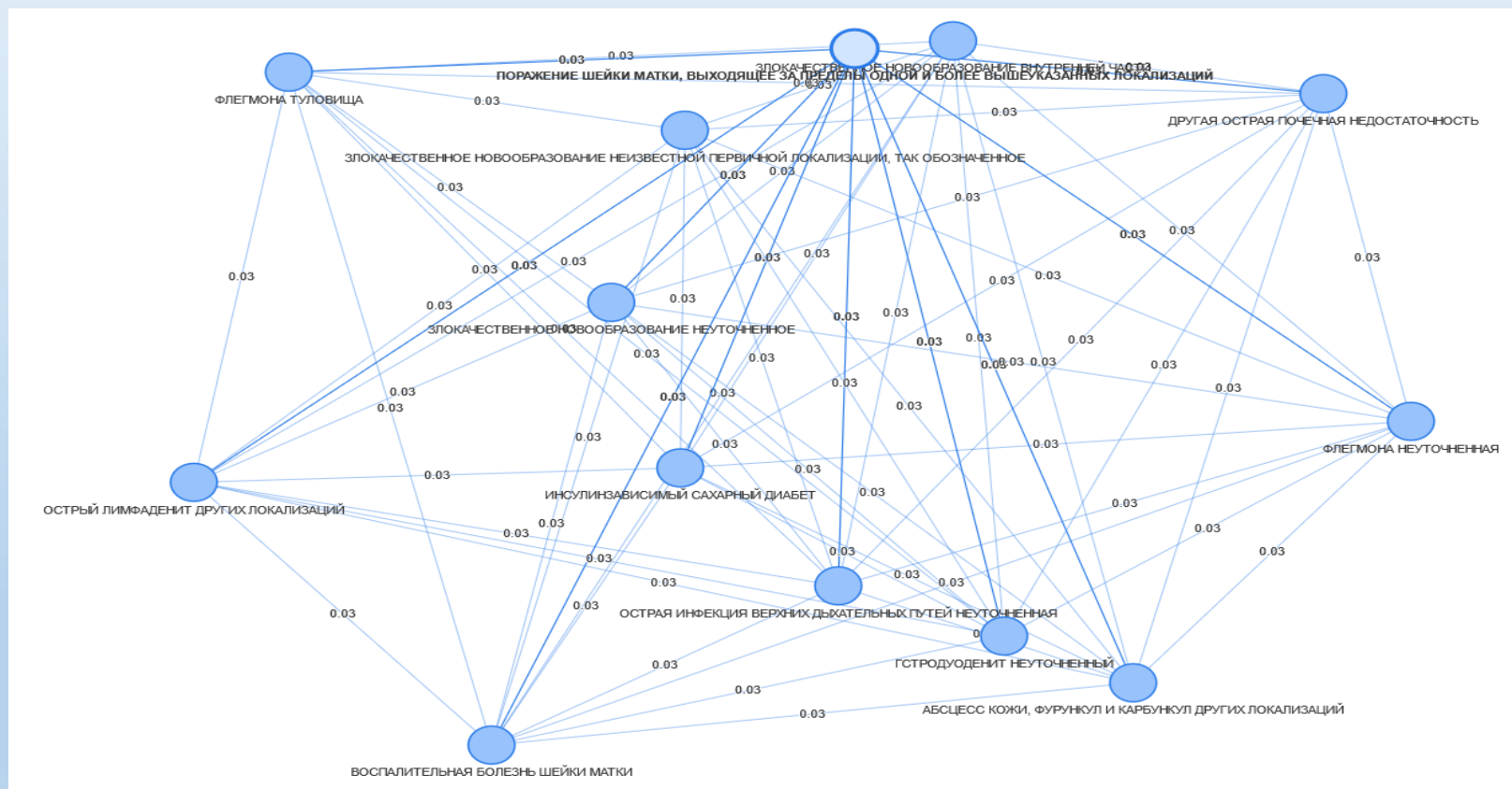
Аспирант РУДН

Публикации:

1. Calculation of a uniform gas flow from the heated tungsten plate surface Bull. Nov. Comp. Center, Comp. Science, 44 (2020), 1–8 © 2020 NCC Publisher
2. Математические подходы к отбору показателей потребности в медицинской помощи / Черкасов С.Н., Федяева А.В., Исянова А.М., Шевченко М.А., Лобанов А.В. // Вестник медицинского стоматологического института (принята к печати, **ВАК**)

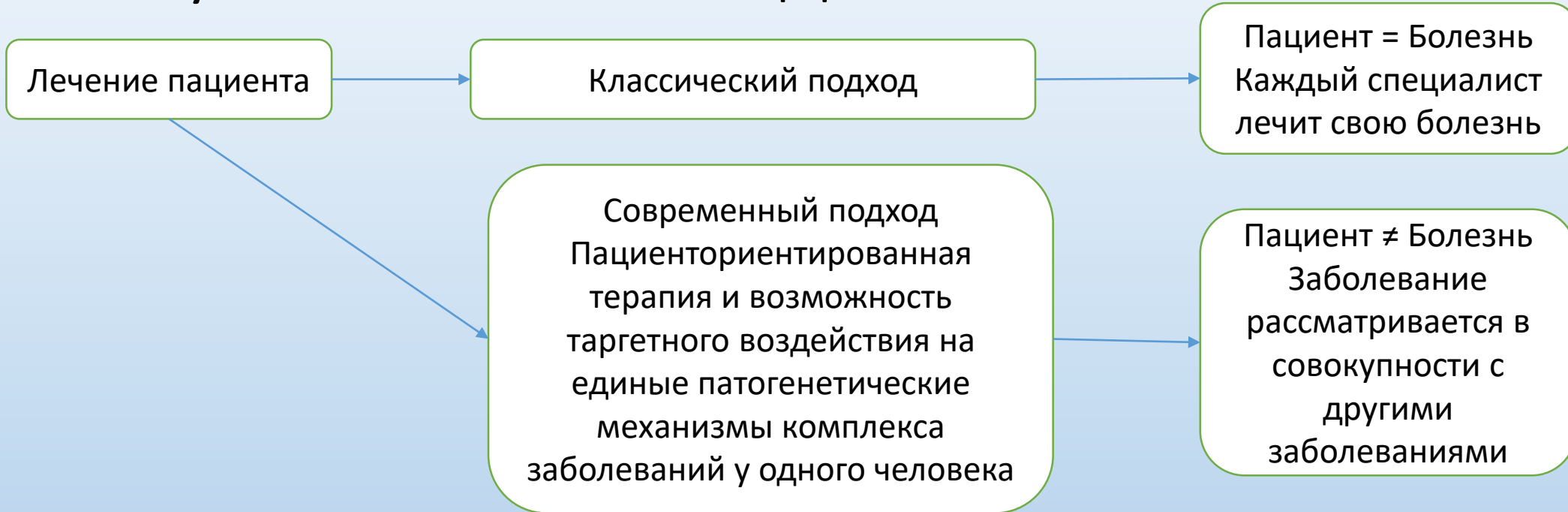
Задача № 6

- Нахождение мультиморбидных групп заболеваний.

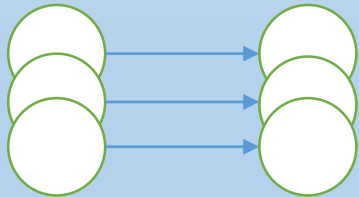


Мультиморбидность – это наличие двух или более хронических (или рецидивирующих) состояний/заболеваний.

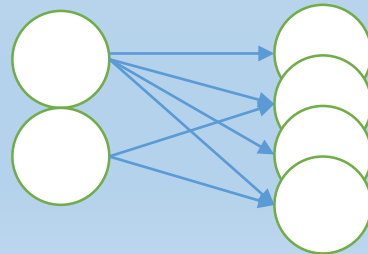
Актуальность исследования



*Классический подход к
лечению*



Пациентоориентированная терапия



Объекты исследования

- Анонимизированные данные обращаемости пациентов.
- Демографические данные.

id	Возраст	Код по МКБ-10	Дата начала случая	Дата окончания случая
2	30	G51.9	16.12.2016	26.12.2016
2	30	M79.2	14.12.2016	23.12.2016
3	35	Z12.3	03.02.2016	03.02.2016
...				



id	A1.1	A1.2	A1.3	...	Z38.5	Z 38.6
0	0	0	0	...	3	0
1	1	0	0		0	0
2	2	1	1		3	0
...						

Методы исследования

- Ассоциативные правила

antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
E11.7	M42.1	0,146256123	0,792162351	0,123863	0,84688995	1,069086	0,008004	1,35744
E11.9	I11.9	0,13663401	0,617039888	0,098845	0,7234315	1,172423	0,014537	1,384684
E11.9	J06.9	0,13663401	0,523792862	0,078726	0,57618438	1,100023	0,007158	1,123619
E11.9	M42.1	0,13663401	0,792162351	0,115115	0,8425096	1,063557	0,006879	1,319685
J06.9, E11.7	I11.9	0,0810007	0,617039888	0,059482	0,73434125	1,190103	0,009501	1,441549
I11.9, E11.7	J06.9	0,108117565	0,523792862	0,059482	0,55016181	1,050342	0,002851	1,058619
E11.7	J06.9, I11.9	0,146256123	0,343072078	0,059482	0,40669856	1,185461	0,009306	1,107241
M42.1, E11.7	I11.9	0,123862841	0,617039888	0,091847	0,74152542	1,201746	0,015419	1,481616
I11.9, E11.7	M42.1	0,108117565	0,792162351	0,091847	0,84951456	1,0724	0,006201	1,381115

$$supp(x_1 \cup x_2) = \frac{\sigma(x_1 \cup x_2)}{|T|}$$

$$lift(A \rightarrow C) = \frac{confidence(A \rightarrow C)}{support(C)}$$

$$conviction(A \rightarrow C) = \frac{1 - support(C)}{1 - confidence(A \rightarrow C)}$$

$$conf(x_1 \cup x_2) = \frac{supp(x_1 \cup x_2)}{supp(x_1)}$$

$$levarage(A \rightarrow C) = support(A \rightarrow C) - support(A) \times support(C)$$



Обозначения:

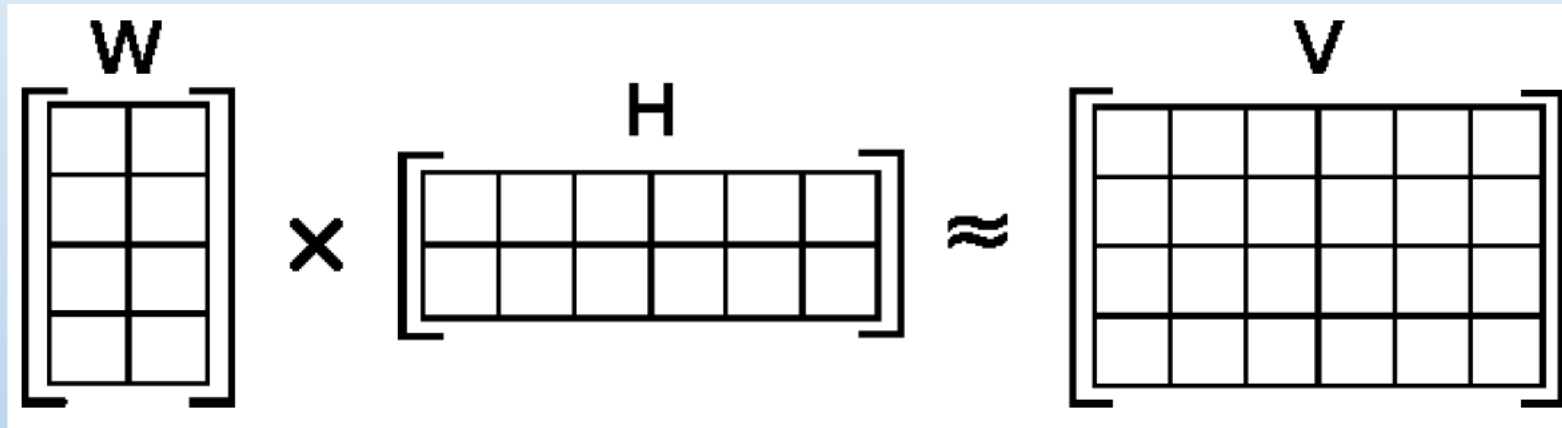
antecedents (A, x_1) – набор заболеваний номер 1

consequents (C, x_2) – набор заболеваний номер 2

T – число строк в исследуемой группе

Методы исследования

- Матричная факторизация



Методы разложения матриц: метод главных компонент (PCA), усеченное сингулярное разложение (TruncatedSVD), неотрицательное разложение матрицы (NNMF).

Методы сравнения произведения разложения и оригинальной матрицы: среднеквадратичная ошибка (MSE), средняя абсолютная ошибка (MAE).

Дополнительные метрики: коэффициент объясненной дисперсии (EVR), разреженность (Sparsity).

Методы исследования.

- Исследование применимости методов кластеризации

K-means

Mini Batch K-Means

Affinity Propagation

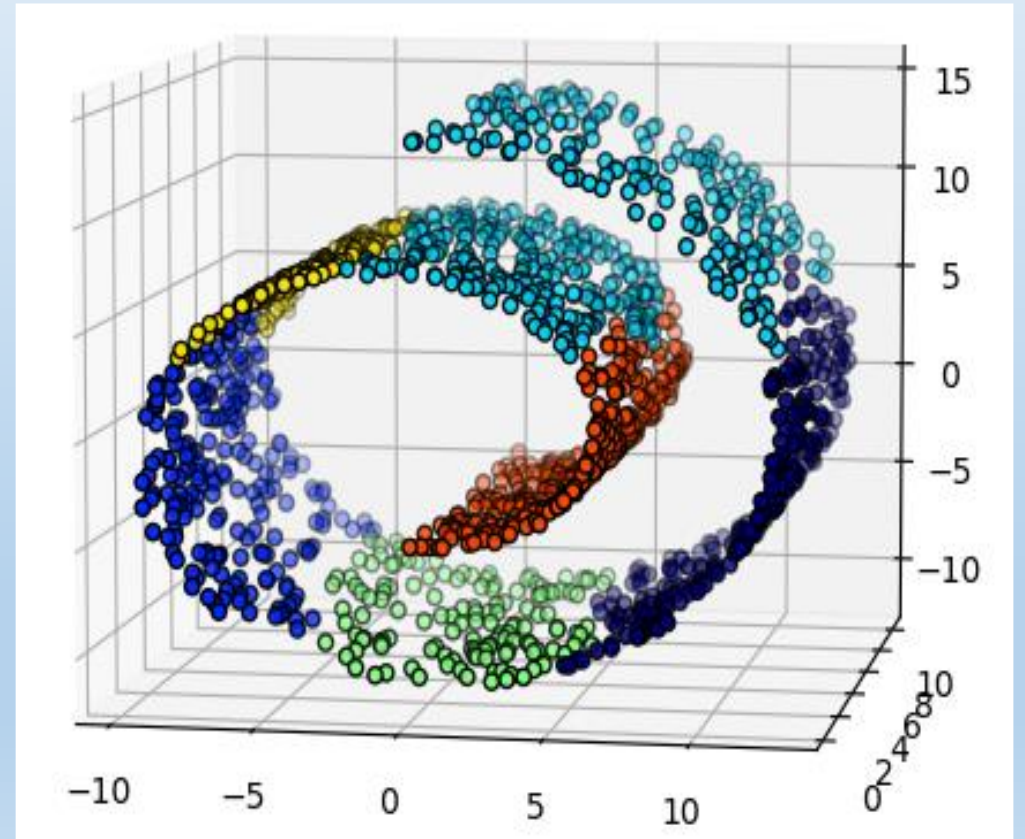
Mean Shift

Spectral clustering

Different label assignment strategies

Hierarchical clustering (как семейство методов)

...



Спасибо за внимание!