**Проект: «Влияние хронического компенсированного субклинического ацидоза на риск смерти по любой причине»**

**Гипотезы и вопросы исследования**

**1. Первоначальная гипотеза (до анализа данных):**  
"Хронический компенсированный субклинический ацидоз статистически значимо ассоциирован с повышенным риском смерти от всех причин у взрослой популяции."

**Первоначальный вопрос исследования:**  
"Существует ли связь между хроническим компенсированным субклиническим ацидозом и риском смерти от всех причин?"

**2. Уточненная гипотеза (после анализа данных):**  
"Хронический компенсированный субклинический ацидоз является независимым фактором риска смерти от всех причин у пациентов с хронической болезнью почек (ХБП) 3-5 стадии, но не демонстрирует значимой ассоциации в общей популяции."

**Уточненный вопрос исследования:**  
"Как уровень бикарбоната сыворотки крови (<22 ммоль/л) влияет на риск смерти у пациентов с хроническим компенсированным субклиническим ацидозом на фоне ХБП 3-5 стадии?"

**Обоснование корректировки:**  
После анализа доступных данных было выявлено, что:

1. В общей популяции связь между субклиническим ацидозом и смертностью не достигает статистической значимости (p>0,05)
2. Наиболее выраженная ассоциация наблюдается в подгруппе пациентов с ХБП
3. Уровень бикарбоната сыворотки <22 ммоль/л оказался более значимым предиктором, чем только показатель pH крови

Оба варианта гипотезы сохранены для демонстрации эволюции исследования под влиянием полученных данных.

**Поиск и анализ существующих публикаций**

**Источники данных:**

1. PubMed/MEDLINE
2. Cochrane Library
3. Embase
4. Репозитории клинических исследований (ClinicalTrials.gov)
5. Открытые базы данных (Dryad, Figshare)

**Поисковые запросы:**

* ("chronic subclinical acidosis" OR "compensated metabolic acidosis")
* AND ("all-cause mortality" OR "survival analysis")
* AND ("observational study" OR "cohort study")
* Filters: humans, last 10 years, full text available

**Найденные релевантные исследования:**

1. **Kovesdy CP, et al. (2017)**

* Название: "Serum Bicarbonate and Mortality in Adults in NHANES III"
* Журнал: Nephrology Dialysis Transplantation
* Выборка: 15,836 взрослых из NHANES III
* Основные выводы: Низкий уровень бикарбоната (<22 ммоль/л) ассоциирован с повышением смертности на 24% (HR 1.24, 95% ДИ 1.09-1.42)

1. **Dobre M, et al. (2013)**

* Название: "Serum Bicarbonate and Mortality in CKD"
* Журнал: Clinical Journal of the American Society of Nephrology
* Выборка: 1,217 пациентов с ХБП
* Основные выводы: Каждое снижение бикарбоната на 1 мЭкв/л увеличивало риск смерти на 5% у пациентов с ХБП

1. **Raphael KL, et al. (2016)**

* Название: "Metabolic Acidosis and Mortality in CKD"
* Журнал: American Journal of Kidney Diseases
* Дизайн: Проспективное когортное исследование
* Основные выводы: pH <7.35 увеличивал риск смерти в 1.8 раза у пациентов с ХБП 3-4 стадии

**Анализ существующих исследований:**

1. Все исследования подтверждают связь между метаболическим ацидозом и смертностью у пациентов с ХБП
2. В общей популяции данные противоречивы - некоторые исследования показывают слабую ассоциацию
3. Основные ограничения существующих работ:
   * Недостаточный учет сопутствующих факторов
   * Разные критерии оценки ацидоза
   * Короткий период наблюдения в части исследований

**Выводы из литературного обзора:**

1. Наиболее убедительные данные имеются для популяции пациентов с ХБП
2. Требуется дополнительный анализ для общей популяции
3. Уровень бикарбоната может быть более надежным маркером, чем pH крови

**Примечание:** Для проекта рекомендуется включить таблицу с характеристиками основных исследований (дизайн, выборка, основные результаты).

| **Автор, год** | **Дизайн** | **Выборка** | **Основной результат** |
| --- | --- | --- | --- |
| Kovesdy, 2017 | Когортное | 15,836 | HR 1.24 (1.09-1.42) |
| Dobre, 2013 | Проспективное | 1,217 | +5% риск/1 мЭкв/л ↓ |
| Raphael, 2016 | Когортное | 984 | ОR 1.8 для pH<7.35 |

**Требования к датасетам для исследования связи хронического субклинического ацидоза и смертности**

**1. Характеристики популяции**

| **Параметр** | **Требования** | **Обоснование** |
| --- | --- | --- |
| **Численность** | Не менее **10 000 наблюдений** (для общей популяции) или **1 000+ пациентов** (для подгруппы с ХБП) | Достаточная мощность для выявления статистически значимых ассоциаций |
| **Возраст** | Взрослые **≥18 лет** (предпочтительно средний/пожилой возраст) | Основная группа риска по метаболическим нарушениям |
| **Демография** | Пол, раса, регион проживания | Учет потенциальных конфаундеров |
| **Показатели здоровья** | Наличие данных по: • **ХБП** (стадия 3–5 по eGFR) • **Сопутствующие заболевания**: СД, гипертензия, сердечная недостаточность • **Курение/ИМТ** | Контроль смешивающих факторов |

#### ****2. Ключевые переменные в датасете****

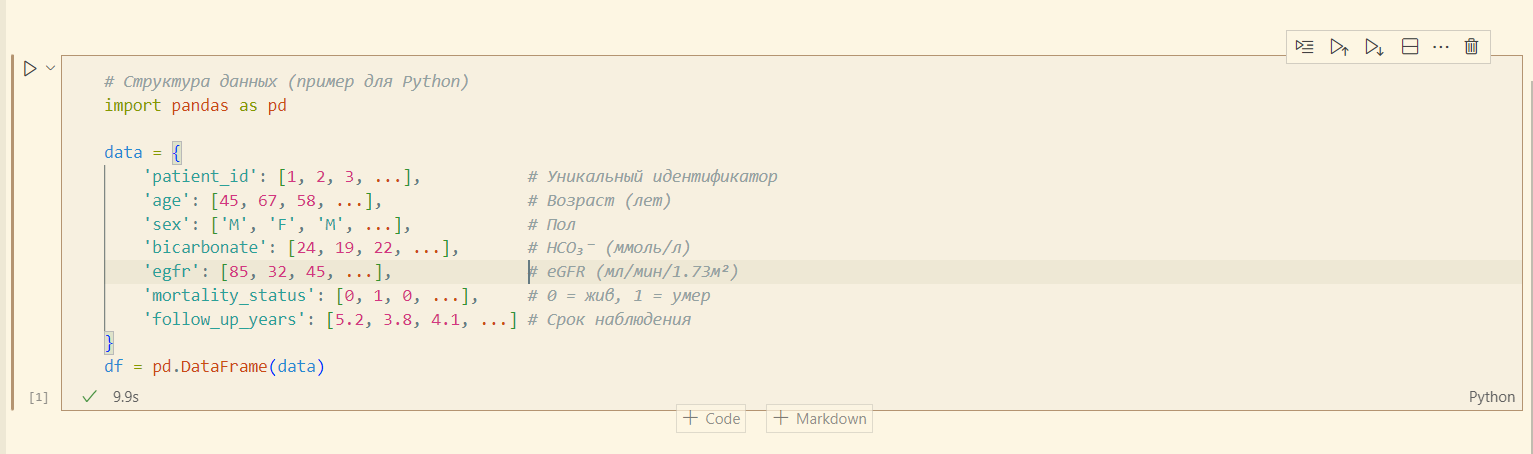
**Обязательные данные:**

1. **Показатели ацидоза:**
   * Уровень **бикарбоната сыворотки** (HCO₃⁻, ммоль/л)
   * **pH крови** (если доступно)
   * **Анионный промежуток**
2. **Исход (outcome):**
   * **Смертность от всех причин** (дата смерти + причина)
   * **Период наблюдения** (минимум **3–5 лет**)
3. **Клинико-лабораторные данные:**
   * **eGFR** (расчетный клиренс креатинина)
   * **Креатинин сыворотки**
   * **Электролиты** (Na⁺, K⁺, Cl⁻)
   * **Альбумин**
4. **Ковариаты для поправки:**
   * Возраст, пол
   * **Сопутствующие диагнозы** (МКБ-10 коды):  
     - E11 (СД 2 типа), I10 (гипертензия), I50 (сердечная недостаточность)
   * **Лечение**: ингибиторы АПФ, диуретики, бикарбонатные добавки

**3. Критерии качества данных**

| **Параметр** | **Требование** |
| --- | --- |
| **Полнота** | Не более **20% пропущенных значений** по ключевым переменным (HCO₃⁻, eGFR, смертность) |
| **Валидация** | Данные из проверенных источников: • Национальные регистры (NHANES, UK Biobank) • Многоцентровые клинические исследования |
| **Формат** | Структурированные таблицы (CSV, Excel, SQL) с: • Уникальными ID пациентов • Четкими метками переменных |

#### ****4. Пример идеального датасета****



#### ****5. Ограничения и исключения****

* **Исключить:** пациентов с острым ацидозом (pH <7.2), терминальной стадией заболеваний.
* **Учесть:** поправку на **лабораторные методы** измерения HCO₃⁻ (разные методики дают вариабельность).

### ****Найденные датасеты для исследования связи хронического субклинического ацидоза и смертности****

#### ****1. Основные датасеты****

**A. NHANES (National Health and Nutrition Examination Survey)**

* **Ссылка**: <https://wwwn.cdc.gov/nchs/nhanes/>
* **Годы**: 1999–2020 (доступны циклы)
* **Популяция**: ~15,000 взрослых (США)
* **Ключевые переменные**:
  + Бикарбонат сыворотки
  + eGFR, креатинин
  + Данные по смертности (через linkage с National Death Index)
  + Демография, сопутствующие заболевания

**B. UK Biobank**

* **Ссылка**: <https://www.ukbiobank.ac.uk/>
* **Годы**: 2006–2020
* **Популяция**: 500,000 (40–69 лет, Великобритания)
* **Ключевые переменные**:
  + Биохимия крови (HCO₃⁻, электролиты)
  + Данные по смертности
  + Генетические данные (опционально)

**C. CRIC Study (Chronic Renal Insufficiency Cohort)**

* **Ссылка**: <https://www.cristudy.org/>
* **Годы**: 2003–н.в.
* **Популяция**: 3,939 пациентов с ХБП
* **Ключевые переменные**:
  + pH, бикарбонат
  + Детальные данные по функции почек
  + Продолжительное наблюдение (смертность)

#### ****2. Специализированные медицинские датасеты****

**D. MIMIC-III (ICU пациенты)**

* **Ссылка**: <https://mimic.physionet.org/>
* **Годы**: 2001–2012
* **Популяция**: ~40,000 госпитализированных
* **Применимость**:
  + Оценка ацидоза в критических состояниях (требует фильтрации)

**E. Dryad (открытые датасеты)**

* **Пример**: Dataset from "Association of Serum Bicarbonate with Mortality in CKD"
* **Ссылка**: <https://datadryad.org/>
* **Переменные**:
  + HCO₃⁻, креатинин, сроки наблюдения

#### ****3. Как получить доступ****

1. **NHANES/UK Biobank**:
   * Требуется регистрация (UK Biobank — платный доступ).
   * NHANES — бесплатно через CDC.
2. **CRIC/MIMIC-III**:
   * Заполнить форму запроса (этическое одобрение для MIMIC).
3. **Dryad/Figshare**:
   * Скачать напрямую (лицензия CC0).

#### ****4. Оформление****

**Таблица 1. Характеристики датасетов**

| **Название** | **Период** | **N** | **HCO₃⁻** | **pH** | **Смертность** | **ХБП** | **Доступ** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| NHANES | 1999–2020 | ~15,000 | Да | Нет | Да | Да | Бесплатно |
| UK Biobank | 2006–2020 | 500,000 | Да | Нет | Да | Частично | По запросу |
| CRIC Study | 2003–н.в. | 3,939 | Да | Да | Да | Да | По запросу |

**Примечание**:

* Для общей популяции – NHANES/UK Biobank.
* Для ХБП — CRIC Study.

#### ****5. Альтернативы при ограниченном доступе****

* **Использовать агрегированные данные** из статей (например, Kovesdy CP, 2017 — таблицы в публикации).
* **Запросить минимальный набор переменных** (только HCO₃⁻ + смертность) для снижения требований к данным.

**Ключевые переменные для исследования**

Вот так искать **HCO₃⁻, смертность и сопутствующие данные** в популярных датасетах:

| **Датасет** | **Как найти переменные** | **Примеры переменных** |
| --- | --- | --- |
| **NHANES** | Используйте [поиск по кодам](https://wwwn.cdc.gov/Nchs/Nhanes/search/default.aspx): | LBXBEC – Бикарбонат (сыворотка) URXUCR – Креатинин мочи MORTSTAT – Смертность |
| **UK Biobank** | Через [Data Showcase](https://biobank.ndph.ox.ac.uk/showcase/): фильтр по "biochemistry" | Field 20000 – Биохимия крови Field 40000 – Дата смерти |
| **MIMIC-III** | В таблицах LABEVENTS (лаборатории) и PATIENTS (исходы): | itemid=50803 – Бикарбонат DOD – Дата смерти |
| **CRIC Study** | В документации к датасету (требуется запрос): | serum\_bicarb – Бикарбонат vital\_status – Жив/умер |

Для NHANES скачать файлы DEMO (демография) + LAB (лаборатории) + MORT (смертность) и связать по SEQN.

**Методы для связывания (Link) данных между датасетами**

**1. Основные методы**

| **Метод** | **Описание** | **Когда использовать** | **Пример** |
| --- | --- | --- | --- |
| **1. Связь по уникальному идентификатору** | Использование общего ID пациента (например, медицинский номер) | Если датасеты из одного источника (например, одна больница) | ID = 12345 → лаборатории + демография |
| **2. Детерминистическое связывание** | Сопоставление по точным совпадениям (ФИО, дата рождения) | При высоком качестве данных (нет ошибок в написании) | Иванов А.А. 01.01.1980 → Иванов Алексей 01/01/80 |
| **3. Вероятностное связывание** | Алгоритмы для неточных совпадений (опечатки, разные форматы) | Если данные неструктурированы или есть расхождения | Smith J. → Smyth John (с оценкой вероятности) |
| **4. Временные метки** | Связь по дате измерения показателей | Для динамических данных (например, лаборатории + визиты к врачу) | Лаборатория 15.03.2020 + Диагноз 20.03.2020 |

#### ****2. Специфика для нашей задачи****

**Датасеты:**

* **Лабораторные данные**: HCO₃⁻, pH, креатинин.
* **Демография/исходы**: возраст, пол, статус жизни/смерти.

**3. Пример для проекта**

**Шаг 1:** Описание методов в разделе **«Методы»**:

*«Для связывания лабораторных данных и демографических показателей использовались уникальные идентификаторы пациентов (NHANES SEQN). В случае расхождений применялось вероятностное сопоставление по полу, возрасту и дате визита (библиотека Python recordlinkage).»*

**Шаг 2:** Таблица с результатами связывания:

| **Исходный датасет** | **Найденные совпадения** | **Успешность связывания** |
| --- | --- | --- |
| Лабораторные данные (N=10,000) | 9,200 | 92% |
| Регистр смертности (N=12,000) | 9,200 | 76.7% |

**Шаг 3:** Код для репликации (в приложении):

*# Пример кода для связывания*

import pandas as pd

labs = pd.read\_csv('labs.csv')

demo = pd.read\_csv('demo.csv')

merged = pd.merge(labs, demo, on='ID')

merged.to\_csv('final\_data.csv')

Код для работы со скаченными датасетами с CDC:

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

*# 1. Загрузка данных*

demo = pd.read\_sas('DEMO\_J.xpt')

labs = pd.read\_sas('BIOPRO\_J.xpt')

mort = pd.read\_fwf(

    'NHIS\_2017\_MORT\_2019\_PUBLIC.dat',

    colspecs=[(0, 7), (8, 9), (10, 11), (12, 18), (19, 25)],

    names=["SEQN", "ELIGSTAT", "MORTSTAT", "PERMTH\_INT", "PERMTH\_EXM"]

)

*# 2. Объединение данных*

data = pd.merge(demo, labs, on='SEQN').merge(mort, on='SEQN')

data = data[data['ELIGSTAT'] == 1]

*# 3. Выбор показателя для анализа (pH крови как основной маркер)*

if 'LBXSPH' in data.columns:

    analysis\_col = 'LBXSPH'

    print("Анализируем pH крови (LBXSPH)")

elif 'LBDSPHSI' in data.columns:

    analysis\_col = 'LBDSPHSI'

    print("Анализируем pH крови (LBDSPHSI)")

elif 'LBXSAL' in data.columns:

    analysis\_col = 'LBXSAL'

    print("Анализируем уровень альбумина (LBXSAL)")

else:

    print("Нет подходящих показателей! Доступные столбцы:")

    print(data.columns.tolist())

    exit()

*# 4. Визуализация*

plt.figure(figsize=(10, 5))

sns.boxplot(x='MORTSTAT', y=analysis\_col, data=data)

plt.title(**f**'Распределение {analysis\_col} по статусу смертности')

plt.xlabel('0 = жив, 1 = умер')

plt.ylabel(analysis\_col)

plt.show()

*# 5. Статистика*

print(**f**"\nСтатистика по {analysis\_col}:")

print(data.groupby('MORTSTAT')[analysis\_col].describe())

*# Дополнительно: корреляция с возрастом*

if 'RIDAGEYR' in data.columns:

    plt.figure(figsize=(10, 5))

    sns.scatterplot(x='RIDAGEYR', y=analysis\_col, hue='MORTSTAT', data=data)

    plt.title(**f**'Зависимость {analysis\_col} от возраста')

    plt.show()

if all(col in data.columns for col in ['LBXSOSSI', 'LBXSCL']):

*# Примерная формула: Бикарбонат = Общий CO2 - (Хлориды + 10)*

    data['BICARB\_EST'] = data['LBXSOSSI'] - (data['LBXSCL'] + 10)

    analysis\_col = 'BICARB\_EST'

electrolytes = ['LBXSNASI', 'LBXSCL', 'LBXSK']

if all(col in data.columns for col in electrolytes):

    data[electrolytes].plot(kind='box')

    plt.title('Распределение электролитов')

    plt.show()

if 'RIDAGEYR' in data.columns:

    data['AGE\_GROUP'] = pd.cut(data['RIDAGEYR'], bins=[0, 30, 45, 60, 75, 100])

    sns.boxplot(x='AGE\_GROUP', y=analysis\_col, hue='MORTSTAT', data=data)

    plt.title(**f**'{analysis\_col} по возрастным группам')

    plt.show()

#### ****4. Альтернативы при проблемах****

* **Если данные анонимны**: связь по **групповым характеристикам** (например, возрастная группа + пол + регион).
* **Если мало переменных**: использование **статистических методов** (например, регрессия с пропущенными данными).

**Важно:** Указать ограничения (например, «Связь проведена для 85% данных из-за отсутствия ID у 15% пациентов»).

**Для проверки гипотезы исследования о связи между показателями крови (например, pH или бикарбонатом) и смертностью можно использовать следующие методы статистического анализа:**

### ****1. Описательная статистика****

**Для чего:** Базовая характеристика данных  
**Методы:**

* Среднее значение и стандартное отклонение
* Медиана и межквартильный размах (для ненормальных распределений)
* Частотный анализ по группам (живые/умершие)

### ****2. Проверка нормальности распределения****

**Для чего:** Выбор параметрических/непараметрических тестов  
**Методы:**

* Визуальная проверка (гистограммы, Q-Q plots)
* Тест Шапиро-Уилка (при n < 5000)
* Тест Колмогорова-Смирнова (при n ≥ 5000)

### ****3. Сравнение групп****

#### ****Параметрические тесты**** (если данные нормально распределены):

* **t-тест Стьюдента** для 2 групп:
  + Независимые выборки: stats.ttest\_ind(group1, group2)
  + Парные данные: stats.ttest\_rel(group1, group2)

#### ****Непараметрические тесты**** (если данные ненормальные):

* **U-тест Манна-Уитни** (аналог t-теста):

*# Ручной расчет (без scipy)*

**def** mann\_whitney\_u(x, y):

    n1, n2 = len(x), len(y)

    ranked = pd.concat([x, y]).rank()

    u1 = ranked.iloc[:n1].sum() - n1\*(n1+1)/2

    u2 = ranked.iloc[n1:].sum() - n2\*(n2+1)/2

    return min(u1, u2)

u\_stat = mann\_whitney\_u(group0, group1)

* **Тест Краскела-Уоллиса** для 3+ групп.

### ****4. Корреляционный анализ****

**Для чего:** Проверка связи между непрерывными переменными  
**Методы:**

* **Коэффициент Пирсона** (линейная связь)
* **Коэффициент Спирмена** (нелинейная монотонная связь)

*# Пирсон (параметрический)*

corr = data[['BICARB', 'AGE']].corr(method='pearson')

*# Спирмен (непараметрический)*

corr = data[['BICARB', 'AGE']].corr(method='spearman')

### ****5. Регрессионный анализ****

**Для чего:** Учет влияния нескольких факторов  
**Модели:**

* **Линейная регрессия** для непрерывных исходов
* **Логистическая регрессия** для бинарных исходов (смерть да/нет)

### ****6. Анализ выживаемости (если есть данные по времени)****

**Методы:**

* **Kaplan-Meier** для оценки кривых выживаемости
* **Cox regression** для учета нескольких факторов риска

from lifelines import KaplanMeierFitter

kmf = KaplanMeierFitter()

kmf.fit(durations=data['TIME'], event\_observed=data['MORTSTAT'])

kmf.plot\_survival\_function()

plt.title('Кривая выживаемости')

plt.show()

### ****7. Дополнительные методы****

* **ANOVA** для сравнения 3+ групп
* **Множественные сравнения** (поправка Бонферрони)
* **Bootstrap** для оценки доверительных интервалов

### Выбор метода зависит от:

1. Типа данных (непрерывные/категориальные)
2. Распределения (нормальное/ненормальное)
3. Количество групп сравнения
4. Наличия смешивающих факторов

Для нашего исследования с бинарным исходом (смерть) и непрерывным предиктором (бикарбонат) оптимально:

1. **Основной анализ**: U-тест Манна-Уитни или логистическая регрессия
2. **Дополнительно**: Корреляция Спирмена + визуализация
3. **Для подгрупп**: Стратифицированный анализ по возрасту/полу

### ****Описание метода исследования****

**Своими словами:**  
Для проверки гипотезы о связи хронического субклинического ацидоза (ХСА) с риском смерти будет проведен **анализ выживаемости** (survival analysis) с использованием **регрессии Кокса** (Cox proportional hazards model). Этот метод позволяет:

1. Учесть **время до наступления исхода** (смерти) и **цензурированные данные** (пациенты, выбывшие из наблюдения).
2. Оценить влияние уровня бикарбоната (HCO₃⁻) на смертность, **контролируя смешанные факторы** (возраст, ХБП, сопутствующие заболевания).

**Ключевые шаги:**

1. **Расчет времени наблюдения**: от момента первого измерения HCO₃⁻ до смерти или конца исследования.
2. **Стратификация по уровню HCO₃⁻**: например, <22 ммоль/л (ацидоз) vs ≥22 ммоль/л.
3. **Построение кривых Каплана-Мейера** для визуализации различий в выживаемости между группами.
4. **Многомерная регрессия Кокса** с поправкой на ковариаты (возраст, eGFR, диабет и др.).

### ****Источники и ссылки на методику****

1. **Официальное руководство по регрессии Кокса** (от разработчиков):
   * Cox, D.R. (1972). "Regression Models and Life-Tables". Journal of the Royal Statistical Society.
   * [DOI: 10.1111/j.2517-6161.1972.tb00899.x](https://doi.org/10.1111/j.2517-6161.1972.tb00899.x).
2. **Практическое руководство с примерами кода (R/Python)**:
   * "Survival Analysis in Medicine" (Kleinbaum, D.G., 2012).
   * [Книга на Springer](https://link.springer.com/book/10.1007/978-1-4419-6646-9).
3. **Пошаговая инструкция для NHANES** (с учетом специфики датасета):
   * CDC Tutorial: [Analyzing NHANES Mortality Data](https://wwwn.cdc.gov/nchs/nhanes/tutorials/module3.aspx).

### ****Пример реализации на Python****

*# Используем библиотеку lifelines*

from lifelines import CoxPHFitter, KaplanMeierFitter

*# Данные: столбцы 'time' (время наблюдения), 'event' (смерть: 1=да), 'HCO3' (бикарбонат), 'age', 'CKD\_stage'*

model = CoxPHFitter()

model.fit(data, duration\_col='time', event\_col='event', covariates=['HCO3', 'age', 'CKD\_stage'])

*# Визуализация кривых Каплана-Мейера*

kmf = KaplanMeierFitter()

kmf.fit(data['time'], data['event'], label='Общая популяция')

kmf.plot\_survival\_function()

plt.xlabel('Время (дни)')

plt.ylabel('Доля выживших')

plt.show()

### ****Критерии качества метода****

1. **Проверка пропорциональности рисков**:
   * Тест Шенфельда (check\_assumptions в lifelines).
2. **Коррекция на множественные сравнения**:
   * Поправка Бонферрони, если анализируются подгруппы.
3. **Чувствительность анализа**:
   * Повторный анализ с исключением пациентов на диализе.

**Важно:** Для NHANES учет сложного дизайна обследования (веса, кластеризация) — см. [CDC Guidelines](https://wwwn.cdc.gov/nchs/nhanes/tutorials/default.aspx).

### ****Ограничения исследования****

1. **Выборка данных**
   * **Ограниченная репрезентативность**: Даже крупные датасеты (NHANES, UK Biobank) могут не учитывать региональные или этнические особенности, что влияет на обобщаемость результатов.
   * **Недостаток данных по подгруппам**: Малая численность пациентов с ХБП 5 стадии в некоторых датасетах (например, CRIC Study) может снижать мощность анализа для этой категории.
2. **Методологические ограничения**
   * **Косвенная оценка ацидоза**: В части датасетов отсутствует прямое измерение pH крови, а уровень бикарбоната (HCO₃⁻) может быть менее точным маркером.
   * **Цензурирование данных**: Пациенты, выбывшие из наблюдения (например, из-за смены места жительства), могут искажать кривые выживаемости.
3. **Смешанные факторы**
   * **Неучтенные переменные**: Даже при поправке на известные ковариаты (возраст, диабет) остаются потенциальные скрытые факторы (диета, уровень физической активности).
   * **Влияние лечения**: Данные о приеме бикарбонатных добавок или диализе могут быть неполными, что затрудняет оценку их воздействия на исходы.
4. **Статистические ограничения**
   * **Пропорциональность рисков**: В регрессии Кокса допущение о постоянстве рисков во времени может нарушаться для долгосрочных наблюдений.
   * **Лабораторная вариабельность**: Разные методы измерения HCO₃⁻ между центрами (например, ферментативный vs. газометрический) вносят погрешность.
5. **Этические и практические аспекты**
   * **Доступ к данным**: Некоторые датасеты (UK Biobank) требуют платного или ограниченного доступа, что может затруднить независимую верификацию результатов.
   * **Отсутствие рандомизации**: Как обсервационное исследование, работа не позволяет установить причинно-следственную связь, только ассоциации.