

Паспорт набора данных

1. Идентификатор набора данных

Будет уточнен.

2. Версия набора данных

1.0

3. Наименование набора данных

«Всероссийская база данных диагностики запора у детей»

4. Описание набора данных.

4.1. Аннотация

Датасет содержит клинико-anamнестические данные 2250 детей, собранные на основании анкетирования. Данные представлены в виде структурированной таблицы и содержат следующую информацию: возраст, пол, распределение по городам проживания на территории РФ, наличие или отсутствие различных симптомов хронического запора, признаков осложнения запора, данные семейного анамнеза, данные проводимого обследования и лечения, использование диеты, влияние запора на качество жизни детей, данные инструментальных исследований (УЗИ ОБП, ЭГДС, ректороманоскопия, периферическая электрогастроэнтерография (ПЭГЭГ) (у 60 пациентов).

1.2. Клиническая задача

Актуальной клинической задачей для здравоохранения является быстрая диагностика и своевременное лечение запоров, особенно у детей. Это связано с высокой распространенностью заболевания среди детей и его влияния на качество жизни самих пациентов и всей семьи. Также следует отметить немаловажную роль возникающих осложнений запора в случае отсутствия лечения. Система автоматической стратификации симптомов запора позволит своевременно выявить таких детей и вовремя оказать им персонализированную помощь.

1.3. Назначение набора данных

Датасет предназначен для обучения и тестирования моделей машинного обучения, решающих задачу отнесения пациента к группе больных запором или к группе риска по развитию осложнений запора на основании информации, получаемой к окончанию первичного приёма у педиатра или гастроэнтеролога.

1.4. Нозологии

Z00 - Общий осмотр и обследование лиц, не имеющих жалоб или установленного диагноза

K59.0 – Запор

K59.8 - Другие уточненные функциональные кишечные нарушения

K59.9 - Функциональное нарушение кишечника неуточненное

K60.0 - Острая трещина заднего прохода

K60.1 - Хроническая трещина заднего прохода

K62.5 - Кровотечение из заднего прохода и прямой кишки

R15 - Недержание кала

5. Владелец набора данных

ФГАОУ ВО Первый МГМУ им. И.М. Сеченова Минздрава России
(Сеченовский Университет)

5.1. Контактные данные

rectorat@staff.sechenov.ru

5.2. Авторы

Мацукатова Белла Одиссеевна

Эрдес Светлана Ильинична

6. Порядок предоставления доступа к набору данных

Будет уточнен.

7. Тэги

ДАТАСЕТ, ЕДИНАЯ ГОСУДАРСТВЕННАЯ ИНФОРМАЦИОННАЯ СИСТЕМА В СФЕРЕ ЗДРАВООХРАНЕНИЯ, МЕДИЦИНСКАЯ ИНФОРМАЦИОННАЯ СИСТЕМА, РАЗМЕТКА ДАННЫХ, ПЕДИАТРИЯ, ГАСТРОЭНТЕРОЛОГИЯ, ЗАПОР, ЭПИДЕМИОЛОГИЯ, РАСПРОСТРАНЕННОСТЬ, ОСЛОЖНЕНИЯ, ДИАГНОСТИКА, ЛЕЧЕНИЕ, НЕЙРОННЫЕ СЕТИ

8. Источник данных

ФГАОУ ВО Первый МГМУ им. И.М. Сеченова Минздрава России (Сеченовский Университет)

8.1. Глубина набора данных

Использованы данные анкетирования детей или их родителей во время визита к участковому педиатру в детских поликлиниках Центрального, Северо-Западного, Южного, Сибирского и Уральского Федеральных Округов Российской Федерации, а также находящихся на стационарном лечении в гастроэнтерологическом отделении Клиники детских болезней Сеченовского центра материнства и детства, на период с 01.2010 по 01.2015 гг.

8.2. Объем набора данных

Всего в состав датасета включены данные анкет 2250 детей, дополнительно у 60 детей, находившихся на стационарном лечении, имеются данные инструментальных обследований.

8.3. Критерии включения пациентов и исследований

Критерии включения:

- подписание информированного согласия на участие в исследовании;
- возраст от 0 до 18 лет;
- дети, пришедшие с родителями или одни с визитом к участковому педиатру в детских поликлиниках.

Дополнительные критерии включения для стационарных пациентов:

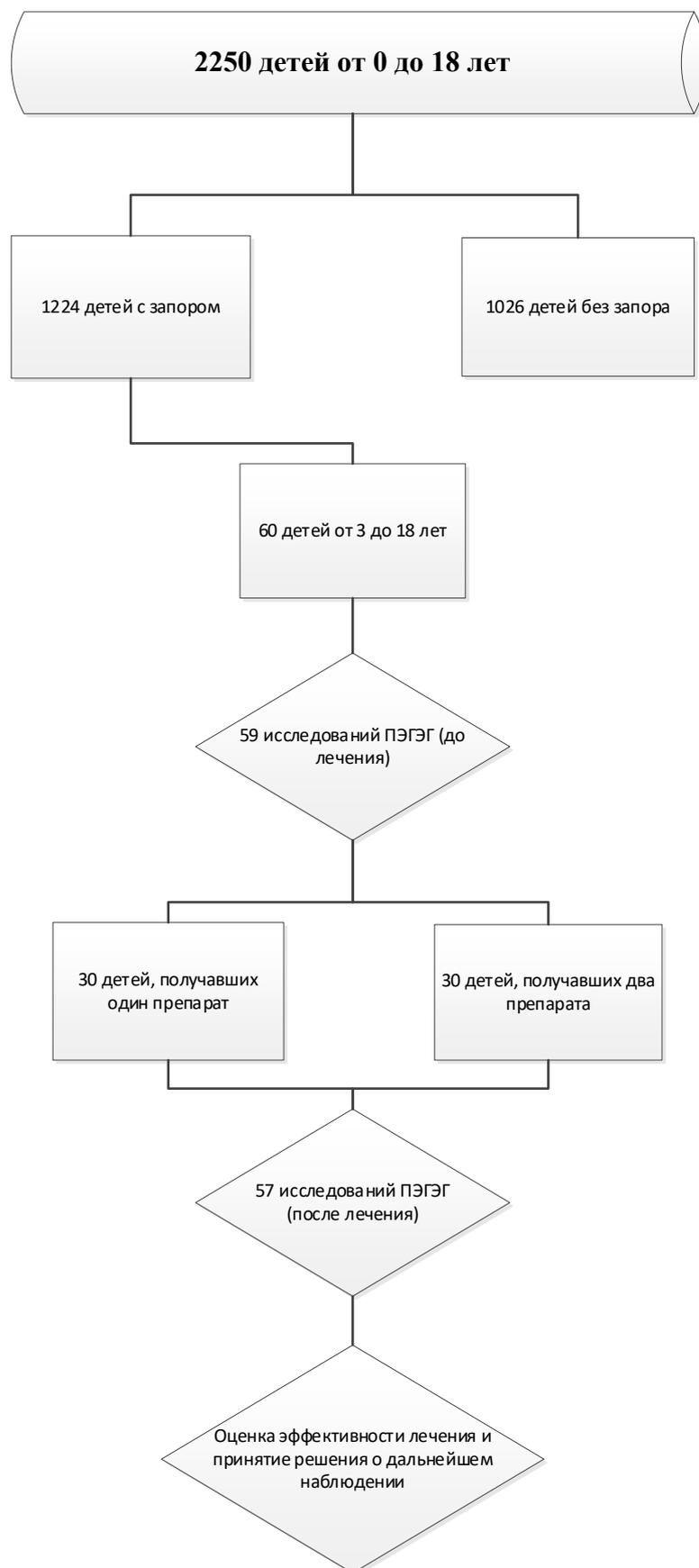
- возраст от 3 до 18 лет;
- сочетание двух и более жалоб:

1. менее 3 дефекаций в неделю;
 2. затруднение в процессе дефекации в более чем 25% случаев;
 3. вздутие живота или твердый кал в более чем 25% случаев дефекации;
 4. ощущение незавершенной эвакуации в более чем 25% случаев дефекации;
 5. чувство аноректальной обструкции/блокады в более чем 25% случаев дефекации;
 6. мануальная помощь в более чем 25% случаев дефекации;
- отсутствие органических, метаболических и эндокринных расстройств.

Критерии исключения:

- пациенты, не давшие согласия на участие в исследовании;
- пациенты старше 18 лет.

8.4. Особенности проведения исследований



8.5. Половозрастные характеристики набора данных

1. Соотношение пациентов по полу: мальчики - 1057, девочки - 1193

2. Возраст:

- минимальный – 1 месяц
- максимальный – 18 лет
- медиана – 6,5 лет

8.6. Характеристики оборудования

В случае стационарного наблюдения за пациентами можно использовать аппаратуру для проведения УЗИ органов брюшной полости (аппараты ультразвуковой диагностики), эндоскопических исследований (тубусные и гибкие фиброскопы), суточной периферической электрогастроэнтерографии (гастроэнтеромонитор ГЭМ-01 «Гастроскан-ГЭМ»).

8.7. Дата публикации набора данных

Будет уточнена.

8.8. Дата обновления набора данных

Будет уточнена.

9. Модели машинного обучения

Нейронные сети.

10. Расположение набора данных

Будет уточнено. Набор данных расположен в Единой государственной информационной системе в сфере здравоохранения.

11. Структура набора данных

11.1. Схема структуры директорий и файлов

Датасет представлен одним файлом электронной таблицы Microsoft Excel 2007 формата .xls размером 473 КБ.

11.2. Описание структуры и принципа наименования файлов и директорий

Файл содержит данные в «сыром виде» в текстовом формате с разбиением на строки и столбцы.

12. Обзор разметки данных (ретроспективная разметка данных, предоставленных в «сыром виде»). («Эталонный датасет»).

12.1. Особенности формата разметки

Ретроспективная разметка пациентов, пришедших к участковому педиатру в детских поликлиниках. Перечень метаданных: идентификационный номер, возраст, пол, город проживания, симптомы запора до лечения, симптомы осложнения запора, данные по питанию и лечению запора в анамнезе, влияние запора на качество жизни, данные инструментальных исследований, наличие сопутствующих заболеваний, данные ПЭГЭГ до лечения, симптомы запора после лечения, данные ПЭГЭГ после лечения.

12.2. Классы разметки

Будут уточнены.

12.3. Принципы разметки

Будут уточнены.

12.4. Принципы верификации разметки

Будут уточнены.

12.5. Статистика использования лейблов и классов

Будут уточнены.

13. Правила использования и распространения

13.1. Копирайт

ФГАОУ ВО Первый МГМУ им. И.М. Сеченова Минздрава России
(Сеченовский Университет)