

Паспорт набора данных

1. Идентификатор набора данных: будет уточнено.
2. Версия набора данных: 1.0.
3. Наименование набора данных: «Набор клинических и лабораторно-инструментальных данных о пациентах прооперированных по поводу местно-распространенных и осложненных форм рака дистального отдела желудка».
4. Описание набора данных:

Аннотация: для создания дата сета использовались данные из медицинской документации, содержащие гендерно-возрастные показатели, объективизированную оценку соматического статуса пациента, лабораторные и инструментальные методы диагностики, стадирование основного заболевания, факторы прогноза, варианты лечения и результаты патоморфологического исследования пациентов с местно-распространенными и осложненными формами рака желудка, а также их влияние на непосредственные и отдаленные результаты. На лимфоузлах размечены области здоровой ткани, области метастатического поражения и области артефактов.

Клиническая задача: в соответствии с клиническими рекомендациями, при лечении осложненных форм рака желудка, нет четких показаний при лечении пациентов на продвинутых стадиях болезни с учетом тяжести состояния, коморбидности и вида осложнений. Исходя из вышеперечисленного, на основании накопленных данных есть необходимость создания персонализированного алгоритма лечения для данной когорты пациентов. Системы поддержки принятия врачебного решения (на основе искусственного интеллекта и машинной обработки полученных данных) могут позволить автоматически определять наиболее оптимальный метод лечения и прогноза, в зависимости от конкретной клинической ситуации. В перспективе, такое решение должно положительно повлиять на выбор хирургической тактики и алгоритма лечения по каждому пациенту.

Назначение набора данных: датасет предназначен для разработки ML-решений и иных методов математической обработки данных для построения наиболее эффективного алгоритма лечения пациентов с местно-распространенными и осложненными формами рака желудка. Датасет разбит на два набора: основной (обучающий) набор данных и фиксированный набор данных.

Нозологии: рак желудка (C16.3 преддверия привратника; C16.4 привратника; C16.5 малой кривизны желудка неуточненной части; C16.6 большой кривизны желудка неуточненной части; C16.8 поражение желудка, выходящее за пределы одной и более вышеуказанных локализаций).

5. Владелец набора данных: ФГАОУ ВО Первый МГМУ им. И. М. Сеченова Минздрава России (Сеченовский Университет).

Контактные данные: будут уточнены.

Авторы: Спартак Алексей Андреевич, Хоробрых Татьяна Витальевна, Ивашов Иван Валерьевич, Андриянов Андрей Сергеевич.

6. Порядок предоставления доступа к набору данных: будет уточнен.
7. Тэги: датасет, единая государственная информационная система в сфере здравоохранения, медицинская информационная система, разметка данных, персонализированный подход в лечении, хирургическое лечение, онкология, рак желудка, осложненные формы рака желудка.

8. Источник данных

Глубина набора данных: выгрузка из архива Университетских клинических больниц (Сеченовский Университет) до 2020 г. включительно.

Объем набора данных: выборка из 100 пациентов, разделенных на две группы (группа лапароскопического доступа — 46 пациентов; группа открытого доступа — 54 пациента).

Критерии отбора клинических параметров и пациентов им соответствующих, а также виды лабораторно-инструментальных исследований

Критерии включения в исследование:

1. Письменное информированное согласие пациента на участие в исследовании;
2. Пациенты с местно-распространенным и диссеминированным раком дистального отдела желудка, и наличием одного или нескольких осложнений, требующих хирургического лечения (опухолевый стеноз, рецидивирующее опухолевое кровотечение);
3. Пациенты с раком дистального отдела желудка и наличием потенциально резектабельных метастатических очагов или поражений (печень, единичные очаги на париетальной брюшине, лимфоузлы 3 порядка);
4. Патоморфологическая верификация первичной опухоли;
5. Возраст пациентов – от 18 лет до 85 лет;
6. Общее состояние пациента по шкале ECOG – ВОЗ не более 3, по шкале Карновского > 50%; по шкале анестезиологического риска (ASA) ≤ 3 .

Алгоритм обследования пациента состоял из следующих пунктов:

1. Физикальное обследование, ректальное исследование;
2. Патоморфологические и цитологические исследования;
3. Лабораторные методы исследования: клинический анализ крови; биохимический анализ крови; коагулограмма; клинический анализ мочи; определения группы крови; резус-фактора и фенотипа; определение уровня онкологических маркеров;
4. Инструментальные методы исследования:
 - a) Эндоскопические: эзофагогастродуоденоскопия с биопсией и татуажом границ опухоли; колоноскопия;
 - b) Рентгенологические методы: рентгенография органов грудной клетки; рентгенологическое исследование желудка с водорастворимым контрастом; мультиспиральная компьютерная томография органов грудной клетки, брюшной полости и малого таза с внутривенным контрастированием и построением 3D-модели сосудистой архитектоники; МРТ органов брюшной полости, малого таза (по показаниям);
 - c) Ультразвуковые исследования: УЗИ органов брюшной полости, забрюшинного пространства, органов малого таза, лимфатических узлов; УЗДГ сосудов шеи; УЗДГ вен нижних конечностей; ЭХО-КГ (всем пациентам старше 45 лет и с сопутствующей сердечно-сосудистой патологией);
 - d) Дополнительные методы диагностики: электрокардиография; суточное мониторирование артериального давления и ЭКГ при наличии показаний; определение функции внешнего дыхания;
5. Консультации врачей других специальностей: терапевта; кардиолога, уролога, гинеколога, эндокринолога, клинического фармаколога (по показаниям).

Особенности подготовки исследований:

Поло-возрастные характеристики набора данных

Основная группа включала 46 пациентов, из которых: 27 (58,7%) мужчины и 19 (41,3%) женщины в возрасте от 37 до 80 лет. Средний возраст составил $64,8 \pm 9,71$ лет.

Группа сравнения состояла из 54 пациентов: 22 (40,7%) мужчин и 32 (59,3%) женщин в возрасте от 33 до 85 лет. Средний возраст составил $65,17 \pm 12,43$ лет.

Характеристики оборудования

Лапароскопические операции выполняли с использованием операционных видеоэндоскопических стоек фирмы «Karl Storz» включающих телекамеру, видеопроцессор, монитор, автоматический инсуффлятор углекислого газа, осветитель, система ирригации, аспирации, электрохирургический блок.

Дата публикации набора данных: 2023 г.

Дата обновления набора данных: 2023 г.

9. Модели машинного обучения: искусственные нейронные сети.

10. Расположение набора данных: набор данных расположен в Единой государственной информационной системе в сфере здравоохранения.

11. Структура набора данных

Схема структуры директорий и файлов: электронная таблица (Excel).

Описание структуры и принципа наименования файлов и директорий.

12. Обзор разметки данных: использовался горизонтальный способ внесения данных в таблицу. Для применения дополнительных методов статистического анализа, использования машинного обучения для прогнозирования тех или иных событий, может быть произведена дополнительная разметка данных совместно с математиками и статистиками.

13. Правила использования и распространения

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