

Прогнозирование исхода онкологических заболеваний при проведении CAR-T клеточной терапии с помощью топологического анализа данных

Выполнила:

Пунько Анастасия Владимировна

Анализ данных в биологии и медицине

Научный руководитель:

Чернышев Всеволод Леонидович

Июнь, 2023

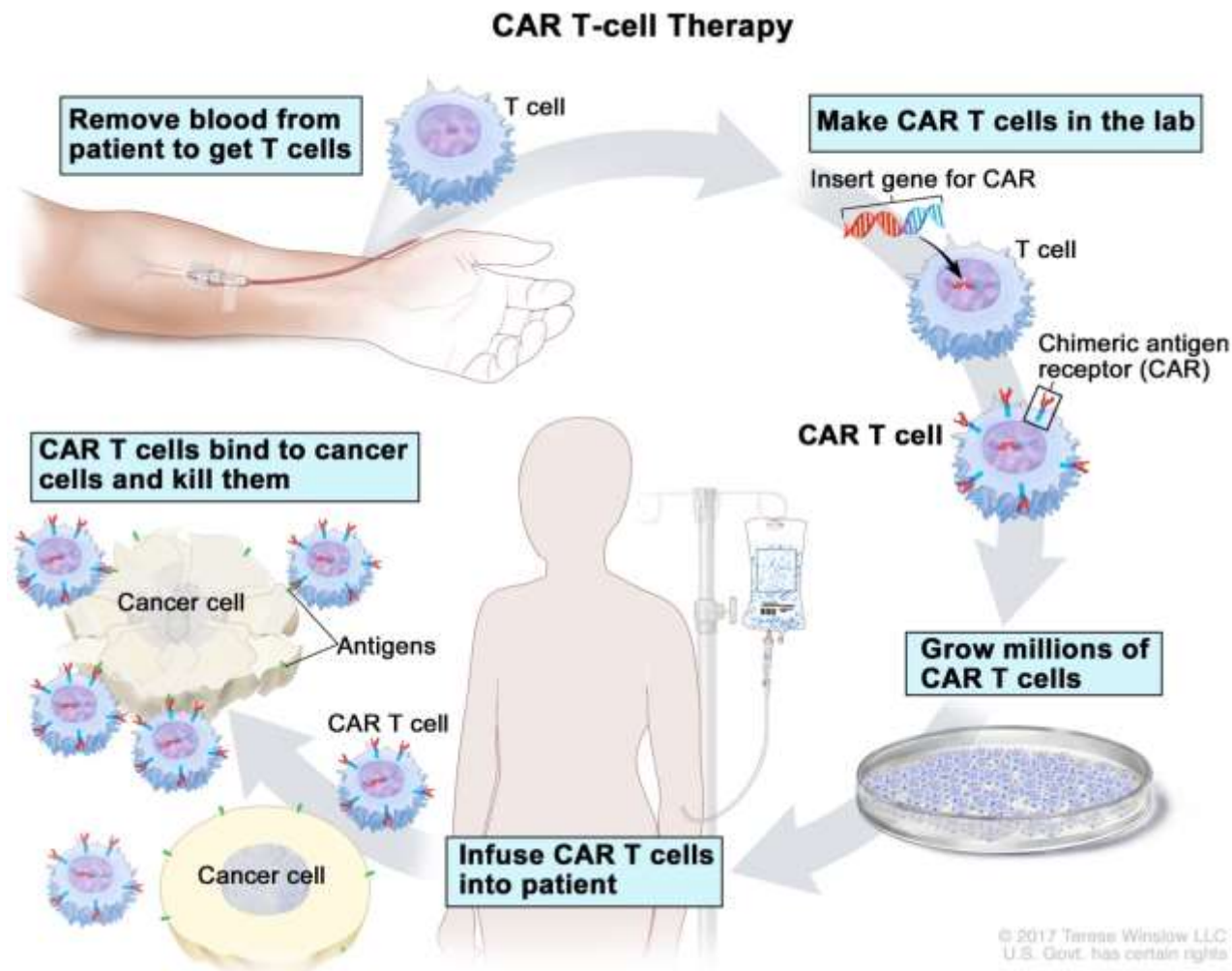
Содержание

1. Актуальность работы
2. Цели работы
3. Описание данных
4. Методы топологического анализа данных
5. Результаты
6. Заключение

Актуальность работы

- ❖ Онкологические заболевания являются главной причиной смерти и серьезным препятствием на пути увеличения продолжительности жизни.
- ❖ Прогнозируется, что в 2040 году во всем мире будет зарегистрировано около 28,4 миллиона новых случаев рака [Chhikara B. S., 2022].
- ❖ На сегодняшний день иммунотерапия является важной частью лечения некоторых видов рака [Abbott M., 2019].
- ❖ Для проведения CAR-T терапии проводится лейкофerez и разделение клеточных популяций. Далее Т-клетки модифицируют, добавляя ген специфического химерного антигенного рецептора. И получают достаточное количество CAR-T клеток, выращивая их в лабораторных условиях. После этого CAR-T клетки вливают пациенту [Sermer D., 2019].

Актуальность работы



Цели работы

- ❖ Применение топологического анализа данных для прогнозирования ответа на иммунотерапию CAR T-клетками у пациентов с гематологическими заболеваниями.
- ❖ Применение топологического анализа данных для прогнозирования синдрома высвобождения цитокинов после проведения иммунотерапии CAR T-клетками у пациентов с острым лимфобластным лейкозом.

Описание данных

2 набора данных

Для прогнозирования ответа
(CR/SD,PD,PR,NR)

Для прогнозирования синдрома
высвобождения цитокинов

Данные пациентов со злокачественными
гематологическими заболеваниями,
которые получили лечение с помощью
CAR T-клеток
(данные о вводимой дозе CAR T-клеток,
индикация, возраст и пол пациентов).

54 пациента

CR (complete response) – полный ответ,
SD (stable disease) – стабильная болезнь,
PD (progressive disease) – прогрессивное заболевание,
PR (partial remission) – частичная ремиссия,
NR (no response) – отсутствие ответа.

Данные пациентов с лейкозом, у
которых после проведения CAR T-
терапии возник синдром высвобождения
цитокинов, а также данные здоровых
доноров, для которых был определен
цитокиновый профиль.

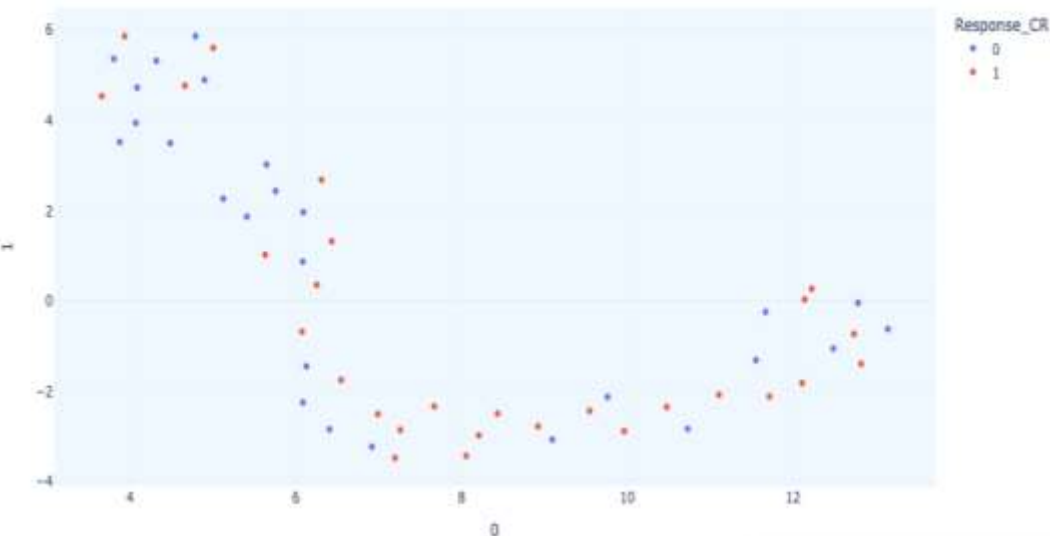
100 пациентов (50 пациентов с CRS и 50 доноров).

Были проанализированы следующие цитокины:
IL2, IL4, IL5, IL6, IL7, IL8, IL10, IL12, IL13, IL15, IL17, sIL2R α ,
MCP1, EGF, IL1RA, sIL1RI, sIL1RII, sIL4R, sIL6R, sgp130,
TNF α , IFN α , IFN γ , MIP1 α , MIP1 β , MIG, GCSF, GM-CSF, IP10.

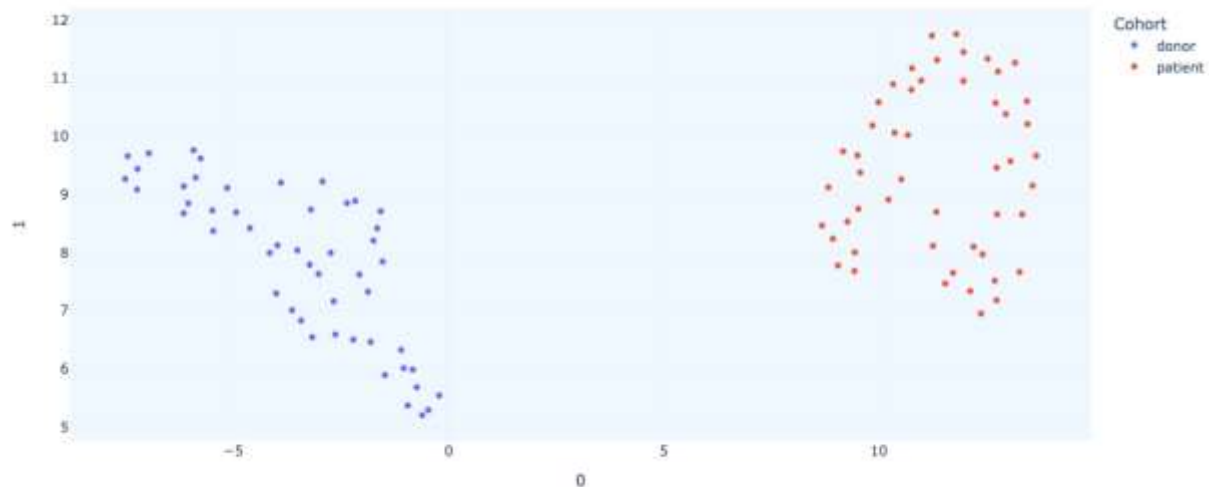
Методы топологического анализа данных

- ❖ Топологический анализ данных включает в себя набор инструментов для визуализации, исследования и анализа данных, основанных на топологии (область математики, изучающая абстрактные понятия формы и связности).
- ❖ Главная идея топологического анализа данных заключается в использовании топологии и геометрии для получения качественной и часто количественной информации о структуре данных.
- ❖ Топологический анализ данных:
 - Метод устойчивых гомологий
 - Метод Mapper
 - Метод RTD-AE

Результаты



Визуализация набора данных
методом UMAP в 2D.



Результаты

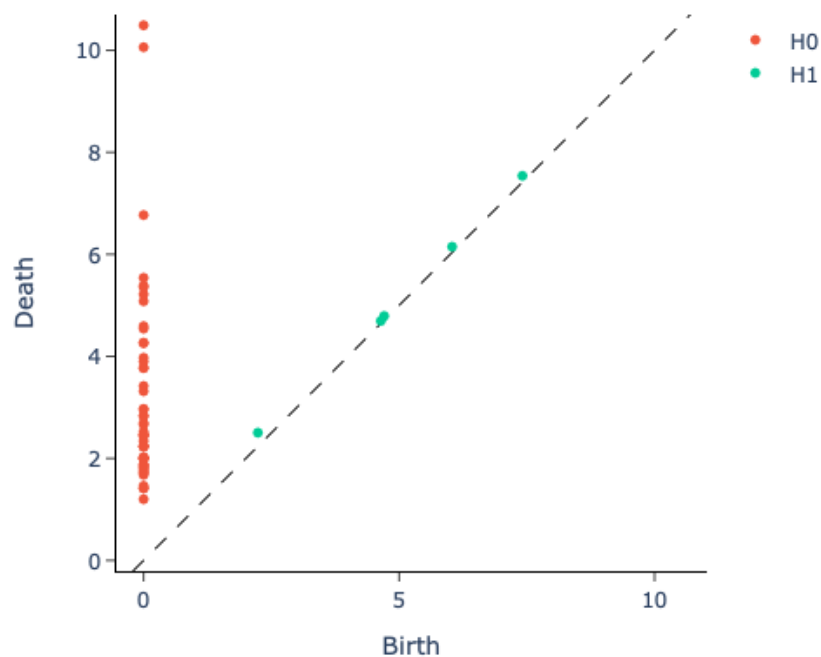


Диаграмма устойчивости для параметра
ответа на терапию.

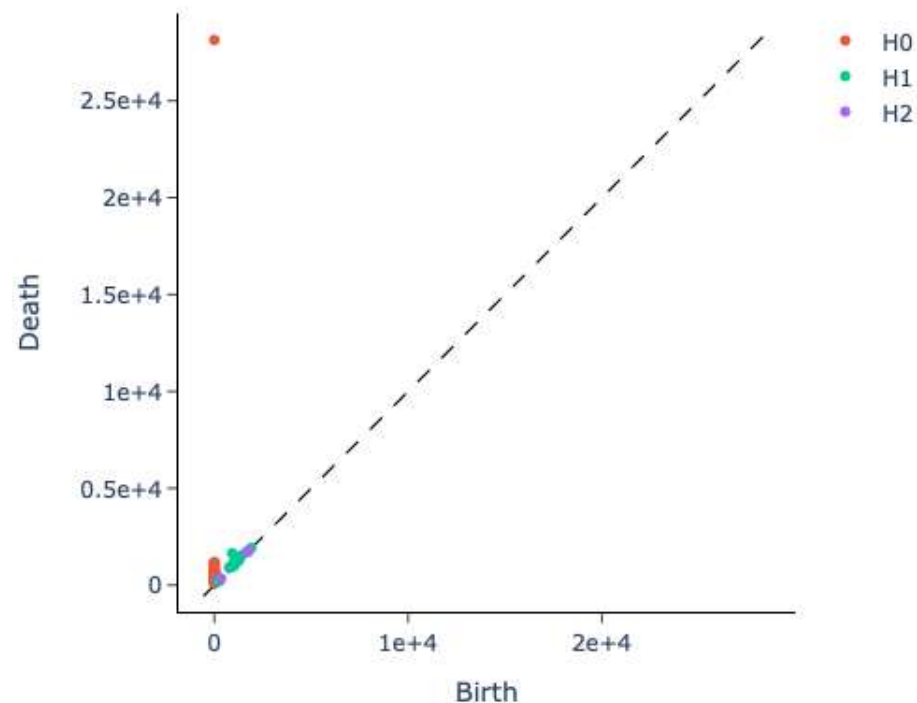
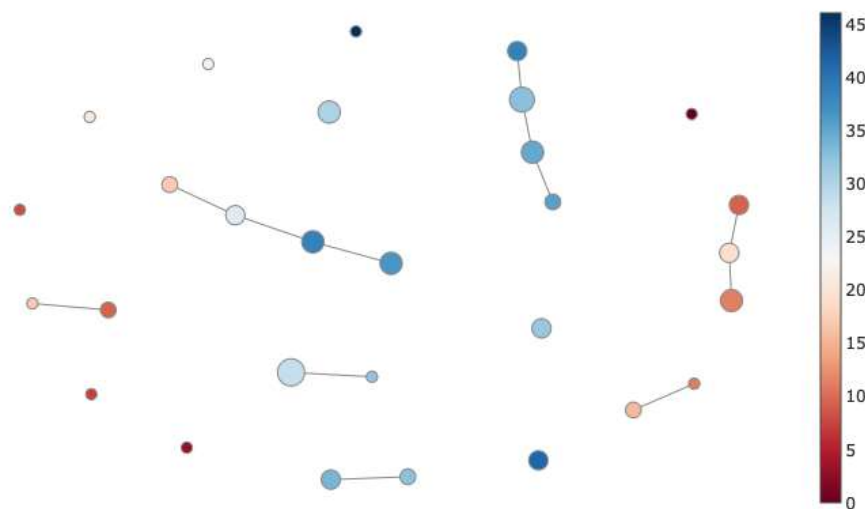
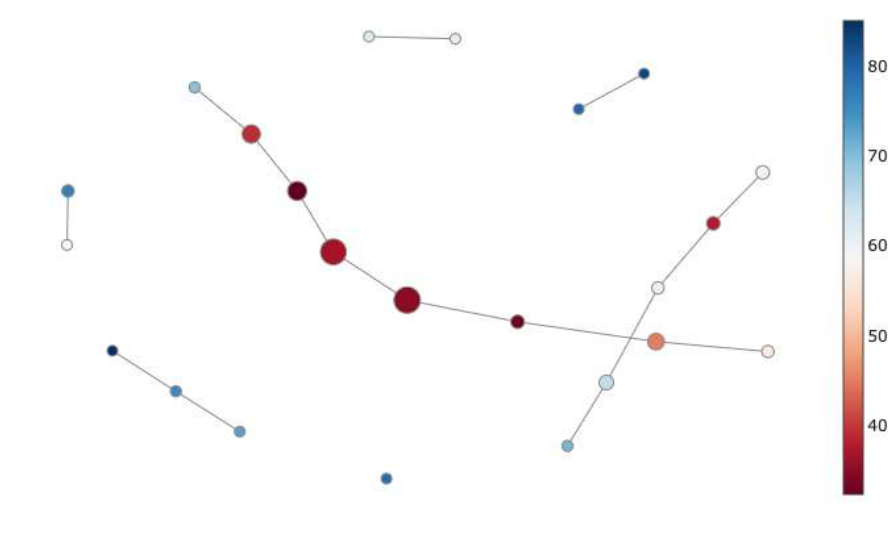


Диаграмма устойчивости для параметра
Когорта (пациент/здоровый донор).

Результаты



Граф, построенный методом Марпер, для кластеризации данных по виду ответа на терапию.



Граф, построенный методом Марпер, для кластеризации данных по цитокиновому профилю пациентов с CRS и здоровых доноров.

Заключение

- ❖ При исследовании доз CAR T-клеток, которые получили пациенты с гематологическими опухолями, топологические структуры не были обнаружены, однако при сравнении метрик модели классификации было показано, что точность классификации с использованием топологических данных составляет 57.1%, а при использовании стандартного алгоритма точность падает до 42.9%.
- ❖ При исследовании цитокинового профиля пациентов с CRS и здоровых доноров было обнаружено, что данные классифицируются на группы с точностью 100% даже стандартными методами.



Результаты

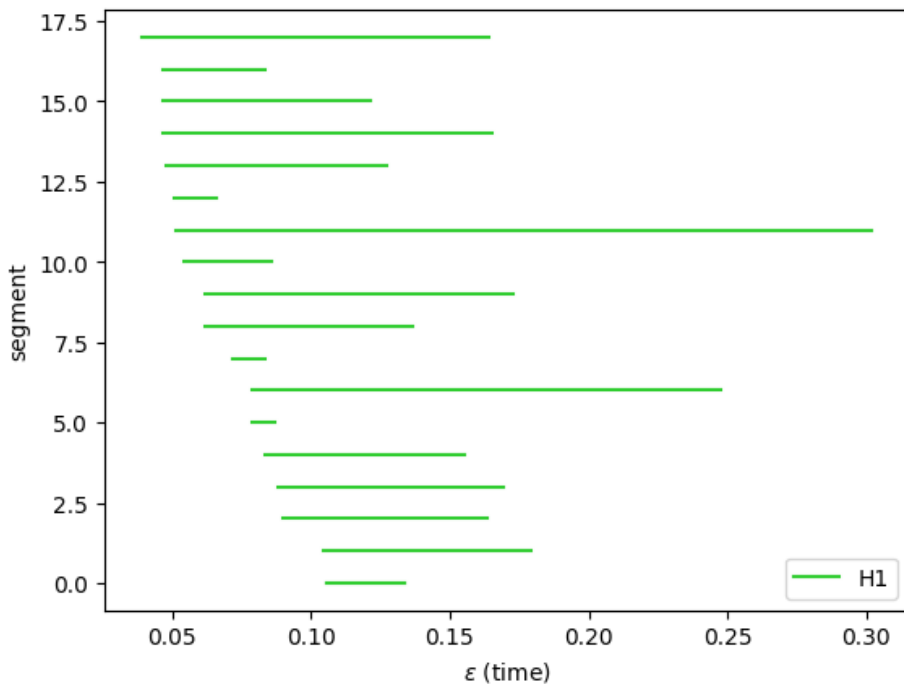


График R-Cross-Barcode для данных ответа на терапию.

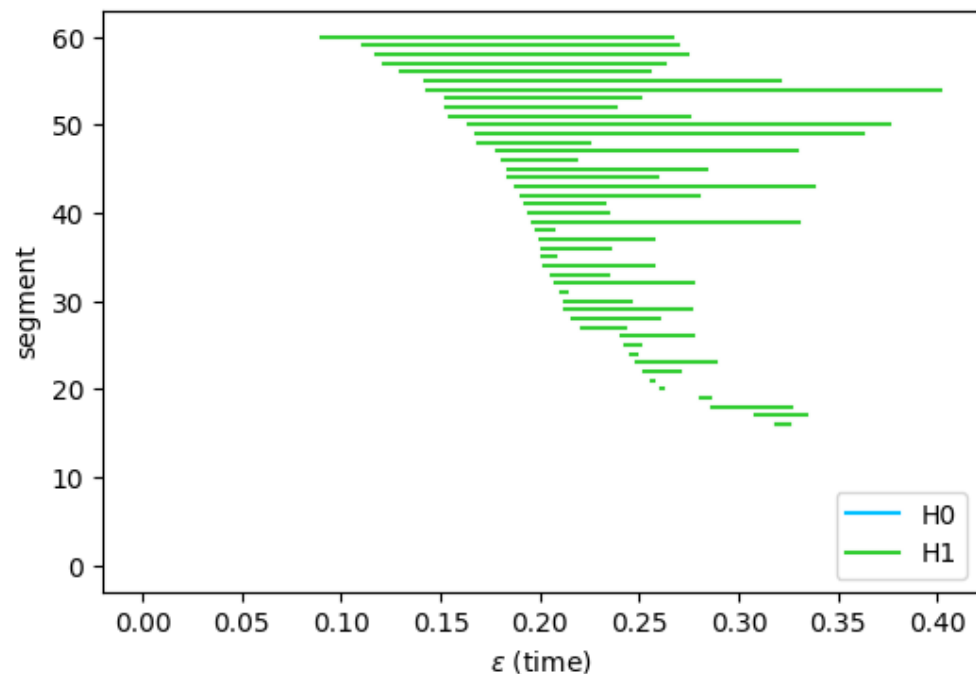


График R-Cross-Barcode для данных цитокинового профиля пациентов с CRS и здоровых доноров.

Выводы и заключение

- ❖ Освоены алгоритмы топологического анализа данных: устойчивые гомологии, Mapper и RTD-AE.
- ❖ Построены диаграммы устойчивости, кривые Бетти, также были сопоставлены метрики модели классификации RandomForest с использованием топологических данных и стандартного алгоритма.
- ❖ Для отображения основных особенностей данных и для кластеризации данных был получен топологический граф с помощью алгоритма Mapper.
- ❖ В результате применения алгоритма RTD-AE построен график R-Cross-Barcode, который измеряет различия топологии двух облаков точек, а также определено значение RTD.