

МОДЕЛИ И МЕТОДЫ АНАЛИЗА КЛИНИЧЕСКИХ И ЛАБОРАТОРНЫХ ПОКАЗАТЕЛЕЙ НА РАЗНЫХ ВОЛНАХ ПАНДЕМИИ НА ПРИМЕРЕ COVID-19

Порунова А.П.¹ Сафронова Е.В.²

¹ НИ ТГУ, ИПМКН, 932228, e-mail: nastya.p-va@yandex.ru

² НИ ТПУ, ИШИТР, аспирант, e-mail: ev.kashcheeva@mail.ru

Введение

Выявленный в 2019 году COVID-19 (Coronavirus disease – 2019) является потенциально тяжелой острой респираторной инфекцией, вызываемой коронавирусом SARS-CoV-2 (2019-nCoV). С 2019 года вирус мутирует и уже известно множество штаммов вируса SARS-CoV-2, эпидемия все еще сохраняет свою остроту [1,2].

Вызываемое коронавирусом SARS-CoV-2 инфекционное заболевание COVID-19 кроме клинических симптомов и синдромов сопровождается изменениями в показателях клинического и биохимического анализа крови. Лабораторные проявления COVID-19 имеют свои особенности и тесно взаимосвязаны с характером течения инфекции. Таким образом, анализ механизмов патогенеза развития COVID-19 по биохимии крови, способствует накоплению новых знаний и дает возможность разработать подходы к комплексному лечению и ранней диагностики новой коронавирусной инфекции и профилактике жизнеугрожающих осложнений.

Основной задачей цифровой медицины является разработка эффективных, быстрых и недорогих альтернативных методов ранней диагностики пациентов с COVID-19. Даже самые опытные врачи могут интерпретировать лишь малую часть информации, содержащейся в обычных результатах лабораторных анализов крови, в том числе в результатах биохимии. Кроме того, на их основе чрезвычайно трудно определить тяжесть состояния пациентов с COVID-19[3]. В этом контексте модели классификации ML, работающие с данными на основе биохимии для определения предварительного диагноза COVID-19, могут быть эффективным инструментом в системах поддержки принятия клинических решений.

Целью данной работы является подготовить и проанализировать данные клинических и лабораторных показателей пациентов с COVID-19 для дальнейшего использования.

Для достижения этой цели были поставлены следующие задачи:

- Проанализировать предметную область – описание признаков;
- Получить описание набора данных;
- Проанализировать с помощью статистических критериев и визуализировать показатели биохимии крови трех волн COVID-19.

Описание работы

Для предварительной оценки показателей биохимии крови трех волн COVID-19 были взяты три выборки анализов пациентов 2, 3 и 4 волны.

Каждая из выборок содержат записи 2500 пациентов. Записи (показатели анализов биохимии крови) состояли из столбца данных времени и 19 признаков: Мочевина, Креатинин, Глюкоза, Общий билирубин, АСТ, АЛТ, ЛДГ, Холестерин, D-димер, А-амилаза, Щелочная фосфатаза, С-реактивный белок, КФК-МВ, КФК, Прямой билирубин, Ферритин, Общий белок, Мочевая кислота.

Была проведена предварительная обработка данных:

- Были удалены строки, полностью состоящие из Null-значений. Так же были удалены строки связанные с тестированием оборудования (строки, содержащие не более трех не Null-значений);
- Были удалены выбросы с помощью межквартильного диапазона.

Попарное сравнение среднего значения показателей крови разных волн было проведено с помощью теста Стьюдента с уровнем значимости 0.05. Результаты сравнения приведены в таблице 1.

Таблица 1

Признаки с одинаковым значением среднего в разные волны пандемии

2 и 3 волна	2 и 4 волна	3 и 4 волна
Нет	А-амилаза, Tstat = -1.786 С-реактивный белок, Tstat = 0.742 Холестерин, Tstat = 1.345	Ферритин Tstat, = 1.225 АСТ Tstat, = 0.247 ЛДГ Tstat, = -0.39

	КФК, Tstat = -1.36 АЛТ, Tstat = 0.245	
--	--	--

Так же были найдены признаки различные в разные волны. На рисунке 1 приведены результаты сравнения двух наиболее значимых показателей биохимии крови – С-реактивного белка и D-димера.

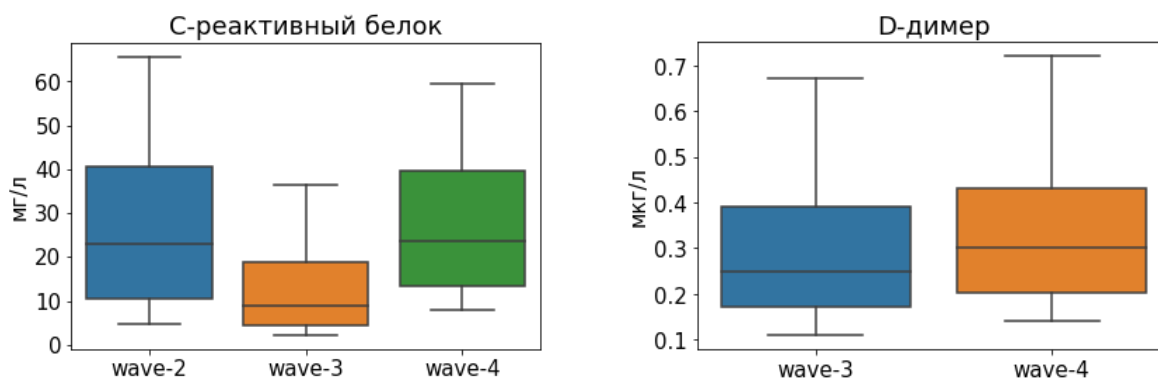


Рис. 1. Сравнения среднего значения С-реактивного белка и D-димера в разные волны пандемии

Далее был проведен анализ динамики изменения показателей в течении 2, 3 и 4 волн COVID-19. Для этого данные были дополнительно обработаны:

- данные времени были приведены к одному виду, так как они были получены с разных записывающих устройств;
- анализы пациентов были отсортированы по дате поступления;
- пропущенные значения показателей были заполнены с помощью линейной интерполяции;
- для более наглядного представления динамики изменения показателей, было произведено сглаживание данных показателей по дням (вычислено скользящее невзвешенное среднее).

Заключение

Была выполнена предварительная оценка показателей разных штаммов COVID-19 с помощью критерия Стьюдента. Наибольшее отличия средних показателей были выявлены между 2 и 3 волнами. Также выявлена динамика изменения клинических и лабораторных показателей на протяжении волн пандемии.

Список использованных источников

1. Zhou C., Increased serum levels of hepcidin and ferritin are associated with severity of COVID-19 / C. Zhou, Y. Chen, Y. Ji, [и др.] // Medical science monitor: international medical journal of experimental and clinical research, 2020. – V. 26.
2. Huyut, M. T. Forecasting of Oxidant/Antioxidant levels of COVID-19 patients by using Expert models with biomarkers used in the Diagnosis/Prognosis of COVID-19 / M. T. Huyut, Z. Huyut // International Immunopharmacology. – 2021. – P. 108127.
3. Zhu, J.S. Deep-learning artificial intelligence analysis of clinical variables predicts mortality in COVID-19 patients / J.S. Zhu, P. Ge, C. Jiang, [и др.] // Journal of the American College of Emergency Physicians Open. - № 6. - 2020. - P. 1364-1373.