

**ОПРЕДЕЛЕНИЕ ВАЖНОСТИ ПРИЗНАКОВ ПАЦИЕНТОВ ДЛЯ ОЦЕНКИ ВЕРОЯТНОСТИ
РЕЗИСТЕНТНОСТИ К ПРОТИВОТУБЕРКУЛЕЗНЫМ ЛЕКАРСТВЕННЫМ СРЕДСТВАМ**Е.В. Кашеева

Научный руководитель: доцент, к.т.н. С.В. Аксёнов

Национальный исследовательский Томский политехнический университет,

Россия, г. Томск, пр. Ленина, 30, 634050

E-mail: ev.kashcheeva@mail.ru**DETERMINATION OF THE IMPORTANCE OF PATIENTS SIGNS FOR EVALUATING THE
PROBABILITY OF RESISTANCE TO THE ANTITUBERCULOSE MEDICINES**E.V. Kashcheeva

Scientific Supervisor: Assistant Professor, Candidate of Technical Sciences S.V. Axyonov

Tomsk Polytechnic University, Russia, Tomsk, Lenin str., 30, 634050

E-mail: ev.kashcheeva@mail.ru

***Abstract.** This article describes the steps involved in preparing and analyzing data about patients with a diagnosis or suspicion of multidrug-resistant tuberculosis. A classifier was built and the importance of indicators was calculated for women and men, as well as for different age groups.*

Введение. Одной из десяти ведущих причин смерти в мире, по данным всемирной организации здравоохранения, является туберкулез [1]. Туберкулез это инфекционное заболевание человека и животных, поражающее лёгкие. Лечение туберкулеза осуществляется с помощью лекарственных средств или хирургическим путем. Однако не каждый препарат способен оказать влияние на вирус, существует такое понятие как лекарственная резистентность, что означает устойчивость организма к лекарственным средствам. Выделяют пять типов лекарственной резистентности к туберкулезу [2], определив наличие резистентности, специалисты медицинских учреждений подбирают соответствующее лечение. Было решено определить важность признаков пациентов по отношению к наличию резистентности к противотуберкулезным препаратам в зависимости от возраста и половой принадлежности пациентов.

Подготовка данных к анализу. Данные для анализа были взяты на белорусском портале по туберкулезу [3]. Информация о пациентах, поступивших в отделение МЛУ–ТБ РНПЦ пульмонологии и туберкулеза с диагнозом или подозрением на туберкулез с множественной лекарственной устойчивостью, хранилась в 412 файлах (отдельно по каждому пациенту). Для проведения анализа данных возникла необходимость в объединении информации из всех файлов в один. Так, было проведено конвертирование структуры данных и объединение информации из 412 файлов в один.

Для анализа решили использовать следующие показатели: результаты предыдущего лечения, результаты анализов мазка и культуры, наличие сопутствующих заболеваний, хирургическое вмешательство, мезенхимальная терапия стволовыми клетками, рост, вес, инвалидность, симптомы туберкулеза, образование, жилье, бывший заключенный, вредные привычки, работа, ВИЧ, аллергия, аномалии сосудов средостения, аномалии легких, бронхиальная обструкция, распространение, уменьшение объема легких, лимфаденопатия, нодикальцинатум, плеврит, пневмоторакс.

Анализ данных. Для начала определили количество пациентов в зависимости от возраста и половой принадлежности (Рис. 1). Выделили 4 возрастные группы, выяснили, что из 150 женщин 77 в возрасте 19 – 35, 49 в возрасте 36 – 60, 16 в возрасте старше 60 и 8 в возрасте 0 – 18. Что касается мужчин, то из 262 пациентов 152 в возрасте 36 – 60, 90 в возрасте 19 – 35, 19 в возрасте старше 60 и 1 в возрасте 0 – 18.

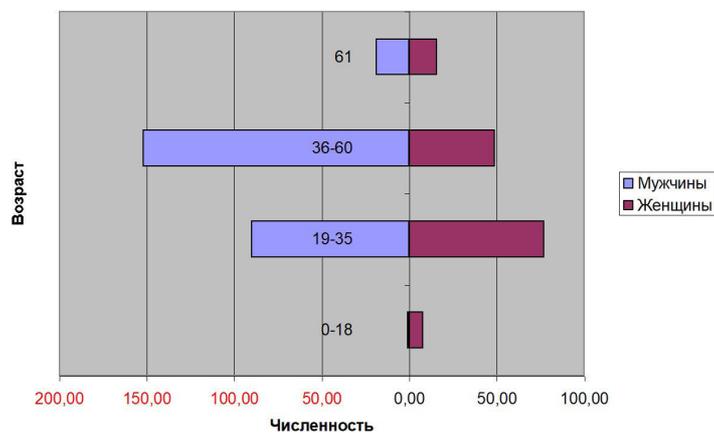


Рис. 1. Диаграмма «Количество пациентов в зависимости от возраста и половой принадлежности»

Затем был построен классификатор RandomForestClassifier и рассчитана важность признаков пациентов, для начала рассчитали важность для мужчин и для женщин (Рис. 2). Выяснили, что наиболее важным для обоих полов является наличие сопутствующих заболеваний, на втором и третьем местах у мужчин находятся наличие инвалидности и вес, а у женщин данные признаки занимают третье и второе места по важности соответственно. Затем следуют такие показатели как результат предыдущего лечения (рецидив) (4 место для обоих полов), стиль жизни (наркоман) и лимфаденопатия.

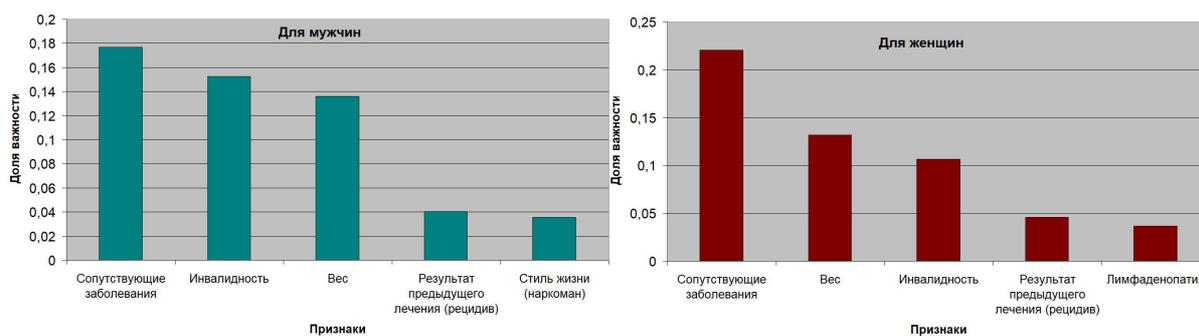


Рис. 2. Диаграммы «Важность признаков пациентов для мужчин и для женщин»

Рассчитали важность признаков пациентов для выделенных ранее возрастных групп (Рис. 3). Для группы от 0 до 18 лет доля важности принимает значения от 0,062 до 0,064, это можно обосновать тем, что в данную возрастную группу входят всего 9 пациентов и на основе такого небольшого количества данных невозможно определить важность признаков для оценки вероятности резистентности. В возрастной группе от 19 до 35 наибольшей долей важности обладает наличие сопутствующих заболеваний, затем идут вес, инвалидность, результат предыдущего лечения (рецидив) и симптомы туберкулеза. Для возрастной группы от 36 до 60 доли важности признаков расположились в порядке

убывания следующим образом: сопутствующие заболевания, вес, инвалидность, стиль жизни (алкоголизм), результат культуры. Что касается последней возрастной группы, в которую были включены пациенты с возрастом старше 60 лет, то наибольшей долей важности обладает такой признак как инвалидность, за ним следуют сопутствующие заболевания, вес, нодикальцинатум и плеврит.

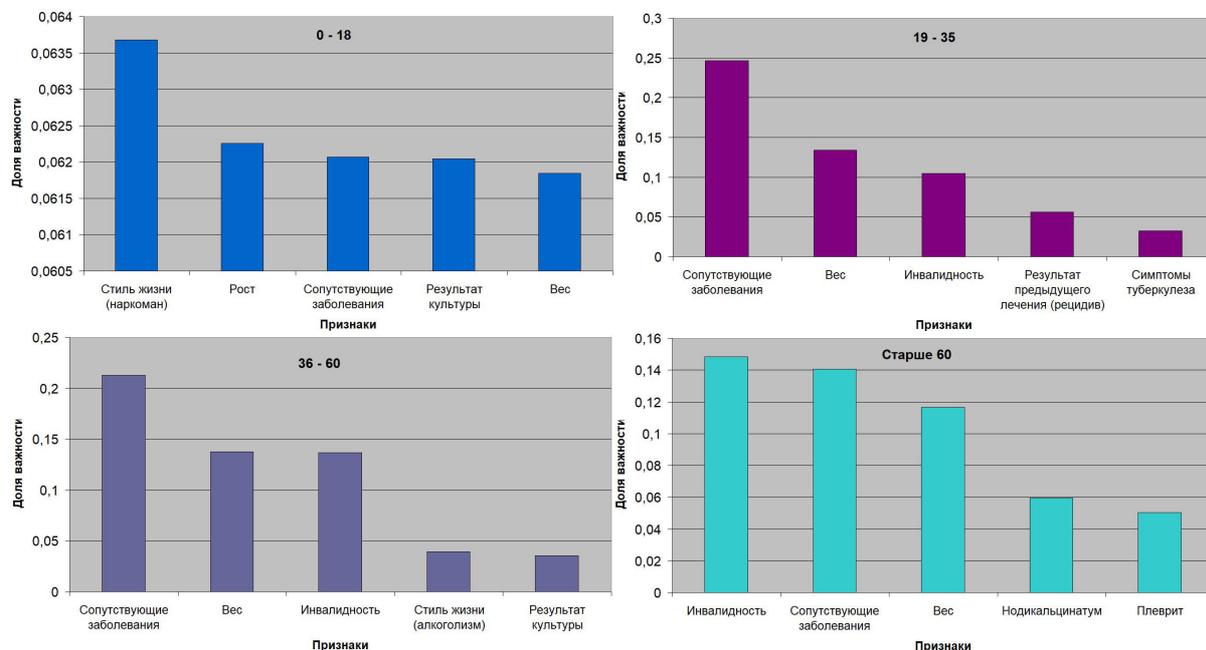


Рис. 3. Диаграммы «Важность признаков пациентов для четырех возрастных категорий»

Заключение. В результате выполнения работы была проведена подготовка данных к анализу, построен классификатор RandomForestClassifier, рассчитаны важности показателей. Выяснили, что для обоих полов для оценки вероятности резистентности к противотуберкулезным лекарственным средствам тремя наиболее важными признаками являются сопутствующие заболевания, инвалидность и вес. Данные признаки находятся в числе первых и для разных возрастных групп. Таким образом, вне зависимости от возраста и половой принадлежности сопутствующие заболевания, инвалидность и вес позволят оценить вероятность резистентности к противотуберкулезным препаратам.

СПИСОК ЛИТЕРАТУРЫ

1. Основные факты о туберкулёзе [Электронный ресурс]: официальный сайт Всемирной организации здравоохранения. – Режим доступа: <https://www.who.int/ru/news-room/fact-sheets/detail/tuberculosis> (дата обращения 9.03.2019).
2. Типы лекарственной устойчивости [Электронный ресурс]: официальный сайт Всемирной организации здравоохранения. – Режим доступа: <https://www.who.int/tb/areas-of-work/drug-resistant-tb/types/en/> (дата обращения 9.03.2019).
3. Данные о пациентах, поступивших в отделение МЛУ–ТБ РНПЦ пульмонологии и туберкулеза [Электронный ресурс]: Белорусский портал по туберкулезу. – Режим доступа: <https://data.tuberculosis.by/> (дата обращения 9.03.2019).