

Raport z danych pacjentów

Abstract

WSTĘP

Rak płuc jest jedną z najczęstszych i najbardziej śmiertelnych chorób nowotworowych na świecie. Według danych Światowej Organizacji Zdrowia (WHO) stanowi on jedną z głównych przyczyn zgonów z powodu nowotworów, a jego rozwój często wiąże się z długotrwałym narażeniem na czynniki środowiskowe i styl życia, takie jak palenie papierosów, zanieczyszczenie powietrza, spożycie alkoholu czy kontakt z substancjami toksycznymi w miejscu pracy.

Celem niniejszej analizy jest zbadanie czynników, które mogą mieć wpływ na ryzyko zachorowania na raka płuc oraz identyfikacja potencjalnych zależności między nimi. Wykorzystany zbiór danych pochodzi z platformy Kaggle i zawiera informacje dotyczące pacjentów, obejmujące: - wiek, płeć, - ekspozycję na zanieczyszczenia powietrza, alkohol, palenie papierosów, - czynniki genetyczne, choroby przewlekłe, dietę, otyłość, - objawy takie jak kaszel z krvią, duszność, utrata masy ciała itp.

Analiza ma na celu zbadanie zależności między wymienionymi czynnikami a występowaniem raka płuc, a także wskazanie zmiennych, które mogą w największym stopniu wpływać na ryzyko zachorowania. W ramach pracy postawiono następujące pytania badawcze:

1. Jakie czynniki najczęściej współwystępują z występowaniem raka płuc?

2. Czy istnieją różnice między płciami w częstości zachorowań?

3. Czy czynniki środowiskowe (zanieczyszczenie, palenie, alkohol) mają istotny wpływ?

*Corresponding author