

MISQ Archivist

A Comprehensive Analysis of Triggers and Risk Factors for Asthma Based on Machine Learning and Large Heterogeneous Data Sources

Wenli Zhang and Sudha Ram

Abstract

Asthma is a common chronic health condition affecting millions of people in the United States. While asthma cannot be cured, it can be managed if we identify and understand triggers and risk factors that cause asthma exacerbations. However, this is challenging because these triggers and risk factors are complex and interconnected, and there are limitations to current mainstream approaches for identifying them. The recent availability of massive amounts of heterogeneous data has opened up new possibilities for asthma triggers and risk factors analyses. In this study, we introduce a data-driven framework, adapt and integrate multiple advanced machine learning techniques, and perform an empirical analysis to (1) derive characteristics of self-reported asthma patients from social media, (2) enable integration and repurposing of highly heterogeneous and commonly available datasets, and (3) uncover the sequential patterns of asthma triggers and risk factors, and their relative importance, both of which are difficult to achieve via retrospective cohort-based studies. Our methods and results can provide guidance for developing asthma management plans and interventions for specific subpopulations and, eventually, have the potential to reduce the societal burden of asthma.

Keywords: Chronic disease management, asthma triggers/risk factors, design science, machine learning, distant supervision, convolutional neural networks, sequential pattern mining, geometric inference, random forest