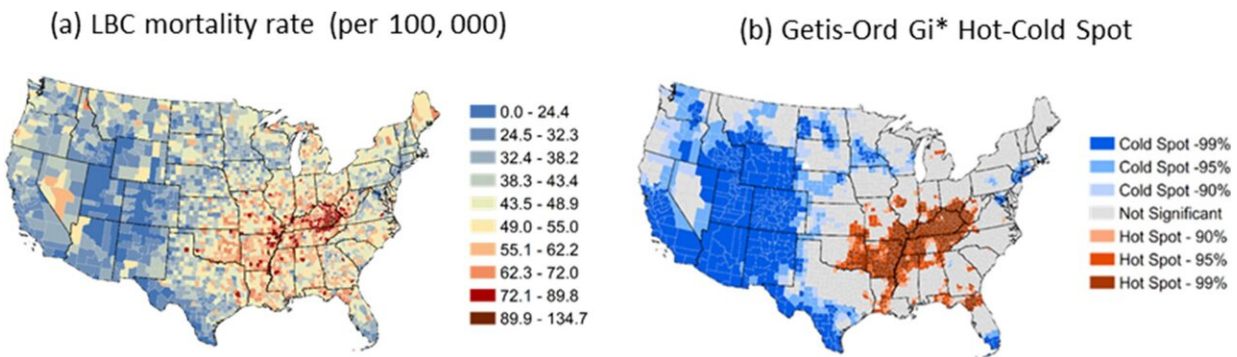# Using artificial intelligence to understand lung and bronchus cancer mortality rates

January 28 2022, by Tessie Mar

(a) County-level 5 years (2013–2017) average annual Lung and Bronchus Cancer (LBC) Mortality Rates, (b) The geographical clusters of counties with significant-high (hot spot—statistically significant positive z-scores, red color) or low (cold spot—statistically significant negative z-scores, sky blue colors) values of the Getis-Ord Gi* statistics for the LBC rate. LBC mortality rates and Getis-Ord Gi* Hot Spot maps were created in ArcGIS Desktop version 10.6.1. Credit: DOI: 10.1038/s41598-021-03198-8

Many people think of robots when they hear the term "artificial intelligence (AI)." However, in the case of a new study on lung and bronchus cancer (LBC) in the U.S., AI refers to various machine learning models stacked together to make high-level predictions about LBC mortality rates.

University at Buffalo researchers Zia U. Ahmed, Kang Sun, Michael

Shelly and Lina Mu authored the new study, which identifies key risk factors of LBC [mortality](link) using explainable [artificial intelligence](link), or XAI. While smoking prevalence, poverty and a community's elevation were most important in predicting LBC mortality rates among the risk factors studied, associations between risk factors and LBC mortality rates were found to vary spatially, and the research explored these geographic differences.

The paper, "Explainable artificial intelligence for exploring spatial variability of lung and bronchus cancer mortality rates in the contiguous U.S.," was published in the journal *Scientific Reports* in December 2021.

The study brought together an interdisciplinary team. Ahmed, Ph.D., is a database/visualization specialist with the UB RENEW Institute; Sun, Ph.D., is a core faculty member with the UB RENEW Institute and an assistant professor of civil, structural and [environmental engineering](link) in the UB School of Engineering and Applied Sciences; Shelly, Ph.D., is an environmental/ecological economist with the UB RENEW Institute; and Mu, Ph.D., MD, is an associate professor of epidemiology and [environmental health](link) in the UB School of Public Health and Health Professions.

Ahmed speaks to the importance of the study and research: "The results matter because the U.S. is a spatially heterogeneous environment. There is a wide variety in socioeconomic factors and education levels—essentially, one size does not fit all. Here local interpretation of machine learning models is more important than global interpretation."

He adds that the results can be useful for public health management and intervention by indicating which areas need support.

"We wanted the model to explain how the known LBC mortality rates and risk factor predictors connect," says Sun.

"The study can be a model for integrating artificial intelligence into an epidemiologic study," Mu says. "It also can serve as an example of the use of prediction models when studying cancer. This can greatly help in identifying high-risk areas where cancer registry is not available."

The study paired ensemble machine learning with explainable algorithms to spatially represent relationships between LBC mortality and risk factors in the U.S., marking an advance in this area of research. AI algorithms work better with more data with multiple models, which is why the stack-ensemble is more useful than any single model.

"XAI in local interpretation is still lacking, especially related to the environment and science," says Ahmed.

AI is a powerful tool because the models learn from data, enabling them to process complex interactions and relationships. The models can "think" by themselves.

Risk factors the study explored accounted for variables relating to lifestyles, socio-economic status, demography, air pollution, and the physical environment. They included cigarette smoking, poverty rate, health insurance, demography, air pollution and biophysical factors.

The study notes that smoking rates were linked with poverty levels and race/ethnicity. The study also notes a strong relationship between socioeconomic status and LBC mortality rates in the U.S.

With regard to air pollution, the researchers examined the pollutants nitrogen dioxide (NO_2), sulfur dioxide (SO2), ozone and particulate matter and their spatial variability in relation to lung and bronchus cancer mortality rates.

  More information: Zia U. Ahmed et al, Explainable artificial

intelligence (XAI) for exploring spatial variability of lung and bronchus cancer (LBC) mortality rates in the contiguous USA, *Scientific Reports* (2021).