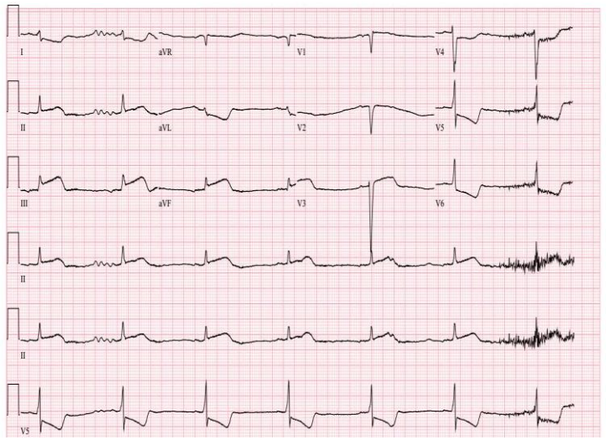
HEALTH
UNIVERSITY OF UTAH

UNLOCKING PATTERNS IN EXPOSURE DATA WITH SHAPELETS

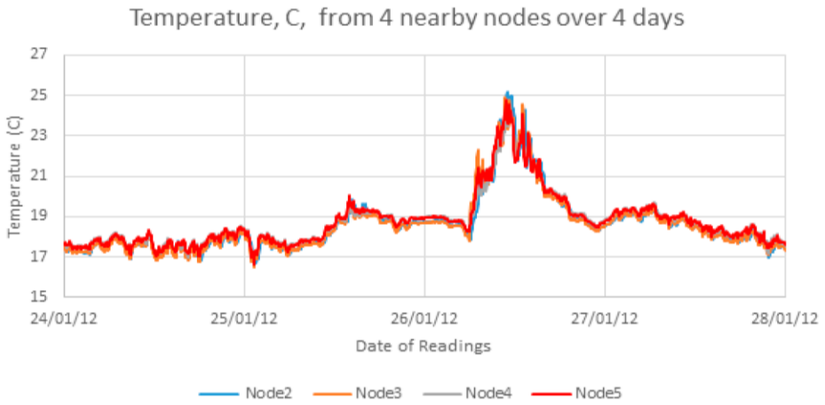
SUKRUT SHISHUPAL
DEPARTMENT OF BIOMEDICAL INFORMATICS

BACKGROUND

- Time series data appears in many domains



Health Monitoring



Sensor Network

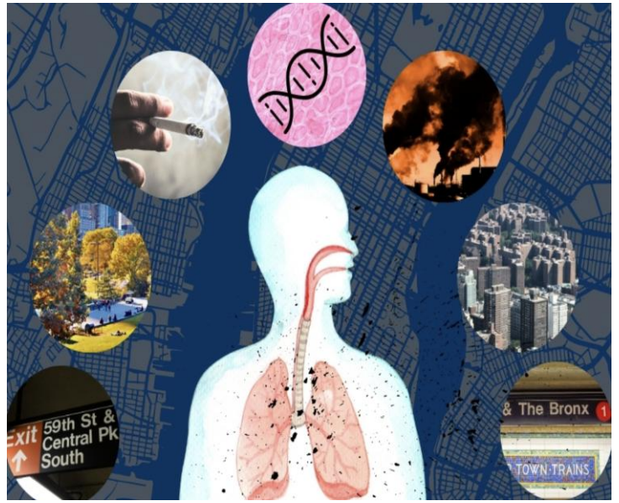
Financial World

HEALTH
UNIVERSITY OF UTAH

BACKGROUND

- Temporal Exposure: Duration & frequency
- Spatial Exposure: Substance & place

+



→

Temporal Exposure

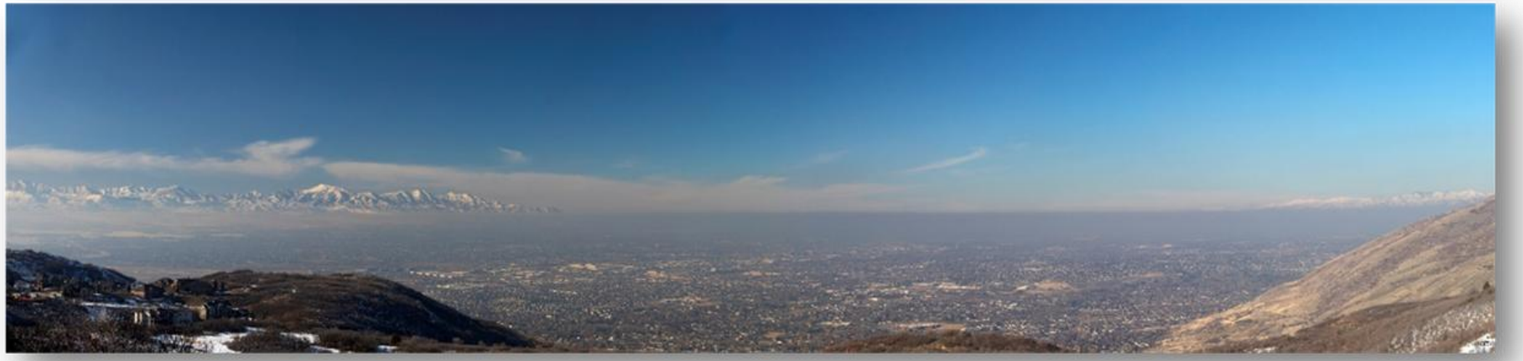
Spatial Exposure

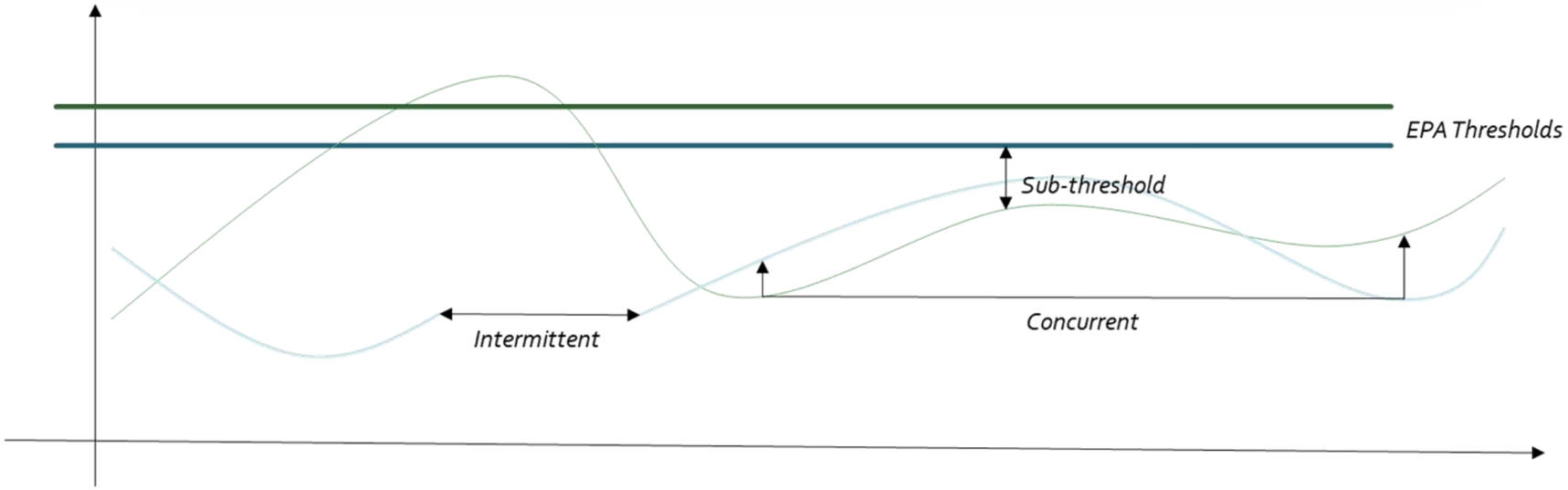
Health Outcome

HEALTH
UNIVERSITY OF UTAH

OBJECTIVES

- Intermittent, sub-threshold effects





OBJECTIVES

ORIGINAL RESEARCH

Intermittent Exposure to Cigarette Smoke Increases Lung Tumors and the Severity of Emphysema More than Continuous Exposure

Naofumi Kameyama[1], Shotaro Chubachi[1], Ahmed E. Hegab[1], Hiroyuki Yasuda[1], Shizuko Kagawa[1], Akihiro Tsutsumi[1], Koichi Fukunaga[1], Masayuki Shimoda[2], Yae Kanai[2], Kenzo Soejima[1], and Tomoko Betsuyaku[1]

[1]Division of Pulmonary Medicine, Department of Medicine, and [2]Department of Pathology, Keio University, School of Medicine, Tokyo, Japan

ORCID ID: 0000-0002-5046-3762 (S.C.).

Abstract

Lung cancer and chronic obstructive pulmonary disease are leading causes of morbidity and mortality worldwide, and cigarette smoking is a main risk factor for both. The presence of emphysema, an irreversible lung disease, further raises the risk of lung cancer in patients with chronic obstructive pulmonary disease. The mechanisms involved in smoke-induced tumorigenesis and emphysema are not fully understood, attributable to a lack of appropriate animal models. Here, we optimized a model of cigarette smoke (CS)–induced lung cancer and emphysema in A/J mice treated with 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, a potent carcinogen. We investigated whether variations in CS exposure patterns with the same total amount and duration of exposure affect tumorigenesis and/or development of emphysema.

Continuous CS exposure for 3 months significantly suppressed 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone–induced development of adenomas and adenocarcinomas; however, emphysema independently developed during this period. Surprisingly, intermittent CS exposure increased the severity of emphysema and resulted in a higher incidence of adenocarcinomas. Furthermore, intermittent CS exposure elicited a marked increase in M2-polarized macrophages within and near the developed tumors. By employing a CS exposure protocol with repeated cycles of cessation and relapse, we provide evidence that intermittent CS exposure enhances tumorigenesis and emphysema progression more than that of continuous CS exposure.

Keywords: lung cancer; emphysema; cigarette smoke; M2 macrophage

Ref: : Naofumi Kameyama, Shotaro Chubachi, Hegab AE, et al. Intermittent Exposure to Cigarette Smoke Increases Lung Tumors and the Severity of Emphysema More than Continuous Exposure. *American Journal of Respiratory Cell and Molecular Biology*. 2018;59(2):179-188. doi:https://doi.org/10.1165/rcmb.2017-0375oc

Published in final edited form as:

IEEE EMBS Int Conf Biomed Health Inform. 2016 February ; 2016: 565–568. doi:10.1109/BHI.2016.7455960.

Predicting Lung Cancer Incidence from Air Pollution Exposures Using Shapelet-based Time Series Analysis

Hong-Jun Yoon,
Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA phone: 865-241-2626; fax: 865-574-6275

Songhua Xu, and
Information Systems Department, New Jersey Institute of Technology, Newark, NJ 07102 USA

Georgia Tourassi [Member, IEEE]
Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA phone: 865-576-4829; fax: 865-574-6275

Abstract

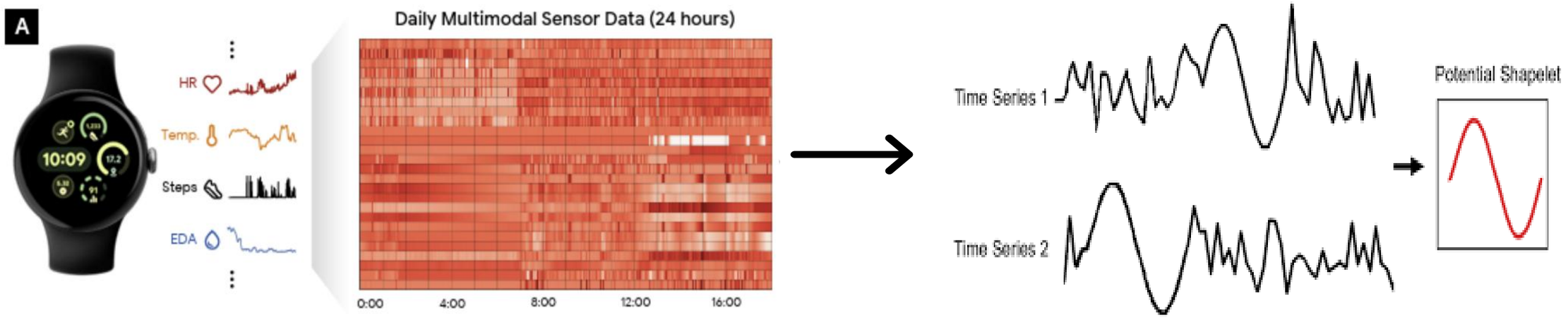
In this paper we investigated whether the geographical variation of lung cancer incidence can be predicted through examining the spatiotemporal trend of particulate matter air pollution levels. Regional trends of air pollution levels were analyzed by a novel shapelet-based time series analysis technique. First, we identified U.S. counties with reportedly high and low lung cancer incidence between 2008 and 2012 via the State Cancer Profiles provided by the National Cancer Institute. Then, we collected particulate matter exposure levels ($PM_{2.5}$ and PM_{10}) of the counties for the previous decade (1998–2007) via the AirData dataset provided by the Environmental Protection Agency. Using shapelet-based time series pattern mining, regional environmental exposure profiles were examined to identify frequently occurring sequential exposure patterns. Finally, a binary classifier was designed to predict whether a U.S. region is expected to experience high lung cancer incidence based on the region's $PM_{2.5}$ and PM_{10} exposure the decade prior. The study confirmed the association between prolonged PM exposure and lung cancer risk. In addition, the study findings suggest that not only cumulative exposure levels but also the temporal variability of PM exposure influence lung cancer risk.

Ref: Yoon HJ, Xu S, Tourassi G. Predicting Lung Cancer Incidence from Air Pollution Exposures Using Shapelet-based Time Series Analysis. *IEEE EMBS Int Conf Biomed Health Inform*. 2016;2016:565-568. doi:10.1109/BHI.2016.7455960

HEALTH
UNIVERSITY OF UTAH

SHAPELETS

- Discriminative sub-sequences of time series data that serve as a feature for classification or regression task



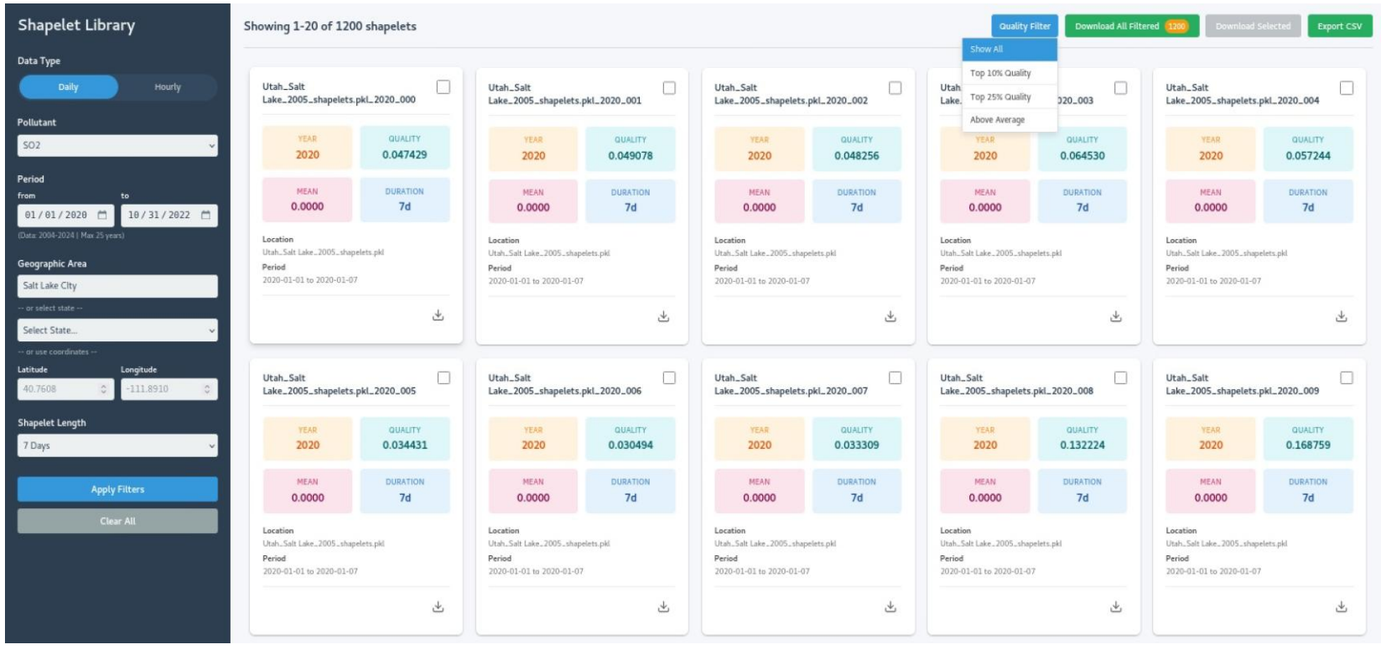
RESULTS

$$600 \times 20 \times 2 \times 100 \approx 2{,}400{,}000$$

(Counties) (pollutants) (Window size) (Number of shapelets / location)

$$2{,}400{,}000 \times 20 \approx 48{,}000{,}000$$

(Years)

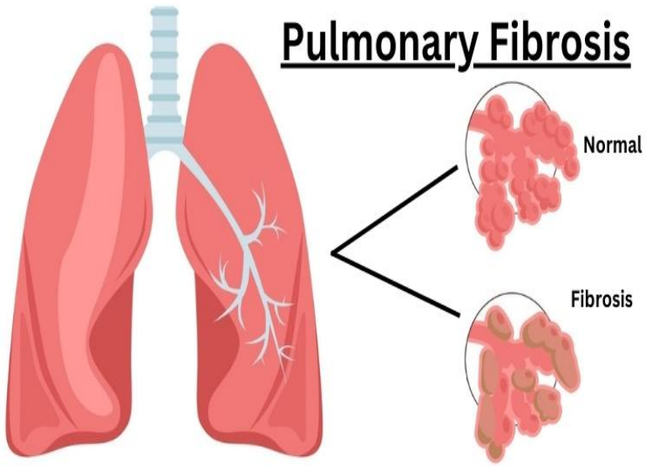


https://ehie-shapelets.ctsi.utah.edu/

HEALTH
UNIVERSITY OF UTAH

NEXT STEPS

- Linking exposure motifs to health outcomes



ACKNOWLEDGEMENTS

This research is supported by the University of Utah 1U4U initiative, NIEHS, 1R24ES036134-01 [SMARTER], NCATS, UL1TR002538, UM1TR004409 [CTSI] and NIBIB, U54EB021973 [PRISMS]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Research Computing provided by the Center for High Performance Computing (CHPC), University of Utah. CHPC resources are partially funded by NIH Research Instrumentation Award S10OD021644.

Thanks Ram Gouripeddi, Julio Facelli, Mollie Cummins, Katherine Sward, Randy Madsen, Rob Smart, Urvi Varma, Amanda Bakian, Noami Riches, John Lin and Tanya Halladay.

HEALTH
UNIVERSITY OF UTAH