IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 6, 2019, accepted December 6, 2019, date of publication December 18, 2019, date of current version January 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2960551

Predicting Asthma Attacks: Effects of Indoor PM Concentrations on Peak Expiratory Flow Rates of Asthmatic Children

DOHYEONG KIM[1], SUNGHWAN CHO[1], LAKSHMAN TAMIL[2], (Senior Member, IEEE), DAE JIN SONG[3], AND SUNGCHUL SEO[4]

[1]School of Economic, Political and Policy Sciences, The University of Texas at Dallas, Richardson, TX 75080, USA
[2]Quality of Life Technology Laboratory, Department of Electrical and Computer Engineering, The University of Texas at Dallas, Richardson, TX 75080, USA
[3]Department of Pediatrics, Korea University Guro Hospital, Seoul 08308, South Korea
[4]Department of Environmental Health and Safety, College of Health Industry, Eulji University, Seongnam 13135, South Korea

Corresponding authors: Sungchul Seo (seo@eulji.ac.kr) and Dae Jin Song (djsong506@korea.ac.kr)

This work was supported in part by the Environmental Health Action Program (Development of Receptor-Based Environment-Induced Diseases Prevention and Management System Using Real-Time Collected Environment and Health Information) under Project 2018001350005, and in part by the Research of Korea Centers for Disease Control and Prevention, South Korea, under Grant 2016ER670300.

ABSTRACT Despite ample research on the association between indoor air pollution and allergic disease prevalence, public health and environmental policies still lack predictive evidence for developing a preventive guideline for patients or vulnerable populations mostly due to limitation of real-time big data and model predictability. Recent popularity of IoT and machine learning techniques could provide enabling technologies for collecting real-time big data and analyzing them for more accurate prediction of allergic disease risks for evidence-based intervention, but the effort is still in its infancy. This pilot study explored and evaluated the feasibility of a deep learning algorithm for predicting asthma risk. It is based on peak expiratory flow rates (PEFR) of 14 pediatric asthma patients visiting the Korea University Medical Center and indoor particulate matter PM10 and PM2.5 concentration data collected at their residence every 10 minutes using a PM monitoring device with a low-cost sensor between September 1, 2017 and August 31, 2018. We interpolated the PEFR results collected twice a day for each patient throughout the day so that it can be matched to the PM and other weather data. The PEFR results were classified into three categories such as 'Green' (normal), 'Yellow' (mild to moderate exacerbation) and 'Red' (severe exacerbation) with reference to their best peak flow value. Long Short-Term Memory (LSTM) model was trained using the first 10 months of the linked data and predicted asthma risk categories for the next 2 months during the study period. LSTM model is found to predict the asthma risk categories better than multinomial logistic (MNL) regression as it incorporates the cumulative effects of PM concentrations over time. Upon successful modifications of the algorithm based on a larger sample, this approach could potentially play a groundbreaking role for the scientific data-driven medical decision making.

INDEX TERMS Asthma, indoor particulate matter, deep learning, peak expiratory flow rates, real-time monitoring.

I. INTRODUCTION

Asthma is known for affecting quality of life of people of all ages throughout the world by restricting social, emotional, and physical aspects of life [1]. Asthma is characterized by hypersensitivity of airways, which is reversible, but requires a constant management of the symptoms that include long term and short-term medication, mostly inhalers. Chronic Obstruction Pulmonary Disease (COPD), for which asthma is one of the risk factors, is characterized by non-reversible airways obstruction. Asthma and COPD are two leading causes of chronic disease burden as measured by DALY (disability-adjusted life years) globally [2], [3]. COPD is the fourth largest cause of mortality in the world at present, and it is forecasted that it will be third leading cause of deaths in westernized countries by 2020 [4]. Both asthma and COPD

The associate editor coordinating the review of this manuscript and approving it for publication was Vincenzo Piuri.

are chronic respiratory conditions that require regular medication and management of patients as their severity may vary with weather and environmental conditions, effectiveness of medication, and other individual factors. Sudden exacerbation because of any reason requires critical medical care and may lead to hospitalization. Asthma and COPD exacerbations are cause of approximately 2 million visits to hospital emergency departments (EDs) per year in the United States [5] and hospital admission rates for asthma patients are 43.8 per 100,000 people in OECD countries on average. Asthma management is much worse than in some countries such as South Korea where hospital admission rates for asthma patients are 98.5 per 100,000 people [6]. Short-term increase in air pollutant concentration has been linked to excess daily emergency room visits and hospital admissions due to asthma and COPD [7], [8].

Environmental health forecasting, one of the least developed branches of forecasting, is a useful tool for measuring environmental risks and provisioning health services by integrating various personal and environmental factors that affect the health status of individuals and populations such as weather and air quality, which are particularly important in forecasting asthma [9], [10]. Given that children tend to spend more time indoors, indoor air quality modeling is particularly important for predicting pediatric asthma [11]. With successful development and implementation of accurate predictive modeling, this provides an advanced notice for the patients to take proper medication or treatment to prevent from falling sick as well as help them to plan their mobility. Ability to forecast risk based on indoor monitoring and personal data is so powerful that it provides people with real-time risk assessment of their asthma exacerbation based on the dynamic patterns of indoor air pollutants and their general health condition.

Despite ample research that demonstrates a link between indoor air pollution and asthma exacerbation [12]–[14], the efforts to forecast personalized risk based on real-time indoor air quality monitoring data are still at their infancy. Most existing studies have used one-time data based on questionnaire survey or monthly/quarterly measurements of indoor air pollutants, rather than hourly or real-time monitoring [15], [16]. Such approach may be useful to identify a static correlation between indoor air pollution and asthma morbidity but cannot be utilized for practical prevention and intervention such as early warning or alert systems tailored to vulnerable households and individuals which requires constant monitoring and analysis of real-time indoor air conditions. The recent advancement of IoT and big data technologies facilitate real-time monitoring of indoor air quality and AI-based forecasting.

Deep learning has been recently used for various health applications, such as translational bioinformatics, medical imaging, pervasive sensing, medical informatics, and public health [17], [18]. However, the level of predictability offered by a deep learning model depends on the availability of big data analytics framework [19] and real-time risk monitoring data via environmental and personal sensors using IoT technology [20]. Compared to other health applications, deep learning has been relatively less applied to asthma or other respiratory illness mostly due to the difficulty in obtaining real-time measurements of pulmonary function variables in non-clinical setting [21]. Most of the existing machine learning algorithms developed to predict asthmatic risks used historic patient records in clinical setting [22] or weekly or daily monitoring data [23]. Upon successful development and implementation of IoT-based mobile pulmonary sensors monitoring and measuring real-time pulmonary function in their daily lives, widespread applications of deep learning algorithms are expected to predict personal risk of developing asthma or COPD by linking with concurrent air pollution and environmental monitoring data collected from their real-time locations. We explore, for the first time, if a deep learning technique can be applied to the real-time records of indoor particulate matter (PM) concentrations to predict asthma risk reflected in peak expiratory flow rates (PEFR). This pilot study could demonstrate a potential role of deep learning approaches as a tool for data-driven medical decision making for asthma care and management.

The paper is structured as following. The next section describes the data and analytic method that was used in this study. Then, the results section reports the findings of the study. The paper concludes with the discussion of implication, limitation and suggestions for future research.

II. MATERIALS AND METHODS

A. PEAK EXPIRATORY FLOW RATE (PEFR) DATA AND PATIENT CLUSTERING

Many organizations such as the National Asthma Education and Prevention Program, and American Thoracic Society recommend the use of pulmonary function test (PFT) to primary care physicians to diagnose and manage respiratory problems [24], but few studies have collected long-term daily records of PFT results mostly due to difficulty in home-based self-tests. PEFR is popularly used to estimate the degree of airway obstruction in patients with asthma because it is relatively easy to measure using an inexpensive small portable device, even by patients themselves at their homes [25]. Recently, a few research groups have used the weekly PEFR records to predict asthma deterioration in children using machine learning models [26], but there is no attempt to collect the daily PEFR records over a long period of time and associate their temporal variations with real-time changes in exposure to indoor environmental risk factors.

As a pilot study to explore the feasibility, we obtained the self-collected PEFR results for a total of 16 pediatric asthma patients, measured twice a day between September 2017 and August 2018, which was approved by the institutional review board at Korea University Guro Hospital (IRB No.2016GR0336). With two cases dropped due to the incompleteness of the records, a set of twice-a-day PEFR records as well as the questionnaire survey responses (e.g. disease and prescription history, demographic variables, etc.) were gleaned and stored for 14 children aged between 6-14. They were then clustered into the two groups via K-means clustering based on the maximum, minimum, mean and standard deviation of the PEFR results of each patient [27]. Also, in order to match PEFR values one-to-one with the indoor air monitoring data, the observed PEFR values recorded twice per day were interpolated to 10-minute intervals for each cluster using 24-hour variation formula of PEFR established by Hetzel [28] and Charleston-Villalobos *et al.* [29]. Since the normal range of PEFR values varies by patients, the interpolated PEFR values were classified into three categories such as ‘Green’ (when a reading is higher than 80% of the best peak flow; normal), ‘Yellow’ (when a reading is between 50% and 80% of the best peak flow; mild to moderate exacerbation), and ‘Red’ (when a reading is lower than 50% of the best peak flow; severe exacerbation), as referenced by Talabere [30].

B. INDOOR PM MONITORING DATA

We installed a continuous measurement device with low-cost monitoring sensors at each patient’s residence during the same period when the PEFR data were collected, which measured particulate matters (PM10 and PM2.5), as well as temperature and relative humidity every 10 minutes. All data were stored at the device as well as transferred to the cloud storage through Wi-Fi networking. We compared the PM levels from our device with those of a reference monitor (DUSTTRAK II AEROSOL MONITOR 8530, TSI Incorporated, Minnesota, USA) in the lab settings to ensure accuracy and validity of our data. The mean ratio of our measurements to the reference values ranged from 0.742 to 0.758, which indicated very similar exposure patterns. The indoor PM data were finally matched with the PEFR data for every patient and every time instant so that the correlated data can be used to detect any potential association or meaningful basis for prediction.

C. DEEP LEARNING ALGORITHM AND PREDICTION

The matched data was analyzed to obtain guidance for predictive modeling. To investigate if there is any correlation between cumulative PM concentrations and PEFR risk categories (green, yellow and red), the mean values of PM10 and PM 2.5 concentrations were calculated for the four different prior periods of 1 hour, 12 hours, 24 hours and 72 hours, for each cluster. The statistical association between PM concentrations and PEFR classifications over time was used to predict the probability of each risk category. The temperature and humidity are incorporated as covariates in both predictive models. We first ran the multinomial logistic (MNL) regression to predict the PEFR risk category for a specific time period for each cluster using only the concurrent PM measurements. We then ran the Long Short-Term Memory (LSTM) deep learning network [31], [32] to predict the same targets and compare the results with those from the conventional regression approach (MNL).

LSTM is a variant of recurrent neural network (RNN) that is designed to overcome the error back-flow (vanishing gradient) problems, through the use of memory cells and several gates, with each of these components being associated with a particular aspect of learning [31]. The memory cell decides what information to memorize about the past that would be required much later in the future. The gates comprise an update gate indicating when to update the memory cell with a new state, a forget gate to indicate when a memory cell state needs to be forgotten and an output gate to decide how to combine update and forget gate information with the memory cell state to compute output activation for the current time step. An appropriate combination of these gates with activations from the previous time-steps helps the network retain information from the distant past and thus longer time-dependent aspects of a sequence can be learnt much more efficiently in a reasonable amount of time [31]. LSTM has been widely used in various sequence-based problems such as natural language processing, but recently used in health applications [32], [33].

While the MNL model does not reflect the state of previous time periods, the LSTM model enables incorporating values from previous time periods. The sensitivity analysis was conducted by comparing the mean concentrations of PM10 and PM2.5 for three PEFR risk categories across the four prior periods (t-6, t-72, t-144, t-432; t=10 minutes) to explore at what cumulative period the PM impact appears noticeable. Based on the results from a series of sensitivity analyses (shown in Table 2), we decided to use the PM measurements collected during the 1-hour (t-6) period before each PEFR record. The Python scikit-learn package was used to implement MNL regression, while the TensorFlow was used for the LSTM modeling. For LSTM, 67% of the matched data were used as a training dataset for model training and the remaining 33% of the matched dataset were used as test dataset. The LSTM structure we employed is illustrated in Figure 1.

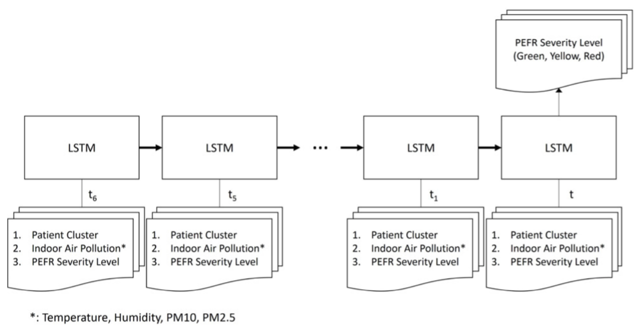


FIGURE 1. Data construction and analysis framework for machine learning.

A cost function was created and used for each machine learning algorithm [34]. The precision score, which is defined as the number of true positives divided by the total number of elements that actually belong to the positive class (also called as positive predictive value) [35], was calculated for each PEFR risk category in each cluster and the results obtained from MNL and LSTM were compared. For the experiment, we used the rectifier activation function (ReLU) as the activation function for the hidden layer [36]. We used this activation function for the output layer to maximize accuracy, which is measured by precision and recall values. In order to optimize the machine learning model, we used the Adam algorithm [37], which has been widely used for machine learning processes because it automatically solves the optimal value by adjusting the learning rate. Both MNL and LSTM models were evaluated by the k-fold cross-validation method by setting k as 10 [38]. This method provides robust evaluation results on model performance regarding data used for experiments. In other words, it repeats the process of creating a training model using the training dataset with k-1 sets and then validates the model on the remaining dataset until it determines the final model with the largest precision score.

III. RESULTS

A. PATIENT CHARACTERISTICS BY CLUSTER

Table 1 given below summarizes the characteristics of 14 asthmatic children in our sample and compares how they differ between cluster 1 (N=10) and cluster 2 (N=4). Although it is difficult to reach conclusions with any statistical significance due to small sample size, it appears that the subjects in cluster 2 are relatively older (more in late puberty) and have more severe symptoms than those in cluster 1, as indicated by a lower mean level of FEV1/FVC and a higher FeNO level. The subjects in cluster 2 also have a higher family history of any allergic disease (atopy, allergic rhinitis or asthma) from both mother and father.

TABLE 1. Characteristics of subjects by cluster.

		Cluster 1 (N=10)	Cluster 2 (N=4)
Age		7-11 years old (average: 8.6)	6-14 years old (average: 9.5)
% Male		30%	75%
% Moderate/Severe Asthma		30%	50%
FEV1/FVC (mean)		0.829	0.818
FeNO (ppb)		5-37 (average: 17.88)	10-165 (average: 52.75)
% puberty II or later		10%	50%
% whose mother or father has any allergic disease		80%	100%
Average PEFR		263.2	218.2
PEFR categories	Red	0.0%	2.9%
	Yellow	49.4%	44.9%
	Green	50.6%	52.2%

On an average level, cluster 2 has a lower mean PEFR than cluster 1 (218.2 vs. 263.2), but Figure 2 indicates a significant fluctuation of average PEFR values over the study period. Overall, the fluctuation of average PEFR looks relatively smaller in cluster 2 than in cluster 1, which partially indicates homogeneity among patients in cluster 2. What seems clear is that there is a substantial difference in temporal patterns of average PEFR values between the two clusters, which validates our cluster-based approach when applying deep learning prediction. The cluster characteristics are found to match better with their PEFR risk categories, implying that cluster 2 includes patients with higher vulnerability; 2.9% of the cluster 2 data belong to "red" category whereas no record belongs to the highest risk category in cluster 1.

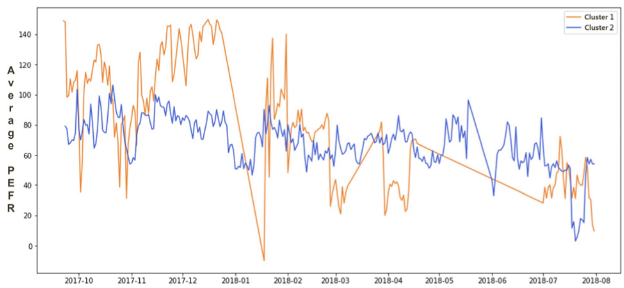


FIGURE 2. Temporal patterns of PEFR by cluster.

B. EXPLORATORY ANALYSIS RESULTS

Table 2 given below demonstrates a potential association between PM concentrations and PEFR risk categories, along with their cumulative impact, for each cluster. For both clusters, the mean concentrations of PM10 and PM2.5 are the smallest for the "green" category in all cumulative periods. The largest mean PM concentration is associated with the "red" category for cluster 2 but with the "yellow" category for cluster 1 (no "red" category for cluster 1). There is a pattern that the PM impact decays over time for both the "yellow" and the "red" categories. The highest mean PM concentrations is at past hour (t-6) for both clusters. This finding supports the rationale for using air pollution data from the past hour in LSTM estimation.

TABLE 2. Mean concentrations of PM10 and PM 2.5 by cluster for each risk category of PEFR during past time periods (1, 12, 24 and 72 hours).

Patient cluster	PEFR risk category	PM10 (mean)				PM2.5 (mean)			
		< t-6 (1 hr)	< t-72 (12 hrs)	< t-144 (24 hrs)	< t-432 (72 hrs)	< t-6 (1 hr)	< t-72 (12 hrs)	< t-144 (24 hrs)	< t-432 (72 hrs)
Cluster 1	Green	34.4	41.3	44.7	53.4	15.3	17.5	19.1	22.6
	Yellow	66.3	53.4	45.7	49.7	25.3	22.0	19.6	21.9
	Red	-	-	-	-	-	-	-	-
Cluster 2	Green	50.5	50.2	50.1	40.2	24.0	27.1	28.4	22.8
	Yellow	69.1	54.8	53.6	56.7	30.9	27.2	27.0	28.4
	Red	100.4	91.4	86.6	89.5	45.9	49.7	49.1	48.0

C. PREDICTION RESULTS

Figure 3 shows a loss function of the model for both training and test datasets during the machine tuning processes. The experimental process was iterated 200 times and the parameters were adjusted for both training and validation datasets at each iteration, along with the loss value of each model. Once the training process was complete, various neural network structures were formed to generate the models through a stepwise process until the final model was determined as that with the largest precision score.

Figure 4 compares the predictive precision scores for PEFR categories by clusters and predictive methods (MNL and LSTM). It is evident that LSTM performance is far superior to

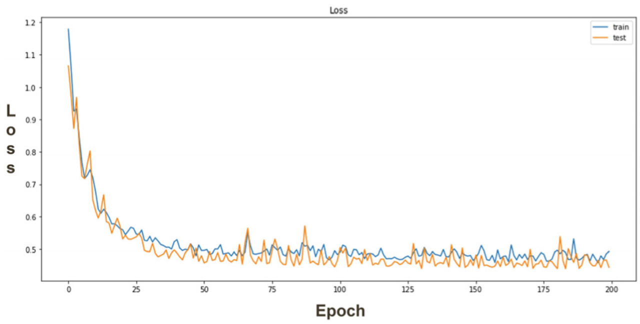


FIGURE 3. Loss function for training and testing data sets.

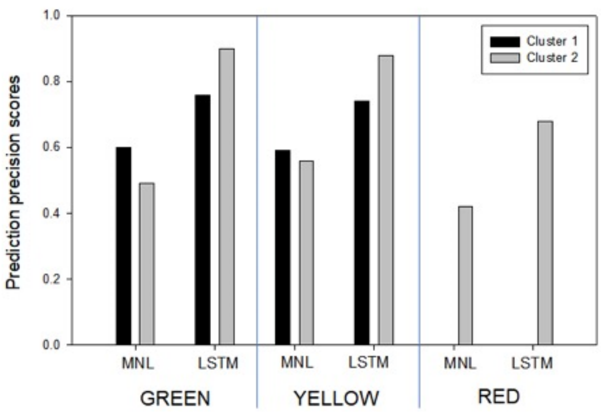


FIGURE 4. Predictive precision scores by cluster: MNL vs. LSTM.

MNL in predicting PEFR category, with 25-27% increase of precision for cluster 1 and 57-84% increase of precision for cluster 2 compared to MNL. Overall, MNL predicts better for cluster 1 than cluster 2 for every risk category, whereas LSTM predicts better for cluster 2 compared to cluster 1. Besides, the predictive performance of LSTM is found to be significantly better when used on clusters than when used on all samples together. This finding implies that the short-term PM accumulation could help improve the prediction of PEFR risk categories, particularly for cluster 2 whose samples are relatively older children with severe asthmatic symptoms. It also confirms etiology of asthma that the intensity of exasperation is normally higher in cases with family history. A deep learning approach such as LSTM is found to be more appropriate than standard statistical models to predict the PEFR risk categories when the cumulative effects of PM 10 and PM 2.5 concentrations are considered simultaneously.

IV. DISCUSSION

Despite ample research on the association between indoor air pollution and allergic disease prevalence or exacerbation, public health and environmental policies still lack evidence for developing a preventive guideline for patients or vulnerable populations mostly due to limited availability of real-time big data and accurate predictive models. The recent popularity of IoT and deep learning techniques could respectively provide enabling technologies for collecting real-time big data and accurately predicting allergic disease risks for evidence-based intervention, but the effort is still in its infancy.

In this study of the year-long PEFR data of 14 pediatric asthma patients matched with the real-time PM 10 and PM 2.5 concentrations, we observed better prediction of asthma risk categories when the short-term accumulations of the PM records were analyzed by a deep learning model. We also found that predictive performance may differ depending on patient characteristics and duration of accumulation. Machine learning techniques have been recently used in distinguishing asthma phenotypes [39] or predicting asthma exacerbations based on telemonitored symptoms [40], but there has been no attempt to predict asthma risk based on real-time monitoring of indoor air pollutants via a deep learning framework. This study is the first of its kind to show the potential of a deep learning approach as a tool for learning patterns between indoor air quality and asthma risk and predicting future risk of asthma exacerbations based on individual patient characteristics.

Although the sample size is not too large, we collected PM and PEFR data at every 10 minutes, which translated into 352,152 training samples for cluster 1 and 140,860 training samples for cluster 2. Considering that only seven input parameters were involved in this study (7), the sample size used in constructing the neural network would be sufficient. The work reported here is only a feasibility study and requires further model development with considerably more samples and other covariates including outdoor air pollutants and individual patient characteristics.

The major limitation of this study, of course, is with regards to the interpolation and clustering of PEFR records in order to match with real-time PM monitoring data. A more thorough sensitivity analysis of those processes would enhance the robustness and feasibility of our effort to overcome the limited measurements of PEFR data. A long-term, comprehensive IoT-based PEFR or PFT monitored data would eventually remove such a limitation. Real-time indoor air monitoring and big data prediction via deep learning would produce valuable information that provides an advanced notice for the patients to take proper medication to prevent from falling sick as well as help them to improve their indoor environments. Such forecasting is also valuable to public health in terms of not only reduced disease burden from asthma but also a more efficient use of limited resources for treating asthma patients [10]. Upon successful refinement of deep learning algorithms based on a larger sample for whom actual real-time data on personal exposure to PM and pulmonary function are available via IoT-based mobile sensors, this approach could potentially play a groundbreaking role for the scientific data-driven medical decision making and prevention activities.

ACKNOWLEDGMENT

The authors would like to declare that they have no actual or potential competing financial interests.

REFERENCES

[1] S. Miadich, "Quality of life in children with Asthma: A developmental perspective," *Amer. J. Pediatric Psychol.*, vol. 40, no. 7, pp. 672–679, 2015.

[2] D. Mannino and A. Buist, "Global burden of COPD: Risk factors, prevalence, and future trends," *Lancet*, vol. 370, no. 9589, pp. 765–773, 2007.

[3] GBD 2016 Mortality Collaborators, "Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015," *Lancet Respiratory Med.*, vol. 5, no. 9, pp. 691–706, 2017.

[4] J. Uriarte, "Nanomechanical assessment of decellularized COPD lung scaffolds," *Amer. J. Respiratory Critical Care Med.*, vol. 199, p. A2557, May 2019.

[5] CDC.gov. *CDC—Asthma*. Accessed: Jul. 15, 2019. [Online]. Available: https://www.cdc.gov/asthma/default.htm

[6] OECD. (2016). *OECD Health Statistics*. [Online]. Available: http://dx.doi.org/10.1787/health-data-en

[7] J. Noh, J. Sohn, S. Cho, Y. Choi, C. Kim, and D. Shin, "Short-term effects of ambient air pollution on emergency department visits for asthma: An assessment of effect modification by prior allergic disease history," *J. Preventive Med. Public Health*, vol. 49, no. 5, pp. 329–341, 2016.

[8] J. Halonen, T. Lanki, T. Yli-Tuomi, M. Kulmala, P. Tiittanen, and J. Pekkanen, "Urban air pollution, and asthma and COPD hospital emergency room visits," *Thorax*, vol. 63, no. 7, pp. 635–641, 2008.

[9] P. Soyiri and D. Reidpath, "Evolving forecasting classifications and applications in health forecasting," *Int. J. General Med.*, vol. 5, pp. 381–389, May 2012.

[10] K. L. Khatri and L. S. Tamil, "Early detection of peak demand days of chronic respiratory diseases emergency department visits using artificial neural networks," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 1, pp. 285–290, Jan. 2018.

[11] K. Isiugo, "Indoor particulate matter and lung function in children," *Sci. Total Environ.*, vol. 663, pp. 408–417, May 2019.

[12] W. Kanchongkittiphon, M. Mendell, J. Gaffin, G. Wang, and W. Phipatanakul, "Indoor environmental exposures and exacerbation of asthma: An update to the 2000 review by the institute of medicine," *Environ. Health Perspect.*, vol. 123, no. 1, pp. 6–20, 2015.

[13] K. Tham, "Indoor air quality and its effects on humans—A review of challenges and developments in the last 30 years," *Energy Buildings*, vol. 130, pp. 637–650, Oct. 2016.

[14] S. Seo, D. Kim, S. Min, C. Paul, Y. Yoo, and J. T. Choung, "GIS-based association between PM10 and allergic diseases in Seoul: Implications for health and environmental policy," *Allergy Asthma Immunol Res*, vol. 8, no. 1, pp. 32–40, 2016.

[15] E. Cleary, M. Asher, R. Olawoyin, and K. Zhang, "Assessment of indoor air quality exposures and impacts on respiratory outcomes in River Rouge and Dearborn, Michigan," *Chemosphere*, vol. 187, pp. 320–329, Nov. 2017.

[16] S. Hwang, S. Seo, Y. Yoo, K. Kim, J. Choung, and W. Park, "Indoor air quality of daycare centers in Seoul, Korea," *Building Environ.*, vol. 124, pp. 186–193, Nov. 2017.

[17] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Centre, B. Lo, and G. Yang, "Deep learning for health informatics," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 4–21, 2017.

[18] C. Wolfe, "Efficacy of a Web-based intelligent tutoring system for communicating genetic risk of breast cancer: A fuzzy-trace theory approach," *Med. Decis. Making*, vol. 35, no. 1, pp. 45–59, 2015.

[19] G. Harerimana, B. Jang, J. W. Kim, and H. K. Park, "Health big data analytics: A technology survey," *IEEE Access*, vol. 6, pp. 65661–65678, 2018.

[20] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, Jan. 2019.

[21] A. Messinger, "Bringing technology to day-to-day asthma management," *Amer. J. Respiratory Critical Care Med.*, vol. 198, no. 3, pp. 291–304, 2018.

[22] Y. Cheng, F. Wang, P. Zhang, and J. Hu, "Risk prediction with electronic health records: A deep learning approach," presented at the SIAM Int. Conf. Data Mining, 2016.

[23] G. Luo, "Predicting asthma control deterioration in children," *BMC Med. Informat. Decis. Making*, vol. 15, no. 84, pp. 1–8, 2015.

[24] J. Choi, "Nationwide pulmonary function test rates in South Korean asthma patients," *J. Thoracic Disease*, vol. 10, no. 7, pp. 4360–4367, 2018.

[25] K. Gulla and S. Kabra, "Peak expiratory flow rate as a monitoring tool in asthma," *Indian J. Pediatrics*, vol. 84, no. 8, pp. 573–574, 2017.

[26] G. Luo, B. Stone, C. Maloney, P. Gesteland, S. Yerram, and F. Nkoy, "Predicting asthma control deterioration in children," *BMC Med. Inform. Decis. Making*, vol. 15, no. 84, pp. 1–8, 2015.

[27] N. Valliammal and S. Geethalakshmi, "Plant leaf segmentation using non linear K means clustering," *Int. J. Comput. Sci. Issues*, vol. 9, no. 3, p. 212, 2012.

[28] M. R. Hetzel, "The pulmonary clock," *Thorax*, vol. 36, no. 7, pp. 481–486, 1981.

[29] S. Charleston-Villalobos, S. Cortes-Rubiano, R. Gonzalez-Camerena, G. Chi-Lem, and T. Aljama-Corrales, "Respiratory acoustic thoracic imaging (RATHI): Assessing deterministic interpolation techniques," *Med. Biol. Eng. Comput.*, vol. 42, no. 5, pp. 618–626, 2004.

[30] L. Talabere, "Development, implementation, and evaluation of an asthma management plan," in *Clinical Asthma*, M. Castro and M. Kraft, Eds. Amsterdam, The Netherlands: Elsevier, 2008.

[31] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[32] P. Chantamit-O-Pas and M. Goyal, "Long short-term memory recurrent neural network for stroke prediction," in *Machine Learning and Data Mining in Pattern Recognition* (Lecture Notes in Computer Science), vol. 10934, P. Perner, Ed. Cham, Switzerland: Springer, 2018.

[33] S. Cho and D. Kim, "Demand-based emergency dispatch system: Role of spatiotemporal machine learning," in *Geospatial Technologies for Urban Health*, Y. Lu and E. Delmelle, Eds. Basel, Switzerland: Springer, 2019.

[34] M. Abadi, P. Barham, J. Chen, Z. Chen, and A. Davis, "TensorFlow: A system for large-scale machine learning," in *Proc. OSDI*, vol. 16. Savannah, GA, USA: USENIX, 2016, pp. 265–283.

[35] D. Powers, "Evaluation: From precision, recall and f-measure to ROC, informedness, markedness & correlation," *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011.

[36] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012.

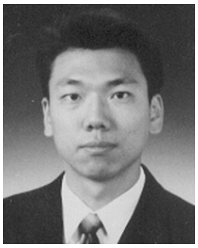
[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980

[38] T.-T. Wong and N.-Y. Yang, "Dependency analysis of accuracy estimates in k-fold cross validation," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 11, pp. 2417–2427, Nov. 2017.

[39] R. Howard, M. Rattray, M. Prosperi, and A. Custovic, "Distinguishing asthma phenotypes using machine learning approaches," *Current Allergy Asthma Rep.*, vol. 15, no. 38, pp. 1–10, 2015.

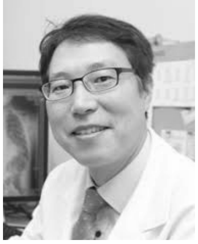
[40] J. Finkelstein and I. Jeong, "Machine learning approaches to personalize early prediction of asthma exacerbations," *Ann. New York Acad. Sci.*, vol. 1387, no. 1, pp. 153–165, 2016.

DOHYEONG KIM received the B.A. and M.S. degrees in public administration from Yonsei University, in 1996 and 1999, respectively, and the Ph.D. degree in health planning from the University of North Carolina at Chapel Hill, in 2007. He is currently an Associate Professor of public policy and geospatial information sciences with The University of Texas at Dallas. He also serves as the Managing Director of the Geospatial Health Research Group. His areas of research are highly interdisciplinary, including global health and safety, environmental health informatics, spatiotemporal big data analysis, and machine learning. Most of his research findings have been published in numerous leading referred journals in public and environmental health fields and presented in over 200 national and international meetings. As of August 2019, his work has been cited 819 times and his H10-index is 16. His multidisciplinary and multinational collaborative research projects have been funded by the U.S. National Institute of Health, the World Health Organization, and the National Research Foundation of Korea. He was a recipient of the 2008 New Investigators in Global Health Award. He currently serves as an Associate Editor of the journals *BMC Public Health* and *Frontiers in Public Health*.



SUNGHWAN CHO is currently a Visiting Scholar with the School of Economic, Political and Policy Sciences, The University of Texas at Dallas. His research focus on geospatial big data analysis for health, transportation, and crime applications.

LAKSHMAN TAMIL (S'81–M'88–SM'05) received the B.E. degree in electronics and communication engineering from Madurai Kamaraj University, India, the M.Tech. degree in microwave and optical communication from IIT Kharagpur, and the M.S. degree in mathematics and the Ph.D. degree in electrical engineering from the University of Rhode Island. He is currently a Professor of electrical and computer engineering with The University of Texas at Dallas (UTD), Richardson, TX, USA, where he is also the Director of the Quality of Life Technology Laboratory. Previously, he was the Founder, the CEO, and the CTO of Yotta Networks Inc., a company that designed and marketed terabit-switching platforms. He has also served as the Senior Manager directing research on advanced optical networks with the Alcatel's Corporate Research Center, Richardson, TX. His invention was the core asset of the Spike Technologies, Inc., that developed and marketed the antennas for the early MMDS systems, a precursor to the modern Wi-Max. He has been a Technology Consultant with Raytheon Company, Electrospace Systems Inc., Spike Technologies, Inc., Lighthouse Capital Partners, and McKool Smith LLP. Over the course of his career, he has contributed to more than 150 research publications and 23 patents and has directed 21 doctoral dissertations. His current areas of research interests are telemedicine, the Internet of Things (IoT), machine learning, and artificial intelligence applications to medicine and healthcare. He is a Fellow of the Optical Society of America and the Electromagnetics Academy and an elected member of the International Radio Science Union (URSI) Commission B and D. He is also serving as the President and the CEO of HygeiaTel, Inc., a start-up that commercializes some of the innovations of Quality of Life Technology Laboratory and as a member of the advisory boards of a few companies. He has also served as an Associate Editor for *Radio science* (Journal Sponsored by the American Geophysical Union and the International Radio Science Union) and as a member of the editorial boards of *Optical Fiber Technology: Material, Devices and Systems* (Elsevier) and *Optical Network Magazine* (Kluwer Academic Publishers).



DAE JIN SONG received the medical degree from the College of Medicine, Korea University, Seoul, South Korea, in 1998, and the M.S. and Ph.D. degrees in medicine from the Korea University, in 2002 and 2007, respectively. From 2005 to 2008, he was a Visiting Scholar with the School of Medicine, University of California at San Diego. He is currently working as a Professor and a Clinical Pediatrician with Korea University Guro Hospital. His research interests include alleviation and prevention of allergy, and childhood respiratory diseases, in particular asthma. Recently, he is working on the development of prevention platform using the IoT devices, including mobile and monitoring system in outdoor. He has a lot of academic articles and publications and received many fellowship awards and honors.

SUNGCHUL SEO received the Doctor of Philosophy degree in environmental health and occupational hygiene with strong background in aerosol science, extensive experiences in project management and laboratory works, exposure assessment for radon and air pollutants, and chemical/biological agents. He is currently a Professor with Eulji University, Seoul, South Korea. He has been the President of the Korea Association of Radon, since 2018. His research interest is in the development of monitoring systems for the level of environmental risk factors (i.e., fine particulate matters) to prevent diseases exacerbations. Recently, he is focusing on establishment of prevention platform using deep learning technology based on big data attributed to real-time IoT devices. He also has many funding and academic articles. He has also served on the executive boards of the Korean Society of Indoor Environment, and for over 20 years, he is dedicated to the improvement of indoor air related to environmental health, including radon, particulate matters, and bioaerosols.