Application of Data Mining Techniques to Predict Allergy Outbreaks among Elementary School Children

Integration of Hourly Air Pollution, Bi-Daily Upper-Air, and Daily School Health Surveillance Systems in Pennsylvania

Ahmed Youssef Agha, PhD  
Dept. of Applied Health Science  
Indiana University  
Bloomington, Indiana, USA  
E-mail: ahmyouss@indiana.edu

David Lohrmann, PhD  
Dept. of Applied Health Science  
Indiana University  
Bloomington, Indiana, USA  
E-mail: dlohrman@indiana.edu

Wasantha Jayawardene, MD  
Dept. of Applied Health Science  
Indiana University  
Bloomington, Indiana, USA  
E-mail: wajayawa@indiana.edu

Gamal El Afandi, PhD  
College of Agricultural, Environmental & Natural Sciences  
Tuskegee University  
Tuskegee, Alabama, USA  
E-mail: gelafandi@mytu.tuskegee.edu

Abstract: Objectives of this study are to determine if a relationship exists between occurrence of allergies among elementary school children and daily upper-air observations (temperature, relative humidity, dew point, mixing ratio) and daily air pollution (particulate matter, sulfur dioxide, nitrogen dioxide, carbon monoxide, and ozone); and, if so, to derive a mathematical model that predicts allergies. Using an ecological study design, school health records of 168,825 students in elementary schools enrolled in “Health eTools for Schools” within 49 Pennsylvania counties were analyzed. Upper-air measurements from ground level to the 850mb pressure level and air pollution measurements were obtained. Appropriate data mining techniques were utilized to validate and integrate three databases. Binary logistic regression used for analysis. A Generalized Estimating Equation model was used to predict the occurrence of more than 13 cases, the daily mean for 2008-2010. Results showed that the prevalence of allergies among school children in Pennsylvania increased over last three years. The primary occurrence of allergies among school children was in August-September, followed by December and April, while the lowest in January and May. Upper-air temperature and mixing ratio, as well as SO2, CO, O3, PM10 were significantly associated with occurrence of allergies (p<0.01). In conclusion, monitoring of upper-air observation and air pollution data over time can be a reliable means for predicting outbreaks of allergies among elementary school children. Such predictions could help parents and school nurses implement effective precautionary measures.

Keywords: air pollutants, upper-air indicators, allergies, school health records, Pennsylvania

I. INTRODUCTION

Hay fever, respiratory allergies, food allergies, and skin allergies are the main types of allergies among children in the US. Percent of children with diagnosed hay fever in the past 12 months is 9.5% (7.1 million), while 11.5% (3.4 million) of children reported as having respiratory allergies. In addition, 4.6% (9.4 million) of children had food allergies in the past 12 months, while 12.6% of children reported were diagnosed with skin allergies [1]. Seasonal allergies are fairly common in children older than age five. According to the American Academy of Allergy, Asthma, and Immunology, about 10-15% of school-age children have seasonal allergies [2].

Seasonality of allergies [3, 4] and their association with asthma [5] are well-known phenomena. Atopic eczema and allergic rhinitis were found to be higher in period September–May. Some evidence support the assertion that non-summer warmth and urban air pollution [6], probably mediated through exposure to common allergens such as dust mites, are possible risk factors for allergies in school-aged children [7]. Meteorological factors greatly influence the occurrence of pollen grains in the air. Dry and hot weather speeds up maturation and the loosening of pollen grains from anthers, and the concentration of pollen grains is considerably higher than in cold and wet weather [8]. Temperature has a positive correlation with pollen count, while relative humidity has an inverse correlation with the pollen count in the atmosphere. Reportedly, the occurrence of allergic diseases is usually associated with increase in the level of CO (carbon monoxide), SO2 (sulfur dioxide), PM10 and PM2.5 (particulate...
matter with a diameter of <10μm and <2.5μm, respectively), and O₃ (ozone), after adjusting for confounders [6].

The current study was undertaken to determine if a relationship exists between occurrence of allergies among school children, based on school health services records, and routine upper-air variables, such as mixing ratio, relative humidity, temperature, and dew point, as well as six air pollutants mentioned above. Although temperature, relative humidity, and air pollutants at ground level have been considered in allergy studies, neither upper-air dew point nor mixing ratio have been linked with allergy or used for prediction of outbreaks. As the amount of humidity is relative to the temperature available to do the work of evaporation, and can fluctuate even if no change occurs in the actual amount of water vapor per unit mass of dry air (mixing ratio), dew point and mixing ratio provide the best assessment of humidity as it relates to some allergies.

School health records have been utilized for allergy-tracking because they are readily accessible through the existing school data collection and storage infrastructure [9]. In many Pennsylvania schools student health records are compiled through a web-based software application portal called “Health eTools for Schools” (hereafter referred to as “eTools”) that was funded by the Highmark Foundation through its Healthy High Five Initiative [10] and developed by InnerLink, Inc., a private, for-profit company. With the utilitarian purpose of providing data for state and local health planning, eTools is offered free of charge to local public, private and parochial schools. It allows nurses to efficiently enter and download student health data, including records of daily clinic visits for allergic conditions, using a hand-held electronic device, and via computerized programming, compile and submit required annual reports to the Pennsylvania State Health Department.

Public health is an information-intensive field, which needs timely, accurate, and authoritative information from a variety of sources [11-13]. According to “The Future of Public Health” report of Institute of Medicine in 1988, which launched a series of public health reforms that continues to this day, the essence of community health assessment is the collection, analysis, interpretation, and communication of data and information from various sources [14].

Consistent with the literature, the current study was designed to investigate the possible relationship between allergic diseases among school children on a particular day and the effect of air pollution indicators, i.e., PM₁₀, PM₂.₅, SO₂, NO₂, CO, and O₃, throughout the day by using data mining techniques. The ultimate purpose was to determine if a mathematical model could be derived to predict daily allergy burden based on the combination of upper air data plus air pollution levels and student health record entries from Pennsylvania schools.

II. METHODS

A. Study Design

Ecological study design was adopted in order to understand the relationship between daily occurrence of allergies in a “whole population” of elementary school students and daily measurements of upper-air observation parameters [15]. In the ecological design, which investigates group-level variables, a geographical region can be analyzed in a cross-sectional manner (once or repetitively) to investigate the variation in a health-related variable (e.g., mean blood pressure, hospitalizations for allergy, and homicide rates) and its associations with regional characteristics (e.g., salt intake, air pollution, handgun laws, and drug policies) [15].

An ecological design has the advantageous ability to control for individual-level variability while at the same time addressing influences at the regional-level. In addition, this study design enables researchers to include all the students of the enrolled elementary schools in the study sample, contrary to one in which each student with an allergic condition serves as his/her own control and excludes other students [15]. Moreover, extracting daily upper-air data from existing environmental databases and retrieving computerized daily health data from existing repositories, such as eTools, is inherently cost-effective in that it requires a very low level of effort.

B. Study Population

School districts located in 49 of the 67 counties in Pennsylvania participated in eTools. From 2008 to 2010, enrolled school districts received eTools services for 168,825 elementary school students. These students constituted the study population. All eTools services for participating school districts were subsidized by the funders of eTools (Highmark Foundation) and covered costs of utilizing eTools through a three-year period. At the individual level, the sole participation eligibility requirement was to be a student with a health record in an elementary school that utilized eTools. Student records from these schools were excluded if they contained incomplete or inaccurately entered health record data. In addition, no race-, ethnicity-, or income-based bias existed in the enrollment of schools districts, schools, or students in eTools. All of the above factors are important in respect to generalizability of study results to elementary school students at the state level or beyond the state level [15].

C. Data Collection

Within the 49 Pennsylvania counties, 168,825 records of elementary school students were identified for this study. Data on allergies were originally noted in records maintained by school nurses as the type of treatment given to a student on a particular day. Treatment options for school health nurses were based on the prescribed medication provided for the student. For purposes of this study, having allergy was defined as “any case managed with antihistamines”. In fact, school nurses recorded the trade name of all the medications administered,
which were then later categorized into functional groups (i.e., antihistamines, analgesics, etc.). It was assumed that the treatment option noted by a school nurse correctly represented the disease (figure-1).

Data analyzed for this study were provided to the authors by InnerLink, Inc. which had custodial responsibility under a common Statement of Understanding and Service Level Agreement with all participating school systems (table-1). The school year in Pennsylvania usually begins in August, ends in June and typically includes an extended winter break. Therefore, surveillance data were commonly unavailable for the last three weeks of June, all of July, the first three weeks of August, and the second half of December every year. According to Pennsylvania State regulations, the healthcare provider should state whether the child is qualified and able to self-administer the medication. The number of visits by middle school and high school students to school health nurse’s office, therefore, may not accurately represent the number of allergy occurrences. Hence, only data from elementary school students’ records were included in this study (figure-1).

The upper-air data used in this study were downloaded from the Wyoming University web site [16]. Measurements were obtained from the ground level to 1500 meters level (equivalent to 850 mb pressure). These upper-air observation values cover an area with a radius of about 800 kilometers, around Pittsburg, PA. As this coverage area includes the entire state of Pennsylvania, upper-air measurements, unlike ground level air-pollution, were consistent for all eTools schools regardless of geographical location. Upper-air observation values are obtained twice every day: at 0Z and at 12Z. The Z-time is the basis for synoptic meteorology, which requires collection of all measurements at the same time every day and, thereby, produces a snapshot of the state of the atmosphere worldwide. The 0Z time in the US is 6 p.m. in Eastern Standard Time (EST) and 7 p.m. in Eastern Daylight Time (EDT), while 12Z time in the US is 6 a.m. in EST and 7 a.m. in EDT. For each day of the three years (2008-2010) the measured values at time 12Z were obtained for analysis. As the 0Z (6 p.m. EST and 7 p.m. EDT) is irrelevant to asthma exacerbations that occur during elementary school hours, authors had to choose the 12Z time for the analysis (figure-1).

Hourly data on six air pollutants were extracted from EPA measurement stations throughout the state. Therefore, unlike in upper-air variables, the geographical distribution of schools within the state affected the study results derived from air-pollution variables (Figure-1). These were obtained for school days only, although the data are available for all days throughout the period 2008-2010. Data for 1am–3pm period were obtained, because it represents the hours before and during school time. After school hours were ignored as the pollutant levels during these hours do not contribute to allergies that occur at school (figure-1).

![Diagram](image1.png)

**Figure 1:** Data Mining Algorithm to Create a GEE Model for Prediction of Allergy Outbreaks in School Children

[T=upper-air temperature; DP=upper-air dew point; RH=upper-air relative humidity; MR=upper-air mixing ratio; EPA=Environmental Protection Agency; GEE=Generalized Estimation Equation; NO2=nitrogen dioxide; SO2=sulfur dioxide; CO=carbon monoxide; O3=ozone; PM10 and PM2.5=particulate matter with <10 and <2.5 μm respectively]

**D. Analysis**

We performed the allowed character checks, cardinality check, consistency checks, data type checks, limit checks, logic check, spelling and grammar check in the process of data validation (figure-1). In addition, missing values were deleted, duplicate values were eliminated, and multivariate outliers were deleted. Data transformation was performed when
required. Multi-collinearity diagnostics were also performed with SAS.

To minimize the effect of the increase in asthma reporting over three years, the ratio of daily and weekly allergy cases to the annual average rather than the absolute number of daily allergy cases was utilized for analysis. Similarly, the ratio of weekly mean of asthma exacerbations to the annual average was used instead of the absolute number of asthma exacerbations per week in order to minimize the effect of the distortion (as indicated above) in the weekly pattern of asthma exacerbation occurrences.

We used binary logistic regression modeling to analyze data from three large databases using SAS. Logistic regression is a useful method utilized widely in data mining applications [17]. It is more advanced than the chi-square test, which allows the analysis of only two categorical variables. It allows the researcher to use both continuous and categorical predictors in modeling. Binary logistic regression converts the binary response into a logit value (the natural logarithm of the odds of the event occurring or not) and then establishes maximum likelihood estimates (MLE). The assumptions of this analysis are particularly applicable to the current study, as we used large databases with large sample asymptotic property. This means that the reliability of the estimates is high when a large number of cases for each observed combination of X is available [17]. In addition, in binary logistic regression, changes in the response itself are not modeled; instead changes in the log odds of the response are modeled. Another important feature is that binary logistic regression assumes neither the linearity of the relationship between predictor variables, nor homoscedasticity and normality of residuals. Model appropriateness was confirmed with the Wald statistic that tested the significance of individual parameters [17].

Generalized Estimating Equation (GEE) model was used to predict the probability of occurrence of cases greater than or equal to 13 (a cutoff value equal to the 2008-2010 mean). The repeated subject (cluster) in the model is the “matched date,” for example, March 8th in 2007, 2008, 2009, and in 2010 is a subject. GEE needs at least 100 clusters to study 5-12 exploratory variables, however, a great deal of confidence is assured with 200 clusters, [18, 19] the circumstance in in the current study. “Logit” link function on binary distribution was used on a binary dependent variable. “P” represents DailyAllergyCases≥13 and “1-P” represents DailyAllergyCases<13. The independent variables were the upper-air observation variables.

To avoid the complexity of obtained GEE model, we adhered to a model with a binary dependent variable (DailyAllergyCases<13 and DailyAllergyCases≥13), rather than having three levels (low, average, and high) of occurring daily allergy cases. In our future study, which will also include pollen data, we will develop separate models for pollen season and non-pollen season. In the future study, we will also incorporate machine learning approaches, in addition to statistical methods.

Table-1: Demographic Characteristics of Students

<table>
<thead>
<tr>
<th>Variable</th>
<th>Student Category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Males</td>
<td>50.80%</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>49.20%</td>
</tr>
<tr>
<td>Race</td>
<td>White Alone</td>
<td>74.78%</td>
</tr>
<tr>
<td></td>
<td>African-American Alone</td>
<td>8.30%</td>
</tr>
<tr>
<td></td>
<td>Hispanic Alone</td>
<td>7.91%</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>3.10%</td>
</tr>
<tr>
<td></td>
<td>Multi-Race</td>
<td>4.27%</td>
</tr>
<tr>
<td>Free-Lunch Eligibility</td>
<td>Eligible (Low-Income)</td>
<td>38.36%</td>
</tr>
<tr>
<td></td>
<td>Ineligible</td>
<td>61.64%</td>
</tr>
<tr>
<td>Urban-Rural Distribution</td>
<td>Rural</td>
<td>39.22%</td>
</tr>
<tr>
<td></td>
<td>Suburban</td>
<td>42.65%</td>
</tr>
<tr>
<td></td>
<td>Urban</td>
<td>18.13%</td>
</tr>
</tbody>
</table>

III. RESULTS

Demographic characteristics of 168,825 elementary school students are shown in table-1 [20, 21]. Offices of school health nurses in Pennsylvania had 259,951 visits of elementary school students for various health problems during the three year period, 2008-2010. A gradual increase of allergy cases was observed from 2008 to 2010, both in elementary schools and in schools with higher grades. This is partially explained by the increase in the number of schools participating in the allergy surveillance system. The total number of schools (elementary, middle, and high) increased from 2008 to 2009, with a slight decrease from 2009 to 2010. However, a boost from 0.85 to 1.77 (208% increase) in the ratio of daily average of allergy cases to number of schools (preK-5) partially explained the real increase in allergy prevalence among school children, because that particular increase is not influenced by an increase in the number of schools participating in the allergy surveillance system. It may, however, be influenced by increased reporting of school nurses, since any surveillance system takes a few years to reach the full functional capacity. Geographical differences in new school enrolment may also have influenced this 208% increase.

Use of antihistamines among school children as reported by school health nurses has an almost identical pattern in each year during 2008-2010. Peak of antihistamine use lies in period from late August to early September, although there are noticeable increases in December (early winter) and in April (early spring) too. The lowest use of antihistamine is reported in January and in May. The summer season compared to winter, had 17 times higher tendency of exceeding daily mean of allergy cases (7 cases), which was followed by fall (2.6 times higher), and spring (2.1 times higher). In other words, 68% of days in summer exceeded daily mean of allergy cases for three years, followed by fall (38%), and spring (35%).

Upper-air temperature peaks in July (hot) and has its minimum in January (cold). There were significant differences between temperature means for each season (p<0.001), with the highest in summer, followed by spring, and then fall. Similarly, upper-air dew point peaks in July (wet) and reaches...
its minimum in January (dry). Significant differences existed between seasons (p<0.001), with the highest in summer, followed by spring, and then fall. Significant differences between mixing ratio means for each season were also statistically significant (p<0.001), with the highest in summer, followed by spring, and fall. Variation in upper-air relative humidity is complex and it’s difficult to identify a specific pattern. Differences between relative humidity means for each season were statistically significant (p<0.02), with the highest in summer, followed by winter, and then fall. The lowest relative humidity was recorded in spring. It was revealed that upper-air temperature and mixing ratio were important in determining whether or not the number of allergy cases is greater than the daily mean of allergy cases for the last three year period (>13 cases).

Peak hours of NO2 were 4–6 am, followed by the remaining morning hours (1–3 am, 7–9 am, and 10–12 noon). Conversely, SO2 levels were highest towards noon (10–12 noon), followed by 7–9 am and 1–3 pm with the lowest levels recorded in the early morning (1–6 am). Highest levels of CO were recorded for the morning hours when students leave for school (7–9 am), followed by earlier hours in the morning (1–6 am). As with NO2, the CO levels decreased towards noon and further decreased during the afternoon. The peak for PM2.5 occurred at 7–9 am, while the afternoon hours had the lowest. PM10 levels did not vary significantly compared to other pollutants and levels gradually decreased from 4 am to 3 pm. PM10 data for 1–3 am and 1-3 pm time intervals were not available.

Table 2: Modeling the Probability that the Number of Allergy Occurrences ≥ 13 among preK-5 Students on a Particular Day Based on Pollutant and Upper-Air Parameters: Analysis of GEE Parameter and Empirical Standard Error Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Error</th>
<th>Z</th>
<th>Prob. &gt;Z</th>
<th>I</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.337</td>
<td>0.256</td>
<td>-9.1</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>0.087</td>
<td>0.031</td>
<td>2.8</td>
<td>0.0054</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T *O3_DAY&gt;CO2_3AM</td>
<td>0.018</td>
<td>0.004</td>
<td>4.7</td>
<td>&lt;0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MR*PM10_7AM&gt;SO2_3AM</td>
<td>-0.003</td>
<td>0.001</td>
<td>-3.9</td>
<td>&lt;0.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

O3_DAY = mean O3 concentration of the previous day
SO2_3AM = mean SO2 concentration for the 1 am – 3 am time period of the previous day
CO2_3AM = mean CO concentration for the 1 am – 3 am time period of the previous day
PM10_7AM = mean PM10 concentration for the 7 am – 9 am time period of the previous day
MR = Upper-air mixing ratio (in g/kg) of the previous day
T = Upper-air temperature (in Celsius) of the previous day

SO2, CO, O3, and PM10 were significant in determining whether or not the number of allergy cases is greater than the daily mean of allergy cases for the last three year period (>13 cases). However, none of these air pollutants predicted allergy occurrences in their univariate relationships with allergy cases. In other words, the interaction of O3 DAY and CO2_3AM with temperature was significant, while the interaction effect of PM10_7AM and SO2_3AM with mixing ratio was significant in predicting allergy cases. Therefore, the results indicate the importance of integrating upper-air and air pollution databases.

Using the GEE model above (table-2), the following equation was adopted for determining the probability odds ratio (y) of having ≥13 (daily-mean for a period of three years + one standard deviation) on the next day:

\[ y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 (x_1 * x_2) \]

\[ \log P/(1-P) = -2.3372 + 0.0874 (T) + 0.0180 (T * O3_DAY * CO2_3AM) - 0.0031 (MR * PM10_7AM * SO2_3AM) \]

As an example, if the upper-air temperature and upper-air mixing ratio were 9°C and 3 g/kg respectively and the levels of O3, SO2, CO, and PM10 were 19.54, 5.34, 0.28, and 20.55 respectively on a particular day, the probability (P) of having more than 13 allergy cases within the surveillance system on the next day is 68%.

IV. DISCUSSION

The surveillance of allergy among school children in Pennsylvania has improved during the last four years in both aspects: the school enrollment in the surveillance system and the disease reporting by school health nurses, a proven, reliable data source [22]. The increased exposure to allergens and changes in other socioeconomic and sociodemographic factors among school children may be the main reasons for increases in allergies in Pennsylvania. The finding of the current study that peaks (main outbreaks) of allergy occurrence were reported in August-September and in April was compatible with evidence from other empirical studies. However, having the highest prevalence in summer months
may be associated with characteristics of pollen that cause allergic symptoms [2].

Annual averages of upper-air observations (temperature, relative humidity, dew point, mixing ratio) did not differ between years. As explained previously, a higher temperature (greater than three year median) is associated with occurrence of more allergies among Pennsylvania school children. However, this study considered only the upper-air temperature, not the surface level ambient temperature. In addition, a higher upper-air mixing ratio was associated with occurrence of more allergies. As discussed previously, the concept of relative humidity is complex, because relative humidity is relative to the amount of energy (measured by temperature) available to do the work of evaporation, and it can be changed without having any change in the actual amount of water vapor in the air (absolute humidity). Therefore, the actual mixing ratio in conjunction with temperature provides a better assessment of humidity, which is more relevant to allergies. The interaction of O₃_3AM and CO₂₃AM with temperature was significant, while the interaction effect of PM10_9AM and SO₂_3AM with mixing ratio was significant in predicting allergy cases. This encourages the integration of multiple databases in the process of prediction.

Put another way, the routinely measured and publicly available upper-air indicators provide a good estimate and a reliable tool for forecasting allergy burden on the school health system today or on a future day. This output of analysis with GEE is based on a three-year mean of allergies (13 cases) and does not restrict the utilization of the above equation beyond interpolation. Therefore, the extrapolation for forecasting is also realistic given the current analysis, based on the average of allergies for the 2008-2010 period. At last, this model should be updated monthly using new data that will be obtained from three databases.

However, there are several limitations in this study. There is a possibility that other allergy triggers, such as pollen and respiratory infections, may confound the relationship between upper-air variables and allergy occurrences. Allergy surveillance in schools is not being conducted during winter break and summer vacations, as well as during weekends and school holidays. The relatively less data available for the summer season due to school vacation, may somewhat distort the relationships caused by unequal representativeness across seasons.

In addition, it is difficult to exclude the possibility that some students who experience worsening of symptoms at night or early in the morning did not attend the school so their allergies are not included in school health records. It was also impossible to estimate the number of students with allergy symptoms who did not receive medications in school, because some students prefer taking medications before going to school or after school. In addition, there is also a possibility that some elementary school students may have taken allergy medications without informing the school nurse. Therefore, we cannot exclude the possibility that use of antihistamine medications as recorded by school nurse somewhat under-represented the occurrence of allergy symptoms at school. Unavailability of antihistamine medications usage data at school level as well as absence of data on demographic characteristics (e.g., ethnicity, income) of students in each school or school district were also limitations.

The number of schools using eTools changed each year with some schools dropping off and others joining, causing fluctuations in the total number of available student data records. Additionally, each year some students were leaving their school or school district. Nevertheless, the total number of student data strings provided for any one year was sufficiently robust, as was the number of data strings available for multi-year comparisons, to generate reliable results.

V. CONCLUSION

Monitoring of upper-air observation data and air pollution data over time using data mining techniques can serve as a reliable means for predicting increased occurrence of allergies among elementary school children. The new mathematical model derived from statistical integration of routine environmental observations and school health records may be used to scrutinize the complexity of allergic diseases as a dynamic outcome determined by multiple environmental parameters. It’s also possible to assess the risk for future allergy outbreaks based on fluctuation analysis of a long time series of atmospheric function for taking more effective precautionary measures. Appropriate data mining techniques are proven to be useful tools in this process of prediction.

Predicting of allergy outbreaks is important for children diagnosed with allergies as their household members can take more preventive measures to avoid allergy occurrences, regardless of the disease severity. In addition, being informed of a possible allergy outbreak on a particular day or week, school nurses and teachers will be able to pay more attention on early identification of allergy occurrences among children in the classroom. Similarly, doctors and nurses in emergency room or in the family practice can take extra precautions and utilize resources more efficiently to serve an increased number of allergy patients.

With accurate prediction of increased number of allergy occurrences, the school health system can play a major role by effectively reallocating both human and physical resources as well as by alerting teachers and vulnerable children to take extra precautions that will manage allergy outbreaks very early. It also enables to identify and quickly treat those allergies that do occur. Mass media and local media can also play an important role in effort. Results of the current study can be used to inform similar programming at the national level and in other states.

ACKNOWLEDGEMENTS

Authors thank InnerLink Inc. for developing eTools and also Highmark Foundation of Pittsburgh, Pennsylvania for
implementing Highmark Healthy High 5 Initiative. The authors report no conflicts of interest in relation to the current study.

REFERENCES


