



EVALUATING THE UTILITY OF SCHOOL ABSENTEEISM DATA 2009-2010 INFLUENZA SEASON

Cheryl Faustino, MPH; Patricia Araki, MPH; Emily Kajita, MS MPH; Megan Jones, MPH; and
Bessie Hwang, MD MPH

BACKGROUND

The epidemiology of influenza has suggested that school aged children play an important role in the acquisition and spread of ILI.¹ During the pinnacle of the 2009-2010 H1N1 influenza pandemic, a principal focus on school absenteeism surveillance emerged — most notably as a non-traditional data source that could allow for earlier outbreak detection of like diseases.² It has been postulated that school absenteeism data may detect various disease outbreaks early under the presumption that disease spreads rapidly in dense school populations. No study to date has been reported on school absenteeism surveillance data in Los Angeles County (LAC), which contains near 90 independent school districts, including the second largest school district in the nation.³

OBJECTIVE

The purpose of this study was to evaluate the utility of LAC school absenteeism data from the largest school district in conjunction with current LAC Department of Public Health (DPH) Acute Communicable Disease Control (ACDC) Automated Disease Surveillance Section (ADSS) influenza-like-illness (ILI) surveillance systems during the 2009-2010 influenza season.

METHODS

Data Collection

LAC school district absenteeism data, collected from school attendance, are negative-based (i.e., absence only) and completed by teachers via an electronic student information system; once per day for elementary schools, once per period for middle/high schools. Any final corrections to daily attendance are made at the end of the school day through an electronic administrative portal. School absenteeism data are received by ACDC ADSS in near real-time on a biweekly basis via Secure File Transfer Protocol. The line listed variables available within the dataset contained: date of school absence, school name, school address and zip code, school sub-district, track number, number of total students enrolled per school per date, and number of students absent per school per date. Reason for absence was not reported by schools. Aggregate percent absenteeism was calculated per date, per school per date, and by school-age groups (elementary/middle [E/M] school and high school) per date.

ILI emergency department (ED) visits and over the counter (OTC) medication sales⁴ are current in-place surveillance systems utilized by ACDC ADSS. School-age stratified ILI ED visits were determined by age; where ages 5-13 were categorized as E/M school and ages 14-17 were categorized as high school. School or age data were not available for either OTC cough/cold medication sales or OTC thermometer sales, thus school-age categories were not created.

Data Analysis

For the purposes of this study, data available from September 1, 2009 through February 28, 2010 were examined. The dataset included 140 schools: 78 E/M schools and 62 high schools. Extreme data points with known explanations for high absenteeism (e.g., days preceding and succeeding major school holidays and winter recess) were removed. Wilcoxon-signed rank tests were performed to measure median differences in school-age percent absenteeism and in number of school-age ILI ED visits. Retrospective time series analyses were conducted to examine the correlations between percent school absenteeism and: (1) ILI ED visits, (2) OTC thermometer sales, and (3) OTC cough/cold medication sales. Cluster analyses were performed to explore levels of significant absenteeism at the school level.



All statistical analyses were conducted with SAS® version 9.2.1 (Cary, N.C.) and spatiotemporal analyses were conducted with SaTScan™ version 9.0.⁵ Statistical significance was set at p-values <0.01.

RESULTS

The study period of September 1, 2009 through February 28, 2010 included pandemic H1N1 influenza, as reported by LAC influenza tracking.⁶ During this time, total percent school absenteeism ranged from 0.2% to 6.2% (median=3.3%; Figure 1). Two school absenteeism peaks were most notable on September 28th, (5.7%) and on February 25th (6.2%). Total ILI ED visits ranged from 571 to 1,596 (median=856), with the highest number of visits incurred on November 2nd. Similarly, OTC thermometer sales ranged from 105 to 866 (median=307), with the highest number sold on November 2nd. OTC cough/cold medication sales ranged from 4,686 to 17,743 (median=13,728), with most number sold on October 30th. Total percent school absenteeism correlated strongest with total ILI ED visits ($r=0.57$) and least with OTC cough/cold medication sales ($r=0.52$) and OTC thermometer sales ($r=0.42$). It has been reported that OTC thermometer sales are a strong correlate of f ILI ED visits.⁷ This is consistent with this study's side analysis, where correlation between OTC thermometer sales and ILI ED visits had the strongest correlation ($r=0.79$).

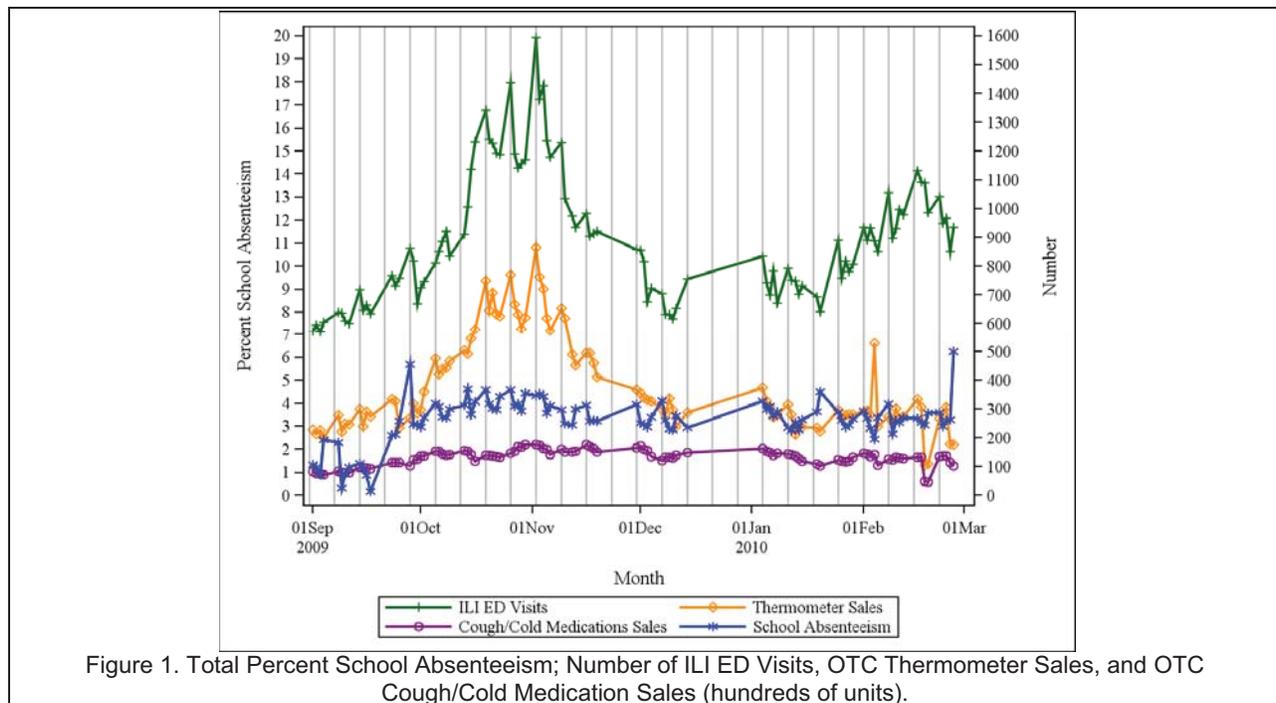


Figure 1. Total Percent School Absenteeism; Number of ILI ED Visits, OTC Thermometer Sales, and OTC Cough/Cold Medication Sales (hundreds of units).

Although a difference in percent school absenteeism between E/M and high school-aged groups has previously been reported², as shown in Figure 2, percent school absenteeism did not differ significantly between these age groups in LAC, with a median of 3.3% for E/M schools and 3.5% for high schools ($p=0.06$). Also, percent school absenteeism peaked similarly for both groups on September 28th (6.6% for E/M and 5.5% for high school). However, during the end of February, percent school absenteeism peaked much higher for the high school-aged group (7.5%) compared to the E/M school-aged group (4.8%).

Figure 3 shows the number of ILI ED visits stratified by school-age groups. Most notably, the E/M school-aged group had significantly more ILI visits to hospital emergency rooms than the high school-aged group (122 median visits versus 34 median visits, $p<0.001$). However, both groups had a similar trend in peak number of ILI ED visits between mid-October to early-November. These ILI ED trends are consistent with influenza tracking within LAC⁶, where pandemic H1N1 influenza largely affected younger age groups.

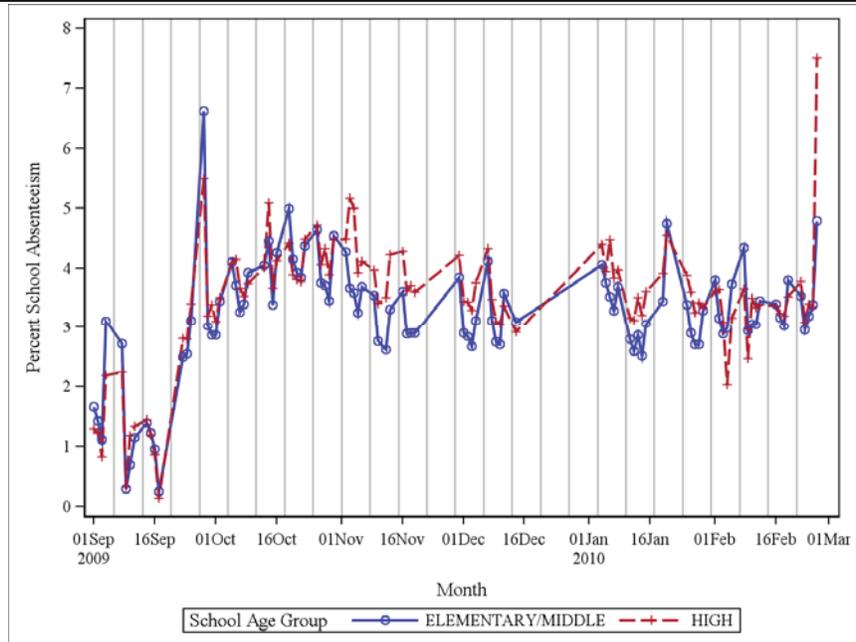


Figure 2. Percent Absenteeism by School-Age Group

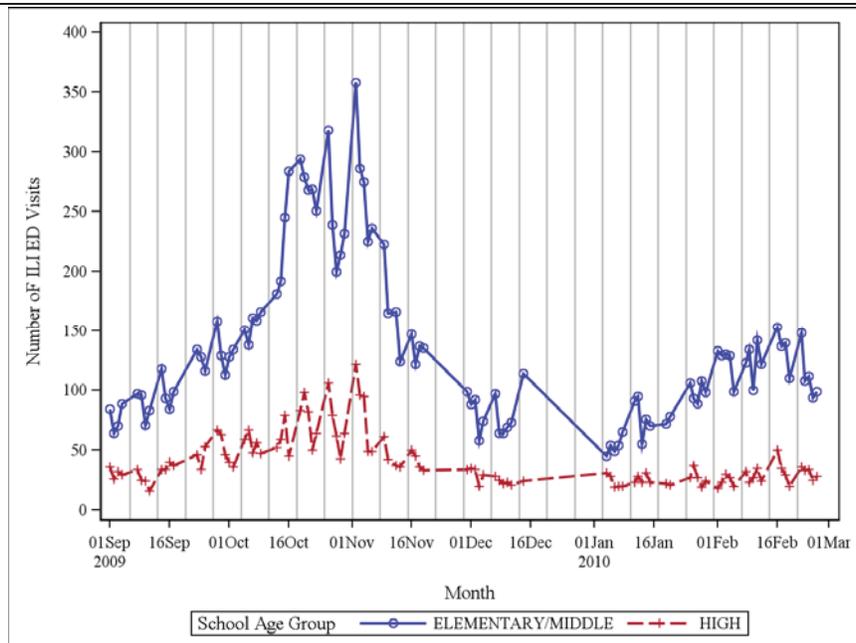


Figure 3. ILI ED Visits by School-Age Group

The correlations between school-age percent absenteeism, school-age ILI ED visits, OTC thermometer sales, and OTC cough/cold medication sales are shown in Table 1. During the study period of September 1, 2009 to February 28, 2010, both E/M and high school absenteeism showed relatively weak correlations to ILI ED visits, OTC thermometer sales, and OTC cough/cold medication sales. Moreover, correlations improved slightly when examined during the peak period of the influenza season, September 1st though December 14th. During this time frame, both E/M and high school-aged percent absenteeism correlated more with OTC cough/cold medication sales, followed by OTC thermometer sales (for high school group) and school-age ILI ED visits (for E/M school group).



Table 1. Pearson Correlation Coefficients by Dates and School-Age Group

	Full study Period 9/1/2009-2/28/2010		Peak Flu Period 9/1/2009-12/14/2009		Late Flu Period 12/15/2009-2/28/2010	
	E/M School	High School	E/M School	High School	E/M School	High School
	School Absenteeism vs. ILI ED visits	0.45	0.36	0.57	0.49	-0.21
School Absenteeism vs. OTC thermometer sales	0.40	0.41	0.55	0.62	-0.22	-0.31
School Absenteeism vs. OTC cough/cold medication sales	0.43	0.55	0.60	0.77	0.03	0.01

SaTScan™ spatiotemporal analysis was used to detect school absenteeism clusters during the peak period of the 2009-2010 influenza season (September 1-December 14), which included pandemic H1N1 influenza. Four statistically significant ($p < 0.01$) school-specific absenteeism clusters were detected. The first cluster was detected at high school A on September 15-17 (observed/expected=15.1). The second cluster was detected at high school B on September 10-11 (observed/expected =23.1). The third and fourth clusters were detected at two different elementary schools but during the same time period of November 2-10 (elementary school A, observed/expected=4.6; elementary school B, observed/expected=2.81). These elementary school clusters coincided with the peak number of ILI ED visits observed in the E/M school-aged group on November 2nd (Figure 3).

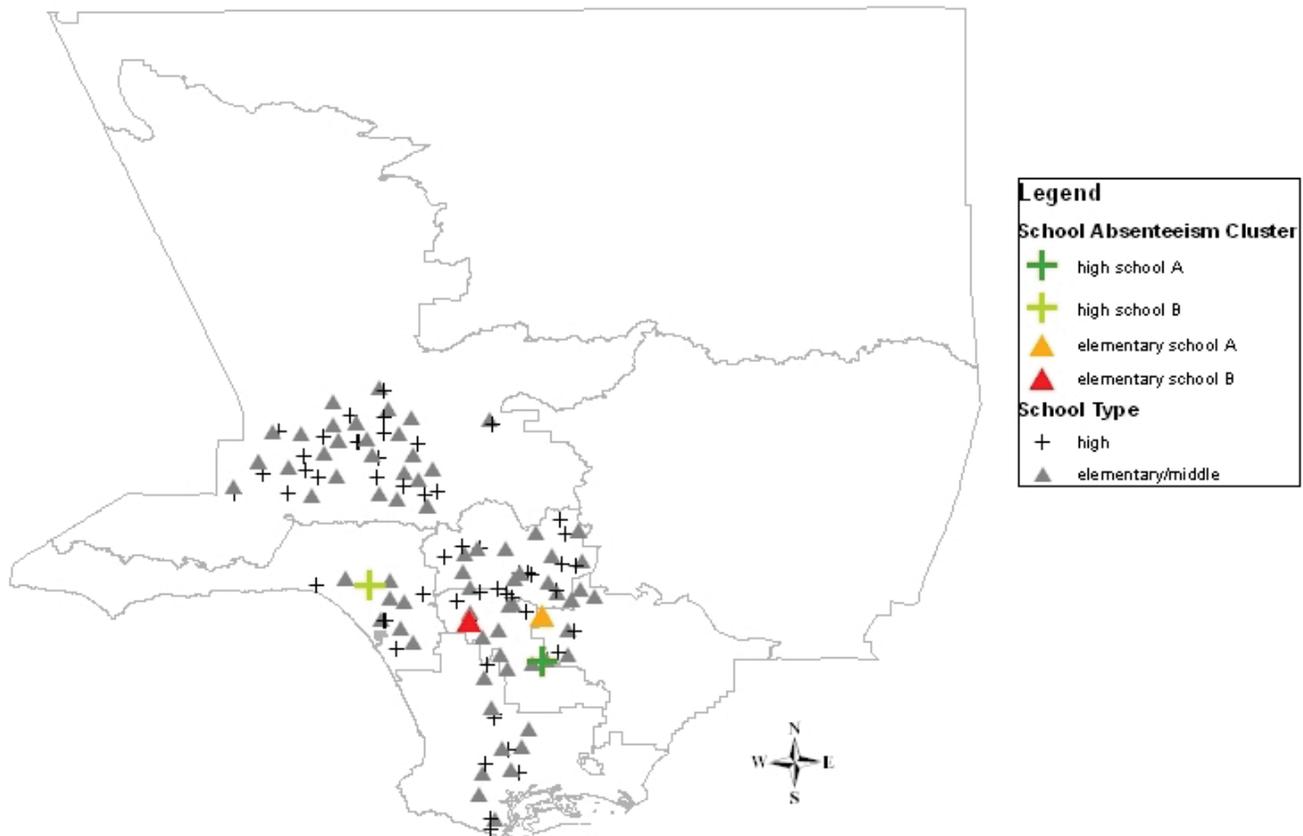


Figure 4. SaTScan™ Map of School Absenteeism Clusters and School Type, Los Angeles County.



DISCUSSION

Prior to establishing and maintaining any new surveillance system, evaluation of its potential utility is essential. From this evaluation of school absenteeism data within LAC, the findings revealed modest utility in conjunction with existing surveillance systems of ILI ED visits, OTC thermometer sales, and OTC cough/cold medication sales. In summary, during the 2009-2010 influenza season, analyses showed total school absenteeism correlated slightly with all three surveillance systems, with the strongest correlation to ILI ED visits. While ILI ED visits were significantly higher for E/M school-aged group, this trend was not paralleled in percent school absenteeism, with no significant difference between E/M and high school-aged groups. In addition to this inconsistency, peak activity within the 2009-2010 influenza season appeared to influence the strength of correlation between school absenteeism, ILI ED visits, OTC thermometer sales, and OTC cough/cold medication sales. However, SaTScan™ spatiotemporal analysis detected schools with high absenteeism, where two clusters were detected at two different elementary schools on the peak days of the 2009-2010 influenza season (November 2-10).

This evaluation of LAC school absenteeism data was not without limitations, including the major limitation of the lack of a “reason for absence” field. As concurred by other studies^{2,8}, providing reason for absence (e.g., ILI-related) improves disease-specific outbreak detection. Several other inherent data limitations included: (1) a 4-day to 4-week lag time of reported dates of absence, (2) the data were only available from Mondays through Fridays, with a likelihood of higher absenteeism on Mondays and Fridays (i.e., day of the week effect), (3) schools were on three different track systems with varying observed holidays/scheduled breaks, (4) only one year of data was available in this study, and (5) only 16% of the targeted LAC schools were represented in this analysis. Despite these limitations, school absenteeism data still afford insight into trends of illnesses in school-aged children that may not be detectable by clinical means. Subsequent to addressing the aforementioned limitations, monitoring aberrant activity in school absenteeism data could serve to assess the need for school closures during school-wide, district-wide and/or county-wide disease outbreaks.

In conclusion, interpreting medical outcomes and time trends from a non-traditional source such as school absenteeism is challenging. Examining school absenteeism during both mild and aggressive influenza seasons may be warranted to fully evaluate its utility of early outbreak detection. In addition, continued assessments of current data capture methods and quality of school absenteeism data within LAC will be addressed before integration into ACDC ADSS' syndromic surveillance systems.

REFERENCES

1. Matlof H, Murray R, Kamei I, Heidbreder G. Influenza in Los Angeles County, 1968–69. HSMHA Health Rep. 1971;86:183–92.
2. Besculides M, Heffernan R, Mostashari F, Weiss D. Evaluation of school absenteeism data for early outbreak detection, New York City. BMC Public Health 2005, 5:105.
3. U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. http://nces.ed.gov/ccd/schoolsearch/school_list.asp?Search=1&DistrictID=0622710.
4. Tsui F, Espino J, Dato V, Gesteland P, Hutman J, Wagner M. Technical description of RODS: a real-time public health surveillance system. Journal of the American Medical Informatics Association. 2003;5:399-408.
5. Space-Time Permutation Model: Kulldorff M, Heffernan R, Hartman J, Assunção RM, Mostashari F. A space-time permutation scan statistic for the early detection of disease outbreaks. PLoS Medicine, 2:216-224, 2005.
6. Influenza Watch Los Angeles County. May 17, 2010;4(29);1-3. <http://publichealth.lacounty.gov/acd/FluSurveillance.htm>.
7. Holtry R, Hung L, Lewis S. Utility of the ESSENCE Surveillance System in Monitoring the H1N1 Outbreak. Online Journal of Public Health Informatics [Online] 2010;2:3. Available: <http://ojphi.org/htbin/cgiwrap/bin/ojs/index.php/ojphi/article/view/3028>.
8. Rodriguez D, Zhang G, Leguen F, O'Connell E, Bustamante M. Using public school absentee data to enhance syndromic surveillance in Miami-Dade County, 2007. Advances in Disease Surveillance. 2007;4:188.