

Predicting influenza using school absences in Allegheny County, PA during the 2010-2014 influenza seasons

Talia M. Quandelacy¹, Shanta Zimmer², Chuck Vukotich², Rachael Bieltz², Kyra Grantz⁴, David Galloway², Yenlik Zheteyeva³, Amra Uzicanin³, Derek A.T. Cummings^{1,4}

1. Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology, Baltimore, MD
2. University of Pittsburgh, Pittsburgh, PA
3. US Centers for Disease Control & Prevention, Division of Global Migration & Quarantine, Atlanta, GA
4. University of Florida, Emerging Pathogens Institute, Gainesville, FL

Background

Increases in influenza infections in the community may accompany increases in school absences among school-aged children. Thus absences may be useful in conducting surveillance for influenza activity in the community. We examined the use of school absences from 9 school districts including 122 schools in predicting county-level virologically-confirmed influenza cases in Allegheny County, Pennsylvania during the 2010 to 2015 influenza seasons.

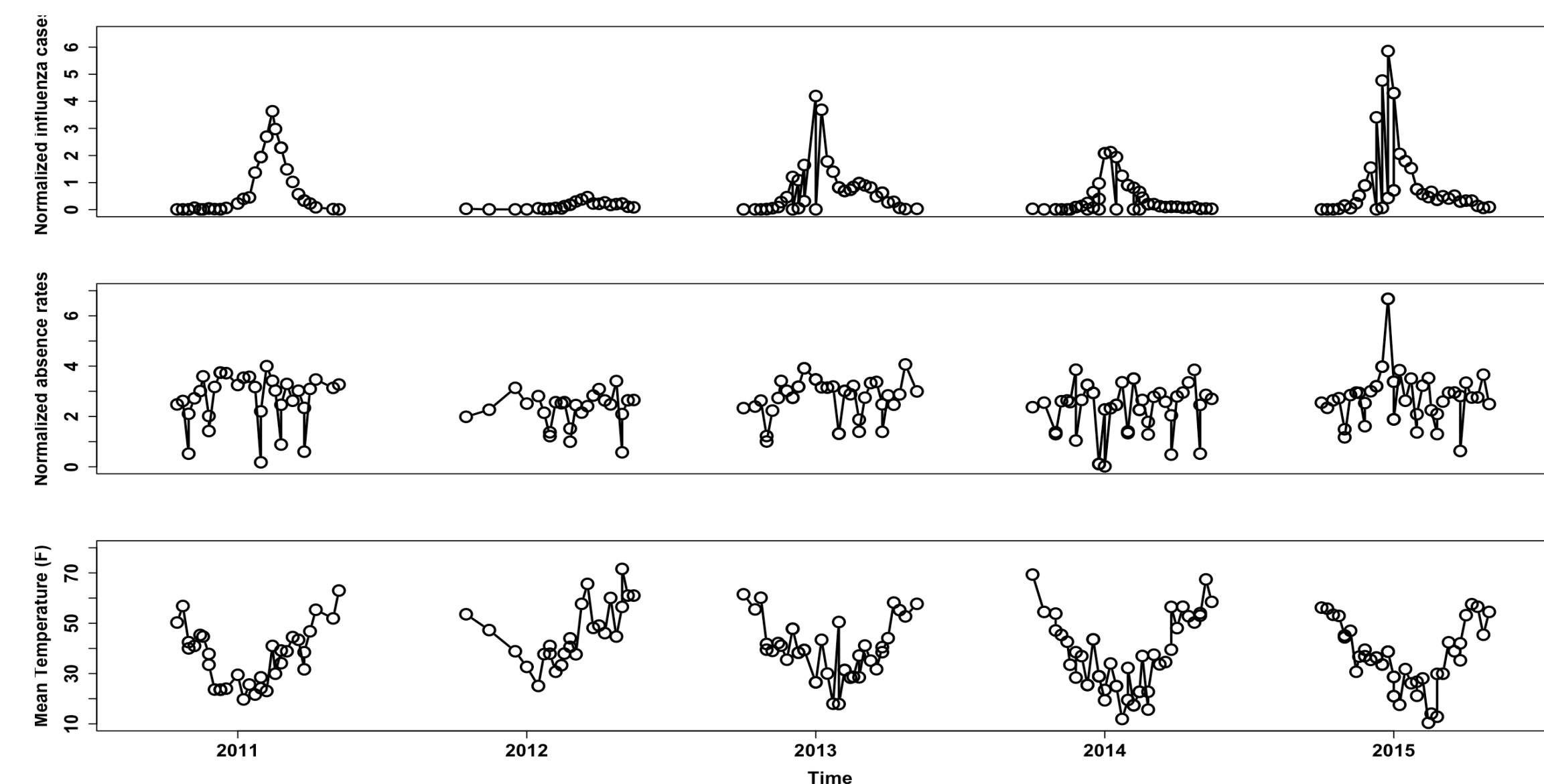


Figure 1. Normalized influenza cases, absence rates and average temperature from 2010-2015 in Allegheny County, PA

Methods

Data

Reported cases of virologically-confirmed influenza from 2010-2015 were provided by the Allegheny County Health Department. Predictors of interest were calendar week, month of the year, weekly average temperature, and weekly reported absences from nine K-12 school districts within Allegheny County.

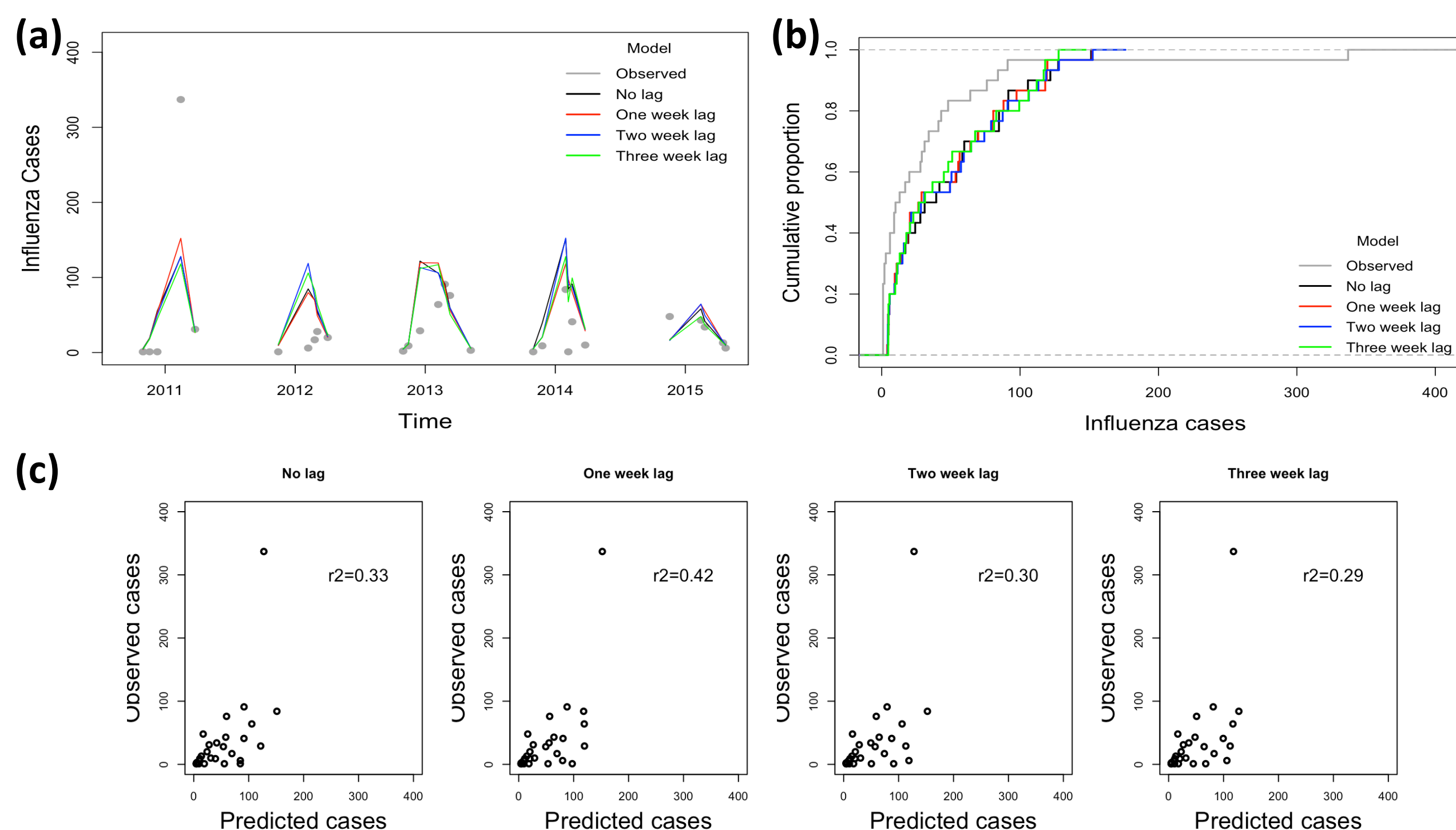


Figure 2. Predictive performance of models using varying absence rate lags, week, & mean temperature (a) over time, (b) comparing predicted and observed influenza cases, and (c) comparing the cumulative proportion of predicted to observed influenza cases

Approach

We built a number of negative binomial regression models using school absences, weeks and temperature to predict county-level influenza isolates. We assessed models of varying time lags (0, 1-, 2-, and 3-weeks), models of school types (i.e. elementary, middle & high school), and models of varying grade absences. We assessed model fits using percent deviance explained, and compared nested and non-nested models using likelihood ratio tests and Akaike's information criterion (AIC). Models were trained on 80% of the data and out of sample validation used the remaining 20%. All-type, school-specific, and grade-specific absence models used 1-week absence lags for comparability.

Table 1. Impact of weekly absence rate lags on the percent (%) deviance explained by absences using absence rates, calendar week and mean temperature models

| Lag | β_{Absence} | P-value | AIC | % Deviance Explained | N | DF |
|--------|--------------------------|---------|------|----------------------|-----|------|
| None | 0.544 | 0.173 | 1329 | 45.8% | 143 | 11.3 |
| 1-week | 0.671 | 0.103 | 1327 | 51.2% | 143 | 10.4 |
| 2-week | 0.416 | 0.287 | 1330 | 45.3% | 143 | 10.8 |
| 3-week | 0.071 | 0.858 | 1310 | 48.8% | 143 | 11.7 |

Each negative binomial model included weekly absence rates at various lags, calendar week, and mean temperature. Abbreviations: AIC: Akaike's Information Criterion; DF: Degrees of freedom

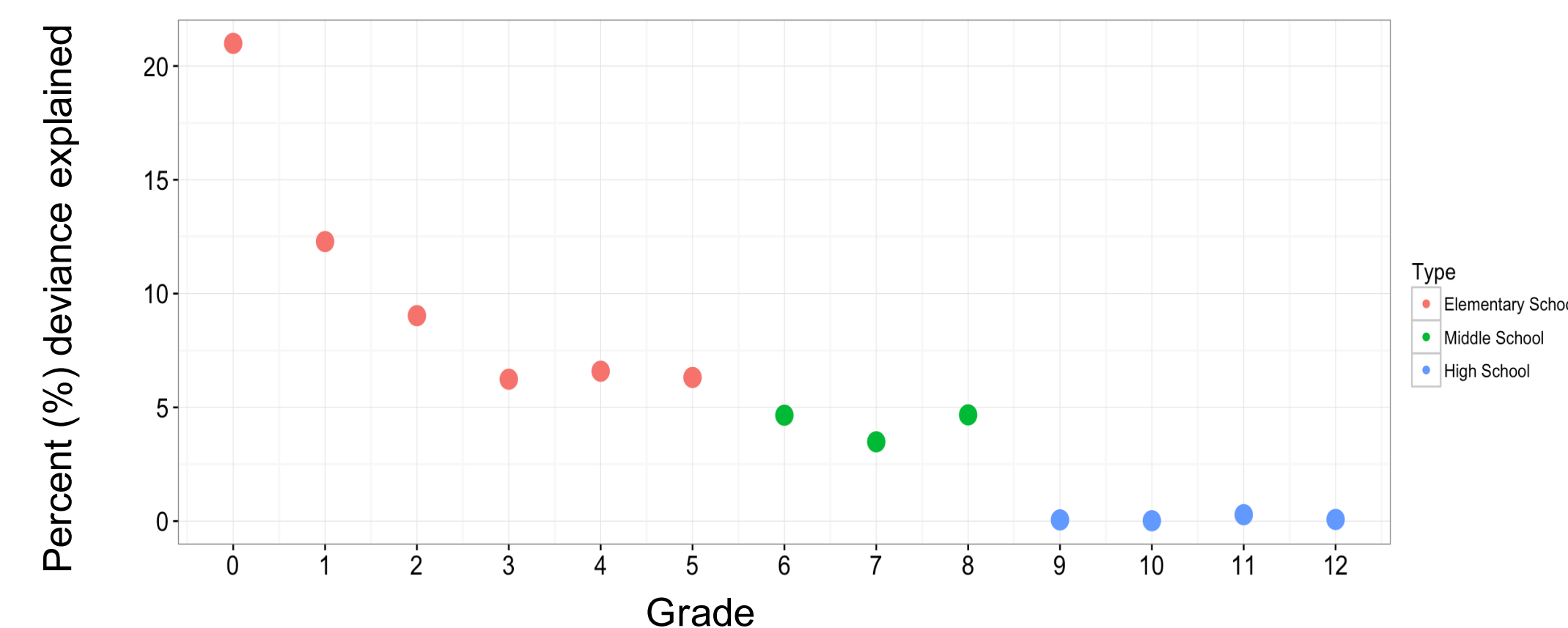


Figure 3. Percent deviance explained by weekly school absences from kindergarten to 12th grade. Colors represent the grades from different types of schools: red indicates elementary school, green indicates middle school and blue indicates high school

Table 2. Impact of weekly absence rate lags on the percent (%) deviance explained by models using kindergarten absences rates, calendar week and mean temperature

| Lag | β_{Absence} | P-value | AIC | % Deviance Explained | N | DF |
|--------|--------------------------|---------|------|----------------------|-----|------|
| None | 1.328 | 0.003 | 1030 | 50.3% | 117 | 11.2 |
| 1-week | 0.830 | 0.031 | 1033 | 49.1% | 117 | 11.6 |
| 2-week | 0.690 | 0.074 | 1157 | 48.7% | 117 | 11.5 |
| 3-week | 0.609 | 0.129 | 1035 | 48.8% | 117 | 11.9 |

Each negative binomial model included weekly kindergarten absence rates at various lags, calendar week, and mean temperature. Abbreviations: AIC: Akaike's Information Criterion; DF: Degrees of freedom

Results

Including school absences significantly improves predictive models of community viral isolates in Allegheny County, PA. Improvements using all absences are modest; however, using only school absences from younger grades (kindergarteners presented here), greatly improves predictive performance. Models utilizing kindergarten absences alone produced reasonably fitting models. When paired with calendar week and temperature, models explained 49.7 % of the deviance in the training set and 43% of the deviance in the out of sample test data sets. Lastly, absences in the previous week and two weeks were significant predictors of flu.

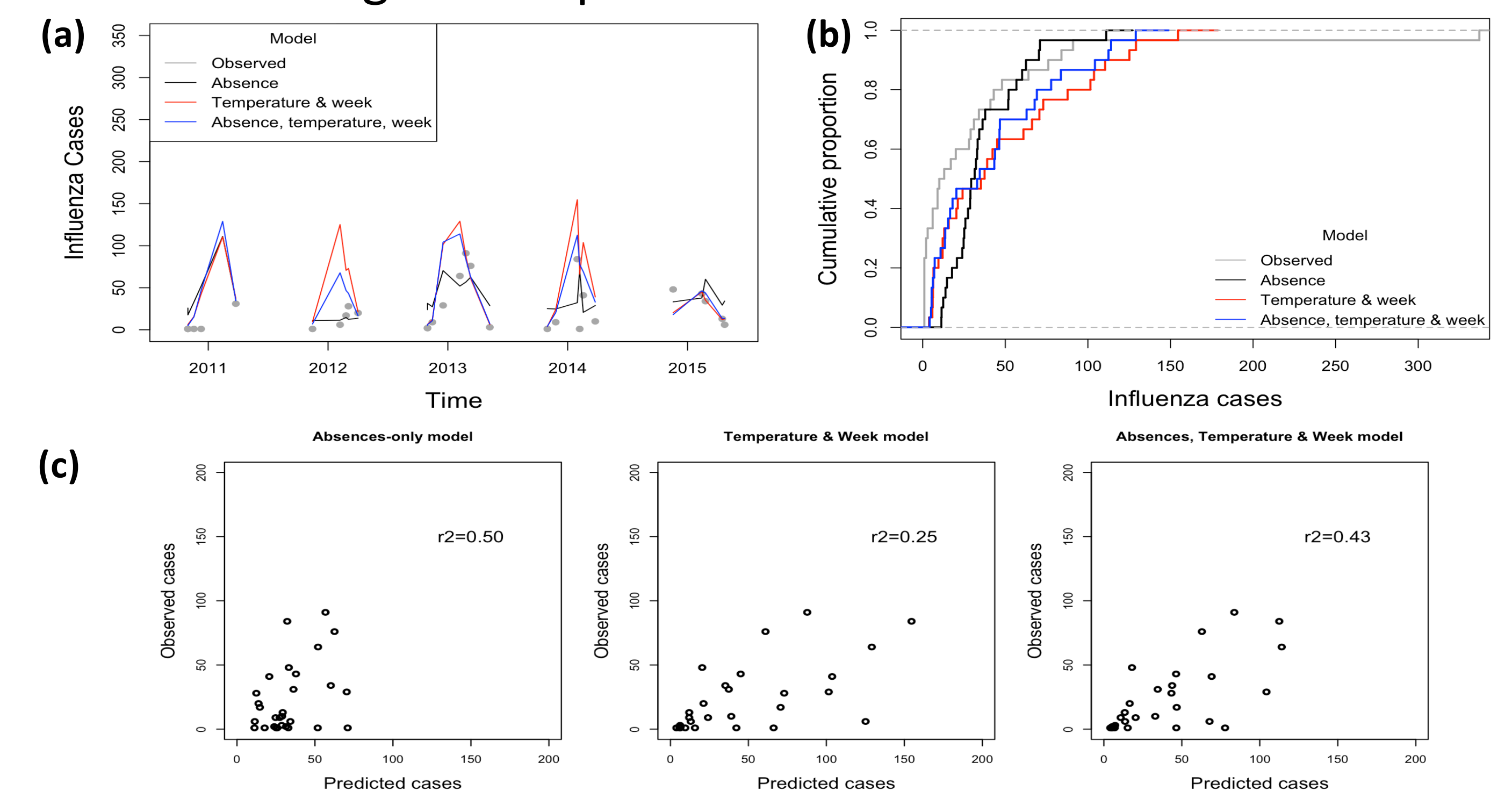


Figure 4. Predictive performance of models using kindergarten absences, week, & mean temperature (a) over time, (b) comparing predicted and observed influenza cases, and (c) comparing the cumulative proportion of predicted to observed influenza cases

Table 3. Model fits using negative binomial models of one-week lagged kindergarten absence rates, calendar week and mean temperature

| Model | β_{Absence} | P-value | AIC | % Deviance Explained | N | DF |
|-------------------------------|--------------------------|---------|------|----------------------|-----|------|
| Absence | 2.867 | <0.001 | 1213 | 20.6% | 132 | 3 |
| Temperature and week | - | - | 1166 | 47.5% | 132 | 10.5 |
| Absence, temperature and week | 1.442 | <0.001 | 1157 | 49.7% | 132 | 9.1 |

Each negative binomial model included weekly kindergarten absence rates lagged by one week, calendar week, and mean temperature. Abbreviations: AIC: Akaike's Information Criterion; DF: Degrees of freedom

Conclusion

We've previously demonstrated that younger students experience higher attack rates of respiratory pathogens compared to older students. Younger students experience non-illness-related absences with reduced frequency compared to older students, which may explain the increased performance using younger student absence data compared to older students. Our results suggest that models including younger student absences in improve the predictive performance of models of influenza incidence, even in models using different structures from the ones considered here.

Funding

This study is supported under the CDC grant 1 U01 CK000337-01.