**SUPPLEMENTARY TEXT**

**Supplemental Methods:**

We compared three commonly used methods that adjust for seasonality, and we found that monthly dummy variables resulted in significantly improved model fit compared to the other two methods (Table S1).

**Table S1 DIC scores of competing seasonality components**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Monthly Dummy | Polynomial Time Trends | Sinusoidal Curves |
| Deviance | 35954 | 38443 | 37671 |
| pD | 195 | 127 | 144 |
| DIC | 36149 | 38570 | 37815 |

The model structure of our hierarchical Bayesian regression is given as:

where denotes the number of all-cause respiratory hospitalizations at time *i*, in age group *j*, and SES group *k*; the expected value of is ; and the variance of is The parameter serves as an overdispersion parameter with indicating that the mean and variance are the same, as in a typical Poisson regression framework. We define the expected value as a function of covariates and random effects such that

(1)

where is the intercept parameter for age group *j* and SES group *k*; represents an intercept term that varies by epidemiologic year, where *g(i)* is a function that maps time to epidemiologic year (defined from July in the previous year to June in the next year); represents a similar intercept term which varies by month, where *m(i)* is a function that maps time to the corresponding month category; describes the group-specific effect of RSV, which associates RSV infections to respiratory hospitalizations; describes the association between influenza infections and respiratory hospitalizations, and varies by epidemiologic year in addition to the groups [28].

We model the coefficients of RSV infections ( parameters) as a multiplicative combination of age and SES effects, such that

where represents the age group effects; represents the SES effects; and accounts for other unexplained variation. A similar structure is applied to the coefficients of influenza infections (), with an additional yearly effect included, such that

where represents the age group effects; represents the SES effects; represents the epidemiologic year effects potentially due to differences in the severity of the circulating strain; and accounts for other unexplained variation. We assign weakly informative prior distributions to the remaining model parameters while ensuring that the remaining intercept parameters (i.e., ) are positive.

Posterior samples were collected using a Markov chain Monte Carlo (MCMC) algorithm. To make posterior inference, three chains of 12,500 MCMC iterations were used following an initial burn-in of 62,500 iterations per chain. The combined set of 37,500 MCMC iterations were then thinned by a factor of 10, resulting in 3,750 less correlated posterior samples from the joint posterior distribution with which to make inference. Convergence was assessed by examining individual parameter trace plots and Gelman-Robin diagnostics [38, 39]. Posterior means and 95% equal-tailed quantile-based credible intervals were calculated using the samples. The model was fitted using the rjags package [40]

The hyperparameters for are as follows:

After we fitted the model, the estimated “true” number of hospitalizations attributable to RSV infections for each time point and stratum were estimated by multiplying posterior samples of by . The average annual incidence of RSV was estimated by dividing the sum of over nine epidemiologic years by the age- and site-specific population. The recording ratios were calculated by dividing the number of recorded ICD-9-CM diagnoses for RSV in each age and SES group by the modeled estimates in the same group. The attributable percent of RSV was calculated by dividing the sum of by the sum of the model-predicted number of all-cause respiratory hospitalizations in each age and SES group ( over the entire study period. (Note that and in each group are almost identical; captures 99.98% of the variability in .) For each measure, we obtained and summarized samples from the posterior distributions of interest.

**Supplemental Results:**

Table S2 ICD-9-CM-coded respiratory hospitalizations in the entire population and RSV hospitalizations among children <2 years in New Jersey, New York, and Washington.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Number of ZIP codes | Respiratory hospitalizations | RSV hospitalizations |
| New Jersey | 581 | 2,555,728 | 170,703 |
| New York | 1715 | 5,482,926 | 328,833 |
| Washington | 512 | 1,379,736 | 100,575 |

Chart, scatter chart

Description automatically generated

Figure S1. Estimated annual RSV-attributable respiratory hospitalization rates by age and SES group in New York, July 2005 - June 2014.

The color texts show the mean estimates of annual RSV-attributable respiratory hospitalization rates in each age and SES group. The error bars indicate the 95% credible intervals of the estimated annual RSV-attributable respiratory hospitalization rates. Color blue, yellow, and red correspond to the estimates in populations from low, medium, and high SES ZIP codes, respectively.

Chart, scatter chart

Description automatically generatedFigure S2. Estimated annual RSV-attributable respiratory hospitalization rates by age and SES group in New Jersey, July 2005 - June 2014.

The color texts show the mean estimates of annual RSV-attributable respiratory hospitalization rates in each age and SES group. The error bars indicate the 95% credible intervals of the estimated annual RSV-attributable respiratory hospitalization rates. Color blue, yellow, and red correspond to the estimates in populations from low, medium, and high SES ZIP codes, respectively.

Chart, scatter chart

Description automatically generated

Figure S3. Estimated annual RSV-attributable respiratory hospitalization rates by age and SES group in Washington, July 2005 - June 2014.

The color texts show the mean estimates of annual RSV-attributable respiratory hospitalization rates in each age and SES group. The error bars indicate the 95% credible intervals of the estimated annual RSV-attributable respiratory hospitalization rates. Color blue, yellow, and red correspond to the estimates in populations from low, medium, and high SES ZIP codes, respectively.