Epidemiol. Infect. (2015), **143**, 2399–2407. © Cambridge University Press 2014
doi:10.1017/S0950268814003276

Using winter 2009–2010 to assess the accuracy of methods which estimate influenza-related morbidity and mortality

M. L. JACKSON[1]*, D. PETERSON[1], J. C. NELSON[1], S. K. GREENE[2], S. J. JACOBSEN[3], E. A. BELONGIA[4], R. BAXTER[5] AND L. A. JACKSON[1]

[1] *Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA*
[2] *Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA*
[3] *Kaiser Permanente of Southern California, Los Angeles, CA, USA*
[4] *Epidemiology Research Center, Marshfield Clinic Research Foundation, Marshfield, WI, USA*
[5] *Vaccine Study Center, Kaiser Permanente of Northern California, Oakland, CA, USA*

Received 12 June 2014; Final revision 31 October 2014; Accepted 12 November 2014; first published online 12 December 2014

SUMMARY

We used the winter of 2009–2010, which had minimal influenza circulation due to the earlier 2009 influenza A(H1N1) pandemic, to test the accuracy of ecological trend methods used to estimate influenza-related deaths and hospitalizations. We aggregated weekly counts of person-time, all-cause deaths, and hospitalizations for pneumonia/influenza and respiratory/circulatory conditions from seven healthcare systems. We predicted the incidence of the outcomes during the winter of 2009–2010 using three different methods: a cyclic (Serfling) regression model, a cyclic regression model with viral circulation data (virological regression), and an autoregressive, integrated moving average model with viral circulation data (ARIMAX). We compared predicted non-influenza incidence with actual winter incidence. All three models generally displayed high accuracy, with prediction errors for death ranging from −5% to −2%. For hospitalizations, errors ranged from −10% to −2% for pneumonia/influenza and from −3% to 0% for respiratory/circulatory. The Serfling and virological models consistently outperformed the ARIMAX model. The three methods tested could predict incidence of non-influenza deaths and hospitalizations during a winter with negligible influenza circulation. However, meaningful mis-estimation of the burden of influenza can still result with outcomes for which the contribution of influenza is low, such as all-cause mortality.

Key words: Influenza, modelling, statistics.

INTRODUCTION

The burden of hospitalizations and deaths caused by influenza is generally estimated using ecological trend studies (e.g. [1–10]). These studies use population-level rates of outcomes that are not specific to influenza, such as all-cause mortality. Outcome rates during time periods when influenza did not circulate are used to predict rates due to non-influenza causes during periods of influenza circulation. During influenza seasons, the difference between observed rates (which are due both to influenza and to non-influenza causes) and predicted rates (due to non-influenza causes only) is then attributed to

* Author for correspondence: Dr M. L. Jackson, 1730 Minor Ave, Suite 1600, Seattle WA, 98101-1448, USA.
(Email: jackson.ml@ghc.org)

influenza. Accurate estimates of the burden of influenza from these studies depend on accurate estimates of the predicted rates of non-influenza outcomes.

A key challenge to estimating non-influenza rates in these ecological studies is confounding by season [11]. Influenza tends to circulate in winter, which is when many other seasonal causes of morbidity and mortality also peak. Ecological study designs must account for this seasonal confounding; a variety of approaches have been used for this (e.g. [2, 4, 7, 8, 10, 12]). Importantly, while the assumptions that underlie these methods have been described [11], few data exist on whether these methods accurately account for seasonal confounding. Because seasonal influenza viruses circulate every winter in temperate regions, winter rates of outcomes in the absence of influenza are usually not observable. So it is unknown whether ecological studies successfully estimate the winter incidence of outcomes due to causes other than influenza.

The emergence and global circulation of the 2009 influenza A(H1N1) pandemic virus (2009 H1N1pdm) provided a unique opportunity to test the accuracy of ecological study designs. The 2009 H1N1pdm virus circulated early relative to the typical influenza season in the United States, with peak circulation occurring in September/October, and with influenza circulation essentially absent by mid-December 2009 [13]. Therefore, observed morbidity and mortality during the winter of 2009–2010 represents winter baselines of these events occurring in the absence of influenza. We used this winter to test the accuracy of common ecological methods used to estimate the winter incidence of outcomes due to causes other than influenza.

METHODS

Study population

We conducted this study within the Vaccine Safety Datalink (VSD) Project, a collaboration between the Centers for Disease Control and Prevention (CDC), America's Health Insurance Plans, and ten geographically diverse healthcare systems ('sites') [14]. In combination, the VSD contains data on site enrolment, healthcare utilization, and mortality for about 3% of the US population [14]. Our study used data from seven VSD sites with complete enrolment, demographic, hospitalization, and mortality data for the study period (1 September 1997 to 31 August 2010): Kaiser Permanente of Northern California (NCK; Oakland, CA); Kaiser Permanente of Colorado (KPC; Denver, CO); Health Partners Research Foundation (HPM; Minneapolis, MN); Marshfield Clinic Research Foundation (MFC; Marshfield, WI); Kaiser Permanente Northwest (NWK; Portland, OR); Kaiser Permanente of Southern California (SCK; Los Angeles, CA); and Group Health Cooperative (GHC; Seattle, WA).

We defined our study cohort as all seniors enrolled in one of the study sites between 1 September 2002 (KPC) or 1 September 1997 (other sites) and 31 August 2010. Seniors began contributing person-time to the study following their first year of continuous enrolment, and continued to contribute person-time until the earliest of death, disenrolment, or the study end date of 31 August 2010. We restricted our study population to seniors (adults aged $\geqslant 65$ years), for two reasons. First, older adults were less susceptible to 2009 H1N1pdm than were other age groups, due to cross-protective antibodies from A(H1N1) influenza strains that circulated prior to 1957 [15]. Thus, seniors would be unlikely to experience delayed health outcomes in winter that could have resulted from influenza infections in autumn. Second, seniors are at high risk of influenza complications and are the most common group for whom influenza burden is estimated (e.g. [12, 16, 17]).

Health outcomes

The primary health outcomes studied were all-cause mortality; hospitalizations due to pneumonia or influenza (PI); and hospitalizations for respiratory or cardiovascular (RC) conditions. We also studied hospitalizations for acute myocardial infarction (AMI) as a secondary outcome. We identified all onset dates of these health outcomes (which aggregate influenza-attributed and non-influenza-attributed events) in study population members during the entire follow-up period. We determined dates of death from administrative records at the participating sites, which combine death data from multiple sources (including hospital discharge records, state mortality records, and enrolment databases).

Hospitalization outcomes were defined by International Classification of Diseases, version 9, Clinical Modification (ICD-9-CM) codes assigned to inpatient visits: codes 480–487 (PI hospitalizations), codes 390–519 (RC hospitalizations), and code 410

https://doi.org/10.1017/S0950268814003276 Published online by Cambridge University Press

(AMI hospitalizations). These codes were chosen for consistency with previous studies [9, 18].

Influenza circulation data

We used data on positive influenza tests from the US World Health Organization and National Respiratory and Enteric Virus Surveillance System (WHO/NREVSS) collaborating laboratories. Publically available WHO/NREVSS surveillance data are stratified by each of the ten Department of Health and Human Services regions of the United States. For each site, we calculated the percent of specimens testing positive for influenza, overall and by type and subtype [A(H1N1), 2009 H1N1pdm, A(H3N2), and B], from the region in which the site is located as a measure of influenza circulation. For each year and region we defined influenza seasons as the consecutive weeks with at least 10% of isolates testing positive for influenza [9].

Analysis

Individual-level data on person-time and counts of health outcomes were aggregated by age group (65–69, 70–74, 75–79, 80–84, $\geqslant$85 years), sex, site, and study week. The aggregated weekly data at each site were merged with weekly influenza data from the corresponding region. We defined the prediction period as the weeks of 14 December 2009 to 1 March 2010, as >95% of influenza infections for 2009–2010 occurred prior to 14 December 2009 [13]. The remaining weeks were the baseline period used to fit the models.

For each of the health outcomes we fit three different statistical models to the data. We first estimated each model's parameters using the baseline period. We then used the baseline model to predict the expected incidence during the prediction period. Because the 2009–2010 winter was effectively influenza-free, the predicted rates are predictions of the outcome rates due to causes other than influenza.

The first statistical model was a cyclic regression model, which was first introduced by Serfling in 1963 [7]. In this approach, data from weeks when influenza circulated are removed from the time series. A cyclic regression model, using sine and cosine terms to represent seasonality, is then fit to the remaining data. Non-influenza incidence during periods when influenza circulates is interpolated from the model parameters, and differences between the observed and predicted influenza season incidence are attributed to influenza. We fit the following Poisson regression model to the data from weeks when influenza did not circulate, modelling the weekly count of events as a function of calendar time:

$$Y_t = \alpha \exp\begin{bmatrix} \beta_{0j} + \beta_1(t) + \beta_2(t^2) + \beta_3(\sin(2\pi t/k)) \\ + \beta_4(\cos(2\pi t/k)) + \beta_5(male) + \boldsymbol{\beta_6}(\mathbf{age}) \end{bmatrix},$$

where Y_t is the number of events during week t, k is the period of the time series ($k = 52{\cdot}177$ for weekly data), β_{0j} is the site-specific intercept (i.e. random intercept model), $\boldsymbol{\beta_6}$ is a vector of parameters for the age strata, and α is the offset terms for log of weekly person-time. Predicted incidence in the total population was calculated by computing a weighted average of the stratum-specific predicted incidences. Because model-based standard errors do not account for the autocorrelation of the data, we calculated 95% confidence limits using seasonal block bootstrapping [19, 20].

The second statistical model we used was the 'virological' regression model, which uses the Serfling model as a foundation and adds data on influenza circulation. This model has been the standard approach used by the CDC for estimating the burden of influenza from ecological studies since 2003 [9, 10, 21, 22]. In the virological regression model, all time points during the baseline period are used, including weeks from both influenza and non-influenza seasons. Parameters are included for percent of tests positive for each influenza type and subtype. For consistency with the standard use of these models [10, 21, 22], we did not include lagged effects of influenza. Incidence of non-influenza outcomes during influenza season is then predicted from the model parameters, setting the influenza covariates to zero. As with the Serfling model, we calculated 95% confidence limits using a seasonal block bootstrap.

The third statistical model we tested was an autoregressive, integrated moving average (ARIMA) time-series model [23]. Use of these models for predicting the burden of influenza has been described in detail elsewhere [2, 11]. In brief, an ARIMA model assumes that the incidence rate at time t (Y_t) is influenced by a 'random shock' (α_t) to the population at time t. The random shock is the cumulative effect of all factors affecting incidence, such as weather, temperature, pathogens, and air pollution. The effect of the random shock may persist for several time periods, so incidence at time t may depend on prior random shocks, α_{t-p} for some values of p. In addition, Y_t may depend on prior incidence, Y_{t-q} for some values of q.

https://doi.org/10.1017/S0950268814003276 Published online by Cambridge University Press

Conceptually, this could be due to depletion of susceptibles during the early stages of an epidemic, leaving fewer susceptibles in later stages. Thus, an ARIMA model has the form:

$$Y_t = \varnothing_1(\alpha_t) + \varnothing_2(\alpha_{t-1}) + \ldots + \varnothing_p(\alpha_{t-p}) + \theta_1(Y_{t-1}) + \ldots + \theta q(Y_{t-q}).$$

ARIMA models may also use differencing to remove trend or drift in the time series. In differencing, the dependent variable is the differenced time series Z_t:

$$Z_t = Y_t - Y_{t-d},$$

where d represents the level, or order, of differencing. ARIMA models can include seasonal lags in α_t and in Y_t to account for cyclic trends in incidence. ARIMA models are described by the orders of autoregressive, differencing, and moving average terms; for example, a (p,d,q)(1,0,0) model refers to a model with a first-order autoregressive term and no differencing or moving average terms; upper-case (P,D,Q) are used to refer to seasonal terms. Finally, ARIMA models can include other time series as independent variables; these models are sometimes referred to as 'ARIMAX' models. In our study, we used ARIMAX methods to model outcome incidence, where we included weekly counts of positive influenza tests as independent variables.

Unlike the cyclic regression models, where the model covariates may be chosen *a priori*, ARIMA and ARIMAX models are built empirically. The modeller identifies values for *p, q,* and *d* that are specific to the time series being modelled. For each outcome and site, we first fit an ARIMA model to the weekly observed data during the entire baseline period. We then added weekly percents of tests positive for influenza (by type and subtype) as predictors if they were significantly associated with the outcome after fitting the initial ARIMA model and if their inclusion in the model did not decrease the fit of the ARIMA model to the data. We used the resulting ARIMAX model to estimate (forecast) weekly outcome incidence rates during the prediction period.

Because ARIMA models are fit to a single outcome time series, an ARIMAX model cannot be adjusted for age or sex. Stratifying by age would have required fitting separate models for each age stratum, a fivefold increase in the number of models to fit. Instead, we fit an ARIMAX model to each outcome time series at each site, aggregated across all age/sex groups. To test whether stratifying by age might improve the accuracy of the ARIMAX model, we fit separate ARIMAX models to each of the five age groups for the PI outcome in one site (NCK), and compared the overall predicted incidence from the unstratified model with the combined predicted incidence from the stratified models. Forecasts from the unstratified model differed from the age-stratified models by <0·5 cases/10 000 person-years (data not shown).

Accuracy endpoints

The study endpoints to assess accuracy of the statistical methods were the errors between the observed and predicted rates of the health outcomes during the prediction period. We compared the predicted incidence of each outcome during the prediction period with the observed incidence in each site. We quantified the prediction accuracy of the statistical models by calculating the difference between the observed and predicted incidence rates. We calculated prediction error as both absolute differences and relative differences as a percentage of the observed incidence. We used US census data to standardize prediction error to event counts in the US population and compared annual predicted deaths in the United States from the virological regression model to recent CDC estimates for adults aged ⩾65 years based on the same model [21] during 1997–1998 to 2006–2007, years for which predictions were available in the present study and the CDC estimates.

This study was approved by the Institutional Review Boards of all participating sites. Analyses were conducted using SAS v. 9.2 (SAS Institute, USA) and Stata version 12 (StataCorp, USA).

RESULTS

We observed a total of 10 947 081 person-years of follow-up time during the study period, of which 31% was in seniors aged 65–69 years, 26% was in seniors aged 70–74 years, 20% was in seniors aged 75–79 years, 13% was in seniors aged 80–84 years, and 10% was in seniors aged ⩾85 years (Table 1). Our study population experienced 408 437 deaths; 149 630 PI hospitalizations; 1 507 965 RC hospitalizations; and 102 810 AMI hospitalizations during the study period. Incidence of all health outcomes fluctuated seasonally (Fig. 1).

In virological regression models, influenza A (H3N2) and B were significantly associated with deaths and with PI and RC hospitalizations. Influenza A(H1N1) was only significantly associated

https://doi.org/10.1017/S0950268814003276 Published online by Cambridge University Press

Table 1. *Distribution of person-time and outcomes by site, age, sex, and influenza year*

Group	Person-years	Deaths	Hospitalizations for		
			Pneumonia or influenza	All respiratory or circulatory conditions	Acute myocardial infarction
Full population	10 947 081	408 437	149 630	1 507 965	102 810
Site					
NCK	4 452 420	163 568	55 756	546 272	39 513
KPC	454 956	16 611	6782	99 851	3049
MFC	486 907	17 432	5472	56 995	3691
HPM	285 380	13 018	8528	85 728	4222
NWK	571 502	25 728	8052	84 014	5424
SCK	4 027 871	142 743	57 148	559 680	40 836
GHC	668 044	29 337	7892	75 425	6075
Age (years)					
65–69	3 429 263	45 346	19 766	262 012	19 351
70–74	2 796 897	60 820	25 627	306 805	21 491
75–79	2 183 608	76 719	31 492	331 352	22 094
80–85	1 471 171	87 529	32 753	300 214	19 733
⩾85	1 066 142	138 023	39 992	307 582	20 141
Sex					
Female	6 099 807	207 873	73 816	740 755	44 948
Male	4 847 274	200 564	75 814	767 210	57 862
Influenza year*					
1997–1998	643 653	23 737	8657	86 127	6480
1998–1999	674 248	24 950	9499	92 752	7143
1999–2000	710 292	27 207	10 502	99 575	7704
2000–2001	744 936	29 400	10 747	106 101	8374
2001–2002	770 184	29 340	11 577	108 363	8206
2002–2003	837 470	31 667	12 212	116 258	8787
2003–2004	895 008	34 883	13 517	125 171	9268
2004–2005	895 982	34 898	13 195	127 889	8658
2005–2006	914 922	35 610	12 649	126 777	8175
2006–2007	925 571	35 069	11 513	125 476	7918
2007–2008	940 641	34 246	12 413	130 387	7798
2008–2009	985 314	33 531	11 866	133 489	7325
2009–2010	1 008 861	33 899	11 283	129 600	6974

NCK, Kaiser Permanente of Northern California (Oakland, CA); KPC, Kaiser Permanente of Colorado (Denver, CO); MFC, Marshfield Clinic Research Foundation (Marshfield, WI); HPM, Health Partners Research Foundation (Minneapolis, MN); NWK, Kaiser Permanente Northwest (Portland, OR); SCK, Kaiser Permanente of Southern California (Los Angeles, CA); GHC, Group Health Cooperative (Seattle, WA).
* Influenza years run from 1 September to 31 August.

with RC hospitalizations, while influenza A(H1N1) pdm was only associated with PI hospitalizations. The exact form of the final ARIMAX models varied by health outcome and by site, but the most typical model included first-order differencing and a first-order moving average as well as a seasonal moving average term [i.e. a (p,d,q)(P,D,Q)(0,1,1)(0,0,1) model]. One or two influenza parameters [typically A(H3N2) or B] were usually included in the final models for death and for PI and RC hospitalizations. Influenza parameters were never significantly associated with AMI hospitalizations in either virological or ARIMA models. Yearly predicted US influenza-related deaths from the virological regression model during 1997–1998 to 2006–2007 ranged from 8531 to 36 972 and were well correlated ($R^2 = 0{\cdot}74$) with CDC's estimated deaths during the same time period, which ranged from 10 800 to 43 727.

During the prediction period, Serfling and virological regression models were more accurate that the ARIMAX models for all health outcomes except

https://doi.org/10.1017/S0950268814003276 Published online by Cambridge University Press

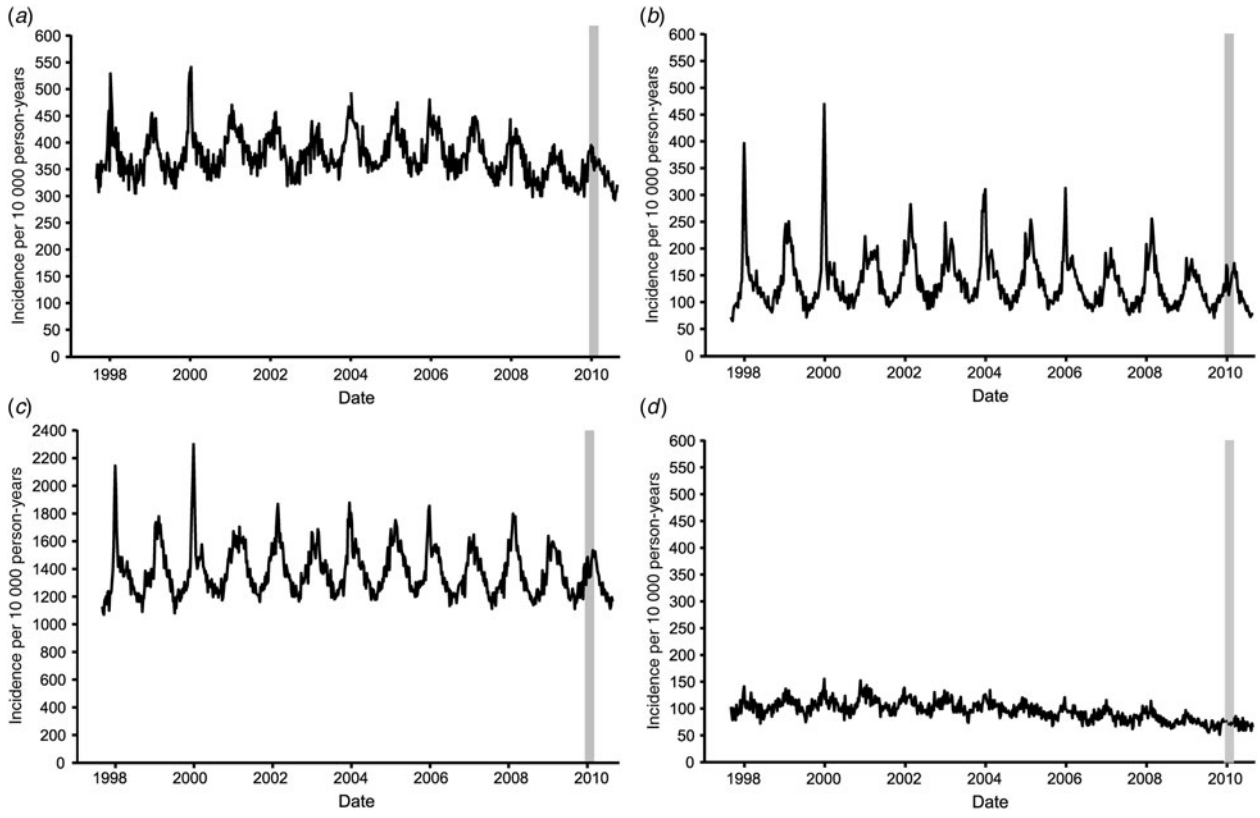

Fig. 1. Observed weekly incidence rates per 10 000 person-years, for (*a*) deaths; (*b*) pneumonia/influenza (PI) hospitalizations; (*c*) respiratory circulatory (RC) hospitalizations; (*d*) acute myocardial infarction (MI) hospitalizations. Grey bars indicate prediction periods.

AMI hospitalizations (Table 2). When averaged across all seven sites, all three statistical methods predicted non-influenza mortality rates during the winter of 2009–2010 with reasonable accuracy (Table 2). All three methods underestimated non-influenza mortality slightly, by 5% (ARIMAX), 2% (Serfling), and 2% (virological). Accuracy for the PI hospitalization outcomes was worse for the ARIMAX model, with 10% under-estimation. Accuracy was high for the RC hospitalization outcome for all three methods. However, confidence intervals were wide for all outcomes and all methods and spanned 0% prediction error. For example, the Serfling method predicted winter mortality incidence was 2% lower than the observed incidence, but the confidence interval on this prediction error was −21% to 17%.

We age- and sex-standardized the virological prediction errors to the US population. In a typical 12-week influenza season, a 2% underestimate of non-influenza mortality corresponded to overestimating deaths due to influenza by 9694 deaths. The 4% underestimate of PI hospitalizations corresponded to overestimating influenza-related PI hospitalizations by 5381.

DISCUSSION

This study provides new insights into the accuracy of methods used in ecological studies of the burden of influenza. In a winter with negligible influenza circulation, we found that all three of the methods predicted winter incidence of the health outcomes with fairly high accuracy. This finding was unexpected. The accuracy of the Serfling method in particular is surprising. This cyclic regression makes the assumption that the winter peak in non-influenza incidence is exactly equal to the summer trough in duration and amplitude [11, 24]. Seasonal changes in mortality and hospitalizations are likely influenced by numerous factors that vary over time in a complex way, such as temperature, weather, air pollution, hours of daylight, and circulation of other pathogens. *A priori*, we did not expect the simplistic cyclic regression function to accurately account for seasonal changes in mortality and in hospitalizations. Our findings suggest that changes in seasonal incidence of mortality and hospitalizations are primarily driven by factors that vary with the same timing and amplitude from year to year, such as hours of sunlight per day (e.g. [25]).

https://doi.org/10.1017/S0950268814003276 Published online by Cambridge University Press

Table 2. *Observed and predicted health outcome rates per 10 000 person-years and prediction errors, with 95% confidence intervals, using three statistical methods, during the influenza-free winter of 2009–2010*

Outcome	Observed rate	Predicted rates			Prediction errors		
		ARIMAX	Serfling	Virological	ARIMAX	Serfling	Virological
Death	369	352 (275 to 442)	361 (290 to 433)	361 (290 to 432)	−5% (−25% to 20%)	−2% (−21% to 17%)	−2% (−21% to 17%)
PI hosp.	140	126 (79 to 195)	137 (109 to 165)	135 (90 to 179)	−10% (−43% to 40%)	−2% (−22% to 18%)	−4% (−36% to 28%)
RC hosp.	1425	1377 (1130 to 1635)	1417 (1190 to 1643)	1428 (1192 to 1663)	−3% (−21% to 15%)	−1% (−16% to 15%)	0% (−16% to 17%)
AMI hosp.	74	76 (42 to 114)	78 (61 to 95)	77 (61 to 93)	2% (−43% to 55%)	5% (−18% to 28%)	4% (−17% to 26%)

PI, Pneumonia/influenza; RC, respiratory/circulatory, AMI, acute myocardial infarction.

We also found that the cyclic regression models, with or without viral circulation data, performed better than the ARIMAX model in predicting winter non-influenza incidence. This is also somewhat surprising, as ARIMAX models were developed specifically for handling some of the unique features of time-series data, such as trends over time, seasonal fluctuations, and autocorrelations [23]. We expected that fitting an ARIMAX model to each individual time series would result in better prediction than applying identical cyclic regression models to each time series. The fact that cyclic regression models appear to predict non-influenza incidence as well as ARIMAX models is also evidence that the seasonal variation in mortality and hospitalizations can be well modelled by a simple cyclic regression function.

Despite the high accuracy of the Serfling and virological regression methods, our results suggest that caution is needed in using these methods to estimate the burden of influenza. A 2% underestimate in non-influenza mortality corresponds to overestimating US influenza-related deaths by over 9694 deaths per year. A recent study using virological regression estimated that influenza causes an average of 21 098 deaths per year [21]. Thus, our results suggest that nearly half of this estimate could be attributable to prediction error rather than influenza. Because influenza only accounts for a small proportion of all winter deaths, even a small error in estimating non-influenza deaths can lead to large errors in deaths attributed to influenza. By contrast, Thompson *et al.* [9] used virological regression to estimate that influenza causes an annual average of 66 373 PI hospitalization in US adults aged $\geqslant$65 years. In our study, the virological method overestimated PI hospitalizations by 5300, which implies that prediction error only accounts for 8% of the PI hospitalizations attributed to influenza.

A potential limitation of this study is that, in contrast to typical years, influenza viruses were circulating intensely during autumn 2009. It is possible that seniors who would typically have had influenza-related hospitalizations or deaths during winter were affected instead in autumn. This effect in turn might cause models to overestimate 2009–2010 winter mortality on the basis of atypically high mortality in autumn 2009. However, we think this is unlikely. Influenza viruses circulating in autumn 2009 were almost entirely 2009 H1N1pdm, which caused comparatively little morbidity and mortality in seniors [13] due to cross-protective antibodies from influenza A(H1N1) strains

that circulated before 1957 [15]. Thus, the impact of the autumn pandemic wave on deaths in seniors was modest at best, and should not substantively affect model predictions. A second limitation is the VSD population used for this study represents only 3% of the total population of seniors in the United States. However, at the time we began this study in 2011, the mortality and hospitalization data that are commonly used for US burden of influenza studies were not yet available for the winter of 2009–2010. Extrapolating from our study population to the entire US for 1997–1998 to 2006–2007 gave similar mortality estimates as a recent study based on the entire US population [21], which increases our confidence that our sample is representative of the United States as a whole. The smaller sample size may lead to wider confidence intervals than are found in similar studies that use data for the entire US population, although similar studies have often ignored autocorrelation when estimating standard errors [4, 21, 22] or have not reported standard errors [9, 10]. We cannot rule out chance as an explanation for our findings that the methods tend to underestimate non-influenza morbidity and mortality.

Estimating the burden of morbidity and mortality caused by influenza remains a matter of public health importance. Our study suggests that ecological estimates of non-influenza outcome rates may be sufficiently accurate when applied to health outcomes where the contribution of influenza is large. By contrast, these models are probably not sufficiently accurate for outcomes where the contribution of influenza is low.

ACKNOWLEDGEMENTS

This work was supported by a subcontract with America's Health Insurance Plans (200–2002–00732) from the Centers for Disease Control and Prevention.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Alling DW, Blackwelder WC, Stuart-Harris CH.** A study of excess mortality during influenza epidemics in the United States, 1968–1976. *American Journal of Epidemiology* 1981; **113**: 30–43.
2. **Choi K, Thacker SB.** An evaluation of influenza mortality surveillance, 1962–1979. I. Time series forecasts of expected pneumonia and influenza deaths. *American Journal of Epidemiology* 1981; **113**: 215–226.
3. **Clifford RE, *et al.*** Excess mortality associated with influenza in England and Wales. *International Journal of Epidemiology* 1977; **6**: 115–128.
4. **Foppa IM, Hossain MM.** Revised estimates of influenza-associated excess mortality, United States, 1995 through 2005. *Emerging Themes in Epidemiology* 2008; **5**: 26.
5. **Neuzil KM, *et al.*** Influenza-associated morbidity and mortality in young and middle-aged women. *Journal of the American Medical Association* 1999; **281**: 901–907.
6. **Nunes B, *et al.*** Excess mortality associated with influenza epidemics in Portugal, 1980 to 2004. *PLoS ONE* 2011; **6**: e20661.
7. **Serfling RE.** Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Reports* 1963; **78**: 494–506.
8. **Simonsen L, *et al.*** The impact of influenza epidemics on hospitalizations. *Journal of Infectious Diseases* 2000; **181**: 831–837.
9. **Thompson WW, *et al.*** Influenza-associated hospitalizations in the United States. *Journal of the American Medical Association* 2004; **292**: 1333–1340.
10. **Thompson WW, *et al.*** Mortality associated with influenza and respiratory syncytial virus in the United States. *Journal of the American Medical Association* 2003; **289**: 179–186.
11. **Jackson ML.** Confounding by season in ecologic studies of seasonal exposures and outcomes: examples from estimates of mortality due to influenza. *Annals of Epidemiology* 2009; **19**: 681–691.
12. **McBean AM, Babish JD, Warren JL.** The impact and cost of influenza in the elderly. *Archives of Internal Medicine* 1993; **153**: 2105–2111.
13. **Anon.** Update: influenza activity – United States, August 30, 2009-March 27, 2010, and composition of the 2010–11 influenza vaccine. *Morbidity and Mortality Weekly Report* 2010; **59**: 423–430.
14. **Baggs J, *et al.*** The vaccine safety datalink: a model for monitoring immunization safety. *Pediatrics* 2011; **127** (Suppl. 1): S45–53.
15. **Hancock K, *et al.*** Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. *New England Journal of Medicine* 2009; **361**: 1945–1952.
16. **Brinkhof MW, *et al.*** Influenza-attributable mortality among the elderly in Switzerland. *Swiss Medical Weekly* 2006; **136**: 302–309.
17. **Nguyen-Van-Tam JS, *et al.*** Excess hospital admissions for pneumonia and influenza in persons $\geqslant$65 years associated with influenza epidemics in three English health districts: 1987–95. *Epidemiology and Infection* 2001; **126**: 71–79.
18. **Nichol KL, *et al.*** Effectiveness of influenza vaccine in the community-dwelling elderly. *New England Journal of Medicine* 2007; **357**: 1373–1381.
19. **Politis DN.** Resampling time series with seasonal components. *Computing Science and Statistics* 2001; **33**: 639–642.
20. **Weinberger DM, *et al.*** Serotype-specific effect of influenza on adult invasive pneumococcal pneumonia. *Journal of Infectious Diseases* 2013; **208**: 1274–1280.

21. **Anon.** Estimates of deaths associated with seasonal influenza – United States, 1976–2007. *Morbidity and Mortality Weekly Report* 2010; **59**: 1057–1062.
22. **Zhou H, *et al.*** Hospitalizations associated with influenza and respiratory syncytial virus in the United States, 1993–2008. *Clinical Infectious Diseases* 2012; **54**: 1427–1436.
23. **Box GEP, Jenkins GM.** *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day, 1976.
24. **Gilca R, *et al.*** The need for validation of statistical methods for estimating respiratory virus-attributable hospitalization. *American Journal of Epidemiology* 2009; **170**: 925–936.
25. **Kriszbacher I, *et al.*** The time of sunrise and the number of hours with daylight may influence the diurnal rhythm of acute heart attack mortality. *International Journal of Cardiology* 2010; **140**: 118–120.