

Temporal association between the influenza virus and respiratory syncytial virus (RSV): RSV as a predictor of seasonal influenza

A. MÍGUEZ¹, A. IFTIMI^{2*} AND F. MONTES²

¹ DGSP, Public Health Center of Valencia, Epidemiology Department, Valencia, Spain

² University of Valencia, Statistics and Operations Research Department, Valencia, Spain

Received 3 July 2015; Final revision 8 April 2016; Accepted 20 April 2016;
first published online 11 May 2016

SUMMARY

Epidemiologists agree that there is a prevailing seasonality in the presentation of epidemic waves of respiratory syncytial virus (RSV) infections and influenza. The aim of this study is to quantify the potential relationship between the activity of RSV, with respect to the influenza virus, in order to use the RSV seasonal curve as a predictor of the evolution of an influenza virus epidemic wave. Two statistical tools, logistic regression and time series, are used for predicting the evolution of influenza. Both logistic models and time series of influenza consider RSV information from previous weeks. Data consist of influenza and confirmed RSV cases reported in Comunitat Valenciana (Spain) during the period from week 40 (2010) to week 8 (2014). Binomial logistic regression models used to predict the two states of influenza wave, basal or peak, result in a rate of correct classification higher than 92% with the validation set. When a finer three-states categorization is established, basal, increasing peak and decreasing peak, the multinomial logistic model performs well in 88% of cases of the validation set. The ARMAX model fits well for influenza waves and shows good performance for short-term forecasts up to 3 weeks. The seasonal evolution of influenza virus can be predicted a minimum of 4 weeks in advance using logistic models based on RSV. It would be necessary to study more inter-pandemic seasons to establish a stronger relationship between the epidemic waves of both viruses.

Key words: ARMAX, influenza, logistic regression, RSV.

INTRODUCTION

Influenza virus and respiratory syncytial virus (RSV) are responsible for causing both mild and severe respiratory infections in people of all ages, although RSV is known for its propensity to particularly affect young children.

From an epidemiological point of view, there is a prevailing seasonality in the presentation of epidemic

waves of RSV infections and influenza. However, it is often difficult to predict with a high degree of accuracy the peak amplitude and duration, as defined by the shape of the seasonal curve [1]. These seasonal patterns vary annually, due to diverse circulating strains and various environmental factors [2].

Although there is a considerable variation in the intensity and timing of the transmission of these respiratory viruses over a period of years, there is always some degree of predictability. This predictability typically occurs in temperate climates with a clearly defined winter seasons [3].

Thanks to the efforts regarding surveillance of respiratory viruses, which has considerably strengthened in

* Author for correspondence: Dr A. Iftimi, University of Valencia, Statistics and Operations Research Department, 46100-Burjassot, Valencia, Spain.
(Email: iftimi@uv.es)

recent years, in February 2013 Bloom-Feshbach *et al.* [4] published a quantitative study of a systematic review of the seasonal patterns of influenza and RSV worldwide. Information on the seasonality of both viruses was available for 137 different localities throughout five continents. In addition, the weekly routine monitoring of the influenza virus available through FluNet, the WHO influenza surveillance system, allowed more refined time-series modelling in a subset of 85 countries [3].

Consistent with the 2006 recommendation of the European Influenza Surveillance Network (EISN) to incorporate RSV information under European influenza sentinel networks in Spain, a descriptive study [5] analysed the information on the movement of RSV obtained under the sentinel influenza surveillance system in Spain from 2006 to 2014; its usefulness was concluded to be a better characterization of seasonal influenza waves.

In Comunitat Valenciana, an autonomous region in Eastern Spain, influenza virus has been under surveillance since 2004, using the analysis of epidemiological surveillance system (AVE). Since 2007, and from the introduction of the microbiological surveillance network of Valencia (RedMIVA) dissemination of RSV incidence has been achieved, with information published on a weekly basis.

AVE is an electronic system which has been developed for epidemiological surveillance in Valencia since 2004. It permits the collection of real-time data from notifiable disease outbreaks and alerts, the analysis of which is automatically disseminated to users.

Within the Comunitat Valenciana, surveillance activities are carried out in 16 units, according to the region's health disaggregation, which act as the first level of specialized surveillance. The electronic monitoring is operative in all 16 surveillance units and, additionally, covers optional primary and speciality care, enabling the units to complete the clinical information of disease cases detected from socio-demographic data of system ambulatory information [6].

Incorporation of the microbiological results is done automatically via RedMIVA. RedMIVA is a Ministry of Health information system aimed at monitoring and research, and is based on the systematic collection of the results from 27 microbiology laboratories of the Valencian health system [7].

The rationale for this study is that the anticipated information on the influenza virus's epidemiology and seasonality, derived from the data collected by these

RSV record surveillance systems, is crucial to create and guide public health strategies at the community level in order to develop effective disease control measures.

The aim of this study is to quantify the potential relationship between the activity of RSV, with respect to influenza virus, so that the RSV seasonal curve might serve as a predictor of the evolution of an influenza virus epidemic wave.

METHODS

A population study was performed covering all the influenza cases reported in Valencia, by means of the AVE, and those cases of RSV registered in RedMIVA. The study period lasted from week 40 (2010) to week 8 (2014), a total of 177 weeks. The validation of influenza cases was performed according to the case definition of the national epidemiological surveillance network.

The objective of predicting the evolution of influenza from the information provided by RSV in advance, can be achieved by either of two different statistical methods: logistic regression and time series.

Binomial logistic regression

For a first approach to the problem, let us examine the graph of influenza evolution in Figure 1. Two distinct situations can be distinguished: one can be termed 'basal', referring to those weeks when a certain threshold is not exceeded, and the other designated the 'peak', when the weekly influenza cases exceed the threshold. Once the threshold is set, the variable gn can be categorized as a new variable, $estg2$, so that:

$$\begin{aligned} estg2 &= 0 \text{ (basal)} && \text{if } gn \leq u, \\ estg2 &= 1 \text{ (peak)} && \text{if } gn > u. \end{aligned}$$

In this context, the prediction has to be done in terms of probability and the model used is a binomial logistic regression [8] whose expression is,

$$\text{logit}\pi_t = \ln \frac{\pi_t}{1 - \pi_t} = \beta_0 + \sum_{k=m}^n \beta_k vrs_{t-k}, \quad (1)$$

where $\pi_t = P(estg2_t = 1)$ and t denotes the week. Using antilogarithms the following is obtained:

$$\pi_t = \frac{\exp(\beta_0 + \sum_{k=m}^n \beta_k vrs_{t-k})}{1 + \exp(\beta_0 + \sum_{k=m}^n \beta_k vrs_{t-k})}. \quad (2)$$

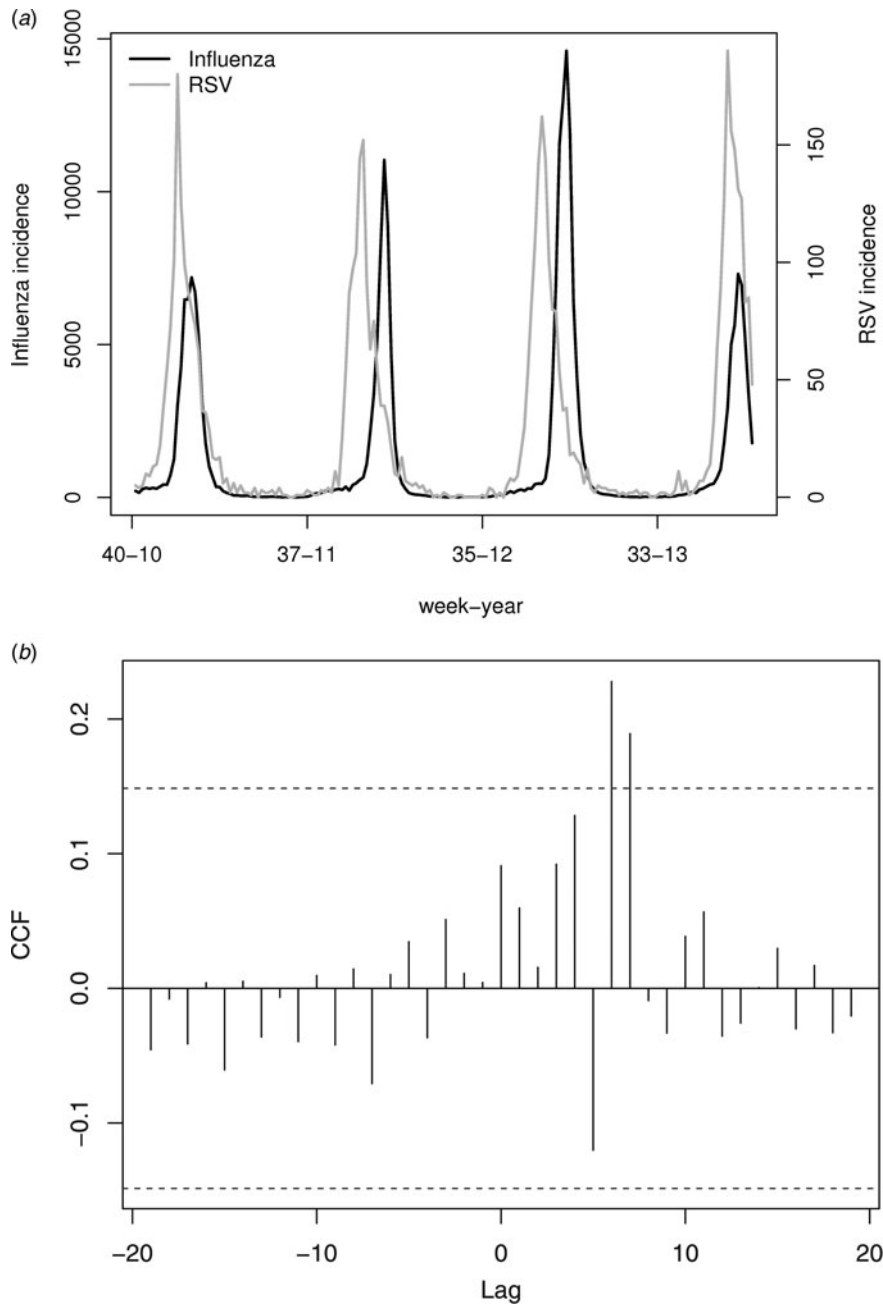


Fig. 1. (a) Weekly evolution of gn and respiratory syncytial virus (RSV) cases, (b) cross-correlation function (CCF) of both series.

Multinomial logistic regression

The above classification can be refined, taking into account the fact that the influenza peak has both a growing phase and a decreasing one. We can define a new variable under three categories,

estg3 = -1 (decreasing peak) if $gn > u$ and the epidemic decreases,
 estg3 = 0 (basal) if $gn \leq u$,
 estg3 = 1 (increasing peak) if $gn > u$ and the epidemic increases.

The logistic model is now multinomial [8] and given that there are more than two categories of dependent variable, the model consists of as many equations as categories minus 1. Each equation relates a category with the one taken as reference. Thus, if by π_i , $i = -1$,

0, 1 we represent $P(\text{est}g_t = i)$, and the reference is for basal category, the two model equations are:

$$\left. \begin{aligned} \log \frac{\pi_{-1}}{\pi_0} &= \beta_{0,-1} + \sum_{k=m}^n \beta_{k,-1} vrs_{t-k}, \\ \log \frac{\pi_1}{\pi_0} &= \beta_{0,1} + \sum_{k=m}^n \beta_{k,1} vrs_{t-k}. \end{aligned} \right\} \quad (3)$$

The relationship $\pi_{-1} + \pi_0 + \pi_1 = 1$ and (3) provide direct expressions for the probabilities associated with each category.

$$\left. \begin{aligned} \pi_{-1} &= \frac{\exp(\beta_{0,-1} + \sum_{k=m}^n \beta_{k,-1} vrs_{t-k})}{1 + \exp(\beta_{0,-1} + \sum_{k=m}^n \beta_{k,-1} vrs_{t-k}) + \exp(\beta_{0,1} + \sum_{k=m}^n \beta_{k,1} vrs_{t-k})}, \\ \pi_1 &= \frac{\exp(\beta_{0,1} + \sum_{k=m}^n \beta_{k,1} vrs_{t-k})}{1 + \exp(\beta_{0,-1} + \sum_{k=m}^n \beta_{k,-1} vrs_{t-k}) + \exp(\beta_{0,1} + \sum_{k=m}^n \beta_{k,1} vrs_{t-k})}, \\ \pi_0 &= \frac{1}{1 + \exp(\beta_{0,-1} + \sum_{k=m}^n \beta_{k,-1} vrs_{t-k}) + \exp(\beta_{0,1} + \sum_{k=m}^n \beta_{k,1} vrs_{t-k})}. \end{aligned} \right\} \quad (4)$$

Time series with regressors: ARMAX models

Predicting influenza in a given week t by the use of RSV values in previous weeks, the influence of the previous values of the influenza itself cannot be ignored. A model that includes the influenza time series and the previous RSV values as regressors, is an ARMAX model for which the general expression is

$$gn_t = \sum_{i=1}^p \phi_i gn_{t-p} + \sum_{j=1}^q \theta_j e_{t-j} + \sum_{k=m}^n \beta_k vrs_{t-k} + e_t, \quad (5)$$

where e_t is a white noise (i.e. random variables with zero mean and i.i.d.). In short, we have an ARMA model (p,q) to which a regression component with previous RSV values has been added. This model can be considered a special case of transfer function models, popularized by Box & Jenkins [9].

RESULTS

Throughout the study period 239 321 cases of influenza were reported, of which 125 135 (52.3%)

were female and 113 976 (47.6%) were male. The highest number of cases were in the 25–44 years age group (78195, 32.7%). Of the total number of reported influenza cases, 22 762 samples of respiratory secretions were processed, with 3436 (15.1%) testing positive.

Relative to RSV, 19 676 cases were recorded, of which 5112 (26%) were laboratory confirmed. Of those filed, 2815 (55.1%) were male and 2286 (44.7%) female. The most affected age group, with 4338 cases, was that of children aged <1 year, comprising 85% of RSV infections.

Looking at time series of influenza cases (gn) and confirmed RSV cases a similar behaviour is observed, but presenting with a certain time lag of the former with respect to the latter (see Fig. 1).

The series have been presented with different scales in order to avoid the effect of the large difference between the incidences of the two diseases. To corroborate the relationship between the two series, which Figure 1a seems to demonstrate, their cross-correlation function (CCF) has been obtained.

The ‘prewhiten’ function of the TSA package in R [10, 11] performs an automatic adjustment of an ARIMA model for both series and obtains the CCF of the residuals. Figure 1b shows the estimation of the CCF for gn and RSV series; it can be seen that RSV precedes gn with a positive cross-correlation between RSV values in the previous weeks (6 and 7) and the current week (gn).

Binomial logistic regression

The threshold chosen to divide gn values was $u = 567$ which amounted to an incidence of 11 influenza cases/

Table 1. Adjustment for the binomial logistic models

	β	s.e.	$\exp(\beta)$	95% CI
Constant	-3.4763	0.5753	0.0309	
RSV _{t-3}	0.0865	0.0177	1.0904	1.0532-1.1288
Constant	-4.1856	0.7633	0.0152	
RSV _{t-4}	0.1317	0.0306	1.1407	1.0744-1.2112

CI, Confidence interval; RSV, respiratory syncytial virus.

100 000 people. In order to get a simple model only a single RSV term was included in model (2). We have established a model for each RSV delay from $k = 3$ to $k = 7$. The set of observations has been randomly divided into two sets, the training set with two-thirds of the cases, and the validation set with the remaining third. The best results were obtained for delays 3 and 4. Table 1 shows the results of both adjustments with the estimation of the parameters and their standard errors. If the value of the log-likelihood is considered, the model with RSV_{t-3} is slightly better than the model with RSV_{t-4}.

An indirect measure of the model’s goodness is its ability to classify. Each observation is predicted as a *basal* or *peak*, according to the probability that the model assigns, the default criteria being to classify into the most likely group (0.5 as the cut-off point). The resulting classifications with both models and for both sets, training and validation, are shown in Table 2. The percentages refer to correctly predicted cases in each category and overall.

The receiver-operating characteristic (ROC) curves [12], are obtained by plotting ordered pairs (specificity, sensitivity) for different thresholds or cut-off points. Given its definition, the curve is within the unit square; its end points are (0,0) and (1,1). Figure 2 shows the curves for both models with the first quadrant’s bisector as reference line. A hypothetical ROC curve that coincides with this bisector means that for any threshold specificity and sensitivity would be 0.5. Just as if the classification had been done by tossing a coin, i.e. by pure chance.

A good method of classification is that for which both specificity and sensitivity are high, showing a graph above the first quadrant’s bisector, as seen by the two curves in Figure 2. The best possible method would be that in which the curve passes through point (0,1), coinciding with the left and top sides of the unit square. The area under the ROC curve (AUC) provides a measure to evaluate the quality of one method; only those methods with an AUC >0.5 are of interest.

The AUC for both curves, and respective confidence intervals obtained by bootstrap are shown in Table 3. Clearly both areas are well above 0.5 and close to 1, the maximum possible value.

Although the model RSV_{t-3} performs better both models are very similar, as tests based on their respective AUCs confirms. The corresponding statistic is $D = -1.274$ with $P = 0.207$, and the null hypothesis of equality of AUCs is accepted.

Moreover, observation of the ROC curves suggests the possibility of improving the specificity of both classifications by changing the cut-off point. Indeed, cut-off points that provide sensitivity and specificity values (see the circle in the graph) improve the specificity with minimal sacrifice of sensitivity. Table 4 shows the results of the classification with cut-off points $P = 0.125$ for the RSV_{t-3} model, and $P = 0.135$ for the RSV_{t-4} model.

Multinomial logistic regression

The best result was obtained by a model involving delays 3 and 6 for RSV. Table 5 shows the results of adjustment with estimation of the parameters and their standard errors.

Table 6 shows the result of classification with the multinomial model for training and validation sets. Percentages refer to those cases correctly predicted in each category and overall.

ROC curves cannot now be used to evaluate the model’s behaviour, since they cannot be constructed for a variable with more than two categories. Hand & Till [13] propose an overall measure that, as AUC, measures the ability of the model to separate each class from the remaining classes. It is expressed as follows:

$$M = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^k \hat{A}(i, j),$$

with

$$\hat{A}(i, j) = \frac{1}{2} [\hat{A}(i|j) + \hat{A}(j|i)],$$

$\hat{A}(i|j)$ being the probability that an individual of class j , randomly selected, would have an estimated probability of belonging to a class i less than an individual of class i , also randomly chosen. M also takes values between 0 and 1. Table 7 shows the value of M for the adjusted model and its 95% confidence interval obtained by bootstrap.

Table 2. Classification table for binomial logistic models

Observed	Model with RSV _{t-3}						Model with RSV _{t-4}					
	Training set Predicted			Validation set Predicted			Training set Predicted			Validation set Predicted		
	Basal	Peak	%	Basal	Peak	%	Basal	Peak	%	Basal	Peak	%
Basal	81	3	96.4	41	0	100.0	82	2	97.6	38	2	95.0
Peak	8	23	74.2	6	12	66.7	4	27	87.1	3	15	83.3
Overall %			90.4			89.8			94.8			91.4

RSV, Respiratory syncytial virus.

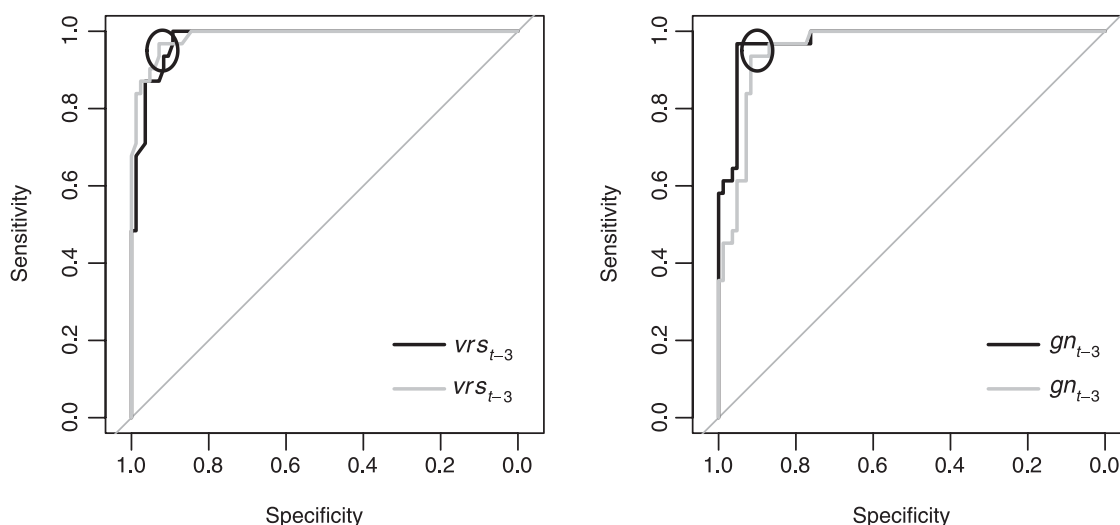


Fig. 2. Receiver-operating characteristic curves for binomial logistic models.

Time series with regressors: ARMAX models

The forecast function of the ‘auto.arima’ package in R [14] allows the simultaneous management of both regression and series adjustment. It was used by introducing delays 3–7 of RSV as regressors; as delay 3 had a non-significant coefficient, it was eliminated in a second adjustment, so that the final expression of model (1) is:

$$gn_t = \phi_1 gn_{t-1} + \phi_2 gn_{t-2} + \sum_{i=4}^7 \beta_i vrs_{t-i} + e_t. \tag{6}$$

Table 8 contains the results of the adjustment of model (6). Figure 3 shows the graph of original and adjusted series and the residuals’ autocorrelation function. It follows from Figure 3(a, b) that the model fits the data well and the residuals are compatible with white noise.

Table 3. Area under curve (AUC) for the two binomial logistic models

Model	AUC	mean.boot	se.boot	95% CI
RSV _{t-3}	0.9793	0.9794	0.0099	0.9585–0.9953
RSV _{t-4}	0.9869	0.9867	0.0072	0.9702–0.9978

CI, Confidence interval; RSV, respiratory syncytial virus.

The above adjustment was obtained using data of weeks 1–171, leaving the remaining 6 weeks for predictions. Table 9 shows these predictions, that are also represented in Figure 4. The graph begins at week 160 in order to be able to observe the previous behaviour of the model. The predictions up to week 174 are close to the observed values, confidence intervals containing the true value, but the predictions for the last 3 weeks overestimate the observed values.

Table 4. Classification table for binomial logistic models with different cut-off points

Observed	Model with RSV _{t-3} (cut-off = 0.125)						Model with RSV _{t-4} (cut-off = 0.135)					
	Training set			Validation set			Training set			Validation set		
	Predicted			Predicted			Predicted			Predicted		
	Basal	Peak	%	Basal	Peak	%	Basal	Peak	%	Basal	Peak	%
Basal	77	7	91.7	39	2	95.1	78	6	92.9	37	3	92.5
Peak	2	29	93.5	1	17	94.4	1	30	96.8	3	15	83.3
Overall %			92.2			94.9			93.9			89.7

RSV, Respiratory syncytial virus.

Table 5. Adjustment for the multinomial logistic model

	<i>b</i>	S.E.	exp(<i>b</i>)	95% CI
log $\frac{\pi_1}{\pi_0}$ Constant	-4.872	0.715		
RSV _{t-3}	0.053	0.018	1.054	1.017-1.093
RSV _{t-6}	0.059	0.017	1.061	1.027-1.097
log $\frac{\pi_2}{\pi_0}$ Constant	-4.568	0.660		
RSV _{t-3}	0.078	0.018	1.081	1.045-1.119
RSV _{t-6}	0.035	0.016	1.036	1.003-1.069
Log-likelihood	102.3			

CI, Confidence interval; RSV, Respiratory syncytial virus.

DISCUSSION

It is difficult to establish truly valid disease morbidity forecasts, when the knowledge of them is the result of a practitioner’s notification following a consultation, but there is a very high correlation between isolations of influenza virus in the laboratory and primary-care notifications of influenza, as can be seen in Figure 5 in which *gnpct* is the percentage of laboratory-confirmed influenza samples.

To the best of our knowledge, this is the first study that statistically quantifies the anticipation of the movement of RSV with respect to the influenza virus. The logistic regression analysis shows good results for both models, and provide a useful tool to predict the evolution of an influenza epidemic from the values confirmed for RSV at least 3 weeks in advance of the event.

Mention should be made of the threshold for the classification of the epidemic situation as basal or peak, denoting an increasing or decreasing phase. Here epidemiological criteria were selected, but it can also be determined graphically or automatically through the so-called threshold models [15, 16].

With regard to the application of time series, a model that fits a dataset well frequently fails to predict. The adjusted model is a good one since all covariates included present statistical significance ($P < 0.05$) and its residuals are compatible with white noise. For weeks 172–174, the predictions are quite good and their confidence intervals contain true values. For last three weeks the predicted values overestimate those actually observed with confidence intervals, the amplitude of which increase as the forecast horizon is further away.

The presence of delays 1 and 2 of the *gn* series in the autoregressive part of model (1), is not a problem when performing predictions above this horizon, because dynamic forecast incorporates the previous predictions for establishing subsequent ones. This is not the case with RSV values because they form part of the model as regressors and must be known in advance. As the model gains fresh information, it must be continuously adjusted to maintain the quality of the predictions.

Using data from two important surveillance systems, we have demonstrated a close temporal relationship between circulating influenza virus and RSV, as the peak of maximum activity of the influenza virus appears at least 3 weeks after the RSV peak. Our results are consistent with those of Gasparini *et al.* [17], which maintains that the RSV annual peak tends to occur in the absence of other respiratory viral pathogens and that influenza epidemics usually occur when RSV decreases.

In the study by Bloom-Feshbach *et al.* [4], the time-series models applied to FluNet data confirmed the presence of latitudinal gradients in seasonal influenza parameters, including timing peak, seasonal amplitude, epidemic duration, and fluctuations in the seasonal pattern from one year to another. They suggest that the global seasonal patterns of influenza and RSV are

Table 6. Classification table for multinomial logistic models

Observed	Training set Predicted				Validation set Predicted			
	Decreasing	Basal	Increasing	%	Decreasing	Basal	Increasing	%
Decreasing	8	3	1	66.7	6	0	2	75.0
Basal	1	79	2	96.3	1	39	0	97.5
Increasing	4	2	13	68.4	1	3	6	60.0
% Global				88.5				87.9

Table 7. M value for the multinomial logistic models

M	media.boot	se.boot	95% CI
0.8355	0.8347	0.0277	0.7794–0.884

CI, Confidence interval.

Table 8. Adjustment for the ARMAX model

	Coefficient	s.e.	z	P
f_1	1.4835	0.0563	26.3442	0.0000
f_2	-0.7073	0.0568	-12.4452	0.0000
RSV_{t-4}	8.8858	3.0682	2.8961	0.0038
RSV_{t-5}	7.8729	2.8336	2.7784	0.0055
RSV_{t-6}	22.1431	2.8563	7.7524	0.0000
RSV_{t-7}	17.6813	3.1347	5.6402	0.0000
Log-likelihood: -1267.8		AIC: 2549.6		

RSV, Respiratory syncytial virus; AIC, Akaike’s Information Criterion.

very similar in temperate locations throughout the Northern Hemisphere, characterized by focused peaks of activity during their respective winters, with a slight advance of RSV over influenza virus [3]. This advance is also observable in the works of Monto in America [3] and Freitas in Brazil [18].

Our results suggest that time-series models can be a useful tool for reporting the relative frequency of viral agents in a given clinical series, for providing a real-time circulation of respiratory viruses during the winter season, while at the same time allowing comparison of patterns of behaviour in different years.

This study has some limitations. On the one hand our models are not applicable in the case of an influenza pandemic since, in those circumstances, the spread of the disease expands logarithmically. On the other hand, due to the lack of laboratory-confirmed RSV infection in the adult population, the actual magnitude of the impact of RSV on adults is not clearly known [19].

Table 9. Predictions for the last 6 weeks

Week	Observed	Prediction	s.e.	95% CI
172	5638	6178.89	546.21	5108.33–7249.44
173	7309	7019.89	977.21	5104.60–8935.18
174	6943	7780.45	1272.94	5285.55–10 275.36
175	5183	7961.67	1423.43	5171.79–10 751.55
176	3341	8605.27	1470.25	5723.64–11 486.90
177	1781	8072.64	1473.34	5184.95–10 960.34

Some authors have described the existence of a biological mechanism called non-specific temporary immunity as a possible cause of the interference induced by the circulation of different respiratory viruses (especially RSV) with the evolution of the epidemic curve of influenza [20–22].

Taking this theory into consideration, we observe in Figure 1 that the seasons with the highest number of recorded cases of RSV (2010–2011 and 2013–2014), coincide with the lowest number of influenza cases. In both seasons, the predominant influenza virus type was A(H1N1) and the time lag between the RSV and influenza epidemic peaks was 3 weeks for the first season and 4 weeks for the second. For the time periods 2011–2012 and 2012–2013, the predominant influenza types were A(H3N2) and B, respectively, and the time lag between the period of maximum activity of both viruses was 6 weeks for the first time period and 7 weeks for the second. This observation supports the hypothesis of the existence of a cross-immunity phenomenon for a short period (3–4 weeks), between RSV infection and influenza virus, with it being more likely that the RSV infection has a short-term protective effect against type A(H1N1) influenza infection.

Since there is evidence of correlation between the progress of RSV and influenza, the public health authorities should utilize this. It would serve to emphasize aspects such as planning the most appropriate

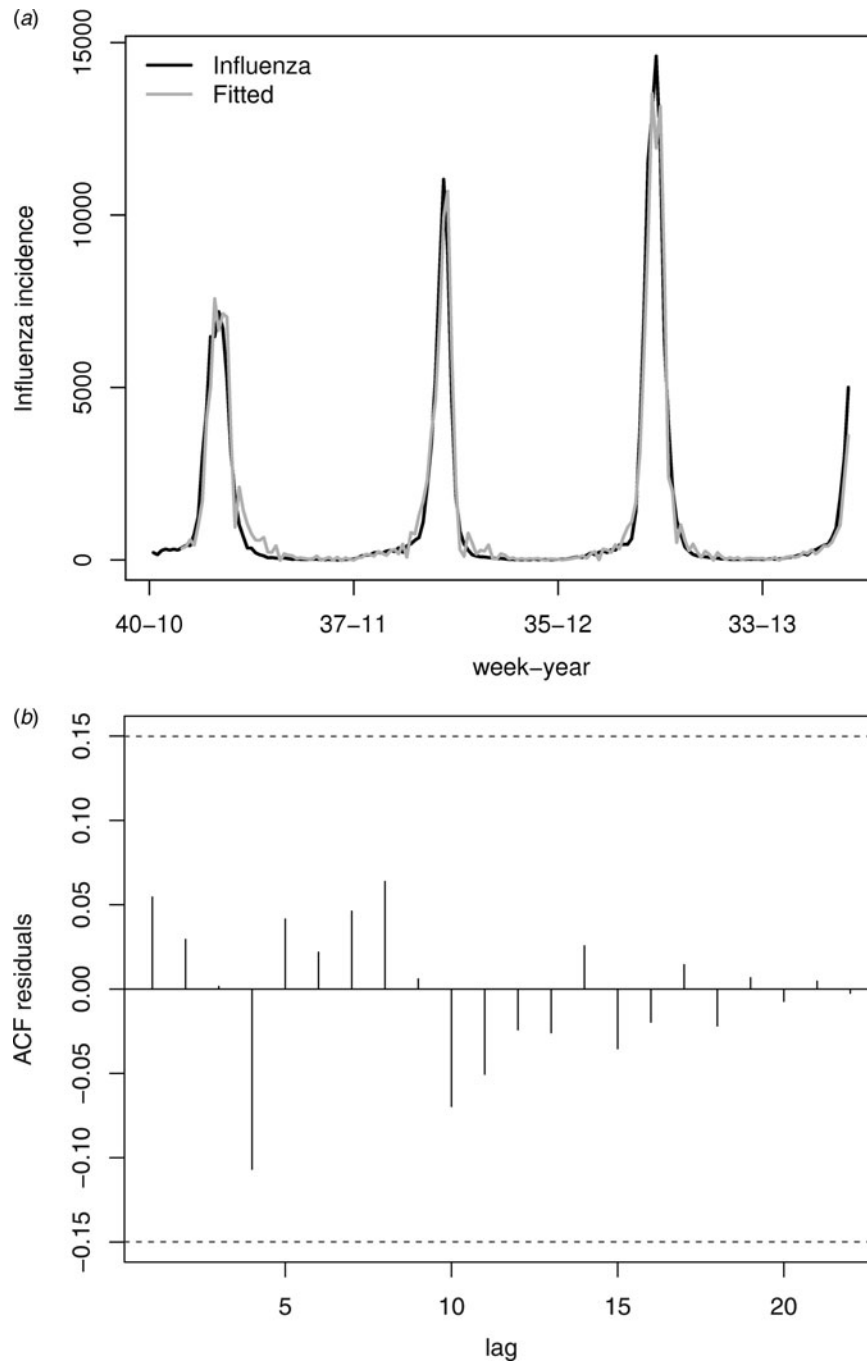


Fig. 3. (a) Original and adjusted series, (b) residual autocorrelation function (ACF).

time to vaccinate certain risk groups against influenza, and to implement other interventions to reduce its annual incidence, increasing the effectiveness of vaccination programmes and influenza prophylaxis.

The demand for outpatient and emergency-room admissions for acute respiratory infection in the winter months follows a parallel course to the incidence of influenza infection [23], so that the prediction of the disease's evolution would help institutions to optimize the

distribution of healthcare resources based on the changing burden of disease in the community. Following this idea, a study by Gilca *et al.* in Quebec [24] showed that the peak of maximum influenza activity precedes by a period of 1–2 weeks the peak in hospitalizations associated with influenza infection. Additionally, providing this information to medical practitioners could contribute to the improvement of the disease's clinical diagnosis.

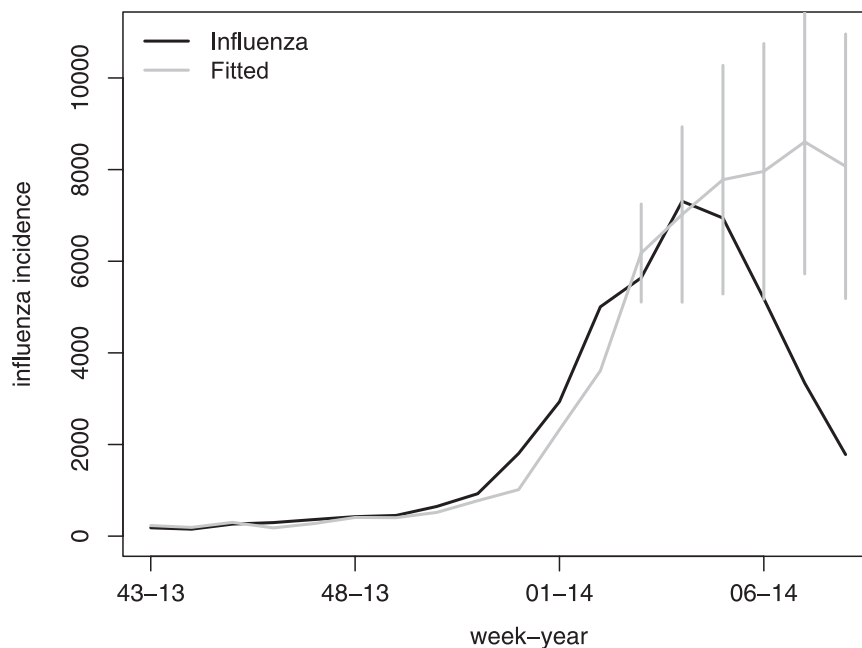


Fig. 4. Predictions and their 95% confidence intervals.

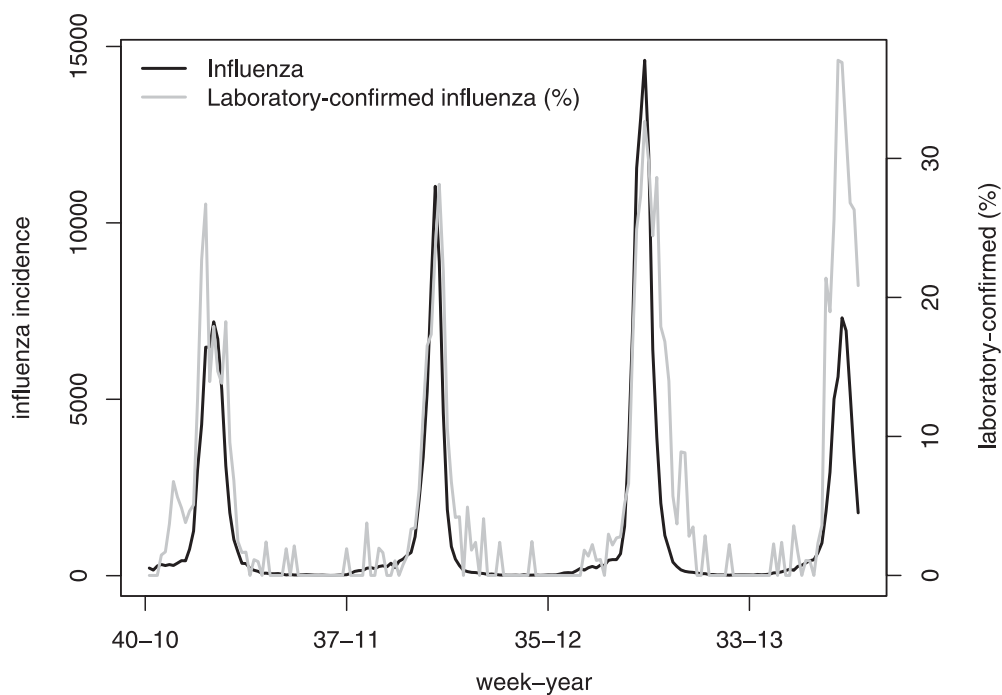


Fig. 5. Weekly evolution of *gn* cases and the percentage of laboratory-confirmed influenza samples, *gnpct*.

An advantage of having advanced knowledge of the evolution of seasonal waves of influenza virus is that taking timely preventive measures in order to achieve a decrease in the incidence of influenza, leads to a decline of serious bacterial diseases such as invasive pneumococcal disease. It has been shown in various

studies that there is a temporal relationship between influenza virus circulation and increased frequency of occurrence of these diseases [25–27].

In conclusion, our objective has been achieved since acceptable predictions of the seasonal evolution of influenza virus from RSV values were established a

minimum of 4 weeks in advance. These predictions are particularly satisfactory for the first 3 predicted weeks 172, 173, 174, although their quality deteriorates as the forecast horizon becomes further way. It would be necessary to study more inter-pandemic seasons to establish a stronger relationship between the epidemic waves of both viruses.

ACKNOWLEDGEMENTS

The authors acknowledge the valuable work done by members of the Servicio de Vigilancia y Control Epidemiológico de la Dirección General de Salud Pública: Francisco González Morán, Miguel Martín-Sierra, Silvia Guiral, Rosa Carbó, Isabel Huertas, Elvira Pérez, Amparo de la Encarnación, Teresa Castellanos and Celia Marín, and the epidemiologists from the Centros de Salud Pública de la Dirección General de Salud Pública, Generalitat Valenciana, Spain.

Adina Iftimi is supported by the Ministerio de Educación, Cultura y Deporte (grant FPU12/04 531) and Ministerio de Economía y Competitividad (grant MTM2013-43917). Francisco Montes is supported by the Ministerio de Economía y Competitividad (grants MTM2013-45381, MTM2013-43917).

DECLARATION OF INTEREST

None.

REFERENCES

1. **Spaeder MC, Fackler JC.** Time series model to predict burden of viral respiratory illness on a pediatric intensive care unit. *Medical Decision Making* 2011; **31**: 494–499.
2. **Lofgren ET, et al.** Disproportional effects in populations of concern for pandemic influenza: insights from seasonal epidemics in Wisconsin, 1967–2004. *Influenza and Other Respiratory Viruses* 2010; **4**: 205–212.
3. **Monto AS.** Occurrence of respiratory virus: time, place and person. *Pediatric Infectious Disease Journal* 2004; **23**: S58–64.
4. **Bloom-Feshbach K, et al.** Latitudinal variations in seasonal activity of influenza and respiratory syncytial virus (RSV): a global comparative review. *PLoS ONE* 2013; **8**: e54445.
5. **Jiménez-Jorge S, et al.** Monitoring respiratory syncytial virus through the Spanish influenza surveillance system, 2006–2014. *Enfermedades Infecciosas y Microbiología Clínica* 2016; **34**: 117–120.
6. **General Directorate of Public Health. Regional Ministry of Health. Valencian Autonomous Government.** Analysis of Epidemiological Surveillance (AVE). Mandatory Notification System. Epidemiology and Health Surveillance (www.sp.san.gva.es).
7. **General Directorate of Public Health. Regional Ministry of Health. Valencian Autonomous Government.** Microbiological Surveillance Network. Epidemiology and Health Surveillance (www.sp.san.gva.es).
8. **Agresti A.** *Categorical Data Analysis*, 2nd edn. New Jersey: John Wiley and Sons, 2002.
9. **Hyndman RJ, Athanasopoulos G.** *Forecasting: Principles and Practice*. OTexts, 2014.
10. **Kung-Sik C, Ripley B.** TSA: time series analysis. R package version 1.01, 2012 (<http://CRAN.R-project.org/package=TSA>).
11. **R Development Core Team.** R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2010.
12. **Pepe MS.** *The Statistical Evaluation of Medical Tests for Classification and Prediction*. New York: Oxford University Press, 2003.
13. **Hand DJ, Till RJ.** A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learning* 2001; **45**: 171–186.
14. **Hyndman RJ.** forecast: forecasting functions for time series and linear models. R package version 4.8, 2013 (<http://CRAN.R-project.org/package=forecast>).
15. **Hamilton JD.** *Time Series Analysis*. Princeton: Princeton University Press, 1994.
16. **Shumway RH, Stoffer DS.** *Time Series Analysis and its Applications. With R examples*. New York: Springer Science + Business Media, LLC, 2006.
17. **Gasparini R, et al.** Influenza and respiratory syncytial virus in infants and children: relationship with attendance at a pediatric emergency unit and characteristics of the circulating strains. *European Journal of Clinical Microbiology & Infectious Diseases* 2007; **26**: 619–628.
18. **Freitas FT.** Sentinel surveillance of influenza and other respiratory viruses, Brazil, 2000–2010. *Brazilian Journal of Infectious Diseases* 2013; **17**: 62–68.
19. **Mangtani P, et al.** The association of respiratory syncytial virus infection and influenza with emergency admissions for respiratory disease in London: an analysis of routine surveillance data. *Clinical Infectious Diseases* 2006; **42**: 640–646.
20. **Cowling BJ, Nishiura H.** Virus interference and estimates of influenza vaccine effectiveness from test-negative studies. *Epidemiology* 2012; **23**: 930–931.
21. **Anestad G, Vainio K, Hungnes O.** Interference between outbreaks of epidemic viruses. *Scandinavian Journal of Infectious Diseases* 2007; **39**: 653–654.
22. **Linde A, et al.** Does viral interference affect spread of influenza? *Eurosurveillance* 2009; **14**.
23. **Glezen WP.** The changing epidemiology of respiratory syncytial virus and influenza. Impetus for new control measures. *Pediatric Infectious Disease Journal* 2004; **23**: S202–S206.
24. **Gilca A, et al.** Seasonal variations in Clostridium difficile infections are associated with influenza and respiratory syncytial virus activity independently of antibiotic

- prescriptions: a time series analysis in Québec, Canada. *Antimicrobial Agents and Chemotherapy* 2012; **56**: 639–646.
25. **Talbot TR, et al.** Seasonality of invasive pneumococcal disease: temporal relation to documented influenza and respiratory syncytial viral circulation. *American Journal of Medicine* 2005; **118**: 285–291.
 26. **Zhou H, et al.** Invasive pneumococcal pneumonia and respiratory virus co-infections. *Emerging Infectious Diseases* 2012; **18**: 294–297.
 27. **Jansen AGSC, et al.** Invasive pneumococcal and meningococcal disease: association with influenza virus and respiratory syncytial virus activity? *Epidemiology and Infection* 2008; **136**: 1448–1454.