KCE REPORT 204S2

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Center

INFLUENZA VACCINATION: PRIORITIZING CHILDREN OR OTHER TARGET GROUPS?

ESTIMATION OF INFLUENZA ATTRIBUTABLE ADMISSIONS AND DEATHS: A REGRESSION ANALYSIS – SUPPLEMENT 2

KCE REPORT 204S2
HEALTH TECHNOLOGY ASSESSMENT

Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
Belgian Health Care Knowledge Centre

INFLUENZA VACCINATION: PRIORITIZING CHILDREN OR OTHER TARGET GROUPS?

ESTIMATION OF INFLUENZA ATTRIBUTABLE ADMISSIONS AND DEATHS: A REGRESSION ANALYSIS – SUPPLEMENT 2

GERMAINE HANQUET (KCE), FRANÇOISE WUILLAUME (WIV/ISP), FRANCE VRIJENS (KCE), PHILIPPE BEUTELS (UNIVERSITEIT ANTWERPEN), NANCY THIRY (KCE), NIEL HENS (UNIVERSITEIT HASSELT)

COLOPHON

Title: Influenza vaccination: prioritizing children or other target groups?

Estimation of influenza attributable admissions and deaths: a regression analysis – Supplement 2.

Authors: Gemaine Hanquet (KCE), Françoise Wuillaume (ISP – WIV), France Vrijens (KCE), Philippe Beutels (Universiteit Antwerpen), Nancy Thiry (KCE), Niel Hens (Universiteit Hasselt)

Reviewers: Jo Robays, Frank Hulstaert and Raf Mertens.

External Experts: Rik Baeten (VIGEZ), Johan Bots (Gemeenschappelijke Gemeenschapscommissie), Liesbeth Dejaegere (VIGEZ), Ann Malfroot (UZ Brussel), Daniel Reynders (SPF Santé Publique – FOD Volksgezondheid), Béatrice Swennen (Université Libre de Bruxelles), Isabelle Thomas (ISP – WIV), Geert Top (Vlaams Agentschap Zorg en Gezondheid), Patrick Tréfois (Question Santé), Yves Van Laethem (Centre Hospitalier Universitaire St. Pierre), Anne Vergison (Université Libre de Bruxelles), Françoise Wuillaume (ISP – WIV).

Acknowledgements: Yannick Vandendijck acknowledges support from a doctoral grant of the Universiteit Hasselt (BOF11D04FAEC). Lander Willem acknowledges support from an interdisciplinary doctoral grant of the Universiteit Antwerpen (ID-BOF25759), Nele Goeyvaerts is beneficiary of a postdoctoral grant from the AXA Research Fund, Adriaan Blommaert acknowledges support from the Universiteit Antwerpen concerted research action number 23405 (BOF-GOA), Joke Bilcke is supported by a postdoctoral grant from the Fund for Scientific Research Flanders (FWO), Niel Hens acknowledges support from the Universiteit Antwerpen scientific chair in Evidence-Based Vaccinology, financed in 2009–2013 by a gift from Pfizer. Support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy) is gratefully acknowledged.

For the simulations we used the infrastructure of the VSC - Flemish Supercomputer Center, funded by the Hercules foundation and the Flemish Government - department EWI. We thank Geert Jan Bex (Universiteit Hasselt and KULeuven) for support with using the VSC cluster, as well as Frank Van Reeth (Expertise Centre for Digital Media, Universiteit Hasselt).

We are indebted to the following persons for their help mainly through general advice, data delivery, data interpretation or review.

We thank Anthony Newall (University of New South Wales, Sydney, Australia) and Mark Jit (Health Protection Agency, The UK) for support in conducting systematic reviews of economic evaluations.

We are grateful to Samuel Coenen and Niels Adriaenssens (Centre for General Practice & Laboratory of Medical Microbiology, Vaccine & Infectious Disease Institute, University of Antwerp), Curt Brugman (Julius Center for Health Science and Primary Care, UMC Utrecht, the Netherlands), Pierre Van Damme (Centre for the Evaluation of Vaccination, Vaccine & Infectious Disease Institute, University of Antwerp) and Herman Goossens (Laboratory of Medical Microbiology, Vaccine & Infectious Disease Institute, University of Antwerp, and coordinator of Genomics to combat Resistance against Antibiotics in Community-acquired LRTI in Europe – GRACE, DG

Research, 2005) for helping with interpreting reviews and the medication data we collected through our survey on influenza like illness, and sharing antibiotics data of GRACE.

We also thank marketing research company “DayOne” (in particular Koen Van Bulck) for conducting the 2011-2012 influenza season survey on influenza like illness in the general population.

This study was made possible thanks to the collaboration of the Institut Scientifique de Santé Publique/Wetenschappelijk Instituut Volksgezondheid (ISP – WIV) that provided the sentinel laboratory and influenza data, the Reference Laboratory for *Streptococcus pneumoniae* at the UZ Leuven, the Technical Cell for the Minimal Clinical Data, the Federation Wallonia-Brussels, the Vlaams Agentschap Zorg en Gezondheid and the Observatorium voor Gezondheid en Welzijn van Brussel-Hoofdstad – Observatoire de la Santé et du Social de Bruxelles-Capitale for the data on the causes of death. We particularly thank Jan Verhaegen and Jozef Vandeven (UZ Leuven), Françoise Wuillaume, Viviane van Casteren, Geneviève Ducoffre and Isabelle Thomas (ISP – WIV), Déogratias Mazina and Murielle Deguerry (Observatoire de la Santé et du Social de Bruxelles-Capitale – Observatorium voor Gezondheid en Welzijn van Brussel-Hoofdstad), Erik Hendrickx, Anne Kongs and Heidi Cloots (Vlaams Agentschap Zorg en Gezondheid), Virginie Charlier (Fédération Wallonie-Bruxelles), Yves Parmentier (INAMI – RIZIV), France Vrijens, Jo Robays, Carl Devos, Stephan Devriese and Frank Hulstaert (KCE). We also thank Esther Kissling (EpiConcept) and Bianca Cox (University Hasselt) for advices on STATA programming. We also thank Kristel De Gauquier (KCE) for continuous support and management throughout the study.

External Validators: Daniel Brasseur (AFMPS – FAGG, European Medicines Agency and Hôpital Universitaire des Enfants Reine Fabiola), Marc Van Ranst (KULeuven, Hoge Gezondheidsraad – Conseil Supérieur de la Santé), Emilia Vynnycky (London School of Tropical Medicine and Hygiene).

Other reported interests: Marc Van Ranst (MVR) has acted as principal investigator or consultant for projects for which the University of Leuven received grants and funds for research. These grants and funds were directly paid to the University and MVR received no personal remuneration for this work. Yves Van Laethem received funds for research and consultancy fees and fees to speak at conferences from companies involved in influenza vaccines but these activities were not related to influenza, and these fees were paid directly to the Research Unit of his hospital. Anne Vergison received consultancy fees from companies involved in pneumococcal vaccines to participate to scientific advisory boards and to speak at conferences on pneumococcal vaccination. These fees were paid directly to her hospital. Beatrice Swennen received a travel grant to participate at a scientific conference. Patrick Tréfois collaborates to the Vax Info journal which received unrestricted educational grant from GSK.

Layout: Sophie Vaes

Disclaimer:	• **The external experts were consulted about a (preliminary) version of the scientific report. Their comments were discussed during meetings. They did not co-author the scientific report and did not necessarily agree with its content.** • **Subsequently, a (final) version was submitted to the validators. The validation of the report results from a consensus or a voting process between the validators. The validators did not co-author the scientific report and did not necessarily all three agree with its content.** • **Finally, this report has been approved by common assent by the Executive Board.** • **Only the KCE is responsible for errors or omissions that could persist. The policy recommendations are also under the full responsibility of the KCE**

Publication date:	12 July 2013
Domain:	Health Technology Assessment (HTA)
MeSH:	Influenza Vaccines; Influenza, Human; Pregnancy; Comorbidity; Health Personnel; Infant; Child, Preschool; Child; Adolescent; Young Adult; Middle Aged; Aged; Cost-Benefit Analysis; Costs and Cost Analysis.
NLM Classification:	WC 515
Language:	English
Format:	Adobe® PDF™ (A4)
Legal depot:	D/2013/10.273/47

How to refer to this document?

Hanquet G, Wuillaume F, Vrijens F, Beutels P, Thiry N, Hens N. Influenza vaccination: prioritizing children or other target groups? Estimation of influenza attributable admissions and deaths a regression analysis – Supplement 2. Health Technology Assessment (HTA) Brussels: Belgian Health Care Knowledge Centre (KCE). 2013. KCE Reports 204S2. D/2013/10.273/47.

This document is available on the website of the Belgian Health Care Knowledge Centre.

■ SUPPLEMENT REPORT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. SUMMARY

1.1. Background and method

A major problem is that admissions and deaths coded as influenza represent a minority of the true influenza hospitalisation and mortality burden.[1, 2] Indeed, only a minority of cases is confirmed by laboratory testing or recognized as due to influenza. The most common influenza-related severe outcome is pneumonia, which may also be caused by other infections (viral or bacterial). Several studies have addressed this problem by estimating the burden of severe influenza disease with the use of multivariate regression analysis, using the underlying temporal variations of influenza and pneumonia, as well as other co-variates of interest to attribute outcomes to these causative agents.[1-8]

For this reason, we conducted an analysis to estimate the number of hospitalization and deaths that are attributable to influenza in Belgium, using Belgian data and based on two major outcomes:

- Deaths from pneumonia and/or influenza (P+I):
  - Coded as principal (or underlying) cause of death
  - Coded as any of the different causes of death
- Admissions for pneumonia and/or influenza:
  - Codes as principal diagnosis
  - Coded as any diagnosis

We also conducted the same analysis on all respiratory and/or circulatory admissions for the sensitivity analysis.

We conducted a multivariate linear regression analysis using as dependent variable the weekly number of admissions coded as pneumonia and/or influenza (P+I) from the Minimal Clinical Data (MCD) on hospitalisation and the weekly counts of deaths coded as P+I from the death certificates. For both outcomes, we analysed separately P+I as *principal* diagnosis/cause of death and P+I as *any* diagnosis/causes of death. Independent variables were the weekly counts of respiratory pathogens that are the most frequent cause of influenza-like-illness or pneumonia (influenza A and B viruses, *Streptococcus pneumoniae*, adenovirus, RSV, *Mycoplasma pneumoniae,* parainfluenza virus, *Haemophilus influenzae),* provided by various laboratory networks and reference laboratories in Belgium. The inclusion of other parameters that may have an influence on the seasonality of P+I admissions were also tested. Separate models were built for each age group and analysis was restricted to the 5 calendar years 2004-2008, which includes 4 influenza seasons (from 2004-05 to 2007-08), as this was the period in which all data were available.

For all outcomes, a better fit was generally found in models including one influenza parameter by season (instead of a single influenza parameter for the whole period), interactions between pathogens, holidays and break returns, and a population term in some age groups. Lagged variables improved the models in the regression on deaths. Models that adjusted for pathogen surveillance coverage and blood culture trends did not improve the models and provided very similar estimates of influenza attributable admissions and deaths.

For admissions, as not all relevant parameters could be fitted into one final model (i.e. interactions of each pathogen with separate influenza parameters by season) as none of them was clearly superior, we ran two separate models:

- Model 1, with influenza parameterized by season;
- Model 2, with an interaction term between pathogens.

1.2. Prediction of influenza attributable admissions

Final models showed a reasonable goodness-of-fit and a minor level of serial correlation of residuals, except in the 75+ in which the fit was inferior. The number of predicted influenza admissions is very similar in both models 1 and 2. Interestingly, adding or changing some parameters that decreased substantially the level of autocorrelation and improved the fit did minimally change the predicted numbers.

When the outcome is P+I as principal diagnosis, around 2100 influenza admissions are predicted in an average season, representing an admission rate of 20/100 000 persons or a 6% of admissions for P+I as principal diagnosis (Table 1). When the outcome is any P+I admissions, above 3000 influenza admissions are predicted by season, representing an admission rate of about 30/100 000 persons or 4-5% of any P+I admissions. The admission rates vary largely across age groups (range 6-93/100 000 by age group for P+I main).

Table 1 – Predicted admissions, rates and % P+I admissions in an average influenza season, by age group and model

	Average number admissions (range by season)		Admission rate per 100 000		% of P+I admissions (average)	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Model based on P+I as principal diagnosis						
<5 years	540 (255-742)	600 (538-673)	92.7	103.0	8%	8%
5-14 years	287 (178-394)	290 (279-304)	23.6	23.9	11%	11%
15-49 years	309 (147-503)	301 (174-441)	6.2	6.0	7%	7%
50-64 years	201 (134-277)	178 (102-241)	10.5	9.3	5%	4%
65-74 years	234 (104-388)	179 (56-373)	24.8	18.9	5%	4%
75 years +	568 (121-1238)	554 (204-1088)	66.2	64.5	4%	4%
Total	2140	2102	20.3	20.0	6%	6%
Model based on P+I as any diagnosis						
<5 years	661 (338-925)	690 (630-798)	113.4	118.4	8%	8%
5-14 years	348 (208-489)	362 (354-369)	28.7	29.8	11%	12%
15-49 years	462 (277-673)	429 (257-598)	9.2	8.6	6%	6%
50-64 years	356 (198-517)	316 (191-480)	18.7	16.6	4%	3%
65-74 years	386 (137-785)	305 (72-655)	40.8	32.2	3%	3%
75 years +	1043 (345-2323)	1019 (376-2001)	121.5	118.6	3%	3%
Total	3256	3120	31.0	29.7	5%	4%

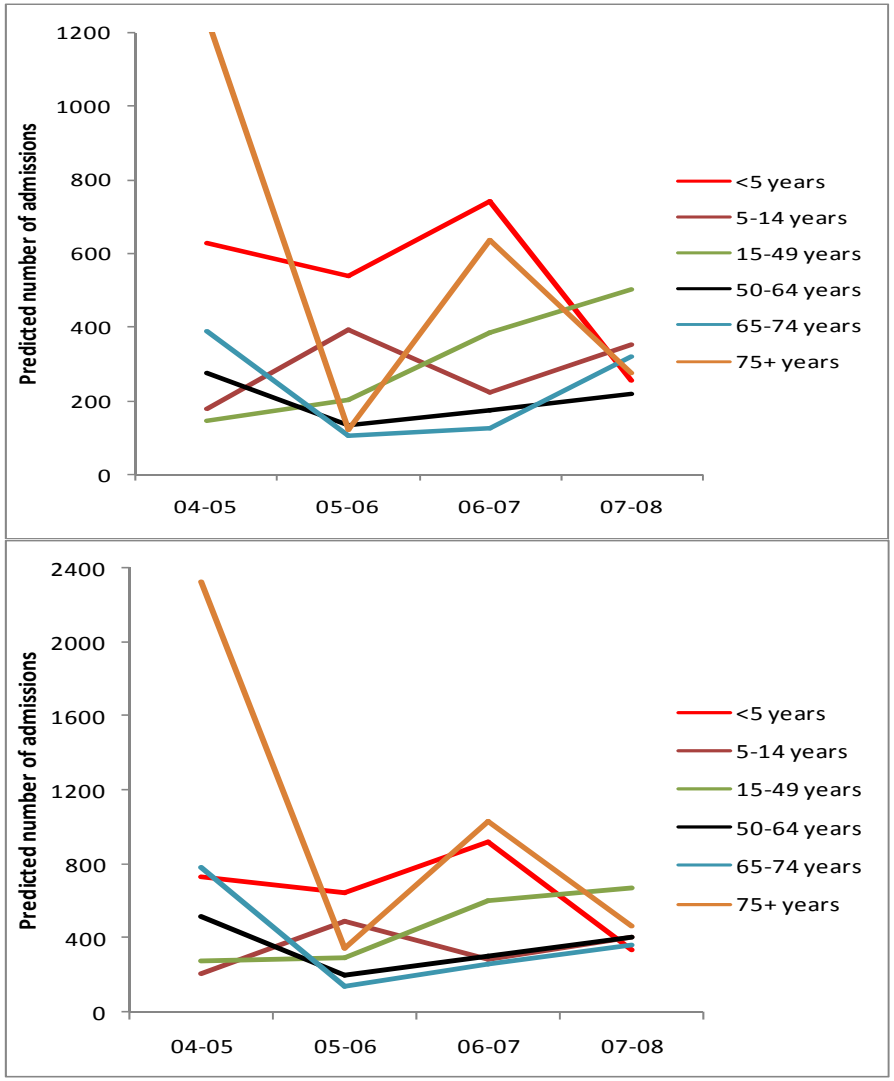
P+I: Pneumonia and/or influenza admissions.

The predictions vary substantially across seasons, with a range of 13-27/100 000 by season for all ages (Table 1 and Figure 1), especially with Model 1 as it parameterizes influenza by season. The four included seasons presented different levels of intensity: the 2004-05 and 2006-07 seasons were characterized as moderate intensity, and the 2005-06 and 2007-08 seasons were considered as low intensity. No season with a high level of intensity could be included in the analysis.

In general, the proportion of P+I admissions that are attributed to influenza by the models is low (6% for P+I main; 4-5% or any P+I) compared to the TIV vaccine efficacy values against ICD-coded P+I admissions (8-32%, see part I). This could be partly due to differences in coding systems, admission patterns and higher intensity influenza seasons in the settings that published efficacy studies (mostly US). The numbers of influenza admissions predicted by the P+I main diagnosis model are on average 40% higher than the number of ICD coded influenza admissions in the MCD dataset, confirming that outcomes coded or diagnosed as influenza are an underestimation of the true influenza burden. The difference is highest in the elderly 75+, in which the models predict 4 times more admissions than the MCD dataset.

A first sensitivity analysis predicted 46% more admissions if our models would only include influenza and RSV as independent parameters. In the elderly 75+ years, this model would predict 70% more admissions than our final selected model. Another analysis considered all respiratory and respiratory coded admissions as outcomes and predicted around 5000 admissions by season, representing 2.3-fold more admissions overall than P+I model 1 (principal diagnosis). This analysis predicted lower numbers of admissions in children but a 3.5-fold higher number in adults ≥50 years of age. However, most of these models showed a poor goodness-of-fit and the results should only be used for sensitivity analysis.

Figure 1 – Influenza admissions by season and age, predicted by model 1 for P+I main diagnosis (top) and any (bottom)



1.3. Prediction of influenza attributable deaths

No model could be run among children when regressing P+I deaths as main cause, and the models in P+I deaths as any cause found no deaths in the 5-14 years. Indeed, the numbers of coded P+I deaths in these groups were extremely low (<10/season). One model was selected in each age group as being clearly superior to others in terms of goodness-of-fit and residual distribution. In the 65+, the best models included a time lag for the dependent variables. All final models showed a good fit, and a minor level of auto-correlation of residuals was only observed in the 75+.

When the outcome is P+I as principal cause of death, 244 influenza deaths are predicted in an average season, representing a death rate of 2.3/100 000 persons and 6% of all P+I deaths as principal cause (Table 2). When the outcome is P+I as any cause of death, 356 influenza deaths are predicted by season, representing a death rate of 3.5/100 000 persons and 3% of P+I deaths as any cause. The death rate is low in young adults and highest among the elderly as expected, at 23.7-31.0/100 000 when regressing on P+I as principal cause or as any cause.

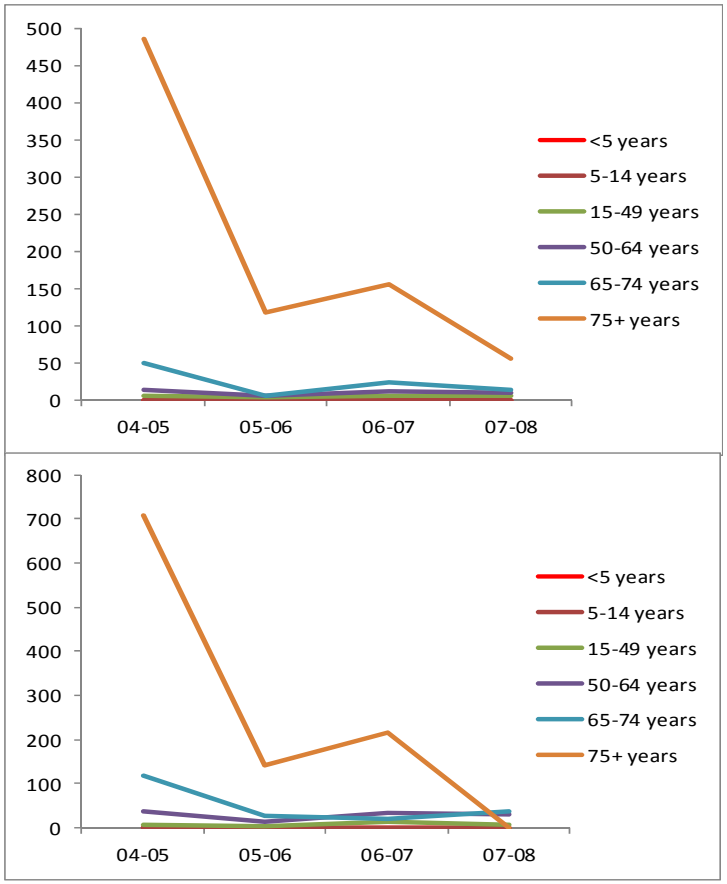
Table 2 – Predicted influenza deaths, death rates and % of P+I deaths by influenza (average across seasons) by age group

Age	Number deaths		Death rate		% of P+I deaths	
P+I cause of death	*Main cause*	*Any cause*	*Main cause*	*Any cause*	*Main cause*	*Any cause*
<5 years	0	2	0.0	0.3	NA	15%
5-14 years	0	0	0.0	0.0	NA	0%
15-49 years	5	8	0.1	0.2	11%	3%
50-64 years	11	30	0.6	1.6	6%	4%
65-74 years	24	51	2.5	5.4	6%	3%
75 years +	204	266	23.7	31.0	6%	3%
Total	244	356	2.3	3.4	6%	3%

The prediction varied substantially across seasons (Table 2), with 86-132 deaths in low intensity seasons and 200-556 deaths in moderate intensity seasons. The numbers of predicted influenza deaths are in average 3-fold higher than the number of ICD coded influenza deaths for both types of outcome (principal or any cause of deaths). A sensitivity analysis predicted the number of deaths in models would include influenza as sole pathogen. We found 40% more deaths in the 75+ for P+I as principal cause of death. When regressing on P+I as any cause of deaths, 80% more deaths would be predicted in the 65+, including 89% more deaths in the 75+.

The results of our final models are comparable with the influenza deaths predicted by one study using a similar outcome and comparable methodology conducted in Australia.[1] However, these deaths represented a higher proportion of P+I deaths than in our study (17-19% vs. 6% in our study). The studies that used broader causes of deaths (respiratory or all cause deaths) found much higher mortality rates.

Figure 2 – Predicted influenza deaths by season and age for P+I as main cause (left) and any cause (right)



1.4. Conclusions

The final models to predict the number of influenza attributable admissions and deaths show a relatively good goodness of fit and residual distribution, with the exception of the elderly ≥75 years. In these four seasons with moderate and low intensity, we predict a range of 2000-3000 influenza admissions and 250-350 influenza deaths by mean influenza season. A high variability of these outcomes across season is observed: the number of admissions in the highest season represented more than the double than those in lowest season, and the number of deaths predicted in the highest season accounts for 6 to 12-fold those predicted in the lowest season. This high variability of the influenza predicted numbers across seasons and age groups confirms the known variability and changing severity of influenza strains.

The estimates of influenza admissions and deaths also vary with the selected outcome. When regressing on any P+I diagnosis, we estimate around 50% more admissions and deaths by season compared to predictions based on P+I as principal cause models. When modeling all respiratory and circulatory admissions as dependent variable, we found a 2.3-fold higher number of admissions compared to estimates from P+I models (principal cause).

These results are difficult to compare with those from other studies, due to differences in outcomes, seasons, independent parameters, indicators reported, type of health system and health seeking behavior. In general, our admission estimates are in line with those from prospective studies, though only recent studies among children were found.[9, 10] Other regression studies predicted overall higher influenza admission and mortality rates in the elderly and lower admission rates in the younger groups.[1, 3, 6-8, 11, 12] However, most seasons covered by these studies were in the nineties when higher intensity seasons were observed, and many studies only involved influenza (and sometimes RSV) as pathogens. When we compared similar seasons and conducted analyses with similar outcomes and pathogens, our estimates were in line with those predicted by these studies.

Our study has two major limitations. One is a remaining level of auto-correlation of residuals, especially in the elderly. However, changes in the model that improved the independence of residuals did hardly affect the predicted influenza-related outcomes. The other limitation is that our study period involves four low to moderate intensity seasons, it is thus not representative of high intensity seasons in Belgium. This likely underestimates the influenza burden on admissions and deaths.

2. BACKGROUND

The objective of this analysis is to estimate the current burden of influenza disease in terms of hospitalizations and deaths in Belgium. Indeed, cost-effectiveness studies on influenza vaccination require reliable data on the total number of admissions and deaths that are caused by influenza, but admissions and deaths coded as influenza (e.g. ICD codes) represent a minority of the true influenza admissions and mortality burden.[1, 2] Indeed, only a minority of cases is confirmed by laboratory testing or recognized as due to influenza. Furthermore, there is no ICD code for influenza-like-illness. The most common influenza-related severe outcome is pneumonia, which may also be caused by other infections (viral or bacterial).

Several studies have addressed this issue by estimating the burden of severe influenza disease with the use of multivariate regression analysis on the temporal variation of reports of influenza and pneumonia, as well as other co-variates of interest.[1-8] This method predicts the number of outcomes (independent variable) that are attributable to influenza, based on influenza indicators and indicators for other factors influencing these outcomes (as dependant variables).

A further difficulty is that influenza hospitalizations are context-specific as they largely depend on health seeking behaviour and health service organization. We thus conducted a similar analysis based on Belgian data for two major outcomes:

- Deaths from pneumonia and/or influenza (P+I),
  - Coded as principal diagnosis
  - Coded as any diagnosis (principal or secondary)
- Admissions for pneumonia and/or influenza,
  - Codes as principal diagnosis
  - Coded as any diagnosis

3. METHODS AND DATA SOURCES

A multivariate linear regression analysis has been conducted. All analyses were run in STATA 12.0. We used generalized linear model regressions (glm) for Poisson, over-dispersed Poisson and negative Binomial regression. We used recent datasets from Belgium, originating from several sources:

- Minimal clinical data on hospitalization (MCD), coded by ICD-9 codes by Belgian hospitals. MCD with the week of admission was only available over 2004-2008.
- Deaths certificates requested from the Flanders, French Community and the Brussels region, coded by ICD-10 codes.
- Sentinel laboratory network registering the number of positive tests for a large number of pathogens in around 70% of Belgian laboratories.
- The National influenza Center which conduct the influenza surveillance in Belgium through a sentinel network of GP and is associated to the Influenza Reference Laboratory where swabs are analyzed.
- The reference Laboratory for *Streptococcus pneumoniae.*

3.1. Parameters considered in regression

The dependent variables are the following outcomes expressed as weekly count, stratified by age group:

- Weekly number of admissions coded as pneumonia and/or influenza (ICD9-coded) from the Minimal clinical data on hospitalization (MCD). We included MCD admissions with ICD-9 codes 480-488 (influenza and all-cause pneumonia or P+I), as primary diagnosis and as any diagnosis.
- Weekly counts of deaths coded as pneumonia and/or influenza (ICD-10 coded) based on death certificates received from the Flanders and French Community and the Brussels region. We included deaths with ICD-10 codes J10-18 (influenza and all-cause pneumonia or P+I), as main/underlying cause of death and as any cause of death.

As MCD with the week of admission and deaths from the French community were only available over 2004-2008, we restricted the analysis to these 5 calendar years.

The following independent variables were considered for inclusion in the model:

- Weekly counts of respiratory pathogens that may cause influenza-like-illness or pneumonia and are reported to the IPH Sentinel Laboratory (SL) network. This network registers the number of positive tests of a large number of pathogens in around 70% of Belgian laboratories. Pathogens considered are influenza (A and B), *S. pneumoniae*, adenovirus, RSV, *M. pneumoniae,* parainfluenza, *Haemophilus influenzae*.
- Weekly laboratory confirmed influenza outpatient counts, derived from the data of the National influenza Centre which conduct the influenza surveillance in Belgium through a sentinel network of GP; swabs are analysed and confirmed by the Influenza reference Laboratory.
- *S. pneumoniae* weekly counts from the KUL reference laboratory for *S. pneumoniae*.
- Holidays (summer and Christmas break) and school begin (after summer and Christmas break).
- A seasonal term.
- Population terms.

Addition of time lag on dependent and independent variables (pathogens) and interaction between pathogens have been considered as well. Age groups have been determined aiming at the best model and a balance between study needs and study power. We tested 2 age stratifications: 6 age groups (<5, 5-14, 15-49, 50-64, 65-74, 75+ years) and 3 age groups (<15, 15-64, 65+ years). One separate model was built for each age group.

Colinearity between independent variables has been assessed by the variance inflation factor (VIF) in standard regression. Independent variables were selected by backward stepwise linear regression for each model. The selection of the main model was initially based on AIC. However, as AIC function using glm in STATA does not take into account the over-dispersion in its calculation, selection of parameters for final models was based on Wald test ($p<0.20$).

3.2. Selection of regression models

Poisson and negative binomial models are the models of choice for count data.[1] We first tested Poisson regression as it is a common starting point for modeling count data, assessed over-dispersion to explore if the Poisson assumption holds, and assessed goodness-of-fit (gof). When a significant ($p<0.05$) test statistic from the gof indicated that the Poisson model was inappropriate, we used negative binomial regression, as this distribution is more appropriate in cases of over-dispersion. As the negative binomial model performed poorly in several age groups, we also tested "overdispersed Poisson regression", which adjusts for over-dispersion by using a Pearson scale dispersion parameter.[4, 13]

However, we found that similar studies on respiratory pathogens used other distribution models such as the standard Gaussian regression.[2, 5] As we have relatively large counts of the dependent variables, we decided to test all distributions (standard Gaussian, overdispersed Poisson and negative binomial) for comparison purposes.

For over-dispersed Poisson and negative binomial regression, an identity link has been selected as we opted for the more realistic assumption of an additive model based on biological grounds: as in other studies, we assume that the number of hospitalizations/deaths increases proportionally with unit increase of pathogen activity.[2, 4, 14] Though we do not assume a multiplicative model, i.e. that the number of hospitalizations/deaths increases exponentially with unit increase of pathogen activity, we also tested a log link in some subgroups to assess whether this link would produce a better model.

For the selection of distribution and link, we only compared the different options by using a "main effects" model. We also tested each model whether it would perform better by merging age groups, to address the potential issue of decreased power by allowing for over-dispersion.

The following variations to the main model have been tested:

1. Stratification by 3 large age groups (<15, 15-64, 65+ years). Some other combinations were also tested for adults.
2. Models with interactions between pathogen variables.
3. Models using separate parameters for influenza A and B, or separate coefficients for influenza by season.
4. Models adding a seasonal term.
5. Models without constant term.
6. Models with holidays and return from breaks.
7. Models with population as covariate, to account for population changes.
8. Models with time lags between pathogen variables and outcome.

The selection of the final models was based on the goodness of fit as measured by the deviance divided by the degrees of freedom (dev/df);[1] this choice was due to the fact that the AIC from over-dispersed Poisson does not integrate the higher uncertainty generated by the corrected scale for over-dispersion. Alternative models have been tested using the over-dispersed Poisson and compared to the main model.

3.3. Analysis of residuals

We used the Pearson residuals for residuals analysis. Residuals were tested for auto-correlation by visual inspection of auto-correlograms, Durbin-Watson test and the Portmanteau test for white noise. Distributions of residuals were also studied by visual inspection of plots against time and against the predicted outcome.

4. DESCRIPTION OF BASELINE PARAMETERS

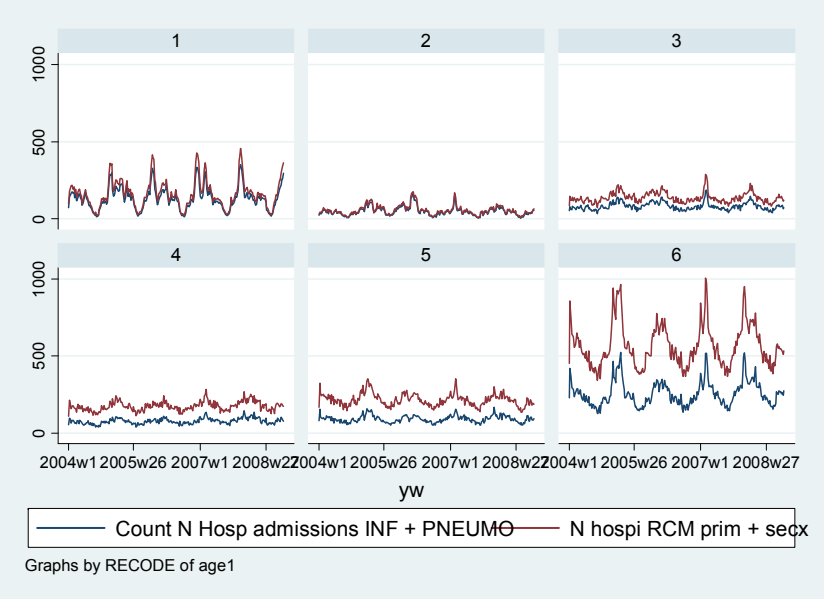
Table 3 – Distribution of parameters by age group (week 1 2004 - week 48 2008)

Count per age group (years)	<5	5-14	15-49	50-64	65-74	75+
Hospit P+I, first diagnosis	35 634	12 737	21 309	20 388	24 002	67 169
Hospit P+I, any diagnosis	43 496	14 741	35 882	45 550	56 093	147 425
% hospit with 1st diagnosis	82%	86%	59%	45%	43%	46%
Deaths P+I, first diagnosis*	16	10	254	845	1925	17489
Deaths P+I, any diagnosis*	48	45	1175	4094	7604	42 176
% deaths with first diagnosis	33%	22%	22%	21%	25%	41%
Influenza SL	3659	776	289	126	79	103
Influenza GP	1113	3958	8457	2980	685	522
Pneumo SL	1331	308	983	931	793	1510
Pneumo RL	1546	324	1409	1356	1148	2185
Haemophilus influenzae	51	7	45	52	49	96
Adenovirus	3838	244	125	31	20	13
RSV	28 245	308	193	78	48	109
Mycoplasma pneumoniae	4573	6515	4592	683	293	294
Parainfluenza	2329	109	59	35	26	20

SL: sentinel laboratories; RL: reference laboratory; GP: general practitioners.
** From week 1 2004 to week 52 2008 because data are complete.*

Table 3 indicates that the number of "non-primary" admission diagnoses is marginal in younger age groups, but increase with age: it represents 54% of all diagnoses in the 75+.

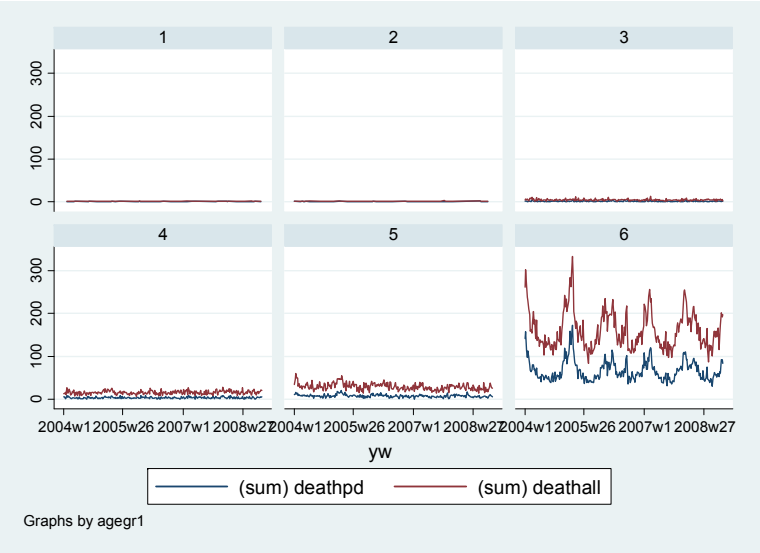
Figure 3 – Hospitalizations due to P+I by week, by age groups (main diagnosis and any diagnosis)



1: <5 years; 2: 5-14 years; 3: 15-49 years; 4: 50-64 years; 5: 65-74 years; 6: ≥75 years.

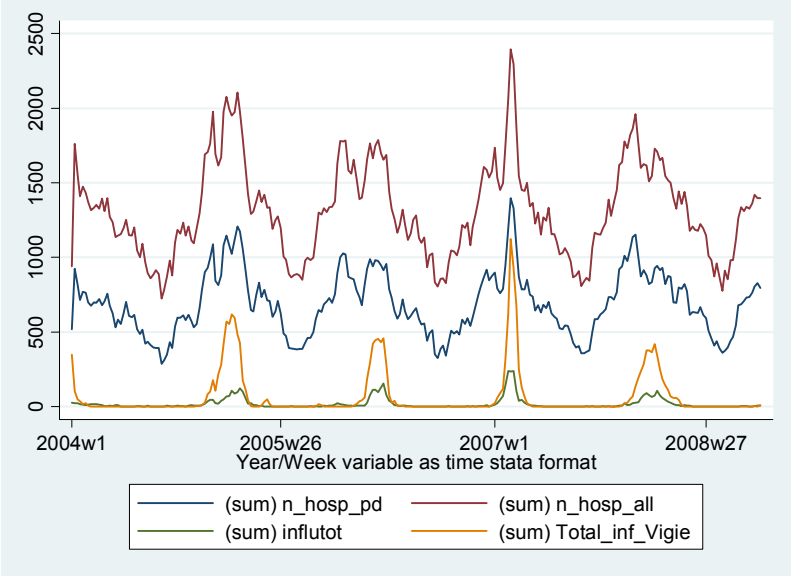
Figure 4 indicate that most P+I deaths occur in the elderly, as expected. The majority of these deaths have P+I coded as non-principal cause of death, and this proportion is lower in the elderly.

Figure 4 – Deaths due to P+I by week, by age groups (main diagnosis and any diagnosis)



1: <5 years; 2: 5-14 years; 3: 15-49 years; 4: 50-64 years; 5: 65-74 years; 6: ≥75 years.

Figure 5 – Hospitalizations and influenza (sentinel labs and GP) by week, all age groups



Influtot is influenza from the sentinel labs. Total_Inf_Vigie is influenza from the GP network.

The seasonal peaks in P+I admissions (Figure 3) and deaths (Figure 4) are concomitant with the peak in influenza detection. Weekly counts of deaths (as first cause of death) are also following a seasonal distribution close to the one of pneumococci (green line in Figure 6).

There is no indication of independent changes in influenza testing during the study period, which could influence the SL influenza data. Comparison between SL and GP data by season show parallel fluctuations in the number of cases by season (Figure 5).

Figure 6 – Deaths and influenza weekly counts, all age groups

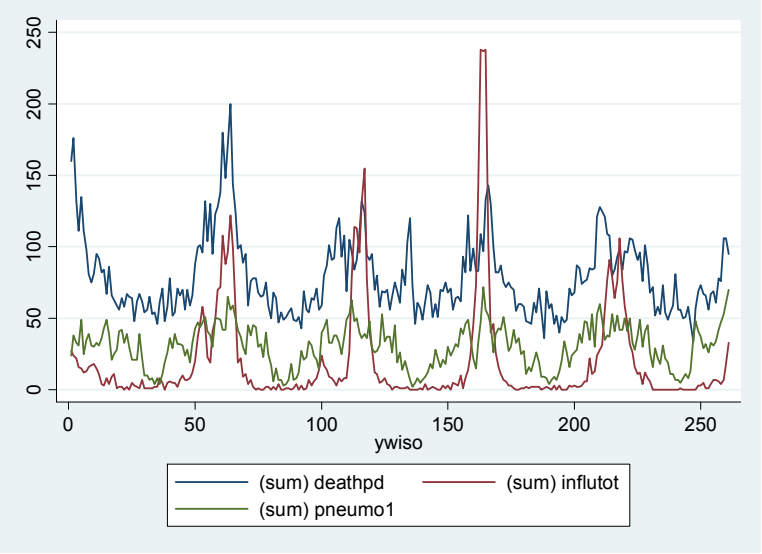
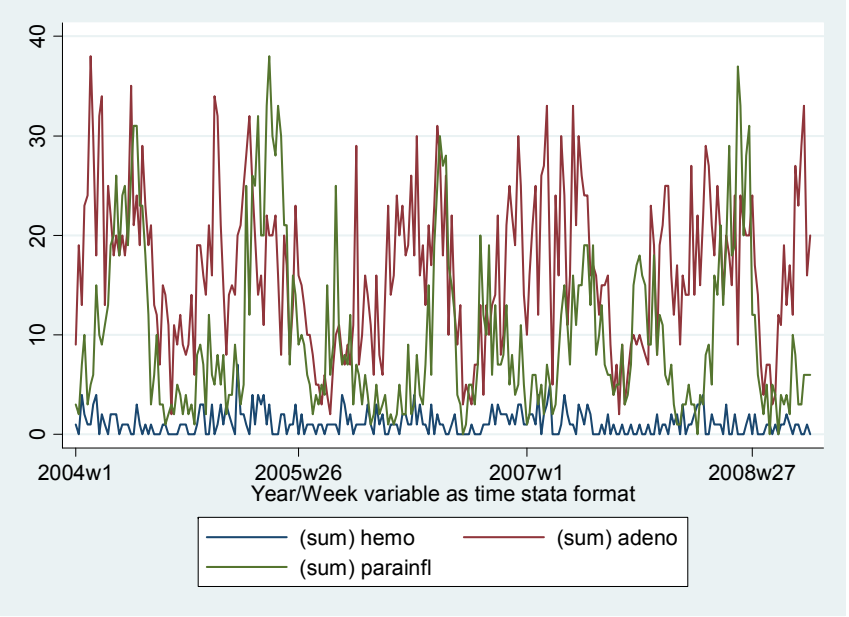
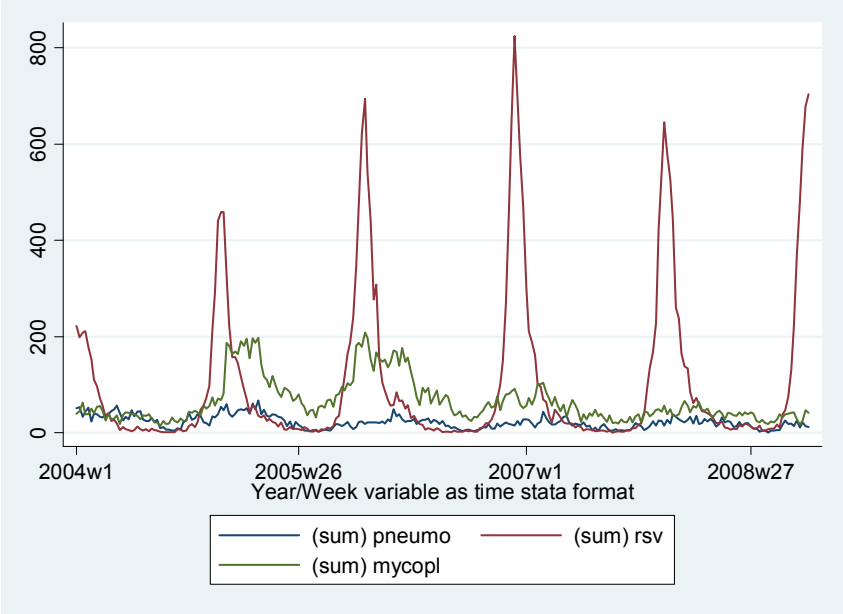


Figure 7 – Adenovirus, *Haemophilus influenzae* and parainfluenza by week, all age groups



Adenovirus, *Haemophilus influenzae* and parainfluenza show small counts of cases by week, but adenovirus and parainfluenza show clear seasonal patterns.

Figure 8 – Mycoplasma, pneumococcus and RSV by week, all age groups



Several dependent variables show a significant correlation pairwise, but this varies by age group (data not shown). No colinearity is suggested as the variance inflation factor (VIF) range 1-2.

5. SELECTION OF MODELS

The main models included only pathogens from the sentinel lab data (except for GP data on influenza), stratified in 6 age groups. The dataset initially used in the main model did not include the week 53 of 2004 as STATA cannot handle ISO weeks.

All alternatives models have also been conducted for all distributions and other outcomes but are not detailed here. The Gaussian regression for admissions is also presented in appendix, for comparison purpose. We present below the results of the different model distribution for admissions, but similar differences between distribution and models were found for deaths.

5.1. Selection of model distribution and link

We confirmed the presence of over-dispersion in all dependent variables: the mean is much smaller than the variance overall, for each age group. This was confirmed by the likelihood ratio test of the over-dispersion parameter alpha by negative binomial regression, and the goodness-of-fit test for Poisson regression: both tests showed that alpha was significantly different from zero ($p<0.05$) in all age groups and outcomes, except for deaths in the 15-64 years. This indicated that the Poisson model was mostly inappropriate. We thus used over-dispersed Poisson models to allow for over-dispersion, using a Pearson scale dispersion parameter (STATA glm command with scale(x2) option).

5.2. Comparing negative binomial and over-dispersed Poisson

Results from the negative binomial and the over dispersed Poisson models are presented below, using the main model for P+I admissions as main diagnosis. Figures comparing the fitted values to the observed values are shown in appendix for both distributions. Compared to negative binomial model, the over-dispersed Poisson allows more significant parameters in the model and influenza coefficients are lower in the <65 years. In the negative binomial distribution, influenza is no longer a significant parameter in the 65+ years (only pneumo remains), which is implausible on epidemiological grounds, while over-dispersed Poisson includes a high

number of pathogens in the best models. Similar results were found for deaths, though no model could be run for children (<15 years).

Table 4 – Summary of findings for negative binomial distribution, main model for P+I admissions (main diagnosis)

Age - years	Model parameters	Dev/df	AIC (using glm)	β influenza (95%CI)*	% influenza admissions
<5	flu SL, pneumo, myco, adeno, parainfl, RSV	0.063	11.54	0.86 (-0.11–1.82)	9%
5-14	flu SL, RSV, myco, pneumo	0.131	9.62	1.90 (0.07–3.73)	12%
15-49	Flu GP, mycopl	0.036	10.78	0.21 (-0.02–0.43)	8%
50-64	Flu GP	0.042	10.72	0.42 (-0.11–0.95)	6%
65-74	Pneumo	0.038	11.04	None (0)	0%
75+	Pneumo	0.043	13.07	None (0)	0%

** Not standardized.*

Table 5 – Summary of findings for over-dispersed Poisson, main model for P+I admissions (main diagnosis)

Age - years	Model parameters	Dev/df	AIC (using glm)	β influenza (95%CI)*	% influenza admissions
<5	Flu SL, pneumo, myco, adeno, parainfl, RSV	6.055	12.5	0.77 (0.59–0.95)	8%
5-14	Flu SL, RSV, myco, pneumo, adeno	4.362	9.86	1.77 (0.38–2.17)	11%
15-49	Flu GP, mycopl, hemo, pneumo, RSV	2.654	8.84	0.18 (0.14–0.21)	7%
50-64	Flu GP, mycopl, pneumo	2.897	9.05	0.30 (0.22–0.39)	4%
65-74	Flu GP, RSV, pneumo mycopl	2.965	9.28	1.05 (0.72–1.38)	3%
75+	Flu GP, pneumo, mycopl, hemo, RSV, parainfluenza	8.047	15.2	4.30 (3.07–5.53)	3%

** Not standardized.*

A log link has been applied to P+I admissions in two age groups only, for comparison purpose (one with poor and one with good fit).

Table 6 – Results from the two distributions with log link, for 2 age groups, P+I admissions (main diagnosis)

Age - years	Distribution	Model parameters	Dev/df	AIC
<5	Neg binomial	Flu SL, pneumo, myco, adeno, parainfl, RSV	0.122	11.60
65-74	Neg binomial	Flu GP	0.039	11.04
<5	Ov Poisson	Flu SL, pneumo, myco, adeno, parainfl, RSV	10.12	16.45
65-74	Ov Poisson	Flu GP, pneumo, rsv, mycopl	3.45	9.75

In these two age groups, the same pathogens are kept in the final model using an identity or a log link, but a much higher deviance/df and higher AIC overall were found with a log link, indicating an inferior model. The log link models also suggest an inferior fit (graphically, see appendix) in non-influenza periods.

Table 7 – Results from the two distributions in three large age groups, P+I admissions (main diagnosis)

Age - years	Distribution	Model parameters	Dev/ df	AIC	β influenza
<15	Neg binomial	Flu SL, pneumo, myco, adeno, parainfl, RSV	0.101	10.56	1.19
15-64	Neg binomial	Flu GP	0.045	10.75	0.287
65+	Neg binomial	Pneumo	0.199	12.21	None (0)
<15	Ov Poisson	Flu SL, pneumo, myco, adeno, parainfl, RSV (all except hemo)	5.490	11.49	0.92
15-64	Ov Poisson	Flu SL, pneumo, myco, adeno, parainfl, RSV (all)	3.027	9.19	0.20
65+	Ov Poisson	Flu SL, pneumo, RSV, hemo*	30.23	36.79	10.12

Using larger age groups does not seem to improve the power, as expected, considering the Wald tests and the number of parameters kept in the model: around the same parameters are kept compared to the smaller age groups. The goodness of fit seems similar in negative binomial but is difficult to compare across age strata. Larger age groups in over-dispersed Poisson show a very poor fit in the 65+ (dev/df>30). In addition, these models make less sense on a biological point of view, as for instance we know that the contribution of RSV is predominant in children <2 years but minimal in children 5-14 years.

Over-dispersed Poisson models perform relatively well in most age groups between 5 and 74 years, as the dev/df range 2.7-4.4, but a lower goodness of fit is found in the <5 and ≥75 years (dev/df >6). The models using negative binomial distribution include less significant parameters than over-dispersed Poisson models, especially in the >50 years. The dev/df are however not comparable across these distributions as the negative binomial in STATA gives a different value scale for dev/df.

The AIC suggests that both models perform similarly, with AIC higher or lower in different age groups according to each distribution. However, the AIC for over-dispersed Poisson is also difficult to compare as it does not take the over-dispersion into account.

In none of the distributions, the models stratified by large age groups perform better than by smaller age group, and this stratification makes less sense on biological grounds. Using a log link produced inferior models than with identity link based on AIC, a multiplicative model is less plausible and interpretation of coefficients is easier with identity link.

We thus opted for an over-dispersed Poisson model with identity link as distribution model, as this model shows a reasonable fit (for a simple model) across age groups and is able to incorporate more parameters.

5.3. Selection of final models

The process for selecting the final models is described below in detail for P+I admissions as main diagnosis. A shorter process has been conducted for other outcomes and summarized in tables of findings. Though final models were separately selected for each outcome as described below, some findings are common.

For all outcomes, a higher fit (based on deviance/df) is generally found in models including one influenza parameter by season, interactions between pathogens, holidays and break returns, and inclusion of a population terms in some age groups.

For admissions, as not all relevant parameters can be fitted into one final model (i.e. interactions of each pathogen with four influenza parameters), we present and compare the results of two models:

- Model 1 is including influenza by season and breaks;
- Model 2 is including interactions between pathogens and breaks.

All final models were based on pneumococcal counts from the reference Laboratory, received in a later phase, as they yield better fit and less autocorrelation than the model using pneumococcal data from the sentinel laboratories. All final models included week 53 of 2004, as a tool to integrate the week 53 2004 in STATA has been found in a later phase.

The pathogens included in each model vary across age groups, which fits with previous knowledge. In children <5 years, a large coefficient for RSV was found in admissions, as expected. In the elderly, a large coefficient for pneumococcus was found for both outcomes as expected. Influenza dataset that produce the best model also varies across age groups: sentinel laboratory (SL) data fits better in <15 years as it has more cases in pediatric ages; GP data fits better with other age groups. However, the largest part of all models is the intercept (constant). A contribution from other pathogens that is constant across seasons makes little sense on a biological point of view as other pathogens causing pneumonia (e.g. chlamydia and rhinovirus) are also seasonal. However, models with no constant or models using a pre-defined seasonal term (sin and cosin function) were not satisfactory.

6. ANALYSIS OF P+I ADMISSIONS

Alternative models are all based on the selected distribution (over-dispersed Poisson and identity link) and the main model, i.e. pathogens from best model, stratified by 6 age groups. Results of regression are described below for the two outcomes of admissions, P+I as principal diagnosis and P+I as any diagnosis.

6.1. P+I admissions as principal diagnosis

6.1.1. Selection of models

Several variations of the main model have been tested to select the final models.

6.1.1.1. Different parameters for influenza

The virulence and severity of influenza disease vary each season with the type of circulating virus. This has been confirmed by the finding that influenza coefficients vary by years and age groups in this dataset. We have tested two options to take this variability into account in the model:

- Include separate influenza parameters by influenza season.
- Include separate parameters for influenza A and influenza B viruses.

As we also found that the calendar year does not capture systematically one influenza season (e.g. 2004 captured no season), we have not considered the calendar year for defining influenza parameters but the influenza seasons (from week 40 to week 39 of the following year).

Influenza parameterized by season

Compared to the model with only one parameter for influenza, these models give a better goodness of fit (lower dev/df) in each age group, as shown in Table 8 for P+I admissions as principal diagnosis.

Table 8 – Model results when influenza is parameterized by season, P+I admissions (principal diagnosis)

Age	Dev/df	Range beta flu per season
<5 years	5.801	0.18–0.90
5-14 years	3.913	0.22–3.09
15-49 years	2.350	0.06–0.26
50-64 years	2.737	0.17–0.56
65-74 years	2.620	0.78–4.16
75 years +	7.646	2.80–6.30

Model with influenza A and B

Table 9 – Model characteristics with influenza A and B parameterized separately, P+I admissions (principal diagnosis)

Age	Dev/df	Beta flu A and B
<5 years	6.031	A: 0.865 – B: 0.313
5-14 years	4.378	A: 1.852 – B: 1.660

A/B data are currently only available for SL data and have thus been tested only in the <15 years (as GP influenza is a better parameter in the 15+), Table 9. These models perform similarly in terms of goodness of fit compared to the main model, but are clearly inferior to the models above with one influenza parameter by season.

6.1.1.2. Interactions between pathogens

The interactions considered were those with biological rationale, i.e. common co-infections or bacterial disease that are known to frequently complicate viral infections. We thus tested the following interactions:

- influenza*pneumo (frequent surinfection) or ip
- influenza*RSV (frequent co-infection) or ir
- influenza*mycoplasma or im

- mycoplasma+pneumo or mp
- rsv*pneumo or rp

We tested the main model for each age group with all possible interactions. Models with significant interactions performed systematically better than the main model, and allowed the influenza parameter to fit much better with the concomitant seasonal peak of admissions. Interestingly, coefficient of interaction terms are negative in all age groups <75 years, which suggest a competition between pathogens. Consequently, all influenza coefficients were larger in a model with interaction terms.

Table 10 – Main models with interactions between pathogens, P+I admissions (principal diagnosis)

Age - years	Interactions with p<0.10	Parameters for best AIC	Dev/df
<5	ip, ir, im*	Flu SL, pneumo, rsv, mycopl, adeno, parainfl, ip, ir, im, rp, mp	5.62
5-14	im, mp*	Pneumo, rsv, mycopl, flu LV, parainfl, ir, im, mp, rp	4.15
15-49	rp, mp*	Flu GP, pneumo, rsv, mycopl, hemo, ip, rp, mp	2.46
50-64	ip*	Pneumo, rsv, mycopl, flu GP, mp, ip, ir, im	2.75
65-74	ip*	Pneumo, rsv, mycopl, flu GP, ip, im, rp	2.80
75+	im	Pneumo, flu GP, parainfl, hemo, rp, im, ip, mp	7.46

6.1.1.3. Seasonality

A seasonal term, defined as to represent the seasonality of admissions that would not be explained by the current pathogens, is often added to similar models to control for the confounding effect of seasonal variation. The term is classically constructed based on cosinus and sinus functions: Cos(2π*week/52) and sin(2π*week/52) are included in the regression.[1, 15]

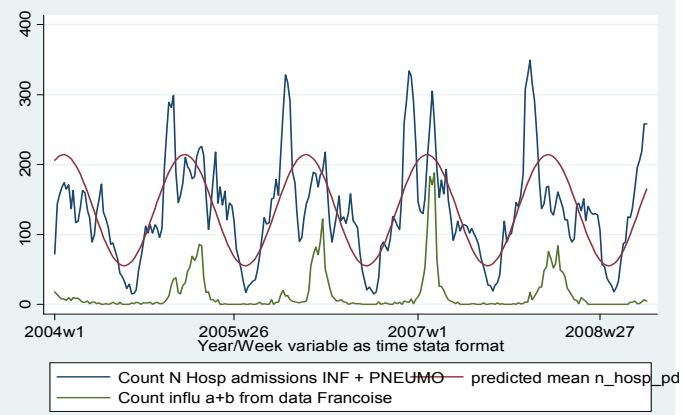
In models including only seasonal terms as independent variables (thus without pathogens), both cosinus and sinus terms are significant in all age groups. The main models with seasonal terms perform better than the main model in all age groups. However, the beta for influenza is systematically (and often substantially) lower because the seasonal term explains part of the flu seasonal variability: the seasonal peaks are overlapping peaks, as seen in Figure 9.

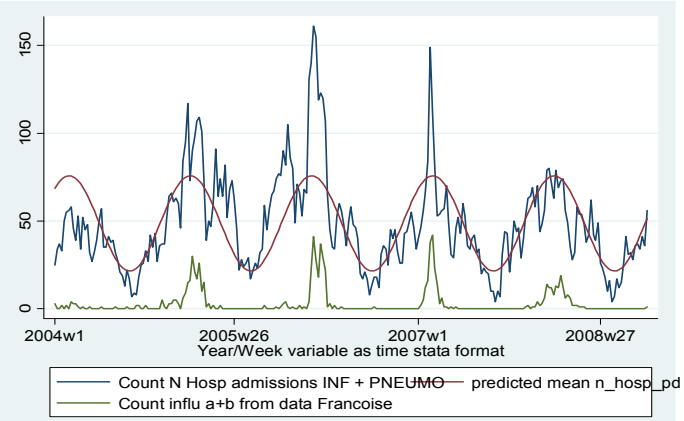
Table 11 – Results of the main model with inclusion of a seasonal term, P+I admissions (principal diagnosis)

Age	Dev/df	Beta flu
<5 years	5.60	0.567
5-14 years*	3.55	1.200
15-49 years	2.36	0.15
50-64 years	2.27	0.18
65-74 years*	1.71	0.65
75 years +	5.21	3.27

** Models keep only pneumo and influenza, while the main model has also RSV and mycoplasma.*

Figure 9 – Model with seasonal term only compared to SL influenza parameter, P+I admissions (principal diagnosis).





Left: <5 years; right: 5-14 years.

Adding a seasonal term that partly overlap the influenza season is thus preventing to determine the true value of influenza attributable admissions, and has thus not been selected. Other methods to parameterize the seasonal variability of admissions that is not related to the model parameters have been explored:

- Schanzer has used a proxy for "other ILI" (ILI laboratory negative counts or ILI -) in its model of influenza deaths after correcting for false negative results, to represent this seasonal variability. We explored the ILI- counts from the GP dataset, but these show a substantial remaining pattern of influenza, in spite of the high sensitivity of the laboratory tests used (PCR), see appendix.
- Other studies have included other parameters that may potentially influence P+I admissions to address this issue (e.g. temperature, humidity, pollution).[14] However, other studies chose to exclude environmental variables as previous analyses have suggested a limited role of these factors and controversial results.[2] For instance, some studies found an association between pneumonia admissions and these factors, while others not.

6.1.1.4. Holidays and school returns

Based on other similar studies, a parameter for the most influential holidays has been tested (Christmas and summer), as well as two separate parameters for school returns (after summer and after Christmas). Indeed, specific viral infections circulate more after returns from these holidays.

When including these variables as sole independent variables, the September return was significant ($p<0.05$) in all age groups, while the January return was significant in the 65+ only (and around 0.10 in 50-64 years). Holidays was a significant parameter in <15 years only. When integrating these 3 variables in the main model, September return was <0.10 in the <15 and 50-74 years; January return was <0.10 in <5 and ≥65 years; holiday was <0.10 in the <50 years. The inclusion of returns and holidays in the main models resulted in better models in all age groups compared to the main model, but the differences in influenza coefficients are moderate compared to the main model.

Table 12 – Results of the main model with holidays and school returns, P+I admissions (principal diagnosis)*

Age - years	Parameters in best model (in addition to pathogens)	Dev/df	Beta flu
<5	Sept and Jan returns, holidays	5.62	0.79
5-14	Sept return and holiday	3.79	1.73
15-49	Holiday	2.63	0.18
50-64	Sept return	2.85	0.30
65-74	Sept and Jan returns	2.88	1.03
75+	Sept and Jan returns	7.63	4.48

** Models with parameters <0.20 for significance; the best AIC has not been searched (yet).*

Compared to the model with only one parameter for influenza, these models give a better goodness of fit (lower dev/df) in each age group, as shown in Table 12 for P+I admissions as principal diagnosis.

6.1.1.5. Population term

As the residual distribution showed a trend in a number of models and age groups, we tested the addition of a population term (see later). This parameter was significant in a number of age groups, mainly adults, in which it improved the model.

6.1.1.6. Intercept

Though it seems unlikely that we can explain 100% of admission seasonality with the included pathogens, the use of a (high) constant term in the best model of some age groups results in models that do not fit well in non-influenza seasons. This may again underestimate the contribution of influenza in admissions. One EU study modeling lower respiratory tract infections and admissions in the elderly has excluded the intercept from the models.[5] We thus also tested the main models without constant term as exploratory analysis.

In the 3 age groups <50 years, the models without constant were substantially inferior to those with constant according to the AIC (Table 13). In the 50+ age groups, the model without constant cannot be computed because all other independent variables become non-significant. This alternative has thus not been kept.

Table 13 – Results of main model, with and without intercept, P+I admissions (principal diagnosis)

Age - years	AIC without intercept	AIC in main model	Coefficient for influenza without intercept[1]	Coefficient influenza in main model
<5	13.00	12.5	0.71	0.77
5-14	12.21	9.85	1.40	1.78
15-49	17.90	8.85	0.089	0.18
50-64	No model	9.04	NA	0.30
65-74	No model	9.27	NA	1.03
75+	No model	15.2	NA	4.30

1: Coefficient not standardized, but the same influenza datasets are used by age group.

6.1.1.7. Time lags

Several similar studies have tested whether time lags for pathogens counts and outcomes would improve the model, and some found significant and consistent time lags.[3, 15] Different time lags for pathogen counts and admissions have been tested. As SL data are mostly from hospitalized cases, a time lag has in principle no justification; for GP influenza data (used in 15+), the IPH surveillance has found that GP cases are admitted within 3 days in average. However, this delay may be more important for e.g. pneumonia admissions in the elderly. We thus constructed cross-correlograms between P+I admissions and the explanatory parameters for the 65+. The majority of cross-correlograms showed the highest correlation at time lag 0. We also used the STATA's VARSOC procedure to determine the appropriate lag length, and the indicated time lags were included in the models. However, the (final) models including these time lags were never better in terms of fit than the model without lag, and the AIC provided by the VARSOC command for the "best" time lag was very close to the AIC of other lags. Several authors also reported that the inclusion of time lags did not enhance the explanatory of the models.[3, 15]

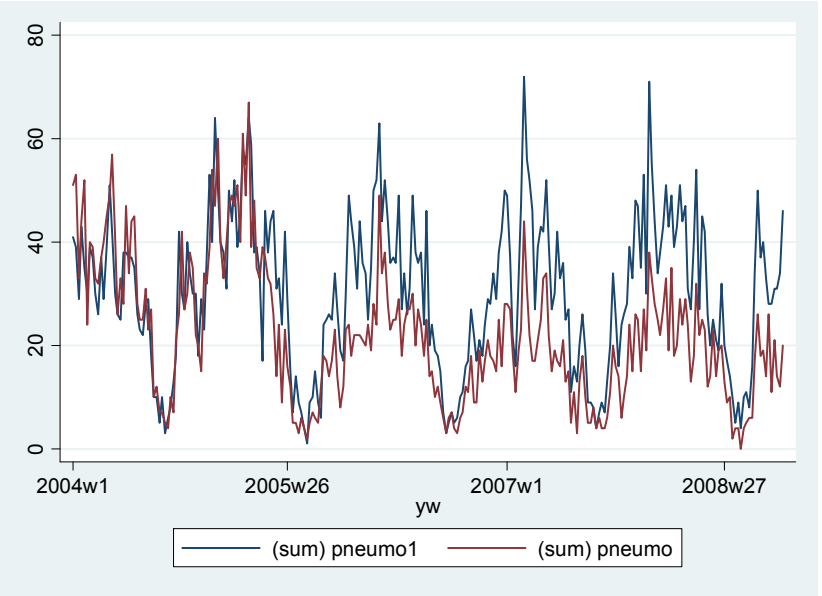
6.1.1.8. Pneumococcal data from the Reference Laboratory (RL)

We requested this additional dataset in a later phase of the study, after residual auto-correlation was found, and only included them in the final models (1 and 2). Both pneumococcal datasets show similar seasonal patterns (Figure 10), but RL counts are slightly higher as this dataset has a higher coverage (estimated at 80% in the <5 years).[16] All models using pneumo data from the Reference laboratory yield a much higher fit than models using data from the sentinel laboratories as the dev/df decreases substantially in all age groups. Residual analysis also shows a lower level of auto-correlation and a more random distribution compared to models using SL data (see later). An example of results based on model 1 (influenza by season and breaks/returns) is presented below.

Table 14 – Goodness of fit (dev/df) by pneumococcal datasets in model 1, P+I admissions (principal diagnosis)

Age	Pneumo data from sentinel labs	Pneumo data from the reference lab
<5 years	5.68	3.57
5-14 years	3.74	2.69
15-49 years	2.13	1.71
50-64 years	2.67	2.00
65-74 years	3.02	2.36
75 years +	9.14	7.76

Figure 10 – Pneumoccal counts by week, sentinel and reference laboratory



Pneumo is SL dataset; pneumo1 is RL dataset.

6.1.1.9. Residual analysis and selection of final models

As not all parameters can be fitted in one final model, we selected these two models:

- Model 1 is including influenza by season and breaks;
- Model 2 is including interactions between pathogens and breaks.

For both final models, auto-correlograms of residuals show some level of positive auto-correlation for the first 1-2 time lags as these are outside the 95%CI for most models. However, these were around 0.2 of auto-correlation (with the exception of the <5 years). The Durbin Watson statistic test confirmed the auto-correlation (Table 15). However, the levels of autocorrelation of most models are considered as minor serial correlation, i.e. lag-1 residual autocorrelation in the range 0.2 to 0.4, or a Durbin-Watson statistic between 1.2 and 1.6 (http://www.duke.edu/~rnau/testing.htm). White noise Portemanteau test is also significant for all models and age groups, except for model 1 for the 5-14 years. Model 1 and 2 indicate a lower degree of auto-correlation than the main model, and model 1 indicate a generally lower auto-correlation than model 2.

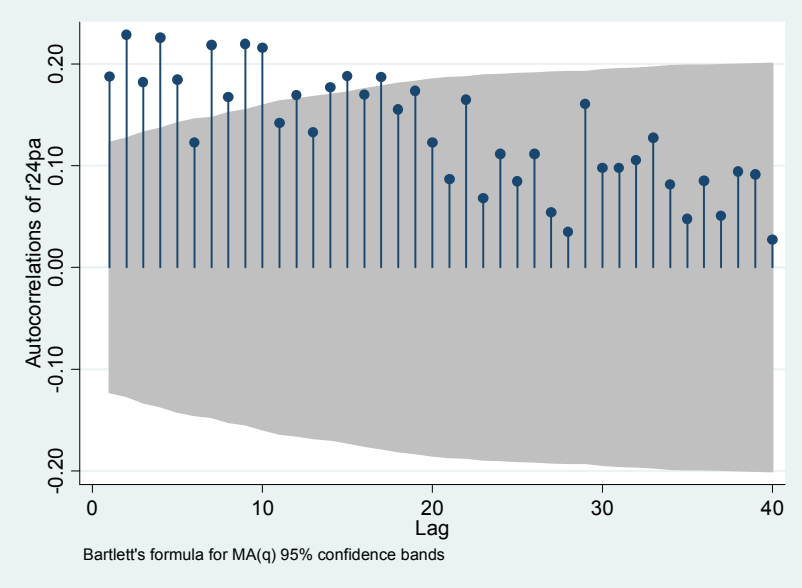
Table 15 – Durbin Watson statistics (value of d), preliminary models 1 and 2*, P+I admission as principal diagnosis

Age groups	Final model 1	Final model 2
<5 years	1.44	1.14
5-14 years	1.35	1.05
15-49 years	1.56	1.35
50-64 years	1.24	1.11
65-74 years	1.35	1.34
75+	1.25	1.30

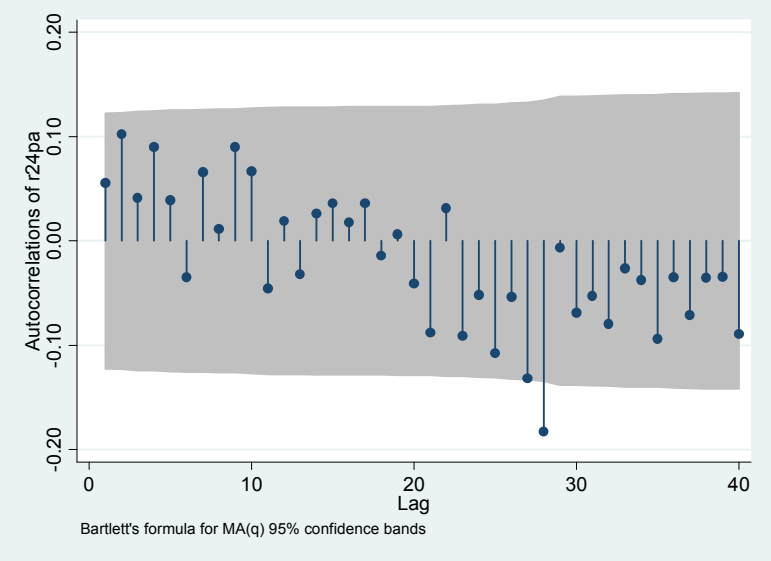
** Models with SL pneumo data and without population term.*
Note: in general, values should be the closest to 2. D<1.0 indicate strong auto-correlation.

The residual analysis also revealed clear time trends in some age groups, especially in the 50-64 years, in which a population increase of 8% was observed over the study period. We found that the inclusion of a population term (mid-year population) in the multivariate models improved the goodness of fit and the distribution of residuals (based on Portmanteau white noise test) in the 50-64 years: in model 2, the Portmanteau test value dropped from values >200 to 50-62 after inclusion of a population term, the residual plot showed no trend and the deviance/df decreased. A mild improvement of fit was obtained in the 75+ using model 1. We thus included a population terms in the 50-64 years of both models, and in the 75+ for model 1, where models fit and autocorrelation substantially improved (see example in Figure 11). Likewise, replacing the pneumo SL dataset by the LR dataset has substantially decreased the level of autocorrelation in all age groups and both models.

Figure 11 – Influence of the inclusion of a population term on residual auto-correlation, P+I in 50-64 years as principal diagnosis



Without population term.



With population term.

Besides auto-correlation, cyclical patterns are also seen in the distribution of residuals over time (see examples in appendix). This was expected as the model showed difficulties to fit well the inter-season dip and the seasonal peaks in admissions. We thus have systematic peaks and dips in residuals in these periods.

Several techniques have been undertaken to attempt to reduce these two problems:[17]

1. Add omitted predictors to the model: We included missing parameters (i.e. parameters that were close to the p=0.20 threshold) that would be plausible predictors. This however did not improve the auto-correlation and lowered the goodness of fit.
2. Include time lags: The lag of the dependant variables decreased tremendously the goodness of fit. The tools for selecting lags provided diverging answers (see before) and the iterative manual search for best lag was more successful. Some lags improved the model fit, but hardly decreased the auto-correlation. In addition, the best fitting lags were less plausible, as they were mostly negative for SL pathogens, while the laboratory result should usually follow the admission date.
3. Add seasonal terms: We re-tested the inclusion of a seasonal term in the final models to attempt reduce the auto-correlation and neutralize the cyclical patterns of residuals. Models with seasonal terms usually improved the goodness of fit but the auto-correlation remained the same or worsened, except for the 50-64 years where it slightly improved.

6.1.2. Findings of final models

The pathogens included in the two final models are relatively similar and all biologically very plausible (Table 16). The final models 1 and 2 show a substantial improvement in the model fit compared to the main model, as the dev/df is much lower (e.g. from 6.1 to 3.6 for <5 years). Most models have a dev/df <3.14 (which is considered as acceptable as the rule of thumb), except for those ≥75+ and to a lower extent those <5 years. The pseudo-R square, calculated manually according to STATA help, is above 0.5 for all models and as high as 0.92 in children <5 years.[18] The residual distribution still showed some level of autocorrelation (Table 17), but improved compared to all previous models, and are of minor importance as autocorrelation did not exceed 0.35 (see appendix 2 and 2.2) and the Durbin-Watson tests were all above 1.3. The models in the elderly >75 years showed lower performance (dev/df above 5 and higher auto-correlation) but acceptable fit according to the R^2. This can be related to lower influenza counts from the current surveillance systems in those ages, due to lower rates of swabbing. Model 1 showed a slightly better goodness-of-fit and lower level of auto-correlation than model 2, but this was not true for all age groups.

Table 16 – Parameters of final model 1 and 2, P+I admissions as principal diagnosis

Age - years	Parameters of model 1	Parameters of model 2
<5	Flu SL (4 seasons), RSV, myco, pneumo, adeno, parainfl, holiday, returns	Flu SL, RSV, myco, pneumo, adeno, parainfl, holiday, RSV*pneumo, influenza*RSV, returns
5-14	Flu SL (4 seasons), RSV, myco, pneumo, hemo, parainfl, holiday, return Sept.	Flu SL, pneumo, RSV, parainfl, myco, pneumo, hemo, holiday, flu*myco, return Sept
15-49	Flu GP (4 seasons), myco, pneumo, holiday, population	Flu GP, myco, pneumo, flu*myco, holiday, return Jan.
50-64	Flu GP (4 seasons), pneumo, population, holiday, return Sept.	Flu GP, pneumo, RSV, hemo, myco*pneumo, RSV*pneumo, flu*pneumo, flu*myco, population, holiday, return Sept.
65-74	Flu GP (4 seasons), RSV, pneumo, holiday, holiday, return Sept.	Flu GP, pneumo, RSV, flu*pneumo, flu*RSV, holiday, returns, population.
75+	Flu GP (4 seasons), pneumo, hemo, mycopl, RSV, population, return Jan.	Flu GP, pneumo, RSV, hemo, myco*pneumo, holiday, return Jan.

Table 17 – Goodness-of-fit and auto-correlation of final models 1 and 2 for P+I admission as principal diagnosis

Age groups - years	Model 1			Model 2		
	Deviance / df	Pseudo-R^{2}**	d of DW*	Deviance / df	Pseudo-R^{2}**	d of DW*
<5	3.56	0.92	1.38	3.62	0.91	1.32
5-14	2.84	0.84	1.52	2.75	0.82	1.32
15-49	1.88	0.77	1.61	2.07	0.73	1.36
50-64	1.97	0.55	1.59	1.89	0.60	1.63
65-74	2.06	0.67	1.49	2.38	0.58	1.57
75+	5.73	0.77	1.31	6.02	0.74	1.32

** Durbin Watson statistics, value of d. In general, values should be the closest to 2; d<1.0 indicate strong auto-correlation.*

*** By calculating the correlation between the observed response and the predicted response and squaring it.*[18]

6.1.3. Predicted influenza attributable admissions

6.1.3.1. Prediction by final models

Table 18 presents the average number of predicted influenza admissions by season, with corresponding predicted admission rates and proportion of P+I admissions for both models. In average, around 2000 admissions (or 20/100 000) are predicted to be due to influenza and this represents around 6% of all P+I admissions as principal diagnosis. The highest rates are predicted in the children, followed by the 75+. These predictions are very similar across the 2 models in spite of the very different calculations: in model 1, influenza variables by season are multiplied by a different coefficient by season and summed, while in model 2 influenza is represented by a single parameter across all seasons, but interaction terms are added. This consistency confirms the robustness of the analysis. The total number of influenza admissions predicted over the study period by the main and the two final models in presented in appendix, but the periods differ as the model 1 contains a smaller number of weeks compared to model 2.

Table 18 – Predicted outcomes by epidemic season (average), for model 1 and 2, for P+I as principal diagnosis

Age groups - years	Number of admissions (average)		Admission rate per 100 000		% of P+I admissions (average)	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
<5	540	600	92.7	103.0	8%	8%
5-14	287	290	23.6	23.9	11%	11%
15-49	309	301	6.2	6.0	7%	7%
50-64	201	178	10.5	9.3	5%	4%
65-74	234	179	24.8	18.9	5%	4%
75+	568	554	66.2	64.5	4%	4%
Total	2140	2102	20.3	20.0	6%	6%

Table 19 and Table 20 describe the number of admissions predicted for each season by model, and confirm the high variability of influenza seasons in terms of numbers and severity: the number of admissions predicted in 2004-05 represents the double of those predicted in 2005-06. Higher numbers are predicted in moderate intensity seasons compared to low intensity seasons. Besides, influenza B viruses predominated in 2005-06 and are known to display a much lower attack rate and severity in the elderly, as reflected by the very low number of admissions in the 65+ and the low number of influenza-coded admissions (MCD data) in this group (Table 20). Model 1, as it included a separate parameter influenza for each season, shows a higher variability across season (Figure 12) and age groups, but model 2 also indicate this variability. However, model 1 predicted the higher number of admissions in school age children during the 2005-06 influenza B season, while model 2 did not.

Table 19 – Prediction of influenza attributable admissions from model 1 by season, P+I as principal diagnosis

Age	2004-05	2005-06	2006-07	2007-08
Season intensity	*medium*	*low*	*medium*	*low*
<5 years	626	537	742	255
5-14 years	178	394	223	352
15-49 years	147	201	386	503
50-64 years	277	134	174	217
65-74 years	388	104	127	319
75 years +	1238	121	637	277
Total	2855	1492	2290	1922

Table 20 – Prediction of influenza attributable admissions from model 2 by season, P+I as principal diagnosis

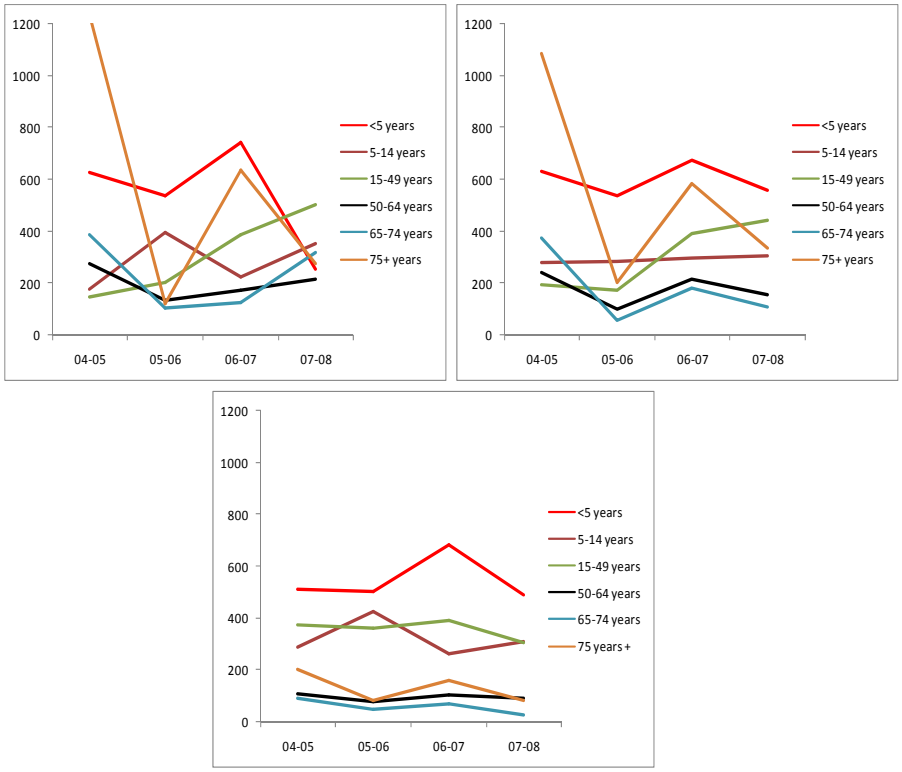
Age	2004-05	2005-06	2006-07	2007-08
Season intensity	*medium*	*low*	*medium*	*low*
<5 years	631	538	673	559
5-14 years	279	282	296	304
15-49 years	196	174	391	441
50-64 years	241	102	214	156
65-74 years	373	56	178	107
75 years +	1088	204	586	337
Total	2807	1357	2338	1906

The predicted numbers of influenza admissions are in average 40% higher than the number of admissions coded as influenza in the MCD dataset (Table 21 and Figure 12). But the difference between influenza predicted and coded data varies across age and season. Predicted numbers are in average similar to ICD coded admissions in children, 15% lower in adults <50 years but represent 2-4 times the number of ICD admissions in the 65+, depending on the model used. Seasonal fluctuations are observed in all datasets, with a lower season in 2005-06 and 2007-08 in the elderly, but these fluctuations are less marked in ICD coded admissions. In the 2005-06 season, the predicted number was similar to the number of ICD coded admissions, while it exceeded it by 80% in the season 2004-05.

Table 21 – Number of admissions coded as influenza in the MCD dataset, principal diagnosis

Age	2004-05	2005-06	2006-07	2007-08
<5 years	514	505	685	492
5-14 years	291	426	262	311
15-49 years	375	362	392	305
50-64 years	107	78	106	93
65-74 years	92	50	72	29
75 years +	204	83	160	85
Total	1583	1504	1677	1315

Figure 12 – Predicted influenza admissions (up, model 1 left and model 2 right) and admissions ICD coded as influenza (down), by age group and season



6.1.3.2. Sensitivity analysis

If only influenza and RSV parameters would be included in the model 1 based on P+I as principal diagnosis, as performed in similar studies,[3, 7] 46% more admissions would be predicted in an average season. The difference would be highest in the 75+. This model has however a poor fit in children and the 75+, with a high level of residual auto-correlation.

Table 22 – Predicted number of influenza admissions in model 1 containing only influenza and RSV pathogens, P+I admissions as principal diagnosis

Age - years	Number of admissions (average per season)			Additional admissions predicted by influenza and RSV models compared to full models
	Model 1, only influenza and RSV parameters	Model 1, only influenza parameters	Model 1, full (all parameters)	
<5	656	Very poor fit	540	21%
5-14	456	475	287	59%
15-49	462	469	309	49%
50-64	267	274	201	33%
65-74	308	310	234	31%
75+	967	Poor fit	568	70%
Total	3115	-	2140	46%

If only influenza is kept in the model, as performed in a few studies, goodness-of-fit is extremely poor in the <5 years of age in which RSV is a major contributor to P+I admissions, but is acceptable in the 15-74 years (dev/df <3.3). These models predict in average 54-73% more admissions by age group compared to the full model.

6.2. Analysis of P+I admissions as any diagnosis

6.2.1. Selection of models

The parameters and goodness-of-fit of the main model for any P+I admissions are shown in Table 23, and are relatively similar to those for P+I as main diagnosis.

Table 23 – Results from the main model for any P+I admissions

Age	Model parameters	Deviance/df
<5 years	All expect hemo	6.39
5-14 years	Flu LV, RSV, myco, pneumo, adeno	4.34
15-49 years	Flu GP, mycopl, hemo, pneumo, RSV	2.60
50-64 years	Flu GP, mycopl, pneumo	3.09
65-74 years	Flu GP, RSV, pneumo mycopl	3.54
75 years +	Flu GP, pneumo, hemo, mycopl, RSV, parainfluenza (but p>0.2)	10.75

Improvements to the main model were also brought by including RL pneumococcal counts, influenza parameterized by season, holidays and break, interaction terms and of a population term for the 50-64 years. The population term did improve the models in the adults.

6.2.2. Findings of final models

The selection of the best model 1 and 2 yielded similar parameters, fit, residual distribution than for P+I as main diagnosis (Table 24). The models show overall a good fit, with the exception of those ≥75 years in which the deviance/df is >7.

Table 24 – Parameters and goodness of fit of model 1 and 2, any P+I admission

Age - years	Parameters of model 1	Deviance / df	Parameters of model 2	Deviance / df
<5	flu SL (4 seasons), RSV, myco, pneumo, adeno, parainfl, holiday, returns	3.67	flu SL, RSV, myco, pneumo, adeno, parainfl, holiday, RSV*pneumo, myco*pneumo, returns	3.82
5-14	flu SL (4 seasons), RSV, myco, pneumo, parainfl, holiday, return Sept.	2.82	flu SL, pneumo, RSV, myco, holiday, flu*myco, return Sept	2.86
15-49	Flu GP (4 seasons), myco, pneumo, hemo, holiday, return Jan., population	1.72	Flu GP, myco, pneumo, hemo, flu*myco, holiday, return Jan.	2.04
50-64	Flu GP (4 seasons), pneumo, population, holiday	1.99	Flu GP, pneumo, myco*pneumo, flu*pneumo, flu*rsv, rsv*pneumo, population, holiday	1.92
65-74	Flu GP (4 seasons), RSV, pneumo, holiday, return Sept., population	2.47	Flu GP, pneumo, RSV*pneumo, holiday, returns, population	2.77
75+	Flu GP (4 seasons), pneumo, hemo, mycopl, RSV, return Jan, population	7.19	Flu GP, pneumo, myco, RSV, hemo, holiday, return Jan, population	7.97

Model 1 has again a better fit than model 2 across all age groups (except in the 50-64 years) and a lower level of autocorrelation. It should be noted that here also, models using pneumo data from the Reference laboratory achieved a much higher fit than models using SL pneumo data. Auto-correlograms of residuals also show positive auto-correlation for the first 1-2 time lags, at around 0.2 of auto-correlation. The Durbin Watson statistic test shows auto-correlation, except in the 15-49 in model 1 and the 50-64 years, in which it is in the inconclusive interval (Table 25). However, this is again considered as minor serial correlation. These models generally present a moderately lower level of auto-correlation than models using P+I as main diagnosis. White noise Portemanteau test is significant in all models except in model 2 in the 50-64 years. Cyclical patterns are also seen in the distribution of residuals over time.

Table 25 – Durbin Watson statistics (value of d), final models 1 and 2 for P+I admission as any diagnosis

Age	Any diagnosis	
	Final model 1	Final model 2
<5 years	1.47	1.34
5-14 years	1.48	1.25
15-49 years	1.66	1.49
50-64 years	1.82	1.80
65-74 years	1.46	1.54
75+	1.49	1.42

Note: in general, values should be the closest to 2. D<1.0 indicate strong auto-correlation.

6.2.3. *Predicted influenza attributable admissions*

The number of predicted influenza admissions is larger than the estimates from models with P+I as main diagnosis, i.e. around 3000 vs. around 2000 admissions for main diagnosis. The proportions of P+I admissions (any diagnosis) that are attributable to influenza (4-5%) are lower than those estimated from P+I as main diagnosis models (6%). The numbers predicted for each season (Table 27 and Table 28) show again the same variability across influenza seasons, and are lower in low intensity season. Again during the season 2005-06, a lower number of admissions is predicted in the 75+ and a higher number among school aged children in model 1. This is also reflected in a low number of the influenza-coded admissions (MCD data) in this group (Table 29).

Table 26 – Predicted outcomes by epidemic season (average), for model 1 and 2, for P+I admissions as any diagnosis

Age - years	Number of admissions (average)		Admission rate per 100 000		% of P+I admissions (average)	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
<5	661	690	113.4	118.4	8%	8%
5-14	348	362	28.7	29.8	11%	12%
15-49	462	429	9.2	8.6	6%	6%
50-64	356	316	18.7	16.6	4%	3%
65-74	386	305	40.8	32.2	3%	3%
75+	1043	1019	121.5	118.6	3%	3%
Total	3256	3120	31.0	29.7	5%	4%

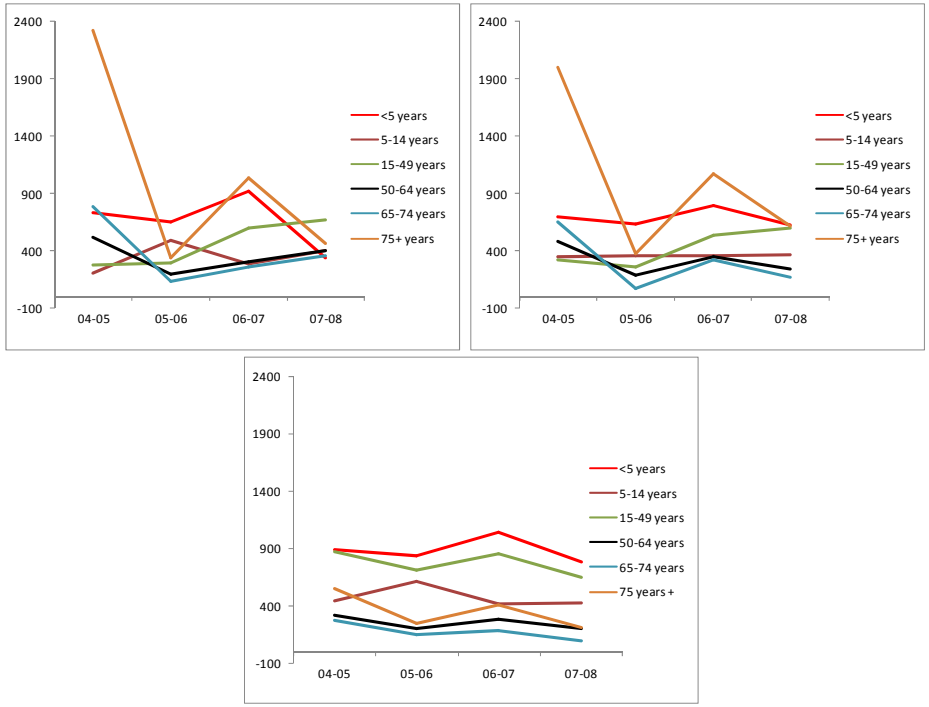
Table 27 – Prediction of influenza attributable admissions from model 1 by season, any P+I diagnosis

Age	2004-05	2005-06	2006-07	2007-08
Season intensity	*medium*	*low*	*medium*	*low*
<5 years	730	651	925	338
5-14 years	208	489	289	407
15-49 years	277	293	604	673
50-64 years	517	198	303	406
65-74 years	785	137	261	359
75 years +	2323	345	1036	469
Total	4840	2113	3418	2653

Table 28 – Prediction of influenza attributable admissions from model 2 by season, any P+I diagnosis

Age	2004-05	2005-06	2006-07	2007-08
Season intensity	*medium*	*low*	*medium*	*low*
<5 years	694	639	798	630
5-14 years	354	361	363	369
15-49 years	319	257	540	598
50-64 years	480	191	353	240
65-74 years	655	72	320	173
75 years +	2001	376	1078	621
Total	4504	1895	3451	2631

Figure 13 – Predicted influenza admissions (left) and admissions coded as influenza (right), any diagnosis, by age group and season



The predicted numbers of influenza admissions are only 12% higher than the number of admissions coded as influenza (any diagnosis) in the MCD dataset (Table 21 and Figure 12), but this proportion varies with age. Here, predicted numbers are around 25% lower than influenza ICD coded admissions in children, 40% lower in adults <50 years but represent 2-3 times the number of ICD admissions in the 65+. Seasonal fluctuations are less marked in ICD coded admissions. In the 2005-06 season, the predicted number is lower than the number of ICD coded admissions, while it exceeds it by 35-45% in the season 2004-05.

Table 29 – Number of admissions coded as influenza in the MCD dataset (any diagnosis)

Age	2004-05	2005-06	2006-07	2007-08
<5 years	891	834	1044	785
5-14 years	443	612	417	427
15-49 years	869	707	857	650
50-64 years	320	205	280	200
65-74 years	269	151	184	93
75 years +	549	246	410	206
Total	3341	2755	3192	2361

6.3. Analysis based on all respiratory and circulatory admissions

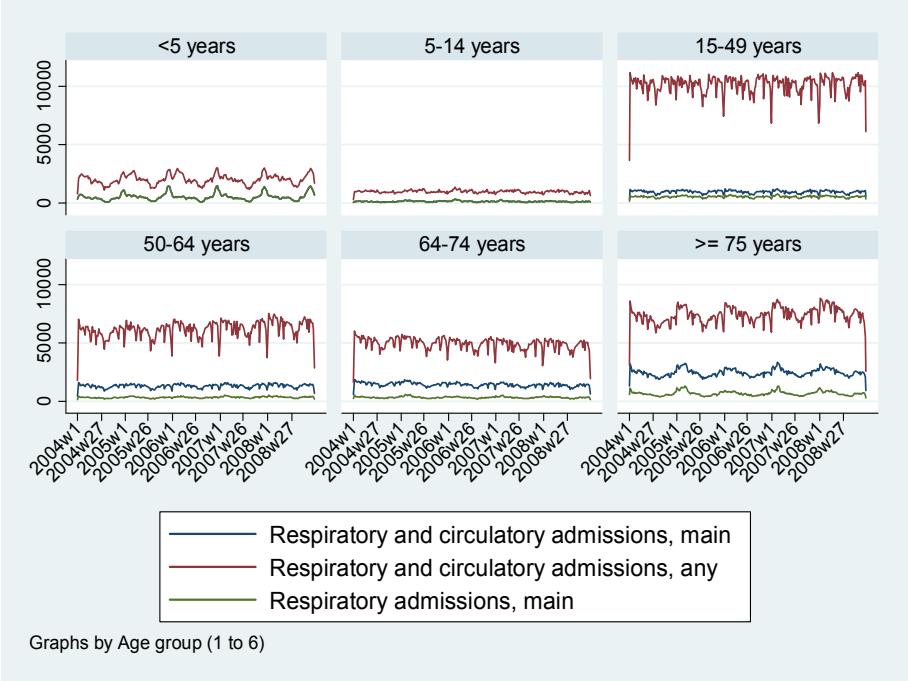
Influenza complications may lead to admissions for other outcomes than influenza/pneumonia only, in particular to admissions for other respiratory or circulatory conditions. This analysis is using the following outcomes: all respiratory admissions (ICD9-codes 460-519) and all circulatory admissions (ICD9-codes 390-459). We only considered the codes reported as principal diagnosis. The admissions for respiratory or circulatory conditions coded as any diagnosis (principal or associated) are also shown below.

6.3.1. Description of data

In average, around 360 000 admissions were coded as respiratory and/or circulatory conditions (R+C), principal diagnosis, each year in the MCD database. Admissions for respiratory conditions amounted to 136 000 by year, representing 38% of R+C admission in average but >95% in children (>95%) and 26% in adults >50 years. P+I admissions represented 27% of the respiratory admissions in average, and the highest proportion was found among the school age children (35% in the 5-14 years) and the oldest (38% in the ≥75 years).

These admissions also display some seasonal patterns (Figure 14). The weekly distribution of respiratory admissions shows overall similar seasonal patterns than the P+I admissions. The distribution of R+C admissions by week show different patterns, with recurrent dips around the summer and Christmas holidays, although this feature is less marked among the youngest and the oldest.

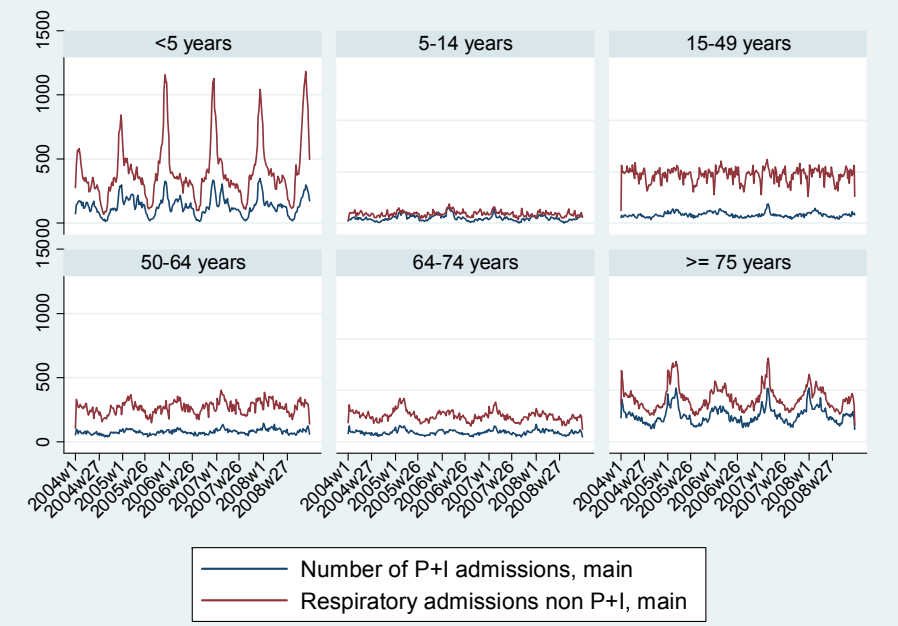
Figure 14 – Weekly distribution of respiratory and circulatory admissions by age group, principal diagnosis



Note: as the number of circulatory admissions in children is very small, the line "Respiratory and circulatory admissions, main" is superimposed by the line "Respiratory admissions, main".

The weekly distribution of respiratory admissions that are not coded as P+I also shows different seasonal patterns compared to P+I in those <50 years of age (Figure 15): large and regular peaks in the <5 years of age are concomitant with the RSV circulation; no specific seasonal patterns can be identified in the 5-49 years. Conversely, the seasonal patterns in those ≥50 are very similar to patterns observed for the P+I admissions.

Figure 15 – Weekly distribution of respiratory and P+I admissions by age group



6.3.2. Selection of models and model findings

Two outcomes were used, using only the principal diagnosis: respiratory admissions only and R+C admissions. The two models used for P+I admissions (model 1 with seasonal influenza and model 2 with interactions between pathogens) showed better fit and lower level of auto-correlation than the model using a single influenza parameter and we compared these two models to select the final model.

For both outcomes, model 2 was generally inferior to model 1 in terms of goodness-of-fit (except in the <5 years) and residual auto-correlation Table 30 and Table 55). In addition model 2 does not include influenza in two age groups, yields negative coefficients for influenza in 2 other age groups for R+C admissions and the intercept (constant) from the regression is negative in 3 to 4 age groups. These findings make interpretation of results complex in terms of predicted influenza admissions. Model 1 has thus been preferred.

When only respiratory admissions are used, the goodness-of-fit is substantially lower compared to models using P+I admissions as outcome. A particularly bad fit is observed in the 75+ (high deviance/degree of freedom). The distribution of residuals is grossly similar to models using P+I admission models. This suggests that the residuals of this outcome do not display a higher amount of non-random patterns. Indeed, non-P+I admissions do not display specific seasonal patterns in the ages 5-49 years (Figure 15), and the seasonal narrow peaks in the <5 years of age are generally well fitted by the RSV parameter.

Table 30 – Goodness-of-fit and residual distribution of models 1 and 2, respiratory admissions only (principal diagnosis)

Age group - years	Model 1			Model 2		
	Deviance / df	Durbin Watson	P white noise	Deviance / df	Durbin Watson	P white noise
<5	7.1	1.46	0.001	6.8*	1.44	0.015
5-14	3.8	1.53	0.061	3.9	1.35	0.005
15-49	4.6	1.91	0.000	5.4*	1.79	0.094
50-64	2.9	1.66	0.158	3.1	1.60	0.093
65-74	4.3	1.42	0.000	4.8	1.50	0.000
75+	11.0	1.31	0.000	11.6	1.32	0.000

** Influenza not included in the final model.*

Models on R+C admissions show a very poor goodness-of-fit (dev/df ranging 6.1-14.6, except in the 5-14 years of age), which was substantially inferior to the respective P+I or respiratory admission models (Table 31). The level of auto-correlation was minor and similar to other admission models (Table 30).

Table 31 – Goodness-of-fit and residual distribution of models 1 and 2, circulatory and respiratory admissions (principal diagnosis)

Age group - years	Model 1			Model 2		
	Deviance / df	Durbin Watson	P white noise	Deviance / df	Durbin Watson	P white noise
<5	14.6	1.47	0.004	6.6*	1.45	0.020
5-14	3.6	1.51	0.060	3.8	1.33	0.009
15-49	6.1	1.86	0.009	7.6*	1.69	0.284
50-64	6.8	1.85	0.000	9.0	1.71	0.003
65-74	6.6	1.70	0.001	9.3**	1.66	0.006
75+	9.1	1.64	0.000	11.9**	1.51	0.000

**: influenza not included in the final model*
***: influenza coefficients are negative*

6.3.3. Predicted influenza attributable admissions

The regression models on respiratory admissions estimated 4300 influenza admissions in an average season (Table 32). This number widely varies across seasons and age groups. The average estimated during medium/high intensity seasons represented twice the average number during low intensity seasons (5776 vs. 2850, respectively). Most predicted admissions were found in the elderly ≥75 years (34%) in average, but this proportion was lower in low intensity seasons (13%), in which the highest burden was among younger adults (37% in 15-64 years).

Table 32 – Prediction of influenza attributable admissions from model 1 by season, respiratory admissions

Age	2004-05	2005-06	2006-07	2007-08
Season intensity	*medium*	*low*	*medium*	*low*
<5 years	0	0	964	497
5-14 years	243	729	467	483
15-49 years	395	557	767	768
50-64 years	1021	550	619	620
65-74 years	1478	171	438	596
75 years +	3534	292	1626	438
Total	6671	2299	4880	3402

The regression models on R+C admissions estimated in average 4900 influenza admissions per season (Table 33) and the same ratio between high/medium and low seasons was observed (ratio 1.9, 6433 vs. 3450 admissions, respectively). The R+C models predicted around the same number of children admissions than the respiratory models; this could be expected as only few circulatory admissions were reported in children. Again, the highest number of influenza admissions were predicted in the oldest in average (45% in ≥75 years), but this was not observed in low seasons, in which 32% of admissions were found in younger adults (15-64 years).

Table 33 – Prediction of influenza attributable admissions from model 1 by season, respiratory and circulatory admissions

Age	2004-05	2005-06	2006-07	2007-08
Season intensity	*medium*	*low*	*medium*	*low*
<5 years	0	0	966	499
5-14 years	239	748	454	491
15-49 years	0	647	852	1029
50-64 years	921	959	803	1148
65-74 years	1823	0	482	0
75 years +	4504	540	1820	840
Total	7488	2893	5378	4007

These predictions represent around the double than those obtained by modeling the P+I admissions only. Surprisingly, these models predicted a lower number of influenza admissions in the <5 years compared to the P+I models (365 vs. 540, respectively). This is probably explained by the fact that influenza was not included in the regression models for 2 seasons (Table 32 and Table 33). Conversely, these models predicted around 3 times more admissions in adults ≥50 years compared to the P+I models.

Table 34 – Predicted admissions and admission rates in an average season for respiratory and respiratory + circulatory admissions as principal diagnosis

Age	Number of admissions (average)		Admission rate per 100 000	
	Respiratory	R+C	Respiratory	R+C
<5 years	365	366	62.6	62.8
5-14 years	480	483	39.6	39.8
15-49 years	622	632	12.4	12.6
50-64 years	702	958	36.9	50.3
65-74 years	671	576	70.9	61.0
75 years +	1472	1926	171.4	224.2
Total	4313	4941	41.0	47.0

R+C: respiratory and circulatory.

Other regression studies involving respiratory admissions have found a significant association with influenza,[1, 2] while most studies using circulatory admissions as outcome did not find significant association with influenza when adjusting for other covariates.[1, 3, 8]

As these models in our study showed an inferior goodness-of-fit, these predictions should be considered with caution. They should not be used for a base case but preferably for a maximum case in a sensitivity analysis.

7. ANALYSIS OF P+I DEATHS

7.1. Analysis of P+I deaths as principal cause

As previously mentioned, over-dispersion in the death datasets was only found in the 65+. However, the same distribution models, correcting for over-dispersion, were used in all age groups for consistency purpose.

In general, regression on P+I deaths as main cause was nearly not possible in children <15 years. The glm command from STATA could not perform the regression for any of the distribution models (Poisson, negative binomial and over-dispersed Poisson), likely due to the very small weekly counts of P+I deaths among children (3 P+I deaths in average). Poisson and negative binomial regression could be run using other commands (not glm) when using a log link, but only a constant was included in the model and the pseudo-R^2 was 0. The merge of the 2 age groups (children <15 years) did not allow for a glm regression neither; a simple negative binomial and Poisson regression could be run (using a log link) but only one independent variable was significant in the model, not influenza, and a very poor fit was found (pseudo-R^2=0.01).

7.1.1. Selection of models

We started with the same main model as used for P+I admissions in the four age groups above 15 years, and tested the performance of alternative models by applying similar changes. We thus tested the addition to the main model of parameters for school breaks and returns, separate influenza parameters by season, interactions between major pathogens, population parameter and lagged variables.

In general, the best models included less pathogen parameters than for admissions. Influenza was systematically significant in all age groups and models, except in the 50-64 years when interaction terms were added. The pathogens kept in the best models were mostly those expected based on clinical knowledge: influenza in all ages; pneumococci, RSV and mycoplasma in the 65+. When influenza was parameterized by season, several seasons were not significant, which can be due to the smaller counts of deaths by season; a model with influenza parameters by season performed only better in the 75+. The population parameter improved the main model in the 65-74 years only.

Table 35 – Goodness of fit of main model and its variations, P+I deaths as main cause

Age - years	Main model	+ holidays/ returns	+ influenza/ season ***	+ interactions pathogens	+ population
<5	NA	NA	NA	NA	NA
5-14	NA	NA	NA	NA	NA
15-49	1.26	1.26*	1.35	1.26	1.26
50-64	1.03	1.02	1.02	1.01**	1.03
65-74	1.34	1.34	1.34	1.31	1.29
75+	3.80	3.60	3.45	3.65	3.80

** Other parameters not significant (only main model is kept).*
*** The influenza parameter is replaced by 2 interaction terms.*
**** Model with influenza by season contains a lower number of observations.*

Table 36 – Goodness of fit and level of autocorrelation across main and lagged models, P+I deaths as main cause

Age - years	Main (unlagged) model		Lag+1 on main		Lag+2 on main	
	Dev/df	Durbin Watson d[a]	Dev/df	Durbin Watson d	Dev/df	Durbin Watson d
15-49	1.26	2.03	1.27	2.02	1.28	2.00
50-64	1.03	2.07	1.00	2.09	1.02	2.09
65-74	1.34	1.55	1.32	1.64	1.23	1.71
75+	3.80	1.01	3.21	1.23	3.08	1.30

Bold: best value across models for this age: dev/df as low as possible and close to 1; Durbin Watson d closest to 2.

[a] Durbin Watson test: values should be the closest to 2. d<1.0 or d close to 4 indicate strong auto-correlation.

Several time lags were also tested, as it is likely that P+I deaths would be delayed in time compared to positive tests for pathogens from the laboratories. Cross-correlograms and other STATA functions indicated that a time lag of +1-2 weeks for deaths had the highest correlation with at least influenza and pneumococcus parameters. Lagging the independent variables did not improve the models; however, lagging the dependent variables improved the model performance in the 65+ years: a lag +2 weeks showed the best goodness-of-fit and the lowest level of autocorrelation (Table 36). This lag is also biologically plausible.

Model performance varied across age groups:

- In the 15-64 years, the main model already showed a high goodness-of-fit (dev/df ≤1.26), no auto-correlation (Durbin Watson d around 2)a and independence of residuals (non-significant Portemanteau test for white noise), see Table 35 and Table 36. Alternative models only minimally changed these characteristics.
- In those aged 65-74 years, the main model showed a high goodness-of-fit but residual auto-correlation. Both characteristics improved substantially when a time lag +2 and a population term were included (Table 35).
- In the elderly aged 75+, the main model showed a moderate goodness of fit (dev/df 3.8), a high level of auto-correlation of residuals (Durbin Watson d 1.0)a and a lack of independence of residuals (significant Portmanteau test for white noise). The addition of returns and holidays, the replacement of influenza by separate parameters by season, interactions terms and time lags did clearly improve the goodness of fit and residual distribution, compared to the main model (Table 35).

As we could not add all valuable parameters to the main model, and the 75+ represents the group with highest mortality burden, we tested and compared several models (Table 37):

- Three models including interactions between pathogens, breaks/returns and different time lags;
- Three models including influenza by season, breaks/returns and different time lags.

Table 37 – Comparison of models for P+I deaths as main cause in the elderly 75+

Model parameters	Dev/df	DW test	p white noise
With interactions, no lag: Influenza, pneumo, RSV, hemo, parainfl,, myco*pneumo, flu*pneumo, rsv*pneumo, return Jan, holiday	3.46	1.21	0.000
With interactions, lag +1: Influenza, pneumo, RSV, hemo, parainfl, myco*pneumo, flu*pneumo, return Jan, holiday	3.01	1.38	0.000
With interactions, lag +2: Influenza, pneumo, RSV, hemo, parainfl, myco*pneumo, flu*pneumo, return Sept	3.04	1.29	0.000
With flu/season, no lag: Influenza by season (4), pneumo, RSV, mycopl, return Jan, holiday	3.40	1.24	0.000
With flu/season, lag +1: Influenza by season (4), pneumo, RSV, mycopl, parainfl, hemo, return Jan, holiday	2.84	1.54	0.016
With flu/season, lag +2: Influenza by season (4), pneumo, RSV, parainfl.	2.87	1.39	0.002

Bold: best value across models for this age: dev/df as low as possible and close to 1; Durbin Watson d closest to 2.

The best model was found with influenza parameterized by season and a lag +1 for deaths, in terms of both goodness of fit and residual distribution (Table 37). However, this model still showed a (minor) level of auto-correlation and no independence of residuals.

The final model selected thus differed by age:

1. In the 15-64 years, the main model (pathogens only) was kept as final model.
2. In the 65-74 years, the main model with +2 time lag and population.
3. In the elderly aged 75+, a model with influenza parameterized by season, a lag +1 for deaths, holiday breaks and returns.

7.1.2. Findings of final models

Influenza was included in the final models of all age groups.

Table 38 – Goodness of fit and residual distribution of final models, P+I as main cause of death

Age - years	Model parameters	Flu beta (95%CI)	Dev/df	DW	White noise
15-49	Influenza, parainfl	0.003 (0.000–0.005)	1.26	2.03	p=0.77
50-64	Influenza, RSV	0.015 (0.01–0.02)	1.03	2.07	P=0.58
65-74	Influenza, pneumo, RSV, mycoplasma, adeno, lag+2, pop	0.150 (0.089–0.211)	1.20	1.77	P=0.67
75+	4 flu seasons, pneumo, RSV, mycopl, parainfl, hemo, return Jan, holiday, lag +1	Range 0.74–2.5 by season	2.84	1.54	p=0.02

DW: Value of d in the Durbin Watson test for auto-correlation.

The goodness of fit was good for all models (dev/df <2.85). In the 65-74 years, the d value of the DW test was in the indecision zone, but the auto-correlogram showed no significant auto-correlation. The model substantially improved in the 75+; residual auto-correlation was still found in this group but to a much lower level than in the main model (Durbin Watson d increased from 1.01 to 1.54) and the auto-correlogram shows that only the first lag is concerned (see appendix).

7.1.3. Predicted influenza attributable deaths

The number of predicted influenza deaths in an average season (Table 39) shows that most predicted influenza deaths (83%) are found in the elderly 75+, as expected. The predicted influenza deaths represent in average 6% of all predicted P+I deaths. This proportion varies largely across seasons (range 2-12% by season) but is stable across age groups, except in the young adults 15-49 years where it is higher (11%).

The number of predicted influenza deaths show higher fluctuations across season compared to admissions (Table 28 and Figure 14). The 2004-05 season shows the highest number of predicted influenza deaths, as also shown for admissions, and very few deaths are predicted in the 2007-08 season.

Table 39 – Predicted influenza deaths in an influenza season (average), for P+I as main cause of death

Age	Number deaths	Mortality rate	% of P+I deaths
<5 years	0	0.0	NA
5-14 years	0	0.0	NA
15-49 years	5	0.1	11%
50-64 years	11	0.6	6%
65-74 years	24	2.5	6%
75 years +	204	23.7	6%
Total	244	2.3	6%

Table 40 – Prediction of influenza attributable deaths by season, for P+I as main cause of death

Age	2004-05	2005-06	2006-07	2007-08
Intensity	***medium***	***low***	***medium***	***low***
<5 years	0	0	0	0
5-14 years	0	0	0	0
15-49 years	6	4	6	6
50-64 years	14	6	13	11
65-74 years	51	6	25	13
75 years +	485	118	156	56
Total	556	132	200	86

The predicted numbers of influenza deaths are in average 3-fold the number of deaths coded as influenza in the death certificates (Table 40 and Figure 14). But the difference between the two datasets also varies substantially with age: predicted numbers are in average 5-9 times higher in adults <75 years but represent the double of ICD influenza deaths in the 75+. Large seasonal fluctuations are observed in both datasets, with a low mortality in the two low intensity seasons (2005-06 and 2007-08), but these fluctuations are not completely parallel (Figure 16).

Figure 16 – Predicted influenza deaths (left) and deaths coded as influenza (right), by age group and season, main cause

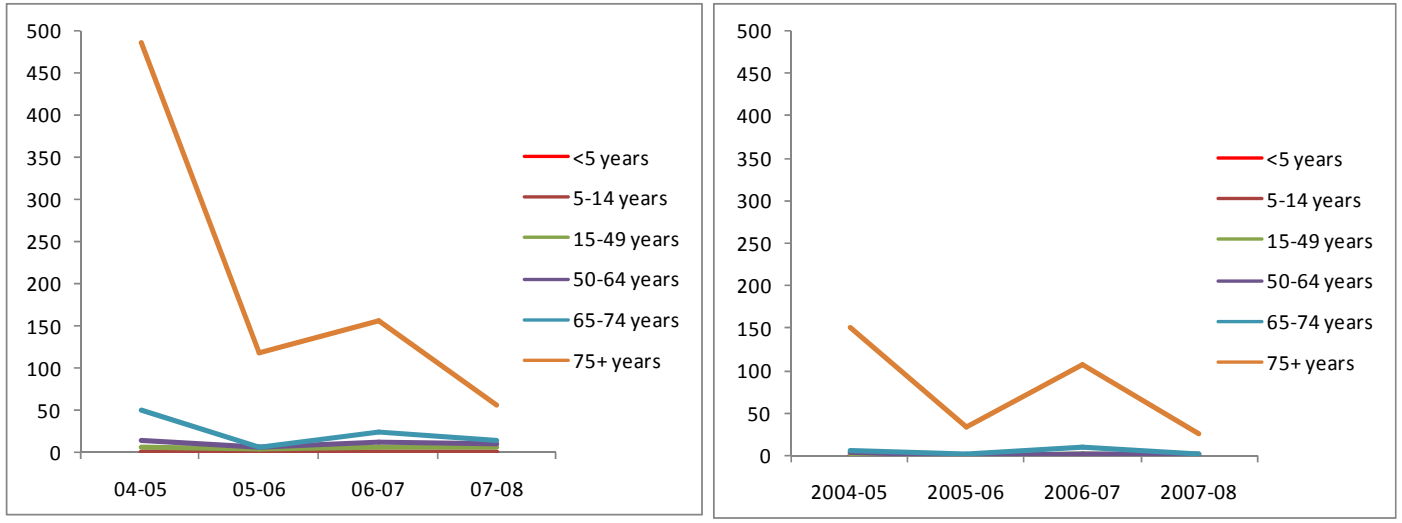


Table 41 – Number of deaths coded as influenza by the communities (main cause)

Age	2004-05	2005-06	2006-07	2007-08
Intensity	*medium*	*low*	*medium*	*low*
<5 years	0	0	0	0
5-14 years	0	0	0	0
15-49 years	2	0	1	0
50-64 years	3	0	1	1
65-74 years	6	2	9	4
75 years +	151	45	106	31
Total	162	47	117	36

7.1.3.1. Sensitivity analysis

If only influenza parameters would be included in the model, as has been performed in similar studies,[3, 7] the goodness of fit is much lower and auto-correlation very high, but the predicted number of influenza deaths is 40% higher (Table 42). This has been performed only in the 75+ years as they concentrate 83% of all predicted influenza deaths: 285 deaths would be predicted in an average season compared to 204 with the best model.

Table 42 – Predicted number of influenza deaths (main) in a model containing only influenza parameters

Age - years	2004-05	2005-06	2006-07	2007-08	Dev/df	Durbin W
75+	696	133	187	123	5.45	0.66

7.2. Analysis of P+I deaths as any cause

When using any P+I deaths, models could be generated in children using glm functions.

7.2.1. Selection of models

Results of the main model for any P+I deaths are shown in Table 43. Influenza was a significant covariate in all age groups except the 5-14 years. Models generally showed a high goodness of fit and a lack of auto-correlation, except in the elderly: auto-correlation was found in the 65+, and a lower goodness of fit was seen in the 75+ years (dev/df=5.21).

Variations of the main model in children could not be run except for the addition of a population term. Holidays and returns were only significant in the 75+, in which it improved the model. Adding interactions did not bring important improvement, adding a population term improved the models in the elderly and separate parameters by season only show a real improvement in the 65+.

Table 43 – Results from the main model for any P+I deaths

Age - years	Model parameters	Deviance /df	d of Durbin Watson
<5	Flu SL, pneumo	0.63	1.94
5-14	Pneumo	0.65	2.00
15-49	Flu SL, myco, adeno	1.07	2.13
50-64	Flu GP, mycopl, pneumo, parainfl	0.99	2.21
65-74	Flu GP, pneumo mycopl	1.90	1.49
75+	Flu GP, pneumo, hemo, mycopl, RSV	5.21	1.08

Table 44 – Goodness of fit (deviance/df) of main model and its variations, any P+I deaths

Age - years	Main model	+ holidays/ returns	+ influenza/ season***	+ interactions pathogens	+ population
<5	0.63	0.63	NA*	NA*	0.63
5-14	0.65	NA*	NA*	NA*	0.65
15-49	1.07	1.06	1.06 / 1.06	1.06	1.07
50-64	0.99	0.99	0.96 / 0.95	0.95**	0.98
65-74	1.90	1.90	1.71 / 1.59	1.86	1.68
75+	5.21	4.87	4.54/ 4.57	5.10	5.18

** Glm model could not run.*
*** The influenza parameter is no longer included, but 2 interaction terms containing influenza are kept.*
**** As this model involves a smaller number of weeks, dev/df of main model for the same weeks is shown first, dev/df of model with influenza by season is shown second.*

Time lags (+ 1 or 2 weeks for the dependent variable deaths) did improve the model performance in the 65+ in terms of goodness-of-fit and residual autocorrelation, with a substantial improvement in the 75+ and a slight one in the 65-74 years; changes in the <65 years were minimal (Table 61). A lag +2 weeks was best and a lag +3 weeks did not improve the models.

Table 45 – Goodness of fit and level of autocorrelation across main and lagged models, any P+I deaths

Age - years	Main (unlagged) model		Lag +1 on main		Lag +2 on main	
	Dev/df	Durbin Watson d[b]	Dev/df	Durbin Watson db	Dev/df	Durbin Watson db
<5	0.63	1.94	NA*	NA*	NA*	NA*
5-14	0.65	2.00	NA*	NA*	NA*	NA*
15-49	1.07	2.13	1.09	2.16	1.08	2.15
50-64	0.99	2.21	0.99	2.16	0.99	2.18
65-74	1.90	1.49	1.78	1.52	1.78	1.58
75+	5.21	1.08	4.45	1.21	4.43	1.23

Bold: best value across models for this age: dev/df as low as possible and close to 1; Durbin Watson d closest to 2 (not taking into account the number of parameters).

In general, variations of the main model did not bring a substantial improvement in the 4 age groups under 65 years. Based on the results from Table 44 and Table 45, a number of models were tested in the 65-74 and the 75+, involving breaks, influenza by season, population and varying time lags for the dependant variable. Models using interactions in the 75+ were also tested but were systematically inferior in to the models with influenza parameters by season using the same lags.

As when using P+I death as main cause, the best models in the elderly were those with influenza parameterized by season and a lag +1 for deaths. Table 46 shows that these more advanced models are clearly superior to the main model in terms of fit and auto-correlation.

[b] Durbin Watson test: values should be the closest to 2. d<1.0 or d close to 4 indicate strong auto-correlation.

Table 46 – Comparison of models for any P+I deaths in the elderly 65+

Model parameters	Dev/df	DW test	p White noise
65-74 years			
Main model*	1.71	1.62§	
With flu/season, no lag: Influenza by season (3), pneumo, pop	1.45	1.93	0.74
With flu/season, lag +1: Influenza by season (4), pneumo, RSV, pop, holiday	1.41	1.91	0.93
With flu/season, lag +2: Influenza by season (2), pneumo, mycopl, RSV	1.46	1.94	0.57
75+ years			
Main model*	4.54	1.21§	
With flu/season, no lag: Influenza by season (4), pneumo, RSV, mycopl, return Jan	4.40	1.26§	0.000
With flu/season, lag +1: Influenza by season (3), pneumo, RSV, mycopl, hemo, return Jan	3.96	1.49§	0.000
With flu/season, lag +2: Influenza by season (3), pneumo, mycopl, RSV	4.00	1.29§	0.000

** Parameters of main model differ from the tables above because a smaller number of weeks is used in models with seasons (end of 2003-04 and beginning of 2008-09 are excluded).*

§ Showing autocorrelation of residuals according to Durbin Watson table for d values.

In the 4 age groups under 65 years, the main model was kept as it was adequate and more parsimonious. The final model selected are thus:

1. In all age groups <65 years, the main model (pathogens only);
2. In the 65-74 years, models with influenza parameterized by season, a lag +1 for deaths, holiday breaks and returns, and a population term;
3. In the 75+ years, models with influenza parameterized by season, a lag +1 for deaths, holiday breaks and returns, and a population term in the 65-74 years.

7.2.2. Findings of final models

All final models <75 years showed a high goodness of fit (dev/df <1.40) and no auto-correlation of residuals according to Durbin Watson and auto-correlograms (Table 47). In the 75+, goodness of fit was acceptable (dev/df = 3.96) but the Durbin Watson test indicated residual auto-correlation. However, fit and residual distribution of the final model was substantially improved compared to the main model, with dev/df decreasing from 5.21 to 3.96 and Durbin Watson d increasing from 1.08 to 1.49. The auto-correlogram shows that the first two lags are concerned but that the level of auto-correlation is minor (autocorrelation around 0.30, see appendix).

Table 47 – Goodness of fit and residual distribution of final models, any P+I death

Age - years	Model parameters	Flu beta (95%CI)	Dev / df	DW	White noise
<5	Influenza SL, pneumo	0.002 (-0.001–0.004)	0.63	1.94	p=0.53
5-14	Pneumo	none	0.65	2.00	p=0.90
15-49	Influenza SL, mycopl, adeno	0.134 (-0.013–0.052)	1.07	2.14	p=0.77
50-64	Influenza, pneumo, parainfl, mycopl	0.042 (0.021–0.063)	0.99	2.21	p=0.14
65-74	4 flu seasons, pneumo, RSV, holiday, pop, lag +1	Range 0.11–0.76	1.40	1.91	p=0.91
75+	3 flu seasons, pneumo, RSV, mycopl, hemo, return Jan, lag +1	Range 1.64–3.07	3.96	1.49	p=0.00

DW: Value of d in the Durbin Watson test for auto-correlation.

The selection of the final models showed relatively similar results of fit and residual distribution as for P+I as main cause of death. Influenza was included in the final models of all age groups, except in the 5-14 years of age – in which P+I mortality is already extremely low.

7.2.3. Predicted influenza attributable deaths

The number of predicted influenza deaths in an average season (356) shows that most predicted influenza deaths (76%) are found in the elderly 75+, as expected. The predicted influenza deaths represent in average 3% of all predicted P+I deaths. This proportion varies across seasons (range 2-8% by season), especially across age groups, and is highest in the <5 years (15%).

The number of predicted deaths also widely varies across season, and is lower during low intensity seasons. This number is particularly low in the 2007-08 season (75 deaths for all age) because the influenza parameter in the 75+ years in that season was above the threshold for parameter selection (p value Wald test p=0.203) and thus no influenza deaths could be calculated in the 75+. This number in 2007-08 is even lower than the predicted deaths by regressing on P+I death as principal cause (85 deaths for all ages). We also calculated the predicted number of influenza deaths if this parameter was retained for the sensitivity analysis (below).

Table 48 – Predicted influenza deaths in an influenza season (average), any P+I death

Age	Number deaths	Mortality rate	% of P+I deaths
<5 years	2	0.3	15%
5-14 years	0	0.0	0%
15-49 years	8	0.2	3%
50-64 years	30	1.6	4%
65-74 years	51	5.4	3%
75 years +	266	31.0	3%
Total	356	3.4	3%

Table 49 – Prediction of influenza attributable deaths by season, any P+I death

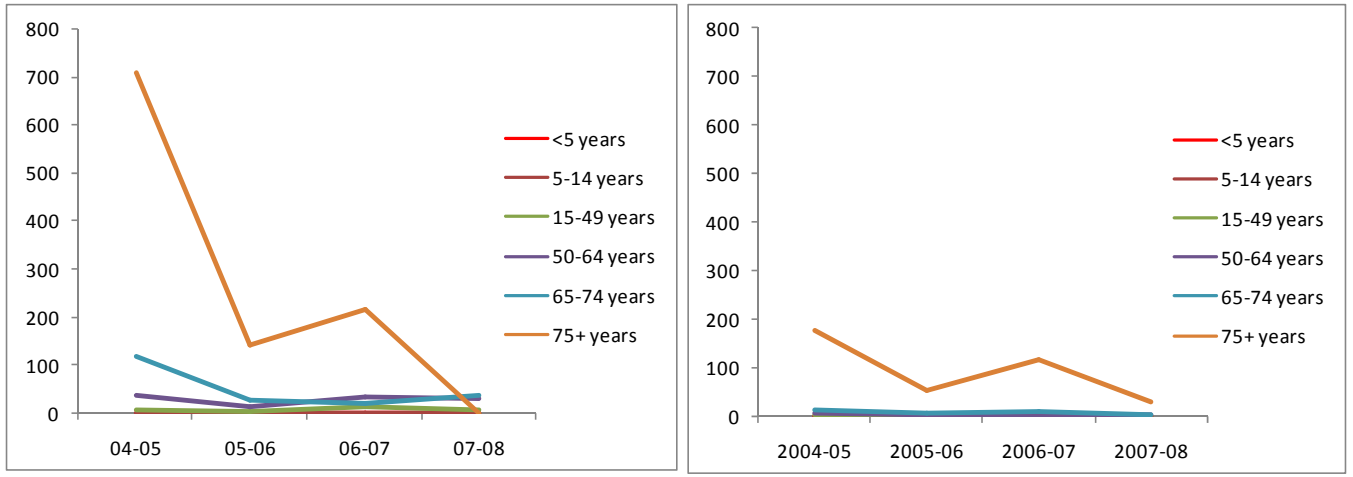
Age	2004-05	2005-06	2006-07	2007-08
Intensity	***medium***	***low***	***medium***	***low***
<5 years	2	1	2	1
5-14 years	0	0	0	0
15-49 years	7	4	13	8
50-64 years	38	15	36	30
65-74 years	120	28	19	36
75 years +	707	142	217	0
Total	873	190	286	75

The predicted numbers of influenza deaths are in average 3.3 higher than the number of deaths coded as influenza (any cause) in the death certificates (Table 50), with the same variation by age as for P+I death as principal cause.

Table 50 – Number of deaths coded as influenza by the communities (any cause)

Age	2004-05	2005-06	2006-07	2007-08
Intensity	*medium*	*low*	*medium*	*low*
<5 years	0	0	0	0
5-14 years	0	0	0	0
15-49 years	2	0	1	0
50-64 years	7	2	2	1
65-74 years	13	7	9	4
75 years +	176	53	116	31
Total	198	63	129	36

Figure 17 – Predicted influenza deaths by the model and deaths coded as influenza (right), by age group and season, any P+I death



7.2.3.1. Sensitivity analysis

If the influenza parameter of the last influenza season is kept in the 75+, the model has a slightly better fit and level of auto-correlation (Table 51). The number of death for the 2007-08 season would amount to 94 instead of 0 deaths, but the average number of deaths by season and for all age groups would only be 8% higher (total of 386 estimated deaths). Confidence intervals would of course be wider around point estimates (not shown).

Table 51 – Predicted number of influenza deaths (any cause) if influenza of the 2007-08 season is forced in the model

Age - years	2004-05	2005-06	2006-07	2007-08	Average season	Dev / df	Durbin W
75+	722	146	221	94	296	3.95	1.50
Total	888	194	291	170	386		

Another sensitivity analysis is also applied to all elderly 65+, as they concentrate 89% of all predicted influenza deaths, by including only influenza parameters in the model. The goodness of fit is good in the 65-74 years but poor in the 75+, and auto-correlation is high in both age groups. All influenza parameters (by season) are significant and the predicted number of influenza deaths is 80% higher compared to the final models (Table 52).

Table 52 – Predicted number of influenza deaths (any cause) in model containing only influenza parameters

Age - years	Model with only influenza			Final model	% additional deaths predicted in models with only influenza
	N admissions	Dev/df	Durbin W	N admissions	
65-74	68	1.74	1.59	51	33%
75+	504	7.82	0.63	266	89%
Total	581			356	80%

8. SUMMARY OF RESULTS AND COMPARISON WITH OTHER STUDIES

The predicted numbers and rates of influenza attributable admissions and deaths are summarized below (Table 53 - Table 55), by model and outcome, for comparison with indicators from other studies. We also included rates predicted by the sensitivity analyses (Table 56).

Table 53 – Predicted number of influenza admissions and deaths per average season in this analysis

Age - years	Admissions P+I				Deaths P+I	
	Principal cause		Any cause		Principal cause	Any cause
	Model 1	Model 2	Model 1	Model 2		
<5	540	600	661	690	0	2
5-14	287	290	348	362	0	0
15-49	309	301	462	429	5	8
50-64	201	178	356	316	11	30
65-74	234	179	386	305	24	51
75+	568	554	1043	1019	204	266
Total	2140	2102	3256	3120	244	356

Table 54 – Predicted rate of influenza admissions and deaths per average season in this analysis (per 100 000)

Age - years	Admissions P+I				Deaths P+I	
	Principal cause		Any cause		Principal cause	Any cause
	Model 1	Model 2	Model 1	Model 2		
<5	92.7	90.4	113.4	118.4	0.0	0.3
5-14	23.6	24.0	28.7	29.8	0.0	0.0
15-49	6.2	5.7	9.2	8.6	0.1	0.2
50-64	10.5	9.4	18.7	16.6	0.6	1.6
65-74	24.8	17.6	40.8	32.2	2.5	5.4
75+	66.2	64.0	121.5	118.6	23.7	31.0
Total	20.3	19.0	31.0	29.7	2.3	3.4

Table 55 – Proportion of P+I influenza attributed admissions and deaths per average season in this analysis

Age - years	Admissions				Deaths	
	Principal cause		Any cause		Principal cause	Any cause
	Model 1	Model 2	Model 1	Model 2		
<5	8%	8%	8%	8%	NE	15%
5-14	11%	11%	11%	12%	NE	0%
15-49	7%	7%	6%	6%	11%	3%
50-64	5%	4%	4%	3%	6%	4%
65-74	5%	4%	3%	3%	6%	3%
75+	4%	4%	3%	3%	6%	3%
Total	6%	6%	5%	4%	6%	3%

NE: Not estimated by the models due to low numbers of outcome.

Table 56 – Predicted rate of influenza admissions and deaths per average season in sensitivity analyses (per 100 000)

Age group - years	Admissions				P+I deaths	
	P+I as principal cause		Respira-tory admis-sions	R+C admis-sions	Principal cause	Any cause
	Influenza and RSV only	Influenza as only pathogen			Influenza as only pathogen	
<5	112.6	Poor fit	62.6	62.8	NE	NE
5-14	37.5	39.1	39.6	39.8	NE	NE
15-49	9.2	9.4	12.4	12.6	NE	NE
50-64	14.0	14.4	36.9	50.3	NE	NE
65-74	32.6	32.8	70.9	61.0	NE	7.1
75+	112.6	Poor fit	171.4	224.2	33.1	58.7
Total	29.6	-	41.0	47.0	-	-

NE: Not estimated by the models due to low numbers of outcome.

8.1. Comparison of influenza attributable admissions

We compared our results with those predicted by published studies using similar regression analyses. As the outcome used as dependant variables substantially influences the prediction of influenza admissions in our study, we conducted this comparison by type of outcome.

Our results were compared to four published studies using P+I admissions as outcome (Table 57). Most of these studies used P+I as principal diagnosis. However, all of these studies included only influenza (and RSV in one) as pathogens in the independent parameters. We thus compared these findings to those of our sensitivity analysis that include only influenza as pathogen.

- Newall et al in Australia found a much higher admission rate based on P+I as principal diagnosis than in our final model, as well as an higher proportion of P+I admissions attributed to influenza.[1] However, the study period involved higher intensity seasons compared to our four influenza seasons and only included influenza and RSV as pathogen parameters. Our models including only influenza and RSV predict for the highest – though moderate – season (2004-05) a rate of 141/100 000 in the ≥65 years, which is close to the 157/100 000 estimate from Newall.
- In the Netherlands, Baltussen found much lower admission rates in the <60 years (3.0 in the <60 years vs. 15 in the <65 years in Belgium).[3] Rates in the elderly cannot be compared as they are stratified by risk status. The analysis was conducted in older seasons (1984-1994).
- Xue at al conducted a similar study over 1998-2006 in Norway but again limited the parameters to influenza and season.[8] The study found a higher mean rate for 1998-2006 in all ages compared to our standard analysis, but a similar rate than in our analysis limited to influenza and RSV parameters (27 vs. 29.6/100 000 in Belgium). The lowest intensity season also showed rates similar to our lowest seasons (around 10 compared to 12.4/100 000 in Belgium).
- Thompson et al also conducted a similar study in the US over 22 seasons, with some very high intensity seasons and only influenza as pathogen.[12] The rates are again higher compared to our findings, but closer to our analysis including only influenza and RSV as pathogens: 29.6/100 000 in Belgium vs. 36.8 for P+I as principal diagnosis. Our analysis found however very different rates by age groups, with higher rate in the 5-49 years (14.7/100 000) and lower rates in the 50-64 years (14.0/100 000) and 65+ (70.6/100 000) (Table 57). Comparison of mean rates between the two studies is also difficult because the periods do not overlap, and estimates in Thompson vary by a factor of 10 between the lowest and highest season rate (all ages).

Pitman et al conducted a similar study on P+I admissions involving the same pathogens, but only raw numbers are predicted, no rates.[2]

Table 57 – Predicted number of influenza admissions in other regression studies, based on P+I admissions

Study	Age group - years	Study design	Rate (/100 000)	% of P+I admis-sions
Newall 2008, Australia (1998-2005)[1]	50-64	Multivariate, P+I admissions (principal). Parameters: influenza A/B, RSV, season and population	33.3	12.3%
	≥65		157.4	12.4%
Baltussen 1998, NL (1984-1994)[3]	0-59	Multivariate on pneumonia admissions. Parameters: year, month, influenza (various parameters)	3.0	NA
	≥65		40 in low risk 185 high risk	NA
Xue 2010, Norway (1998-2006)[8]	All, 98-06	Quasi-Poisson multivariate on P+I admissions (main). Parameters: ILI rate, week and season.	27*	NA
	All, 2004-05		27*	NA
	All, 2005-06		20*	NA
Thompson 2004, US, 1979-2001[12]	All	Poisson multivariate on P+I admissions, primary and any P+I diagnosis. Parameters: influenza, season, trend.	Primary: 36.8 Any: 52.0	Primary: 8.6% Any: 8.0%
	5-49		Primary: 6.8	NA
	50-64		Primary: 37.9	NA
	≥65 **		Primary: 205.0	NA

** Calculated from Table 3 of the publication.*
*** Calculated by Newall et al.*
P+I: Pneumonia and/or influenza.

We also compared our findings from the analyses on respiratory and respiratory and/or circulatory (R+C) admissions to those using similar regression analyses on similar admission outcomes:

- In the US, Zhou et al used the same ICD-9 codes for R+C admissions in a regression analysis involving only influenza and RSV as pathogens over 1993-2008. This study found relatively similar predicted admission rates per 100 000 than in our study in the 5-49 years and 50-64 years: 16.8 vs. 17.9 and 65.6 vs. 50.3, respectively, in the US and Belgium. US rates were higher in the 65+ in average (309 in the US vs. 224 per 100 000 in our study) and corresponded to the rate estimated in the highest season in Belgium (353/100 000 in 2004-05).[19]
- Thompson et al performed an earlier US study on R+C admissions over 1979-2001, including some seasons with very intensity, in an analysis including only influenza as pathogen. It found higher rates than in Belgium but relatively close to the Belgian rates found in the seasons with moderate intensity.[12]
- Newall performed a similar analysis in Australia on respiratory admissions in the adults ≥50 years of age over 1998-2005.[1] Admission rates were higher than in our study, but again the models included only influenza and RSV as pathogens. The Australian mean rates corresponded to the Belgian rates in the highest season: in the 2003-04 (moderate intensity), the rates were estimated at 55.5/100 000 in the 50-64 years and 280/100 000 in the ≥65 years in Belgium, compared to 57.6/100 000 and 282/100 000 in Australia, respectively.
- Pitman et al conducted a similar study in the UK on respiratory admissions over 1996-2004 and involved nearly the same pathogens. Only raw predicted numbers are provided, no rates, but in its discussion, Newall calculated influenza-associated admission rates.[2, 20] Admission rates were lower than in Belgium in the 45-64 years (19.6 vs. 36.9/100 000) but higher in the 65+ (136.9 vs. 118.8/100 000).
- Muller-Pebody performed a multivariate analysis in the elderly 65+ in England, based on unspecified pneumonia, bronchitis, bronchiolitis and COPD admissions.[5] She found a much higher influenza admission rates in that age (229/100 000) compared to 119/100 000 in our analysis restricted to respiratory admissions. However, the intercept was excluded in the models and the study period (1995-98) included influenza seasons with a much higher intensity compared to our study period.[21] This rate is still within the range of Belgian seasonal rates over the 4-year period (26-280/100 000 in the ≥65 years).

- Xue performed a similar analysis on all admissions in Norway, but this outcome cannot be compared to our analyses.[8]

Table 58 – Predicted number of influenza admissions in other regression studies, based on other admission outcomes

Study	Age group - years	Study design	Rate (/100 000)
Zhou 2012, US (1993-2008)[19]	all	Negative binomial multivariate on R+C admissions, excluding influenza and RSV admissions. Parameters: influenza and RSV	63.5
	5-49		16.8
	50-64		65.6
	≥65		309.1
Thompson 2004, US (1979-2001)[12]	All	Poisson multivariate on all R+C admissions, primary diagnosis. Parameters: influenza, season, trend.	88.4
	5-49		20.8
	50-64		83.8
Thomson 2004, estimated by Newall[1]	≥65		445.0
Newall 2008, Australia (1998-2005)[1]	50-64	Multivariate, other respiratory admissions (main). Parameters: influenza A/B, RSV, season and population	57.6
	≥65		282.0
Pitman 2007, US (1996-2003), estimated by Newall[1, 2]	45-64	Standard regression on respiratory admissions. Parameters: RSV, influenza A/B, rhinovirus, adenovirus, *S. pneumoniae*, *M. pneumoniae*, *H. influenzae*	19.6
	≥65		136.9
Muller-Pebody 2006, England (1995-98)[5]	≥65	Multivariate on LRTI admissions unspecified. Parameters: admissions for pertussis, hemo, flu, klebsiella, RSV, pneumo, other streptococci	229
Xue 2010, Norway (1998-2006)[8]	All, 98-2006	Quasi-Poisson multivariate on all admissions. Parameters: ILI rate, week and season.	40 (range 17-66)

LRTI: Lower tract respiratory infections.
R+C: Respiratory and circulatory admissions.

A number of recent cohort studies have also estimated the number of influenza admissions in industrialized countries but these only involved children (Table 59). We selected those that involved more than one hospital:

- Poehling at al. conducted a prospective study among children from three US counties.[9] She found a rate of 90/100 000 children <5 years influenza attributable admissions over 2000-04, ranging from 40 to 150/100 000 per season. These data are fully in line with our estimates of 90-93/100 000 over 2004-08 based on P+I as main diagnosis, with a range of 43-127/100 000 per season.
- Wiegl conducted a similar prospective study in two German hospitals over 1996-2001 and found a rate of 123/100 000 in the <5 years and 22/100 000 in the 5-16 years, which are relatively comparable to our estimates.[10]
- Silvennoinen conducted a *retrospective* study among children <16 years in Finland over 1988-2004 and found an average rate of 93/100 000 in the <5 years, similar to the rates found in our study.[22]

Table 59 – Predicted number of influenza admissions in prospective studies

Study	Age group	Study design	Rate (/100 000)
Poehling 2006, US (2000-04)[9]	<5 years	Prospective study involving 2797 children <5 years	Mean: 90 Range: 40-150
Weigl 2002, Germany (1996-2001)	<5 years	Prospective hospital study involving 3469 children <17 years	123
	5-16 years		22

In conclusion, our admission estimates are in line with those from prospective studies but only studies in children were found. In general, other regression studies tended to predicted higher influenza admission rates in the elderly and lower admission rates in the younger groups. However, there are two major differences with our study: first, most influenza seasons in these studies were in the nineties, involving higher intensity seasons; second, most studies only involved influenza (and RSV) as pathogen and thus tended to over-estimate the proportion of influenza-attributable outcomes. When we compared these findings with our analyses involving similar seasons, similar outcomes and similar pathogens, our estimates were overall in line with those predicted by other studies. These studies also confirm that the use of less specific outcomes as dependant variable, such as all respiratory and circulatory admissions or all-cause admissions, predicts a higher number of outcome – though the association with these outcomes is less frequently significant.

8.2. Comparison of influenza attributable deaths

A few studies have also used a similar regression analysis, with different deaths categories as outcome (Table 60).

The two studies using a similar outcome and methodology found estimates that are closed to our study: Newall in Australia found mortality rates in the 50-64 years of 0.6 vs. 0.6/100 000 in our study, and 17.6/100 000 in the ≥65 years vs. an average of 12.6 and a range of 3.8-29.9 per season in our study.[1] However, these deaths represented a higher proportion of P+I deaths than in our study (17-19% vs. 6% in our study). The regression analysis of Thompson in the US, based on P+I deaths as principal cause, found higher rates by age group but was based on older seasons (1990-1999) with higher intensity due to a more virulent A(H3N2) strain.[11] We found similar rates in the two Belgian seasons with moderate intensity (19.9/100 000 in the ≥65 years vs. 22.1 in Thompson).

Table 60 – Predicted number of influenza deaths in other regression studies

Study	Design	Age group - years	Rate / 100 000	% of P+I deaths
Newall 2008, Australia (1998-2005)[1]	Multivariate, P+I deaths (principal). Parameters: influenza A/B, RSV, season and population	50-64	0.6	18.8%
		≥65	17.6	16.8%
	Multivariate, respiratory causes, same parameters	≥65	91.3	NR
	Multivariate, all causes, same parameters	≥65	116.4	NR
Thompson 2003, US (1976-1999)[11]	Multivariate Poisson on P+I deaths (principal). Parameters: influenza, RSV, population.	All ages	3.1	9.8%
		<1	0.3	NA
		1-4	0.2	

Study	Design	Age group - years	Rate / 100 000	% of P+I deaths
		5-49	0.2	NA
		50-64	1.3	NA
		≥65	22.1	NA
	Multivariate on respiratory and circulatory deaths, same parameters.	50-64	7.5	NR
		≥65	98.3	NR
Schanzer 2007, Canada (1989-98)[7]	Multivariate on all cause deaths. Parameters: flu deaths, months, population	50-64	4.0	NA
		≥65	108.8	NR
Schanzer 2008, Canada (1995-2000)[6]	Multivariate on hospital respiratory deaths. Parameters: flu, RSV, parainfluenza, adeno, other ILI, breaks/returns	<50	0.3 (0.2-0.4)	NR
		≥65	96 (27-159)	NR

NA: Non available; NR: Non relevant.

As expected, the four studies that used broader causes of deaths (respiratory, circulatory or all cause deaths) found much higher mortality rates. Schanzer (2007) performed a regression analysis based on all-cause deaths and influenza as sole pathogen parameter and found a rate of 4/100 000 in the 50-64 years and ; in the ≥65 years, mortality was estimated at 109/100 000 compared to 13 or 18/100 000 if deaths with P+I as main or any cause were considered in our study.[7] Newall also found much higher mortality rates when the analysis was based on all cause deaths.[1] Schanzer (2008) found higher mortality in a regression analysis based on all respiratory deaths.[6] The Thompson study also found higher rates when the regression analysis was based on all respiratory and circulatory deaths.[11]

9. CONCLUSIONS AND DISCUSSION

The final models to predict the number of influenza attributable admissions and deaths show overall a good goodness of fit, although lower in the elderly ≥75 years. The two models, involving influenza parameterized by season or adding interaction across pathogens, provide very similar predictions for influenza admissions although they have a different structure. This suggests that our analysis is quite robust. In these four seasons with moderate and low intensity, we predict a range of 2000-3000 influenza admissions by mean influenza season. The estimates vary largely by season: numbers predicted in the highest season represent around twice those predicted in the lowest season. The highest hospitalization rates are found in children <5 years, which account for 30-40% of all admissions, followed by the elderly ≥65 years, which concentrate around 40% of admissions. We also predict an average of 250-350 influenza deaths per season, with a higher variability across season: the highest season accounts for 6 to 12-fold the number of deaths in the lowest season. This high variability of influenza predicted outcomes across seasons were confirmed by other studies, and reflect the variability and changing severity of influenza strains.[11, 12, 19]

The estimates of influenza admissions and deaths also vary with the selected outcome. When regressing on P+I as principal diagnosis, we estimate around 2000 influenza admissions and 250 deaths by season. When regressing on P+I as any diagnosis, we find an estimated mean of 3000 and 350 influenza admissions by season. When regressing on respiratory and circulatory (R+C) admissions, we estimated around 4500 influenza admissions by season. The same variations was observed in similar studies in other countries, which conclude that determining the most appropriate outcome category is difficult. For instance, estimates based on P+I deaths may underestimate the total burden of influenza deaths because many deaths are caused by other secondary complications.[11, 12]

Our estimates of admissions in children are in line with those found in recent prospective studies. The comparison of our results with those from other regression analyses is complex because most published studies were based on higher intensity seasons from the nineties, included fewer independent parameters and did not always used the same outcomes. Although a number of studies predicted a higher number of admissions and deaths, we consistently found similar estimates when we used similar seasons, outcomes and pathogens.

The predicted proportion of all P+I admissions that are attributed to influenza in the elderly ≥65 years seems low (4-5% by season) when compared to the estimates of TIV vaccine effectiveness in preventing P+I admissions in the same age group (~20%, see Part I). However, our study likely underestimates the influenza burden on admissions and deaths because our analyses are based on four low and moderate intensity seasons, and on specific outcomes. Our analysis on respiratory and circulatory admissions showed a lower fit, especially in some age groups, but the predicted number of admissions is probably closer to the real number of influenza-related admissions.

Our study has a number of limitations. The major one is a remaining level of auto-correlation of residuals, especially in the admissions in the elderly, although it has been substantially reduced by improving the model, using other datasets and adding parameters. The extent to which the lack of independence of the residuals affects our estimates in unknown. However, the changes in the model that greatly improved this independence of residuals have not changed the estimates by more than 5-10 cases by season. Another limitation is that this analysis is valid for low to moderate intensity seasons, and is not representative of high intensity seasons in Belgium. Most of our models include a high intercept, which assumes that the proportion of the model variability that is explained by the unexplained admissions is constant throughout the season. This high intercept tends to underestimate the contribution of influenza. We also tested models without intercepts, and these yielded higher coefficients for influenza, but the goodness of fit and residual distribution were not satisfactory.

APPENDIX: FURTHER DETAILS ON REGRESSION RESULTS

1. COMPARING DIFFERENT DISTRIBUTIONS, ADMISSIONS

1.1. Comparison summary

The tables below compare the parameters, model performances and predicted influenza outcomes for the three main distribution models (main or basic model) based on P+I as principal diagnosis, as an example for all outcomes.

Table 61 – Comparison of parameters and coefficient for influenza in best (main) model across regression distributions, P+I principal

Age - years	Parameters negative binomial	Parameters over-dispersed poisson	Parameters gaussian
<5	All expect hemo	All expect hemo	All expect hemo
5-14	Flu SL, RSV, myco, pneumo	Flu SL, RSV, myco, pneumo, adeno (but p>0.2)	Flu SL, RSV, myco, parainfl, pneumo
15-49	Flu GP, mycopl	Flu GP, mycopl, hemo, pneumo, RSV (but p>0.2)	Flu GP, mycopl, hemo, pneumo
50-64	Flu GP	Flu GP, mycopl, pneumo	Flu GP, pneumo, mycopl
65-74	Pneumo	Flu GP, RSV, pneumo mycopl (but p>0.2)	Flu GP, pneumo, RSV, mycopl
75+	Pneumo	Flu GP, pneumo, hemo, mycopl, RSV, parainfluenza (but p>0.2)	Flu GP, pneumo, hemo, mycopl, RSV, parainfluenza

Table 62 – Influenza attributable admissions predicted by main model, by distributions and age group

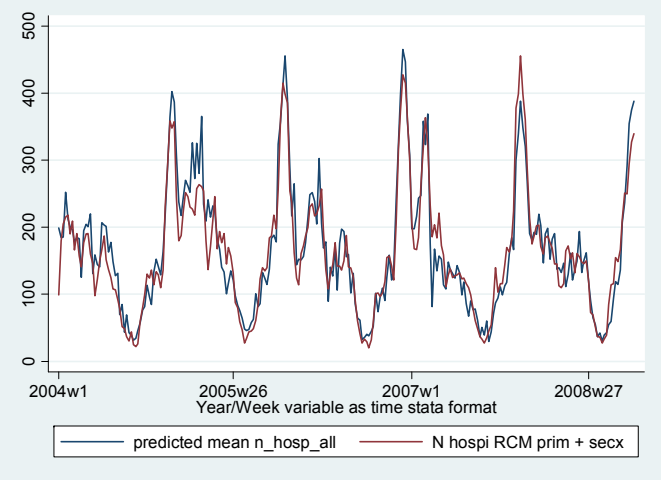
	N predicted influenza admissions			% of all P+I admissions		
Age	Negative binomial	Over-dispersed poisson	Gaussian	Negative binomial	Over-dispersed poisson	Gaussian
<5 years	3033	2727	2672	9%	8%	8%
5-14 years	1462	1370	1305	12%	11%	11%
15-49 years	1710	1495	1479	8%	7%	7%
50-64 years	1227	884	844	6%	4%	4%
65-74 years	0	695	637	0%	3%	3%
75 years +	0	2205	2072	0%	3%	3%
Total	7432	9375	9008	4%	5%	5%

Table 63 – Comparison of model performance for variations from the main model

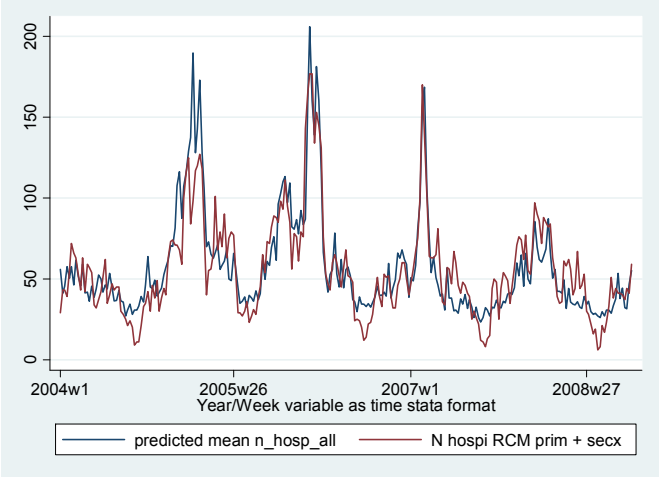
	Deviance/df				AIC (using glm)			
Age	w/ influenza/ season	w/ influenza A/B	w/ interactions	w/ breaks/ returns	w/ interactions	w/ influenza/ season	w/ influenza A/B	w/ breaks/ returns
<5 years	5.801	6.031	5.62	5.62	12.0	12.20	12.45	12.03
5-14 years	3.913	4.378	4.15	3.79	9.619	9.38	9.86	9.29
15-49 years	2.350		2.46	2.63	8.651	8.53		8.80
50-64 years	2.737		2.75	2.85	8.890	8.88		9.00
65-74 years	2.620		2.80	2.88	9.10	9.93		9.19
75 years +	7.646		7.46	7.63	14.61	14.77		14.77

1.2. Model fitting of negative binomial, main model

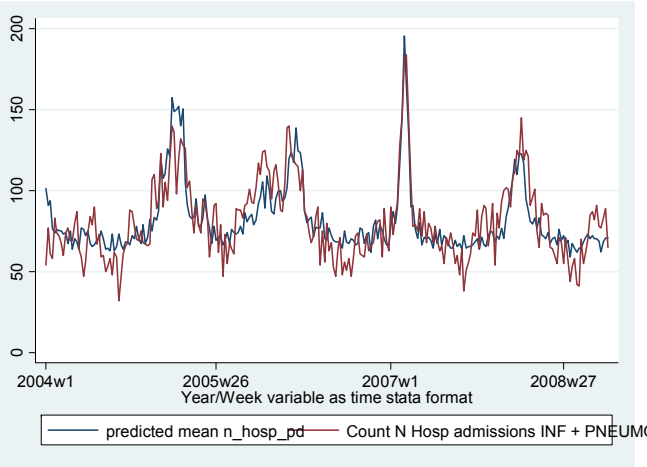
<5 years



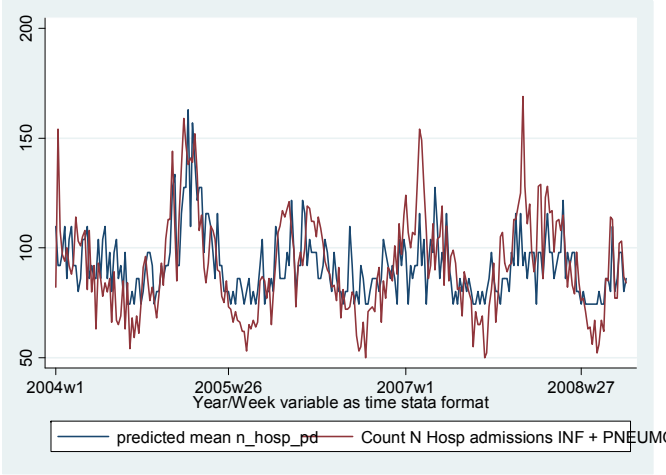
5-14 years



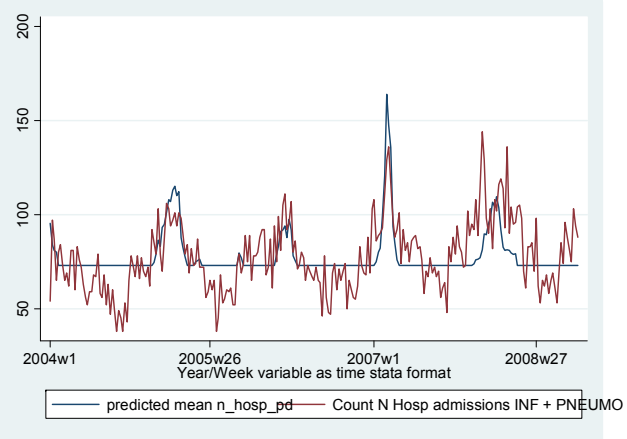
15-49 years



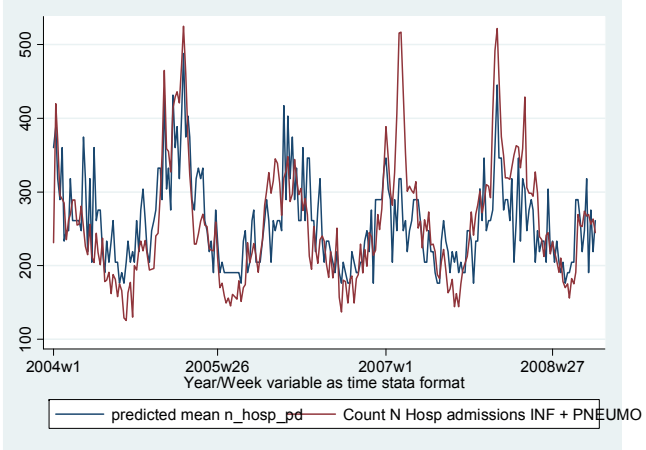
65-74 years, only pneumo



50-64 years, only influenza and intercept = 73

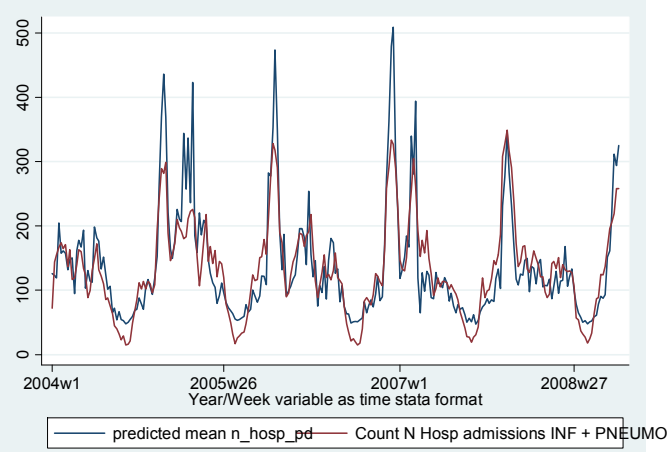


75+ years, only pneumo

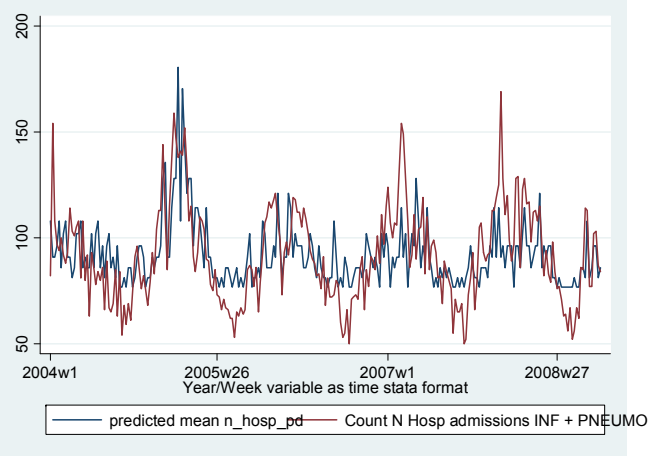


Negative binomial, log link

<5 years



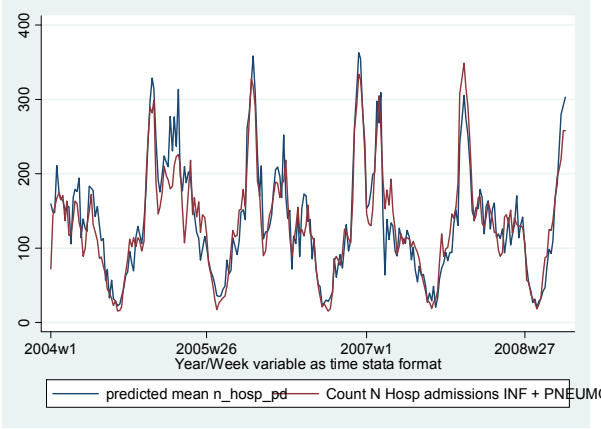
65-74 years



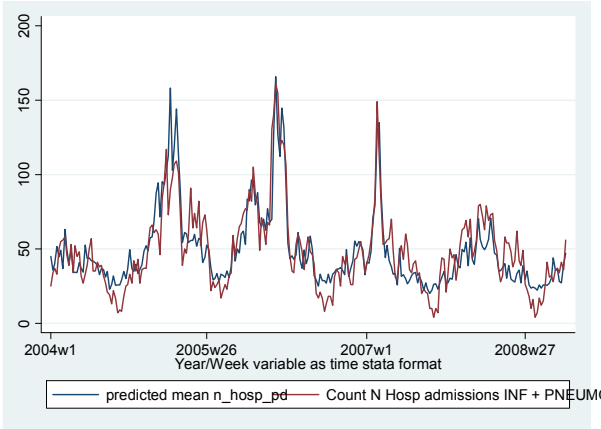
1.3. Model fitting of over-dispersed Poisson regression

Main model

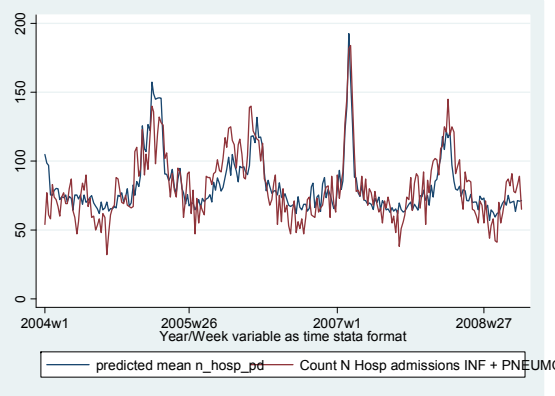
<5 years



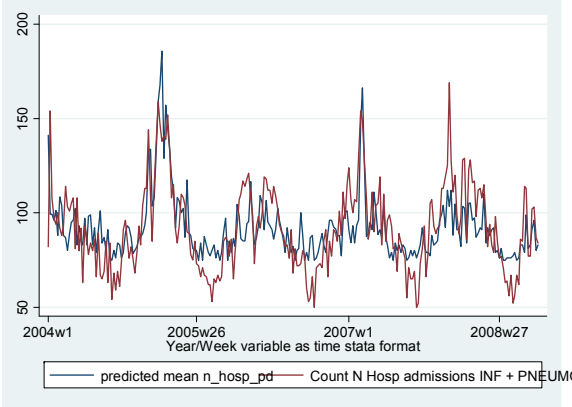
5-14 years



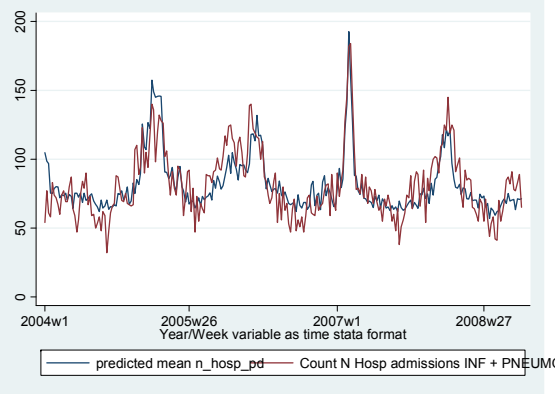
15-49 years



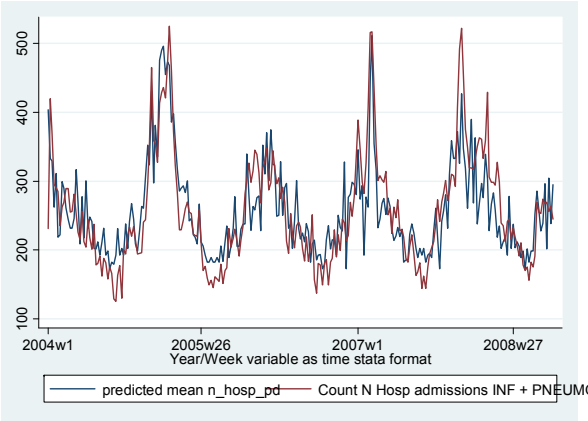
65-74 years



50-64 years

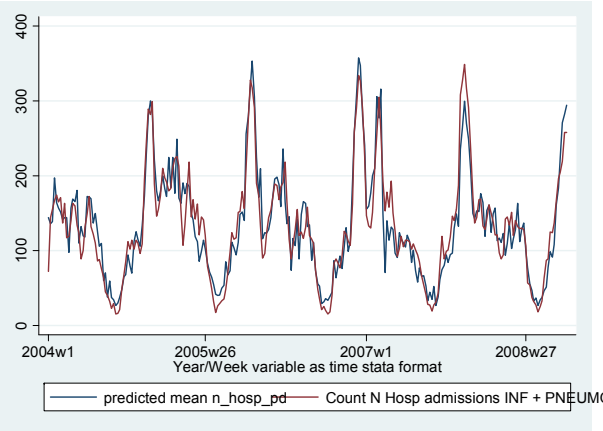


>75 years

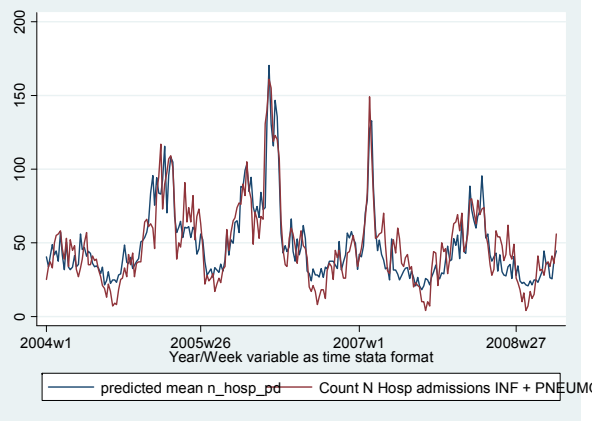


Influenza parameterized by season (for children as example)

<5 years



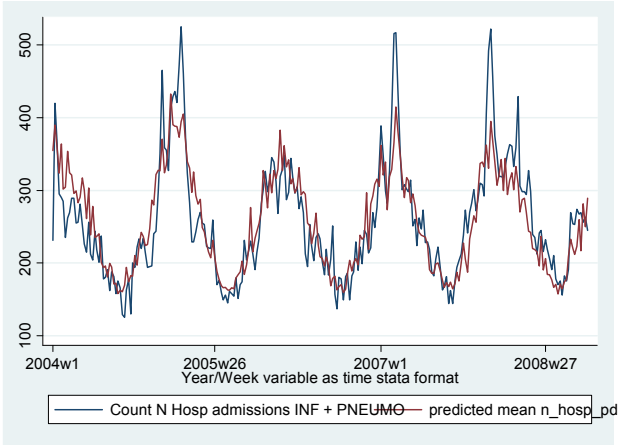
5-14 years



Seasonal terms

Only models in two age groups are presented below for illustration purpose, as these models are not kept.

<5 years



75+ years

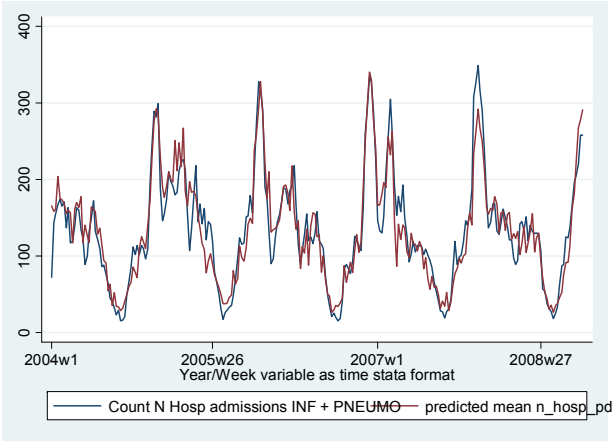


Table 64 – Predicted influenza attributable admissions from main and final models, by age group, P+I admissions as principal diagnosis, entire study period*

Age - years	N predicted influenza admissions			% of all P+I admissions (based on model)		
	Main model	Model 1 wk 40 2004 wk 39 2008	Model 2 wk 1 2004 wk 48 2008	Main model	Model 1	Model 2
<5	2727	1993	2195	8%	7%	6%
5-14	1370	1093	1211	11%	10%	10%
15-49	1495	1123	1194	7%	6%	6%
50-64	776	776	782	4%	5%	4%
65-74	695	981	737	3%	5%	3%
75+	2205	2160	2227	3%	4%	3%
Total	9375	8126	8346	5%	6%	5%

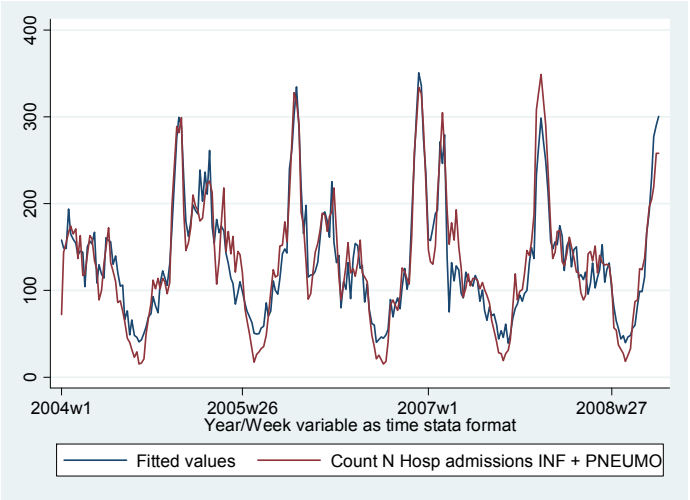
Caution: Study periods differ as model 1 involves a smaller number of weeks as weeks 1-39 of 2004 and weeks 41-48 of 2008 are excluded.

1.4. Standard (Gaussian) regression for admissions

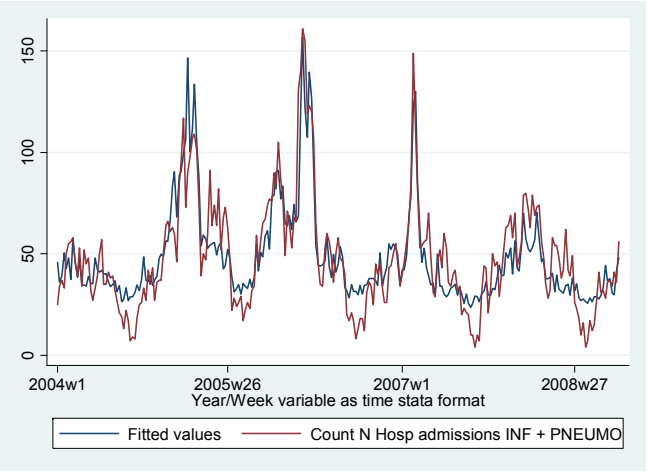
Table 65 – Model performance of the main model in Gaussian regression (P+I admissions, principal diagnosis)

Age - years	Model parameters	Adj. R^2	Stand. beta influenza	Proportion influenza attributable admissions
<5	flu SL, pneumo, myco, adeno, parainfl, RSV (all expect hemo)	0.86	0.29	8%
5-14	flu SL, RSV, myco, parainfl, pneumo	0.74	0.44	11%
15-49	Flu GP, mycopl, hemo, pneumo	0.64	0.55	7%
50-64	Flu GP, pneumo, mycopl	0.35	0.44	4%
65-74	Flu GP, pneumo, RSV, mycopl	0.46	0.35	3%
75+	Flu GP, pneumo, hemo, mycopl, RSV, parainfluenza	0.66	0.32	3%

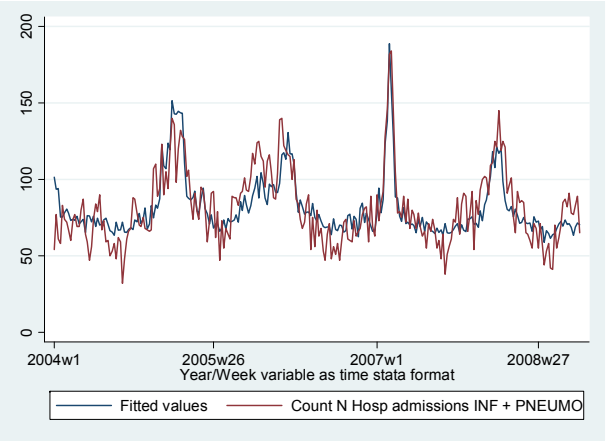
< 5 years



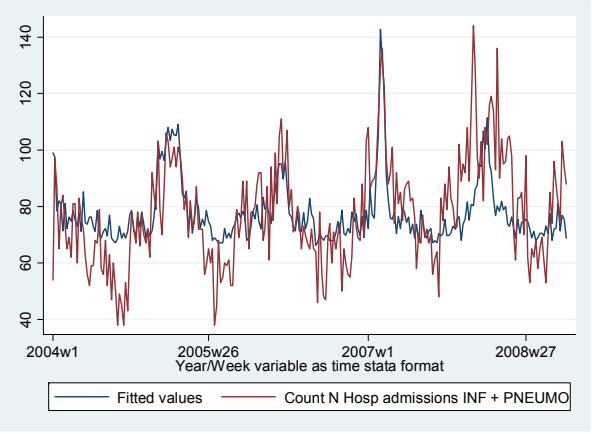
5-14 years



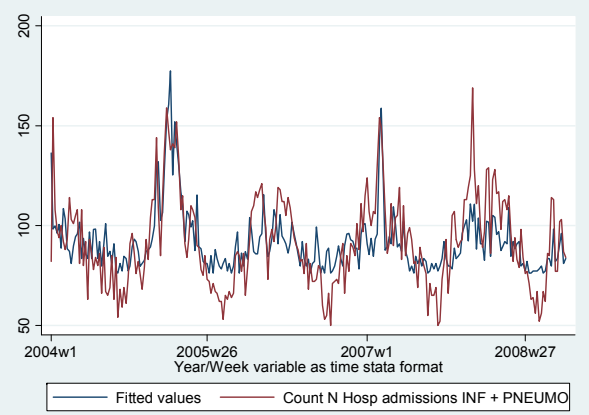
15-49 years



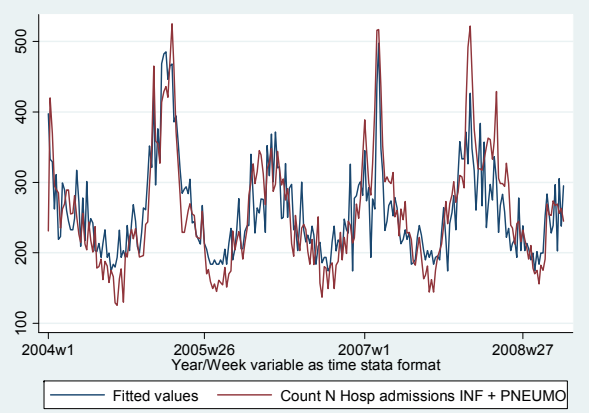
50-65 years



65-74 years



75+



The model does not fit well outside influenza seasons (model too high), especially in children, which seems due to a high constant term.

By large age group

Table 66 – Results of Gaussian regression with larger age groups

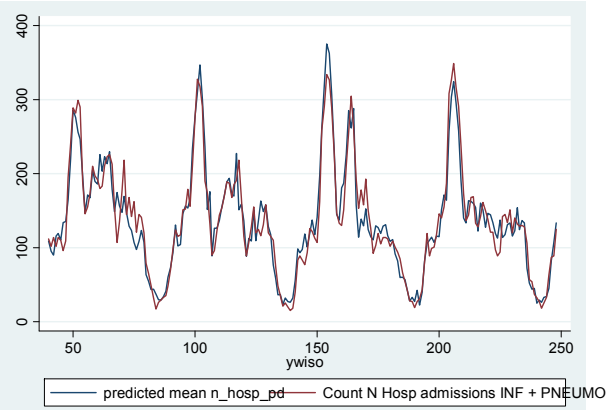
Age	Model parameters	Adj. R^2	Standardized beta for influenza
<15 years	flu SL, pneumo, myco, adeno, parainfl, RSV	0.90	0.25
15-64 years	Flu GP, pneumo, mycopl, hemo, adeno, parainfluenza	0.49	0.51
65 years +	Flu SL, pneumo, hemo, RSV, parainfl, adeno	0.49	0.11

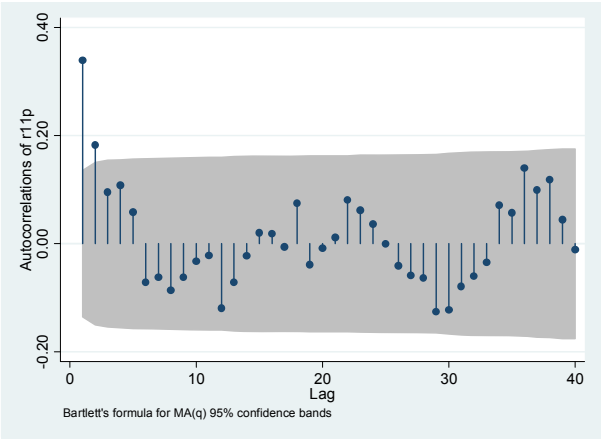
For children <15 years, the model fits better than in each model using smaller age groups, based on R^2. For adults 15-64 years, the model fits rather less well than using 2 separated models by age group.

2. MODELS FITS AND AUTO-CORRELOGRAMS OF DIFFERENT MODELS

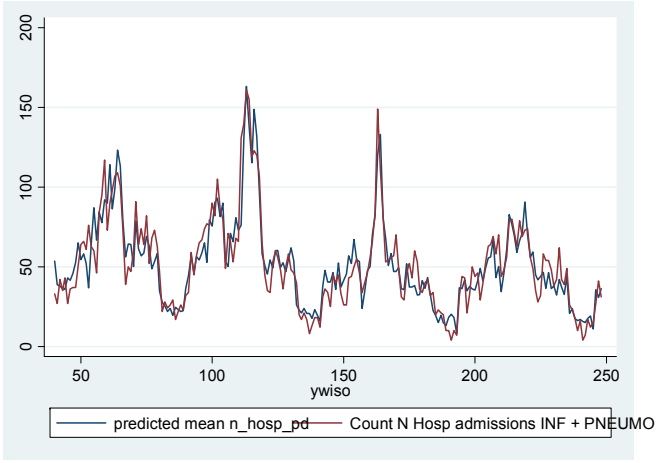
2.1. Model 1, P+I admission as principal diagnosis

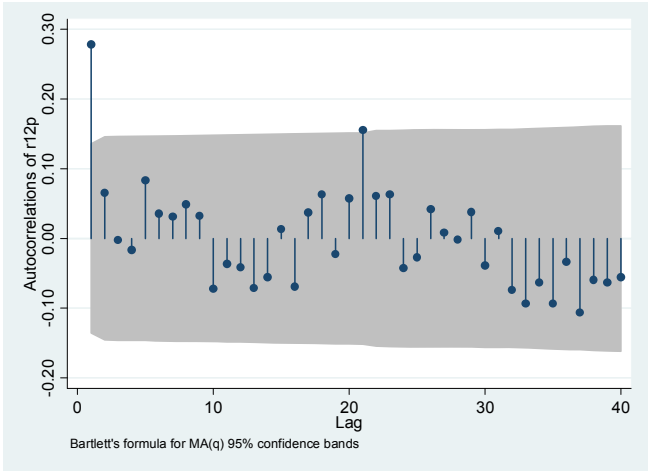
<5 years



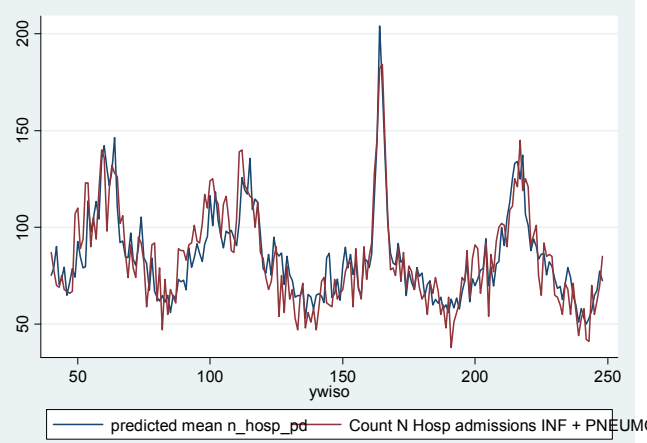


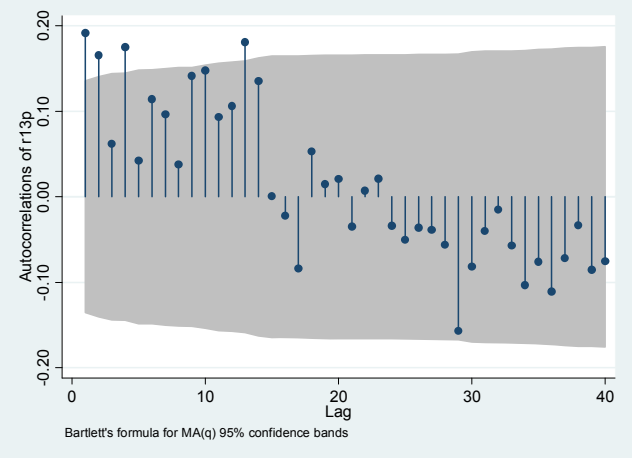
5-14 years



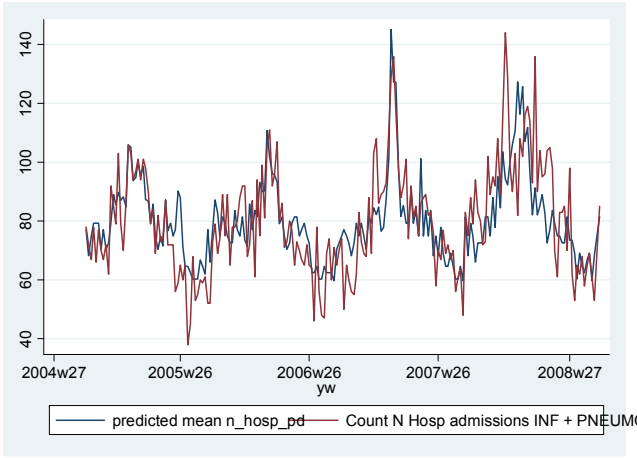


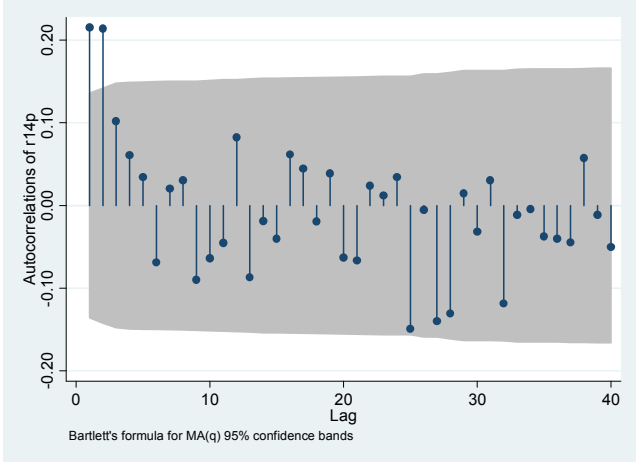
15-49 years



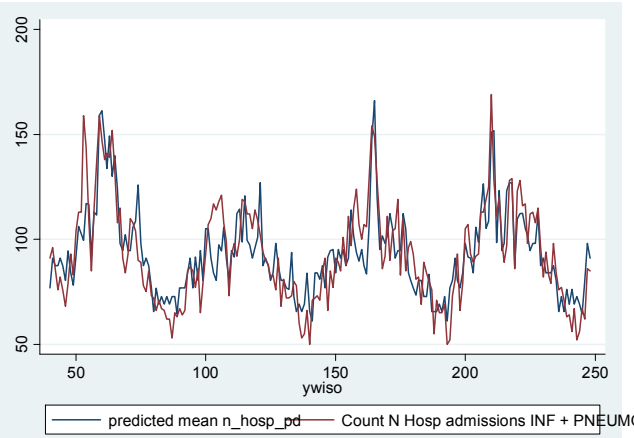


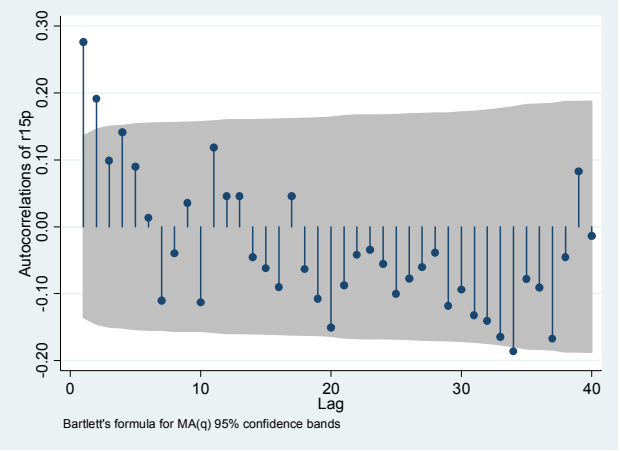
50-64 years



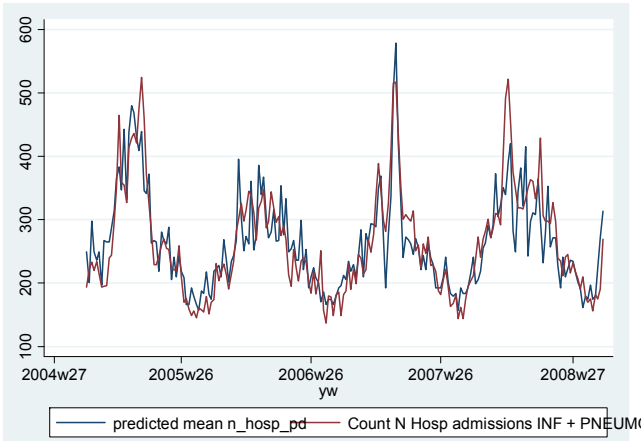


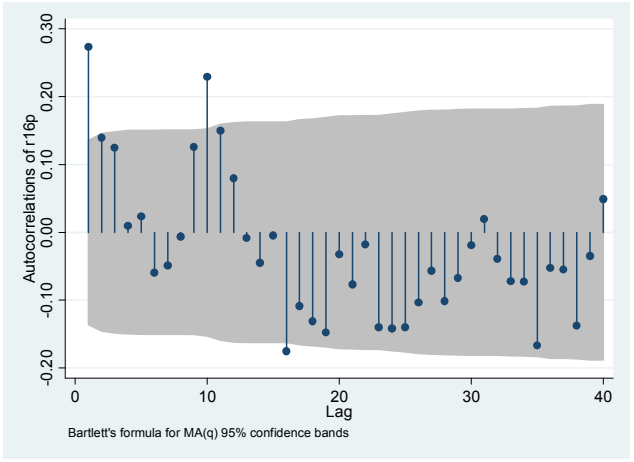
65-74 years





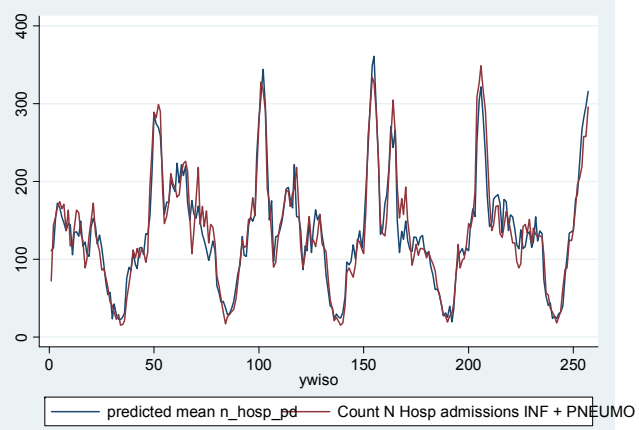
75+

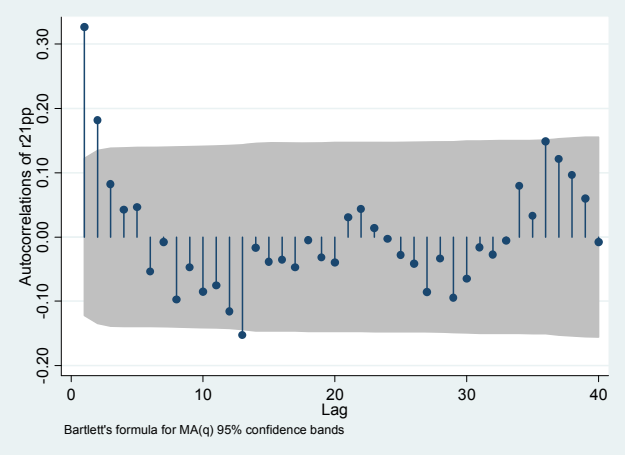




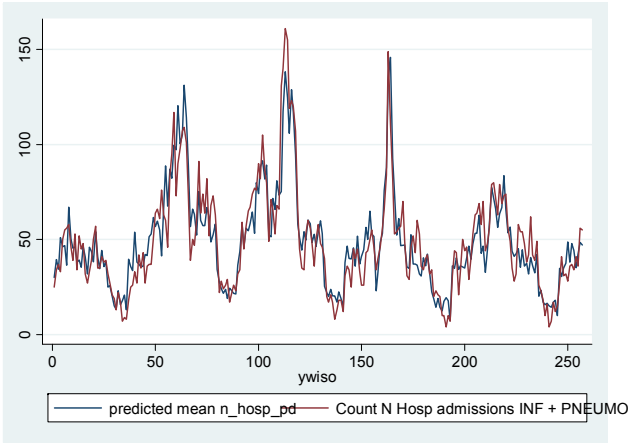
2.2. Model 2, P+I admission as main diagnosis

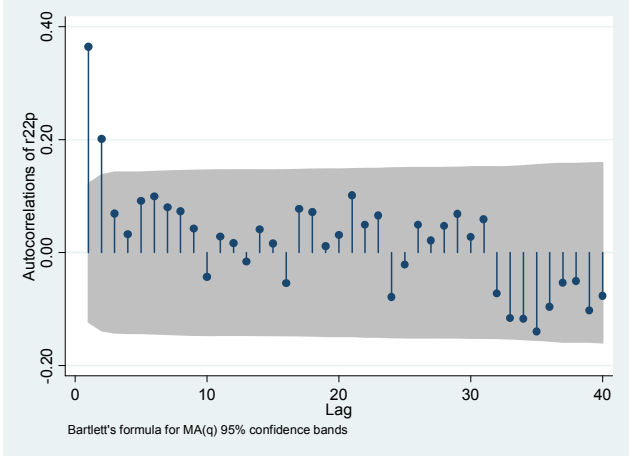
<5 years



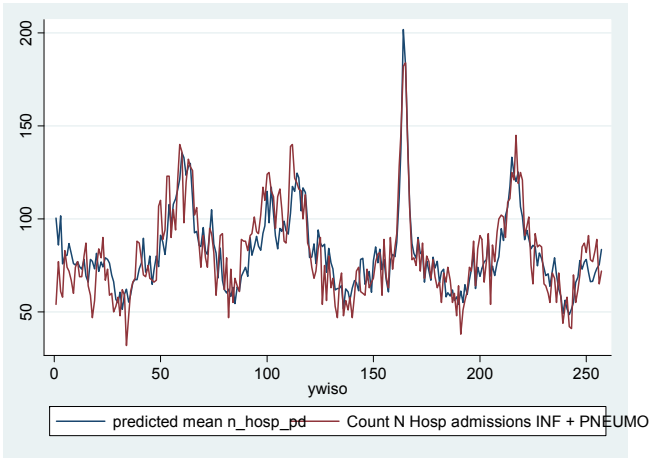


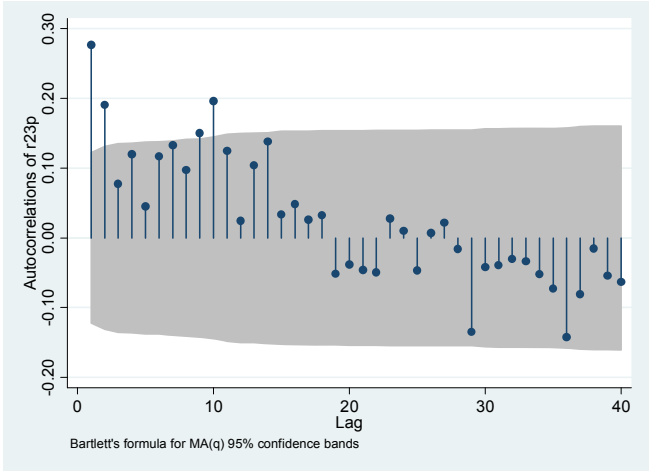
5-14 years



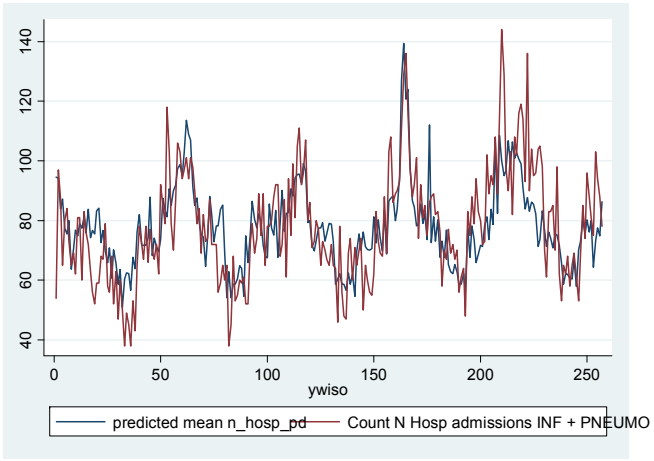


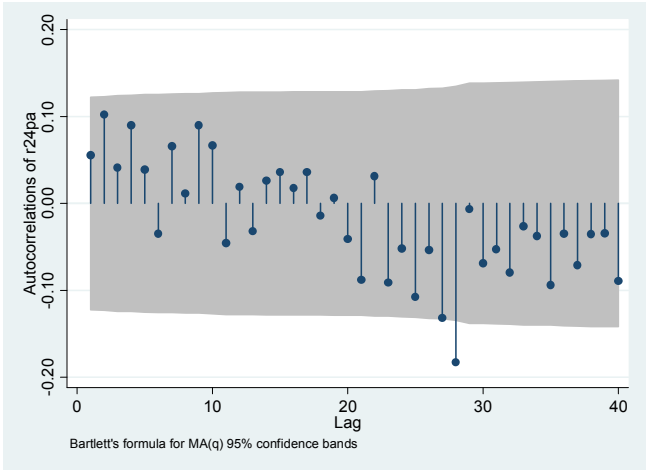
15-49 years



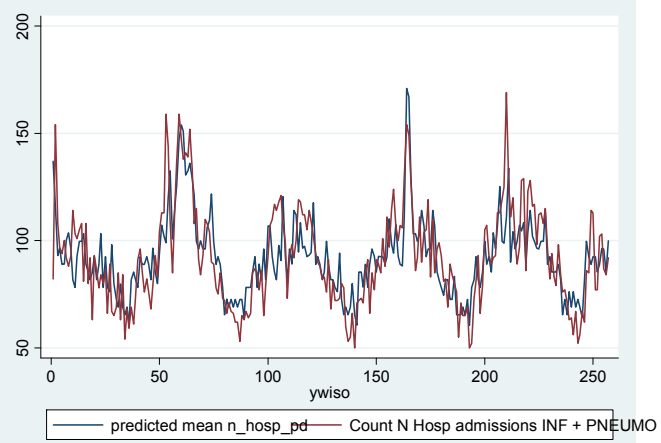


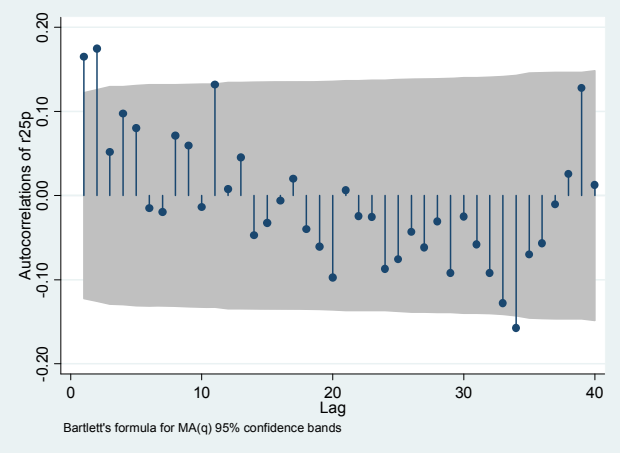
50-64 years



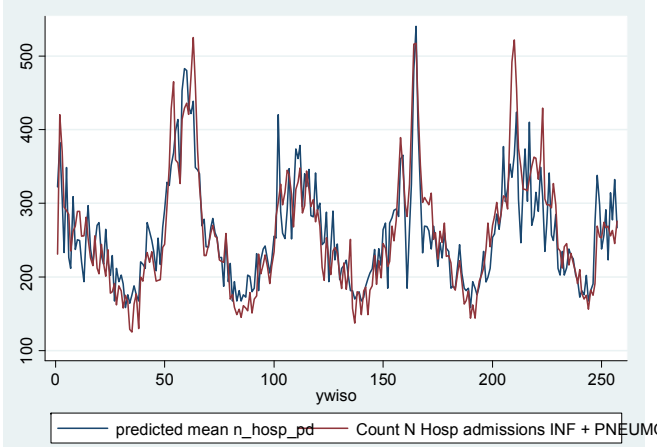


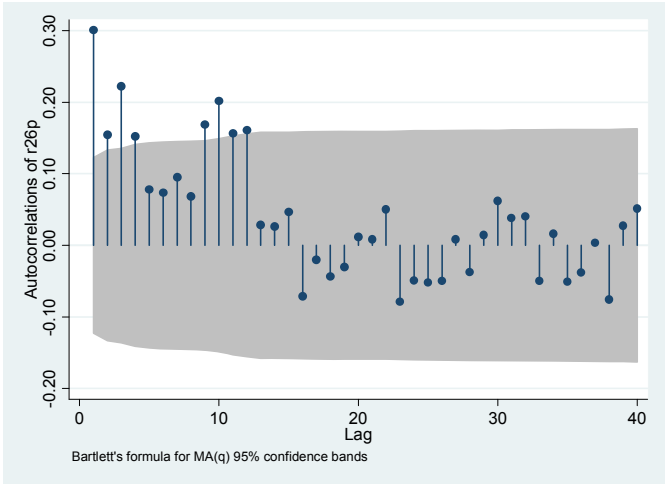
65-74 years





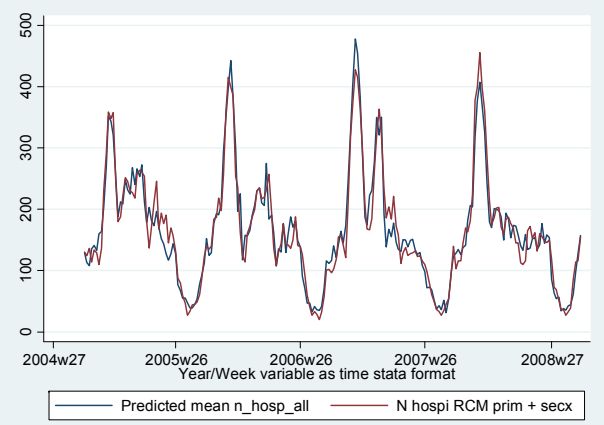
75+

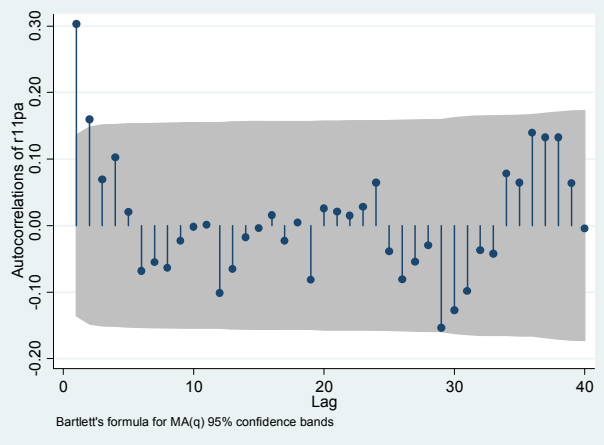




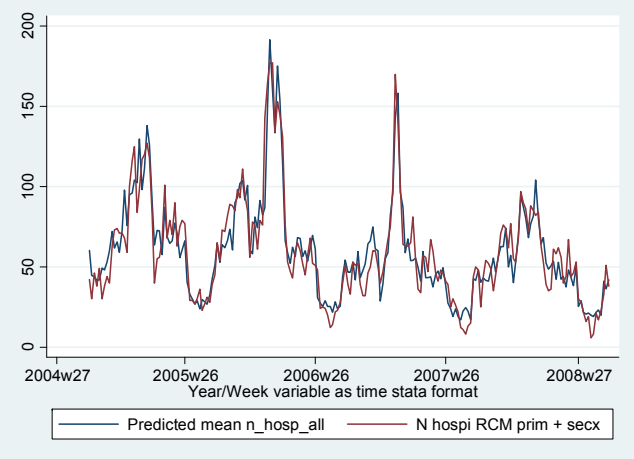
2.3. Model 1, P+I admission as any diagnosis

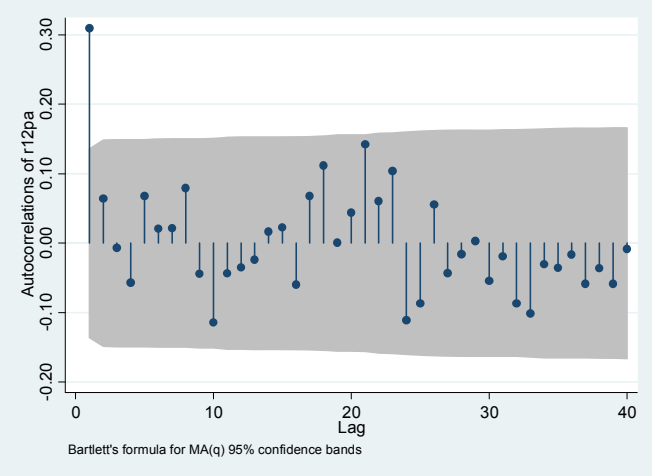
<5 years



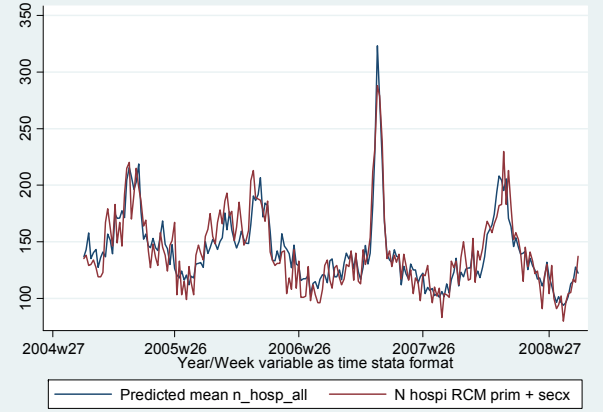


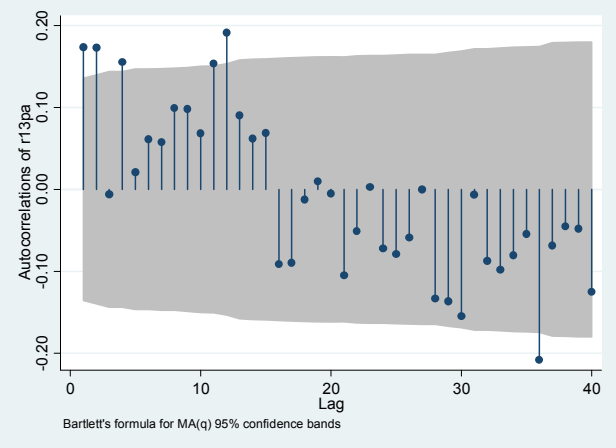
5-14 years



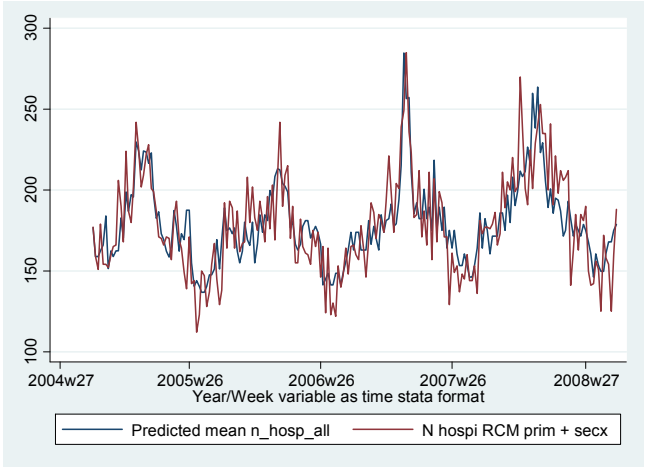


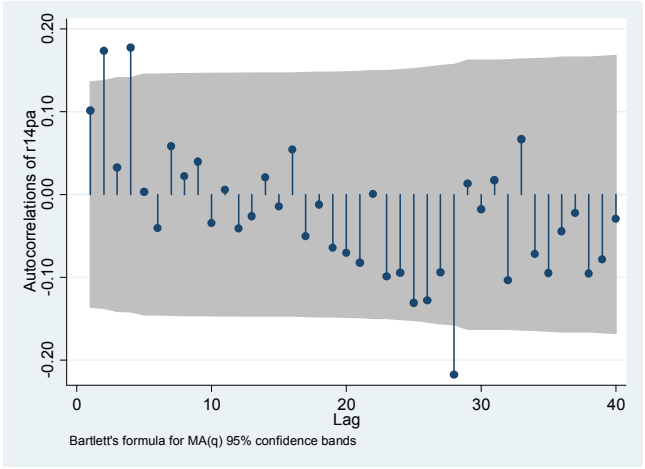
15-49 years



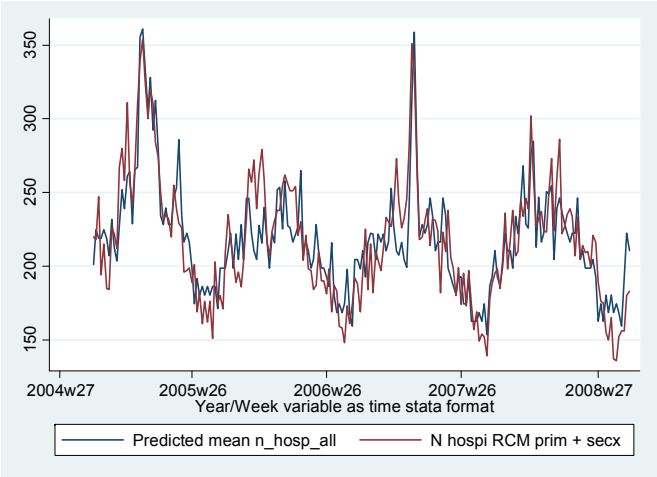


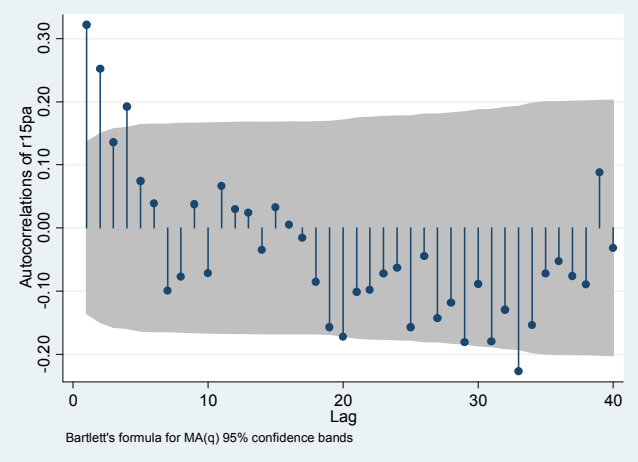
50-64 years



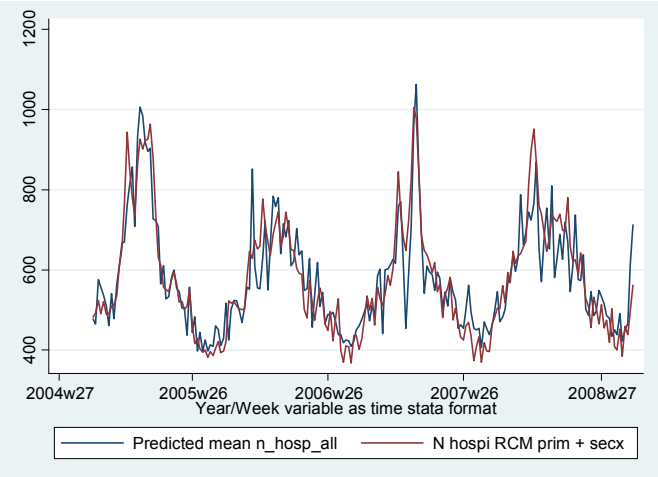


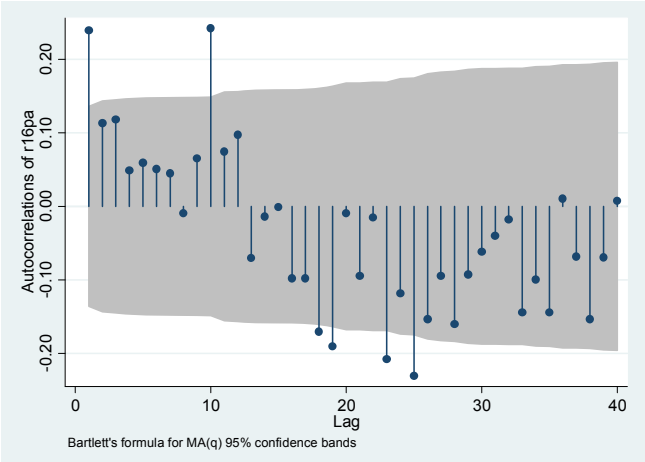
65-74 years





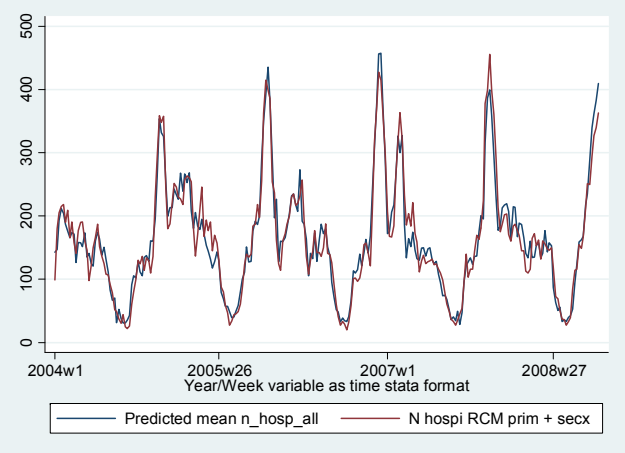
75+

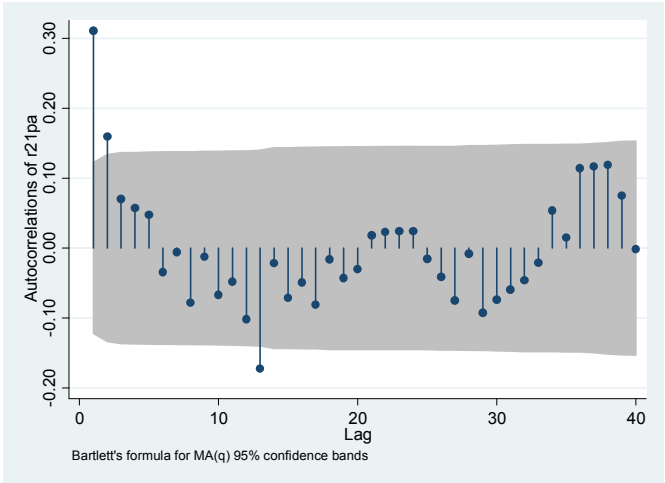




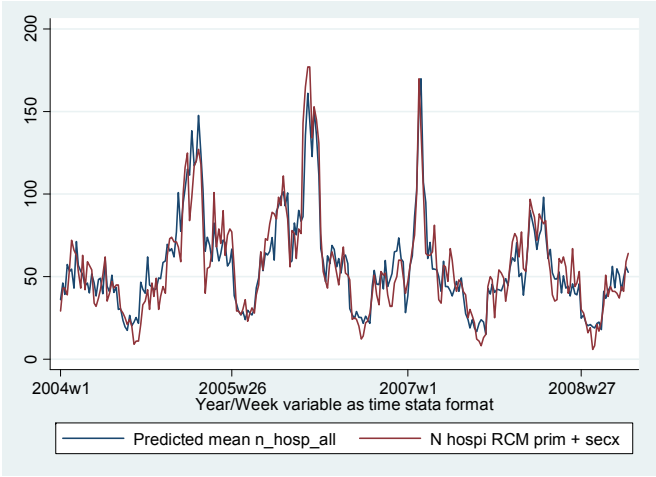
2.4. Model 2, P+I admission as any diagnosis

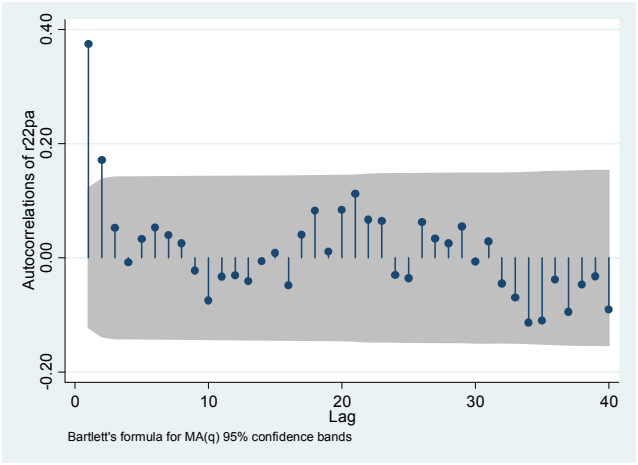
<5 years



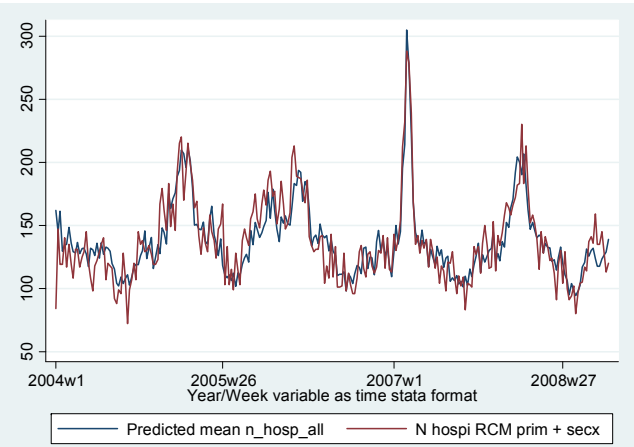


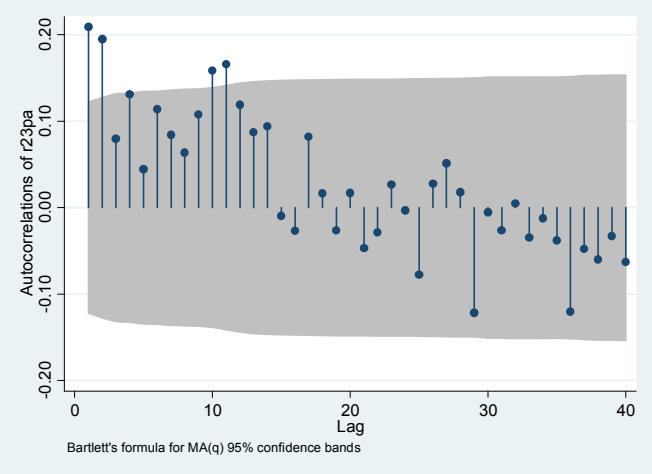
5-14 years



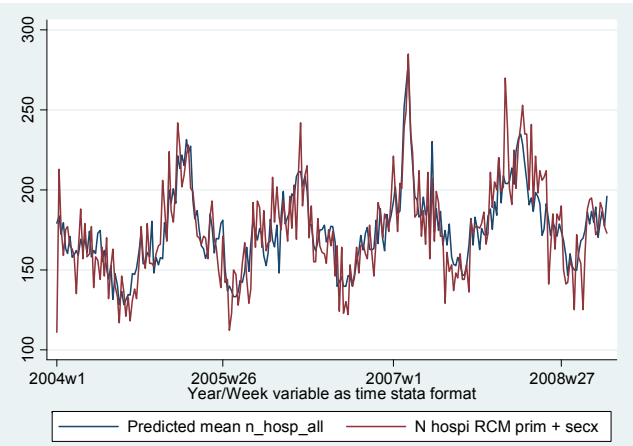


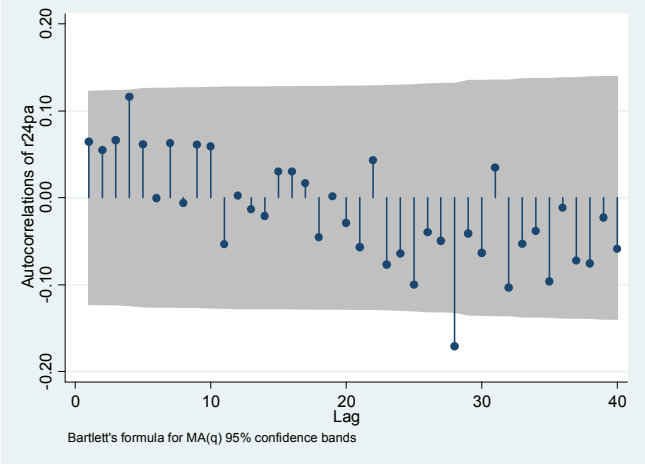
15-49 years



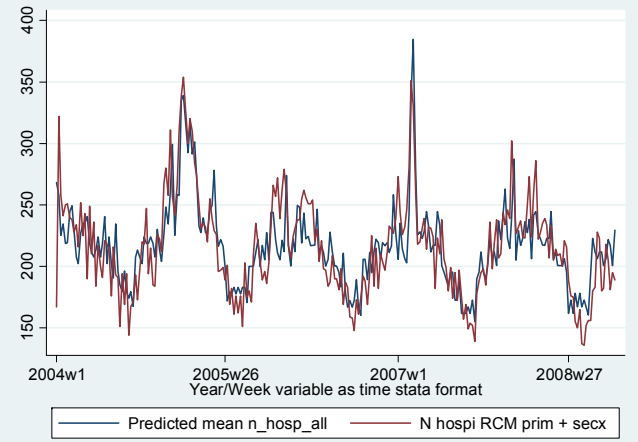


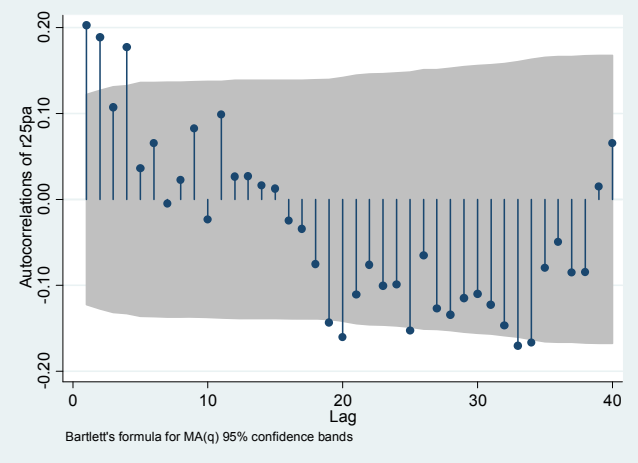
50-64 years



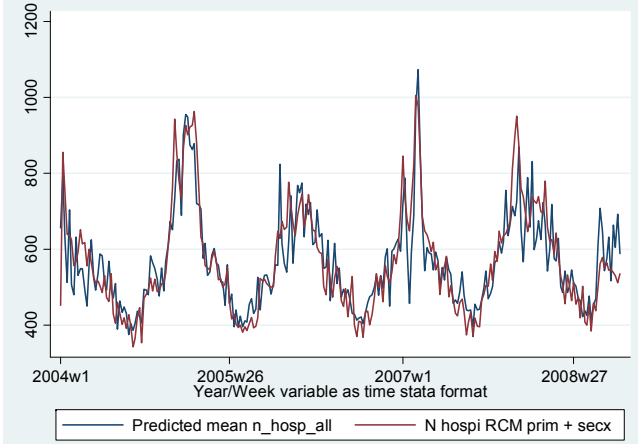


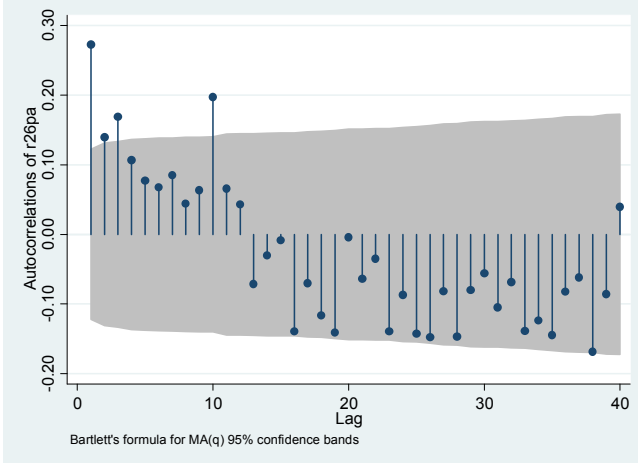
65-74 years





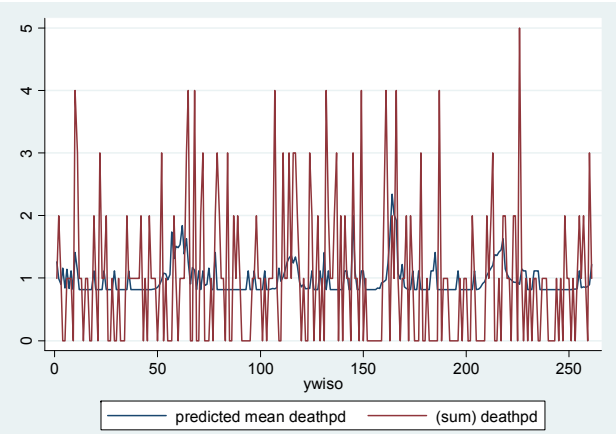
75+

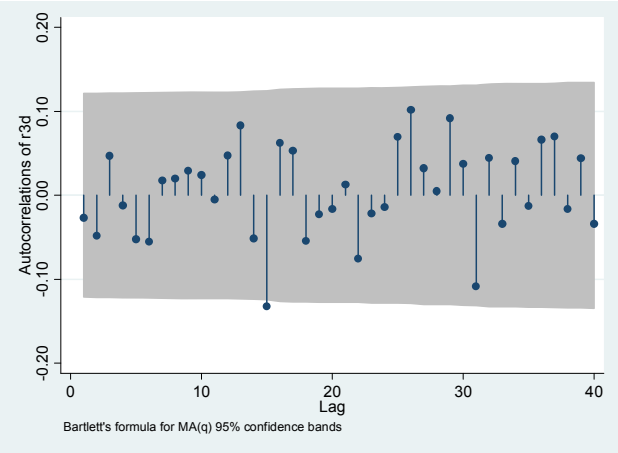




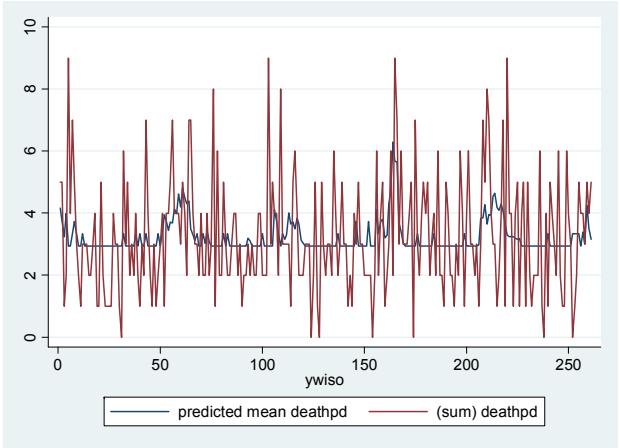
2.5. P+I deaths as principal cause

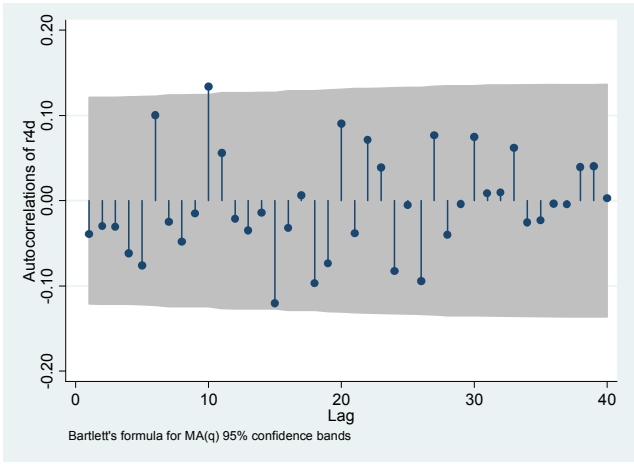
15-49 years



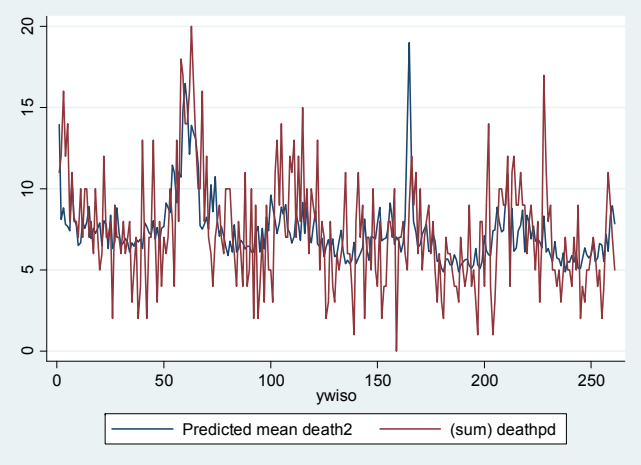


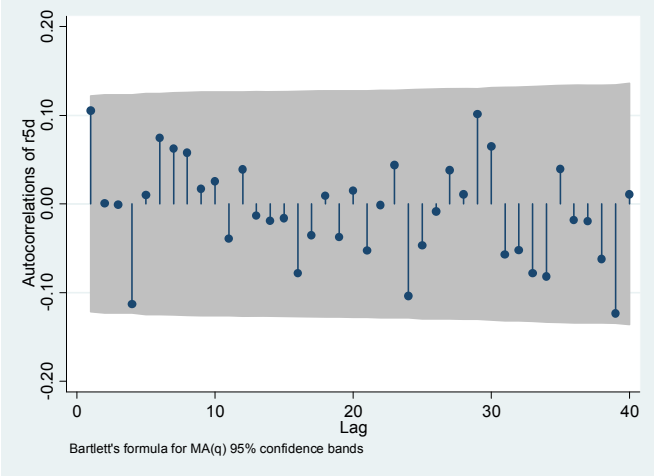
50-64 years



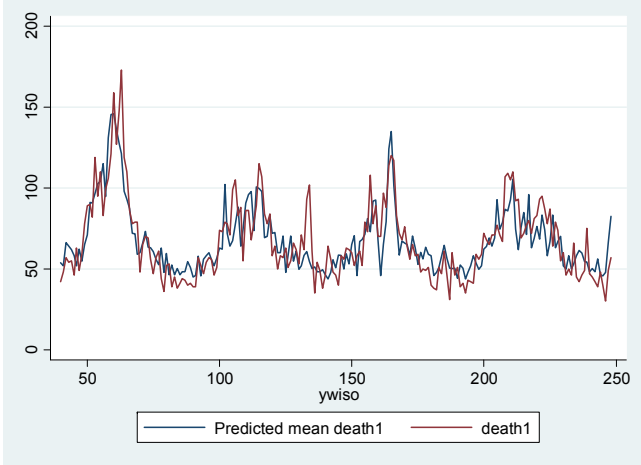


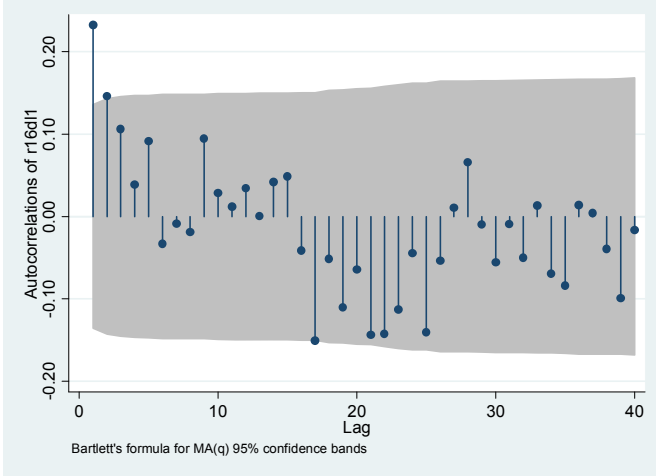
65-74 years





75+ years



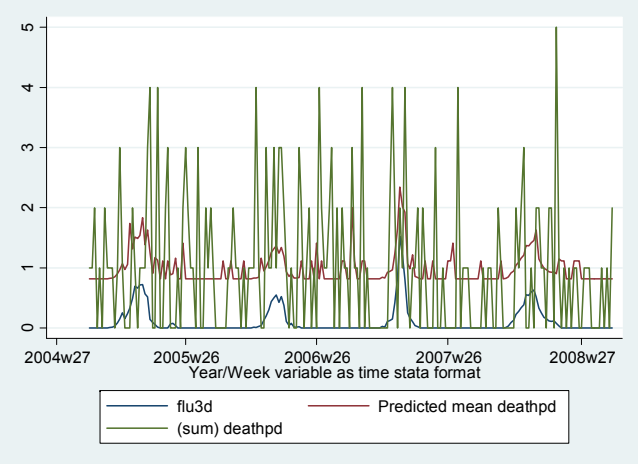


3. ADDITIONAL DATA ON INFLUENZA ATTRIBUTABLE DEATHS

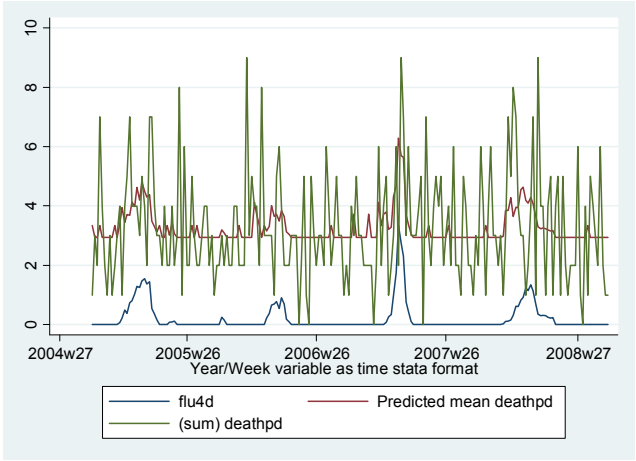
3.1. P+I deaths as main cause

The figures below show the predicted numbers of influenza attributable deaths by week, together with the influenza reported GP cases and the total number of P+I coded deaths.

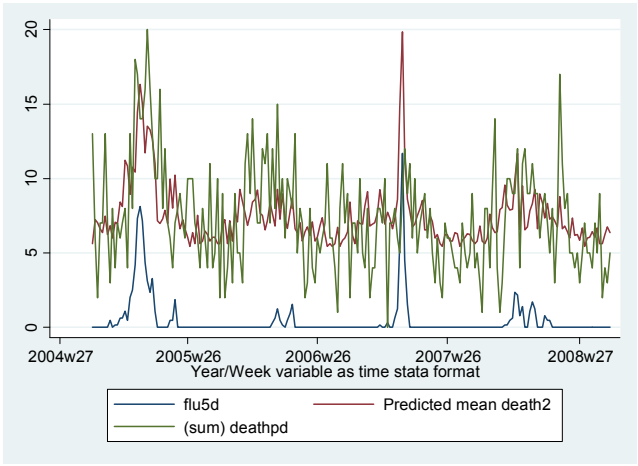
15-49 years



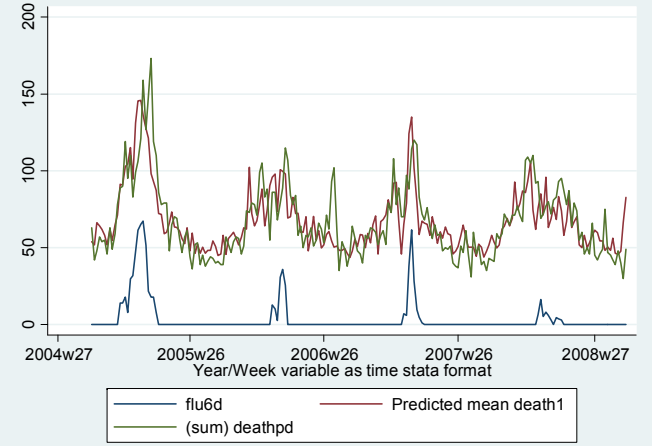
50-64 years



65-74 years



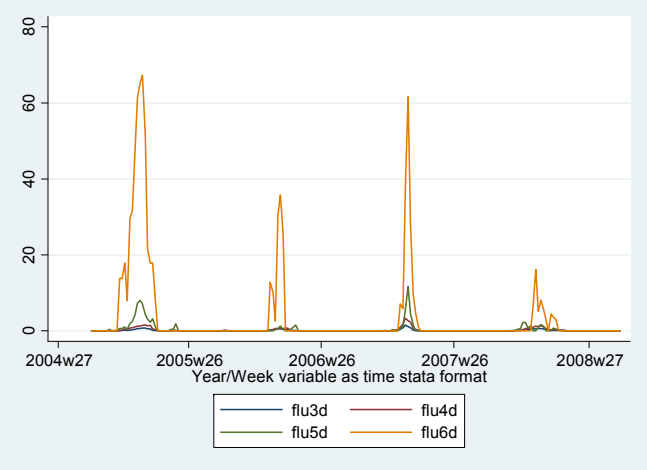
75+



Flu6d: influenza predicted deaths in the 75+ years.

3.2. Attributable influenza deaths by age, P+I as main cause

The figures below show the predicted numbers of influenza attributable deaths by week, and each line represent another age group. As expected, very low numbers of death are predicted in the age groups <65 years.



Flu3d: influenza predicted deaths in the 15-49 years.
Flu4d: influenza predicted deaths in the 50-64 years.
Flu5d: influenza predicted deaths in the 65-74 years.
Flu6d: influenza predicted deaths in the 75+ years.

REFERENCE LIST

1. Newall AT, Wood JG, Macintyre CR. Influenza-related hospitalisation and death in Australians aged 50 years and older. Vaccine. 2008;26(17):2135-41.
2. Pitman RJ, Melegaro A, Gelb D, Siddiqui MR, Gay NJ, Edmunds WJ. Assessing the burden of influenza and other respiratory infections in England and Wales. J Infect. 2007;54(6):530-8.
3. Baltussen RM, Reinders A, Sprenger MJ, Postma MJ, Jager JC, Ament AJ, et al. Estimating influenza-related hospitalization in The Netherlands. Epidemiol Infect. 1998;121(1):129-38.
4. Gilca R, De Serres G, Skowronski D, Boivin G, Buckeridge DL. The need for validation of statistical methods for estimating respiratory virus-attributable hospitalization. Am J Epidemiol. 2009;170(7):925-36.
5. Muller-Pebody B, Crowcroft NS, Zambon MC, Edmunds WJ. Modelling hospital admissions for lower respiratory tract infections in the elderly in England. Epidemiol Infect. 2006;134(6):1150-7.
6. Schanzer DL, Langley JM, Tam TW. Co-morbidities associated with influenza-attributed mortality, 1994-2000, Canada. Vaccine. 2008;26(36):4697-703.
7. Schanzer DL, Tam TW, Langley JM, Winchester BT. Influenza-attributable deaths, Canada 1990-1999. Epidemiol Infect. 2007;135(7):1109-16.
8. Xue Y, Kristiansen IS, de Blasio BF. Modeling the cost of influenza: the impact of missing costs of unreported complications and sick leave. BMC Public Health. 2010;10:724.
9. Poehling KA, Edwards KM, Weinberg GA, Szilagyi P, Staat MA, Iwane MK, et al. The underrecognized burden of influenza in young children. N Engl J Med. 2006;355(1):31-40.
10. Weigl JA, Puppe W, Schmitt HJ. The incidence of influenza-associated hospitalizations in children in Germany. Epidemiol Infect. 2002;129(3):525-33.

11. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, et al. Mortality associated with influenza and respiratory syncytial virus in the United States. JAMA. 2003;289(2):179-86.
12. Thompson WW, Shay DK, Weintraub E, Brammer L, Bridges CB, Cox NJ, et al. Influenza-associated hospitalizations in the United States. JAMA. 2004;292(11):1333-40.
13. Statistical Consulting Group. Analyzing Count Data. In: STATA Library: UCLA: Academic Technology Services. .
14. Yang L, Chiu SS, Chan KP, Chan KH, Wong WH, Peiris JS, et al. Validation of statistical models for estimating hospitalization associated with influenza and other respiratory viruses. PLoS One. 2011;6(3):e17882.
15. van den Wijngaard C, van Asten L, van Pelt W, Nagelkerke NJ, Verheij R, de Neeling AJ, et al. Validation of syndromic surveillance for respiratory pathogen activity. Emerg Infect Dis. 2008;14(6):917-25.
16. Kissling E, Lernout T, Vergison A, Heiden MVd, Hanquet G. Capture-recapture analysis comparing two paediatric studies: Do we need incentives? . In: Proceedings of The 2008 European Scientific Conference on Applied Infectious Disease Epidemiology; 2008 19-21 November 2008; Berlin, Germany; p. 226.
17. Nielsen F. Autocorrelation in time series in STATA. In: Soci709. Chapel Hill: The University of North Carolina; 2006.
18. Cox NJ. Do-it-yourself R-square. In. Durham: STATA; 2003.
19. Zhou H, Thompson WW, Viboud CG, Ringholz CM, Cheng P-Y, Steiner C, et al. Hospitalizations Associated With Influenza and Respiratory Syncytial Virus in the United States, 1993–2008. Clinical Infectious Diseases. 2012;54(10):1427-36.
20. Newall AT, Kelly H, Harsley S, Scuffham PA. Cost effectiveness of influenza vaccination in older adults: a critical review of economic evaluations for the 50- to 64-year age group. Pharmacoeconomics. 2009;27(6):439-50.
21. Fleming DM, ELLIOT AJ. Lessons from 40 years' surveillance of influenza in England and Wales. Epidemiology and Infection. 2008;136(07):866-75.
22. Silvennoinen H, Peltola V, Vainionpaa R, Ruuskanen O, Heikkinen T. Incidence of influenza-related hospitalizations in different age groups of children in Finland: a 16-year study. Pediatric Infectious Disease Journal. 2011;30(2):e24-8.