

EARLY DETECTION OF INFLUENZA-LIKE ILLNESS THROUGH MEDICATION SALES

Maja Sočan¹, Vanja Erčulj², Jaro Lajovic²

¹Centre for Communicable Diseases, National Institute of Public Health, Ljubljana, Slovenia

²Rho Sigma Research & Statistics, Ljubljana, Slovenia

SUMMARY

Monitoring sales of medications is a potential candidate for an early signal of a seasonal influenza epidemic. To test this theory, the data from a traditional, consultation-oriented influenza surveillance system were compared to medication sales and a predictive model was developed. Weekly influenza-like incidence rates from the National Influenza Sentinel Surveillance System were compared to sales of seven groups of medications (nasal decongestants, medicines for sore throat (MST), antitussives, mucolytics, analgo-antipyretics, non-steroidal anti-inflammatory drugs (NSAIDs), betalactam antibiotics, and macrolide antibiotics) to determine the correlation of medication sales with the sentinel surveillance system – and therefore their predictive power. Poisson regression and regression tree approaches were used in the statistical analyses. The fact that NSAIDs do not exhibit any seasonality and that prescription of antibiotics requires a visit to the doctor's office makes the two medication groups inappropriate for predictive purposes. The influenza-like illness (ILI) curve is the best matched by the mucolytics and antitussives sales curves. Distinct seasonality is also observed with MST and decongestants. The model including these four medication groups performed best in prediction of ILI incidence rate using the Poisson regression model. Sales of antitussives proved to be the best single predictive variable for regression tree model. Sales of medication groups included in the model were demonstrated to have a predictive potential for early detection of influenza season. The quantitative information on medication sales proves to be a useful supplementary system, complementing the traditional consultation-oriented surveillance system.

Key words: influenza-like illness, sentinel surveillance system, medication sales, mucolytics, antitussive

Address for correspondence: M. Sočan, Centre for Communicable Diseases, National Institute of Public Health, Trubarjeva 2, 1000 Ljubljana, Slovenia. E-mail: maja.socan@ivz-rs.si

INTRODUCTION

Influenza has an important medical and socio-economic impact. In the northern hemisphere, community-managed illness, outpatient medical visits and hospitalisation rates increase during the cooler months of the year due to intensive circulation of influenza and also other respiratory viruses (1, 2). The most significant contributor to seasonal morbidity and mortality is the influenza virus. Weekly monitoring of consultation rates for influenza-like illness (ILI) in the population, combined with virological data has been the standard influenza surveillance method (3). In principle, influenza surveillance systems should provide timely and accurate information about an emerging influenza epidemic as well as the necessary epidemiological data and virological characteristics of the circulating influenza virus; in particular its match with the virus currently used in the seasonal vaccine (4). The virological surveillance of influenza has developed considerably in the last 60 years. Structured sentinel surveillance started in the mid-sixties (5, 6) and became an established standard for monitoring influenza morbidity in the general population through networks of primary care physicians (7, 8). This approach measures health-care utilisation due to the influenza-like illness. Consultation practice, accessibility and other social factors influence the health care seeking behaviour.

The inherent weakness of such a system is that it does not detect the patients affected by respiratory symptoms who are not ill enough to visit their doctor. It has been previously demonstrated that the number of patients with influenza like illness increases at least a week before the uprising trend has been observed by conventional sentinel surveillance systems (9–12). Therefore, the established sentinel system might not be the best approach to detect the earliest signal of seasonal influenza epidemic (13). Much research has been done to identify new data sources that would serve as an earlier outbreak indicator. Some less conventional methods have been used, like monitoring the number of calls to health help lines and automated analysis of trends in internet searches for influenza related information (14, 15). It has been assumed that a proportion of those who develop a flu-like illness would search for advice on web sites. Researchers correlated the number of clicks on keywords with epidemiological data and found a good correlation with traditional methods of influenza data collecting. The method is extremely rapid (reporting lag of only one day) but misleading spikes of activity appears from time to time e.g. when information about a new influenza virus appeared in media, etc. (12, 16).

The surveillance systems based on automatically collected data are not affected by the above mentioned disadvantages. The data from pharmacies on sales of over-the-counter (OTC) and/

or prescription drugs may be used as a source of timely information about community health (17, 18). The idea is far from being new – Welliver correlated the sales data from pharmacies with laboratory surveillance of influenza virus more than 30 years ago and found that the sales of cold remedies were quite useful for surveillance purposes (19). Later studies yielded variable results regarding the sensitivity, specificity and timeliness of medication sales as an early signal for influenza epidemic or food-borne outbreaks (8, 9, 20).

The present study addressed the usefulness of the data on sales of certain groups of medicines for short-term prediction of the ILI incidence rate detected by a traditional sentinel system.

MATERIALS AND METHODS

Setting and Data Sources

The historical data concerning sales of both OTC and prescription drugs in all forms (tablets, granules and syrups) used to treat common colds and acute respiratory tract infections (including influenza) were provided by the public pharmaceutical institution Gorenjske lekarne. Gorenjske lekarne supplies medicinal products to the inhabitants of 13 municipalities in northern Slovenia (approx. 260,000 people, equivalent to 13% of the total Slovenian population). The institution has been chosen due to its ability to provide data extracted from computer records. The dataset consisted of daily sales from the week 49/2006 – week 39/2009 period.

The following seven groups of medicines were included (ATC codes in parentheses): nasal decongestants (R01AA), medicines for sore throat (MST) (R02), antitussives (R05D), mucolytics (R05C), analgo-antipyretics (N02B), non-steroidal anti-inflammatory drugs (M01), betalactam antibiotics and macrolide antibiotics (J01C, J01D, J01FA).

The weekly ILI incidence rate was obtained from the National Influenza Sentinel Surveillance System (electronic dataset, National Institute of Public Health) for the period extending from the week 40/1999 – week 49/2009. ILI case is defined as any person with sudden onset of fever and with at least one of the following systemic symptoms: malaise, headache or myalgia and at least one of the following respiratory symptoms: sore throat, cough or shortness of breath. The National Sentinel Surveillance System covers the whole country with sentinel sites distributed in all regions. The highest number of reporting out-patient clinics are located in the central region of the country with the highest population density. In addition, there is an intensive daily migration of working population and students to the capital located in the central region. The small size of Slovenia and daily mixing of individuals coming from different parts of the country are reflected in the epidemiological pattern of each seasonal influenza epidemic. The influenza virus spreads through the country very quickly and cases appear practically simultaneously in different parts of the country. This is reflected in the strength of correlation between ILI incidence for Gorenjska region and the whole country (Fig. 1) for the week 49/2006 – week 39/2009 period. Consequently, for the purpose of the present analysis, the national surveillance data were analysed versus the regional sales data (for

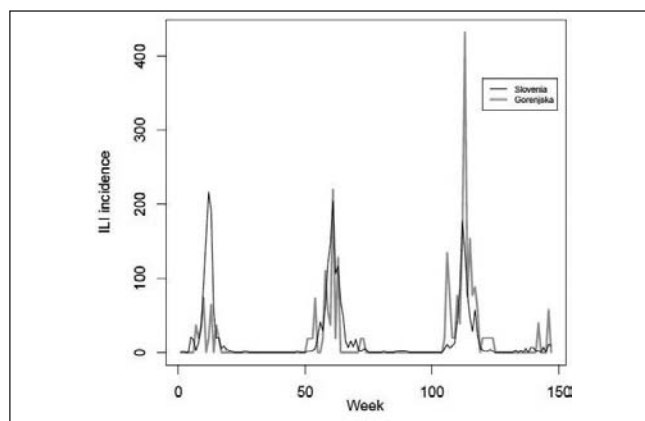


Fig. 1. Comparison of ILI incidence for Gorenjska region and the whole country.

the northern part of the country cared for by Gorenjske lekarne) as it may be reasonably assumed that the sales patterns are the same throughout the country.

Statistical Analysis

The objective of the analysis was to determine whether data on sales of certain groups of medicines (either for each group individually or for several groups in combination) may be used for short-term prediction of the ILI incidence rate and may thus aid (semi-quantitative) detection of critical trends in the incidence rates. Poisson regression and regression tree approaches were used for statistical analyses. Both types of regression models were built based on the ILI incidence rate and the medication sales data for the week 40/2007 – week 39/2009 period, and evaluated using corresponding data for the week 49/2006 – week 39/2007 period.

Time series regression models were also tested. Although ARIMA (0,0,0) (1,1,0) performed very well regarding prediction of the course and magnitude of ILI incidence, it tended to miss the actual increase in ILI incidence for several weeks. Adding medication sales as predictors to time series model did not improve its performance.

Since the sales data for individual groups of medicines were highly correlated (Spearman coefficient >0.90 for all groups) the problem of multicollinearity had to be addressed. The problem was resolved by separately applying each of the following methods: (1) inclusion of a single independent variable (sales of antitussives), (2) inclusion of non-weighted linear combinations of sales data for various groups of medicines (antitussives + mucolytics, antitussives + mucolytics + decongestants, all four groups of medicines [antitussives + mucolytics + decongestants + MST]), (3) inclusion of a new variable (obtained by the principal components analysis) that is a weighted combination of sales data for medicines for sore throat (MST), decongestants and antitussives ($0.91 \cdot \text{MST} + 0.25 \cdot \text{decongestants} + 0.30 \cdot \text{antitussives}$).

Since the possibility of ILI incidence rate prediction was our main point of interest, the Poisson regression modelling was performed using time delay m , i.e. data were time-shifted so as to compare medication sales in Week n with the ILI incidence rate in Week $n + m$. Several (epidemiologically reasonable) time delays were tested; the time delay of 3 weeks proved optimal (prediction

models including 3 weeks time delay performed the best) and was employed in the actual analysis ($m=3$). All analyses were performed using the statistical software package R 2.10 (R 2009); the library part was used for the regression tree analysis (21, 22).

RESULTS

Exploratory data analysis was performed by graphically comparing the similarity of ILI incidence rate time course and medication sales (Fig. 2) as well as of the ILI incidence rate and medication sales of four relevant groups of medicines (Fig. 3) to explore the nature of the relationship between ILI incidence and sales. (Curves on Figure 2 were brought together by a simple transformation (subtraction of a constant) of the sales data to allow for direct graphical comparison and visual assessment of similarity. Therefore, the scale on the ordinate axis is not the same for both curves.)

As expected, the sales of non-steroidal anti-inflammatory drugs (NSAIDs) do not exhibit seasonality and are thus inappropriate for inclusion into the model. Although exhibiting seasonality, antibiotics are also not appropriate: as prescription medicines they require a visit of the doctor prior to being dispensed so they point more towards doctors' than public reactions. Analgo-antipyretics are the third group of medicines inappropriate for modelling: they have a much broader range of indications, not limited to the signs and symptoms associated with influenza/ILI. Among the remaining groups, the ILI curve is the best matched by the sales curves of mucolytics and antitussives; distinct seasonality is also observed with MST and decongestants. Consequently, these four groups were included in the modelling.

Prediction Using the Poisson Regression Model

Association of the ILI incidence rate and the medication sales/dispensing exhibits exponential pattern (Fig. 3). Therefore, Poisson regression was selected for the analysis; this approach is also reported in literature (10). With the aim of developing a model allowing prediction of the ILI time course, various time shifts of data (incidence rate vs. medication sales) were tested in the exploratory phase of the analysis. The 3-week delay (the ILI incidence rate shifted 3 weeks backwards with respect to the medication sales) proved optimal and was employed in the actual analysis, yielding a prediction model for 3 weeks in advance.

The following methods were used to avoid multicollinearity: (1) inclusion of a single independent variable; (2) inclusion of

non-weighted linear combinations of sales of various groups of medicines; (3) inclusion of a new variable obtained by principal components analysis (the new variable proved to explain 98% of the original variables' variance).

The obtained models are summarised in Table 1.

AIC values of all models are very similar and the models can be considered almost equivalent, particularly when taking into account the objective of the present work. However, the model with the new variable and the model with the sales of all four medications performed the best when tested with the 2006–2007 data. The incidence rates predicted by these two models are shown in Figure 4; the actual incidence rate for the period 2006–2007 is superimposed for reference. The performances of the other models were not satisfactory, as the predicted increase of incidence rate lagged behind the actual values by several time points.

Models with (weighted or unweighted) linear combinations of the medication sales perform very well on the test data: the predicted increase is accurate and appropriately matches the actual incidence rate increase. The magnitude of the predicted incidence rate exceeds the actual one but this is irrelevant regarding the objective of this prediction.

Prediction Using the Regression Tree Model

Regression tree analysis was employed as an alternative approach to predictive model development. Figure 5 shows the regression tree obtained using the ILI incidence rate as a dependent variable and the sales of antitussives as an independent variable for the week 40/2007 – week 39/2009 period. Antitussives were selected as the independent variable, as their sales proved to be the optimal predictive variable in this context.

Comparison of the threshold value (478 packages) with the sales of antitussives and the ILI incidence rate indicates that the peak annual ILI incidence rate is achieved within 4–5 weeks after the weekly sales of antitussives exceed the threshold value. The other threshold value (367 packages) represents the first warning sign; with respect to the aforementioned time period it occurs approximately 2 weeks before the first one, i.e. about 6 weeks before the peak ILI incidence rate. From the practical standpoint it is more convenient to express thresholds as relative values; therefore, both threshold values were expressed with respect to the mean antitussives sales during the preceding 52 weeks. Thus it is possible to formulate a rule of thumb: the increase of antitussives sales by 1/3 above the mean antitussives sales during the preceding 52 weeks represents the first warning sign of a possible influenza wave while the increase by (almost) 3/4 predicts the peak incidence in about 4 weeks.

Table 1. Overview of the regression models

Independent variables	Model	AIC
Weighted linear combination of variables (the new variable)	$ILI_{t+3} = \exp(-0.805 + 0.003 \cdot \text{weighted linear combination of the medication sales})$	2248
Antitussives	$ILI_{t+3} = \exp(0.44 + 0.006 \cdot \text{antitussives})$	2248
Cough medicines	$ILI_{t+3} = \exp(0.207 + 0.005 \cdot (\text{antitussives} + \text{mucolytics}))$	2143
Sales of all medicines	$ILI_{t+3} = \exp(-0.954 + 0.002 \cdot \text{all medicines})$	2063
Sales of antitussives, mucolytics and decongestants	$ILI_{t+3} = \exp(-0.686 + 0.004 \cdot (\text{antitussives} + \text{mucolytics} + \text{decongestants}))$	2014

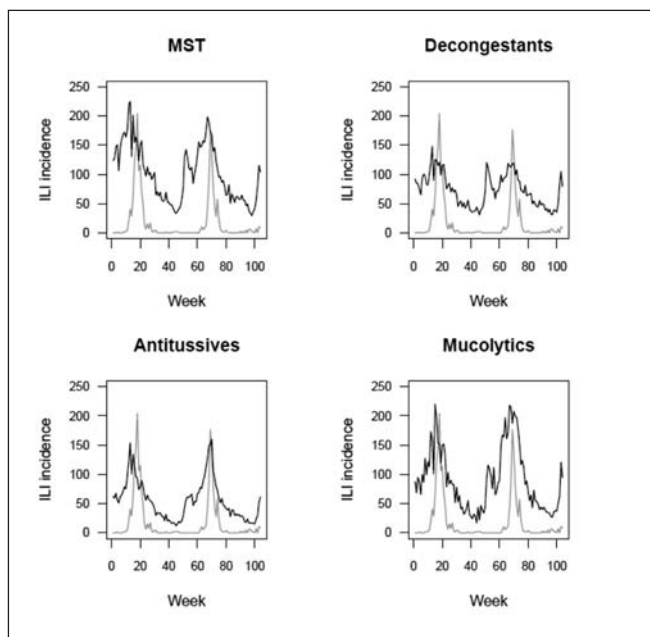


Fig. 2-1. The ILI incidence rate (grey line) and the sales of individual medication groups (black line) in successive weeks of the analysed period (week 40/2007 – week 39/2009). The scale on the ordinate axis is not the same for both curves.

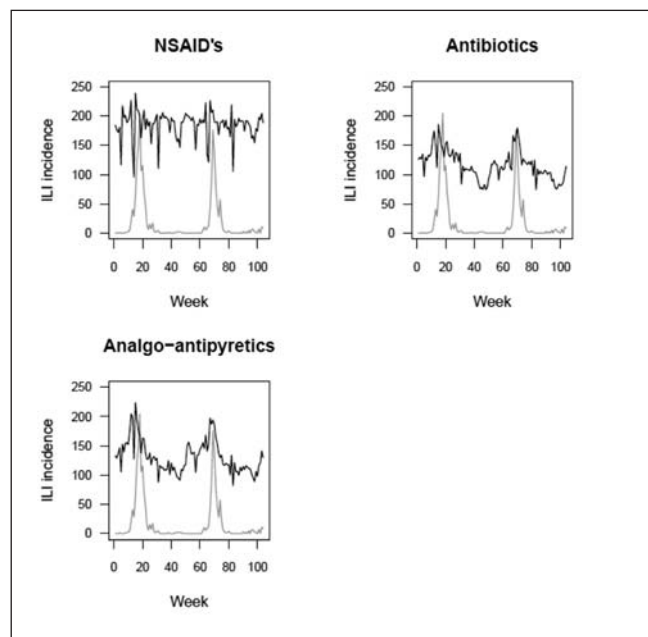


Fig. 2-2. The ILI incidence rate (grey line) and the sales of individual medication groups (black line) in successive weeks of the analysed period (week 40/2007 – week 39/2009). The scale on the ordinate axis is not the same for both curves.

When the rule was evaluated using data for the week 49/2006 – week 39/2007 period (and using the absolute value of 478 because of the lack of sales data for the period before that) it was seen to perform well: the threshold value of 478 was exceeded 6 weeks prior to the ILI peak. The result in this case also included a “false positive” increase 3 weeks before the “correct” signal (Fig. 6).

Variability should be taken into account when interpreting these values and the rule itself, and it should be noted that the relative threshold with respect to the mean preceding 52 week sales may be lower or higher than approx. 70% (rounded to 3/4 in the above rule).

DISCUSSION

The analysis of our data showed that the correlation between the sales of medications used to alleviate influenza symptoms and the ILI incidence differs with respect to medication group. The influenza incidence curve is similar to the sales curves of MST, antitussives, decongestants and mucolytics. The correlation between weekly influenza incidence and these four groups is moderately high. Antibiotics are dispensed on prescription only and are therefore unsuitable for prediction. As shown also by Ganestam, weekly antibiotic prescribing and influenza activity was not correlated – the total quantity of antibiotics dispensed over year remains relatively constant. The antibiotic sales had two distinctive peaks – one earlier in December and one later in March – February (23).

Most syndromic surveillance networks use a standard method whereby trends in consultations for illnesses such as influenza are monitored. However, the majority of people infected by the influenza virus are only mildly ill and would most likely take an over-the-counter medication to alleviate the symptoms and consult the physician if their condition worsens. Therefore, an increase in

influenza-related consultations starts with time lag of one to two weeks with respect to the first symptoms. The studies showed that colds, flu-like symptoms, or diarrhoeal disease are preferentially treated through self-care by taking OTC medications (24, 25). During the outbreak of a *Cryptosporidium* infection, the OTC sales of antidiarrhoeal drugs closely matched the epidemic curve (17). Proctor found the pharmaceutical data were simple to obtain (if computerised), timely and sensitive as such data also include people with mild cases of illness (17). Occasionally, data on drug dispensing revealed much higher burden of an infectious disease than suggested by the routinely collected data (26).

Davies analysed the correlation between local and national OTC sales for flu-like symptoms with hospital admissions over a period of three years. The medication sales exceeded an empirically chosen threshold about two weeks prior to the peak emergency admissions (27). The model was simple to use but did not include other variables, which may affect the OTC sales. A good correlation between medication sales was shown in another study – the model using OTC sales forecasted increase in influenza-like illness in 1–3 week period (10). In contrast, the comparison of OTC cold remedies sales showed low prediction value for increased influenza activity in Japan (20); the inclusion of combined remedies instead of symptom-targeted medications might be an explanation for poor correlation and poor predictive power in the study. The approach to health care differs with respect to social and cultural setting. Easily accessible medical care may encourage earlier visits at the doctor’s office (20). By contrast, in a country with an undeveloped national health insurance system, a majority of patients with relatively mild symptoms would choose self-medication with OTC products.

We developed models for prediction of ILI incidence based on sales data of four groups of medications: MST, antitussives, decongestants and mucolytics as predictors.

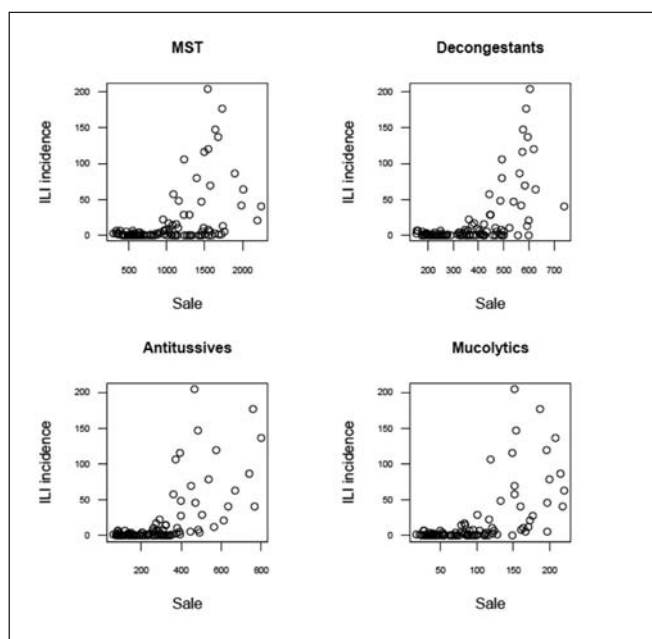


Fig. 3. ILI incidence rate vs. the sales of individual medication groups.

Two types of regression models were developed: the Poisson regression models and the regression tree model. Both types were developed using data for the 40/2007 – week 39/2009 period as the training set and were tested using data for the week 49/2006 – week 39/2007 period as the test set (not used for model building). In this way it was possible to rank models with respect to their predictive performance.

The two best performing Poisson regression models used one of the following as the independent variable: (1) total sales of all four medication groups (MST, decongestants, antitussives, mucolytics) or (2) weighted linear combination of MST, decongestants and antitussives sales (with weights obtained by the principal components analysis). Both provided a very good prediction of both time of the ILI incidence rate increase and the duration of the increase. Although the magnitude of the increase was overestimated by both, this is irrelevant considering the objective of our analysis: the timely identification of a signal indicating the incidence rate increase.

The predictive performance of other Poisson models was unsatisfactory. The “best” among them (with total sales of antitussives, mucolytics and decongestants as independent variables) predicted incidence increase that coincided with the actual incidence increase but the predicted maximum peak was delayed behind the actual one. Other Poisson models (using either antitussives sales or total antitussives and mucolytics sales as the independent variables) performed worst with delays in their predicted incidence increase.

The regression tree model proved to be a rather useful early warning tool. The model permits formulation of a general rule of thumb stating that the peak ILI incidence rate is to be expected within 4–5 weeks after the weekly antitussives sales exceed the mean antitussives sales during the preceding 52 weeks by about 3/4; the early warning sign being the increase of antitussives sales by 1/3 above the mean antitussives sales during the preceding 52 weeks. The early warning sign appears about 2–3 weeks before

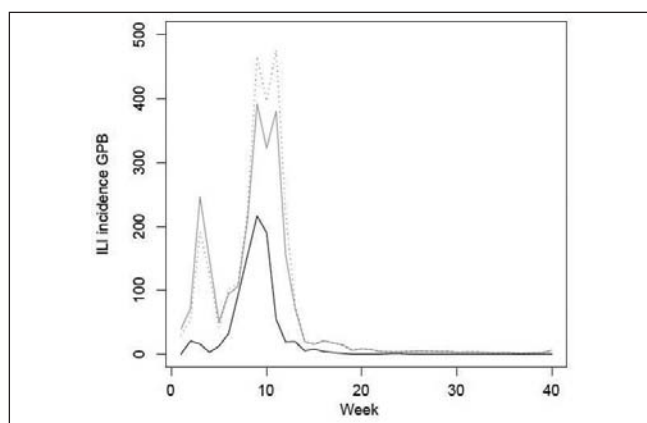


Fig. 4. Actual incidence rates and incidence rates predicted by the best-fitting Poisson regression models.

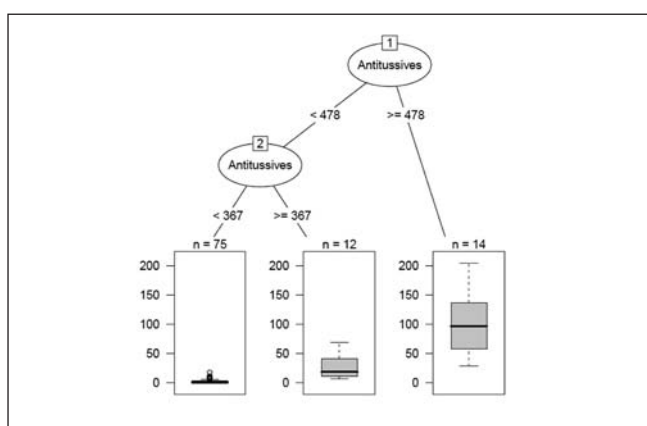


Fig. 5. Regression tree: ILI incidence rate and medication sales (week 40/2007 – week 39/2009).

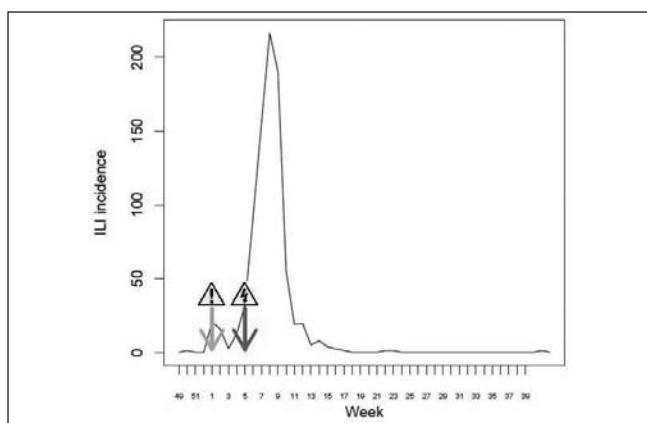


Fig. 6. Prediction of the ILI incidence rate increase using the regression tree model.

the former. Although the finding is of considerable practical interest, one should not disregard the fact that the 3/4 threshold (more accurately: about 70%) may actually be either higher or lower due to the inherent variability of the data. There are different case definitions of influenza and currently there is no gold standard available. The differences in influenza case definitions have an impact on sensitivity and specificity of ILI (28). The findings of

the regression tree model confirmed the high predictive value of cough (specially combined with fever) for influenza as pointed out by Nichol (28).

The predictive utility of the best performing above-mentioned models is very good. Nevertheless, the model built with the 2-year training dataset performed well when tested on the test dataset (not used for the model building). However, the small number of the available/included 52-week periods remains the weak point of the analysis. Therefore, it would be essential to continue data collection and analysis in order to validate the preliminary results; with respect to our findings this would also be epidemiologically reasonable.

Certain limitations of the analyses presented here should also be noted. The medication sales reflect the practice in definite social environments and cannot be generalised (29). The question remains to be answered whether or not the right groups of medications have been chosen. Possibly some remedies (herbal medicines) and food supplements e.g. vitamin C might even better reflect the increasing number of influenza cases in the community. We did not collect data on promotion activities during the study period; however, intensive sales promotion of medicinal products may have an impact on medication purchases. The advice of pharmacist may also influence choice of the medication. Finally, some people purchase antipyretic-analgesics or other medications before going to their holidays and this might also exert impact on OTC sales.

CONCLUSIONS

General conclusions are that data on medication sales proved to be a good predictor of the ILI incidence rate increase. The sales of antitussives alone allow sufficiently accurate predictions of the incidence rate increase, but this requires the antitussives sales to be followed for a longer period (52 weeks). Besides, models based on a single predictor/independent variable are more sensitive to fluctuations of the predictor. For example, in our case a new disease causing increased sales of antitussives would negatively affect the predictive performance of the model. The Poisson regression model with the total sales of antitussives, decongestants, mucolytics and MST as the predictor is not (as) sensible to such influences and also shows a good predictive performance; this is demonstrated by its accurate prediction of the time of incidence rate increase. The model could be further refined to yield more accurate estimates of the magnitude of the incidence rate increase. Considering that both models were built using only 2-year data, their validation and fine tune-in are essential in order to ensure optimal results; this remains the objective of our further work.

Acknowledgements

The research study was founded partly by Slovenian Research Agency (Programme Research in Public Health). We would like to thank Romana Rakovec for providing the data on medication sales from Gorenjske Lekarne.

Conflict of interests

None declared

REFERENCES

1. Pierangeli A, Scagnolari C, Gentile M, Spina MT, Iudicello A, Bertazzoni G, et al. Virological diagnosis of respiratory virus infection in patients attending an emergency department during the influenza season. *Clin Microbiol Infect.* 2010 Apr;16(4):391-3.
2. Dawood FS, Fiore A, Kamimoto L, Bramley A, Reingold A, Gershman K, et al.; Emerging Infections Program Network. Burden of seasonal influenza hospitalization in children, United States, 2003 to 2008. *J Pediatr.* 2010 Nov;157(5):808-14.
3. Fleming DM, van der Velden J, Paget WJ. The evolution of influenza surveillance in Europe and prospects for the next 10 years. *Vaccine.* 2003 May 1;21(16):1749-53.
4. Meijer A, Brown C, Hungnes O, Schweiger B, Valette M, van der Werf S, et al; Virology Task Groups of the European Influenza Surveillance Scheme. Programme of the Community Network of Reference Laboratories for Human Influenza to improve Influenza Surveillance in Europe. *Vaccine.* 2006 Nov 10;24(44-46):6717-23.
5. Fleming DM, Elliot AJ. Lessons from 40 years' surveillance of influenza in England and Wales. *Epidemiol Infect.* 2008 Jul;136(7):866-75.
6. Matlof H, Murray RA, Kamei I, Heidebreder GA. Influenza in Los Angeles County, 1968-69. *HSMHA Health Rep.* 1971 Feb;86(2):183-92.
7. Seligová J, Čulmanová A, Krišťuková Z, Čisláková L, Hudečková H. Changes in surveillance of acute respiratory infections including influenza in the Slovak Republic during 1993-2008. *Cent Eur J Public Health.* 2011 Mar;19(1):20-5.
8. Sočan M, Prosenc K, Nagode M. Differences in clinical predictors of influenza in adults and children with influenza-like illness. *Centr Eur J Med.* 2010;5(1):41-8.
9. Das D, Metzger K, Heffernan R, Balter S, Weiss D, Mostashari F; New York City Department of Health and Mental Hygiene. Monitoring over-the-counter medication sales for early detection of disease outbreaks - New York City. *MMWR Morb Mortal Wkly Rep.* 2005 Aug 26;54 Suppl:41-6.
10. Vergu E, Grais RF, Sarter H, Fagot JP, Lambert B, Valleron AJ, et al. Medication sales and syndromic surveillance, France. *Emerg Infect Dis.* 2006 Mar;12(3):416-21.
11. van Noort SP, Muehlen M, Rebelo de Andrade H, Koppeschaar C, Lima Lourenço JM, Gomes MG. Gripenet: an internet-based system to monitor influenza-like illness uniformly across Europe. *Euro Surveill.* 2007 Jul 1;12(7):E5-6.
12. Valdivia A, Lopez-Alcalde J, Vicente M, Pichiule M, Ruiz M, Ordobas M. Monitoring influenza activity in Europe with Google Flu Trends: comparison with the findings of sentinel physician networks - results for 2009-10. *Euro Surveill.* 2010 Jul 22;15(29): pii: 19621.
13. Thursky K, Cordova SP, Smith D, Kelly H. Working towards a simple case definition for influenza surveillance. *J Clin Virol.* 2003 Jul;27(2):170-9.
14. Cooper DL, Smith G, Baker M, Chinemana F, Verlander N, Gerard E, et al. National symptom surveillance using calls to a telephone health advice service - United Kingdom, December 2001 - February 2003. *MMWR Morb Mortal Wkly Rep.* 2004 Sep 24;53 Suppl:179-83.
15. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. *AMIA Annu Symp Proc.* 2006:244-8.
16. Wilson N, Mason K, Tobias M, Peacey M, Huang QS, Baker M. Interpreting Google flu trends data for pandemic H1N1 influenza: the New Zealand experience. *Euro Surveill.* 2009 Nov 5;14(44): pii: 19386.
17. Proctor ME, Blair KA, Davis JP. Surveillance data for waterborne illness detection: an assessment following a massive waterborne outbreak of *Cryptosporidium* infection. *Epidemiol Infect.* 1998 Feb;120(1):43-54.
18. Hogan WR, Tsui FC, Ivanov O, Gesteland PH, Grannis S, Overhage JM, et al.; Indiana-Pennsylvania-Utah Collaboration. Detection of pediatric respiratory and diarrheal outbreaks from sales of over-the-counter electrolyte products. *J Am Med Inform Assoc.* 2003 Nov-Dec;10(6):555-62.
19. Welliver RC, Cherry JD, Boyer KM, Deseda-Tous JE, Krause PJ, Dudley JP, et al. Sales of nonprescription cold remedies: a unique method of influenza surveillance. *Pediatr Res.* 1979 Sep;13(9):1015-7.
20. Ohkusa Y, Shigematsu M, Taniguchi K, Okabe N. Experimental surveillance using data on sales of over-the-counter medications - Japan, November 2003-April 2004. *MMWR Morb Mortal Wkly Rep.* 2005 Aug 26;54 Suppl:47-52.
21. R Development Core Team. R. The R project for statistical computing [Internet]. Vienna: R Development Core Team; 2009 [cited 2012 May 7]. Available from: <http://www.R-project.org>.
22. Frontline Solvers. XLMiner data mining add-in for Excel [Internet]. Frontline Systems, Inc.; 2010 [cited 2012 May 7]. Available from: http://www.resample.com/xlminer/help/rtree/rtree_intro.htm.

-
23. Ganestam F, Lundborg CS, Grabowska K, Cars O, Linde A. Weekly antibiotic prescribing and influenza activity in Sweden: a study throughout five influenza seasons. *Scand J Infect Dis*. 2003;35(11-12):836-42.
 24. Vingilis E, Brown U, Hennen B. Common colds. Reported patterns of self-care and health care use. *Can Fam Physician*. 1999 Nov;45:2644-6, 2649-52.
 25. Lefterova A, Getov I. Study on consumers' preferences and habits for over-the-counter analgesics use. *Cent Eur J Public Health*. 2004 Mar;12(1):43-5.
 26. Maggini M, Salmaso S, Alegiani SS, Caffari B, Raschetti R. Epidemiological use of drug prescriptions as markers of disease frequency: an Italian experience. *J Clin Epidemiol*. 1991;44(12):1299-307.
 27. Davies GR, Finch RG. Sales of over-the-counter remedies as an early warning system for winter bed crises. *Clin Microbiol Infect*. 2003 Aug;9(8):858-63.
 28. Nichol KL. Heterogeneity of influenza case definitions and implications for interpreting and comparing study results. *Vaccine*. 2006 Nov 10;24(44-46):6726-8.
 29. Magruder SF, Lewis SH, Najmi A, Florio E. Progress in understanding and using over-the-counter pharmaceuticals for syndromic surveillance. *MMWR Morb Mortal Wkly Rep*. 2004 Sep 24;53 Suppl:117-22.

Received October 20, 2011

Accepted in revised form March 2, 2012