

Automated mortality monitoring in Scotland from 2009

A P Wagner (apw40@medschl.cam.ac.uk)^{1,2}, E McKenzie³, C Robertson^{3,4,5}, J McMenamin⁵, A Reynolds⁵, H Murdoch⁵

1. Cambridge Intellectual and Developmental Disabilities Research Group, Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom
2. National Institute for Health Research (NIHR) Collaborations for Leadership in Applied Health Research and Care (CLAHRC) for Cambridgeshire and Peterborough, United Kingdom
3. Department of Mathematics and Statistics, University of Strathclyde, United Kingdom
4. International Prevention Research Institute, Lyon, France
5. Health Protection Scotland, Glasgow, United Kingdom

Citation style for this article:

Wagner AP, McKenzie E, Robertson C, McMenamin J, Reynolds A, Murdoch H. Automated mortality monitoring in Scotland from 2009. *Euro Surveill.* 2013;18(15):pii=20451. Available online: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20451>

Article submitted on 23 April 2012 / published on 11 April 2013

Mortality monitoring systems are important for gauging the effect of influenza and other wide ranging health threats. We present the daily all-cause mortality monitoring system routinely used in Scotland, which differs from others by using two different statistical models for calculating expected mortality. The first model is an extended Serfling model, which captures annual seasonality in mortality using sine and cosine terms, and is frequently seen in other systems. Serfling models fit to summer seasonality well, but not to the winter peak. Thus, during the winter, there are frequent 'excesses', higher than expected mortality, making it harder to directly judge if winter mortality is higher than in previous years. The second model, a Generalised Additive Model, resolves this by allowing a more flexible seasonal pattern that includes the winter peak. Thus, excesses under the second model directly indicate if winter mortality is higher than in previous years, useful, for example, in judging if a new strain of seasonal influenza is more likely to produce death than previous ones. As common in all-cause mortality monitoring systems, the Scottish system uses a reporting delay correction: we discuss the difficulties of interpretation when such a correction is used and possible avenues for future work that may address these difficulties.

Introduction

Mortality surveillance systems are used to monitor for unexpected increases in mortality and are important in monitoring public health. For example, they are used to gauge the effect of influenza [1-3] and to track heat waves and other wide-ranging public health threats [4-7]. Such systems are currently used in: Belgium [8]; England and Wales [9]; Portugal [4]; Sweden [5]; and the United States [1]. A project funded by the European Commission, called the European monitoring of excess mortality for public health action, EuroMOMO, worked to improve the real time monitoring of mortality in Europe [3,6,10]. Among its outcomes was the development of a common consensus system (A-MOMO) for use across Europe, to allow for comparable monitoring between Member States [3,6].

We present the Scottish daily all-cause mortality surveillance system used by Health Protection Scotland (HPS) [11]. This system uses mortality data collated by the National Records of Scotland (NRS) [12] to automatically carry out statistical analysis and produce supporting documentation. Our system differs from others by utilising two statistical models for calculating expected mortality, against which observed levels are compared to detect if mortality is unusually high.

The first model uses sine and cosine terms to model seasonal variation in mortality, extending an earlier model developed by Serfling [13], as is commonly used in other systems [2,4,7]. This model captures summer seasonality well, but does not follow the winter peak closely (Figure 1). Thus, in Serfling models, excess mortality essentially corresponds to mortality above that expected during the summer – during the winter months this is described as 'excess winter mortality' [14]. While influenza may not be the primary driver of excess winter mortality, there is a strong association between them, making excess mortality a useful indicator of influenza [15,16].

The second model uses Generalised Additive Models (GAMs) [17]. GAMs allow for a less restrictive seasonal pattern and so can fit more closely to the winter peak (Figure 1; see [18] for a comparison of approaches to modelling seasonality). Thus, the increase of mortality during the winter is treated as part of the usual seasonal pattern (giving different excess mortality to that of the Serfling model). Consequently, the severity of any seasonal or pandemic influenza (for example, influenza A(H1N1) pdm09 [19]) can more readily be compared with what is usually expected during the winter.

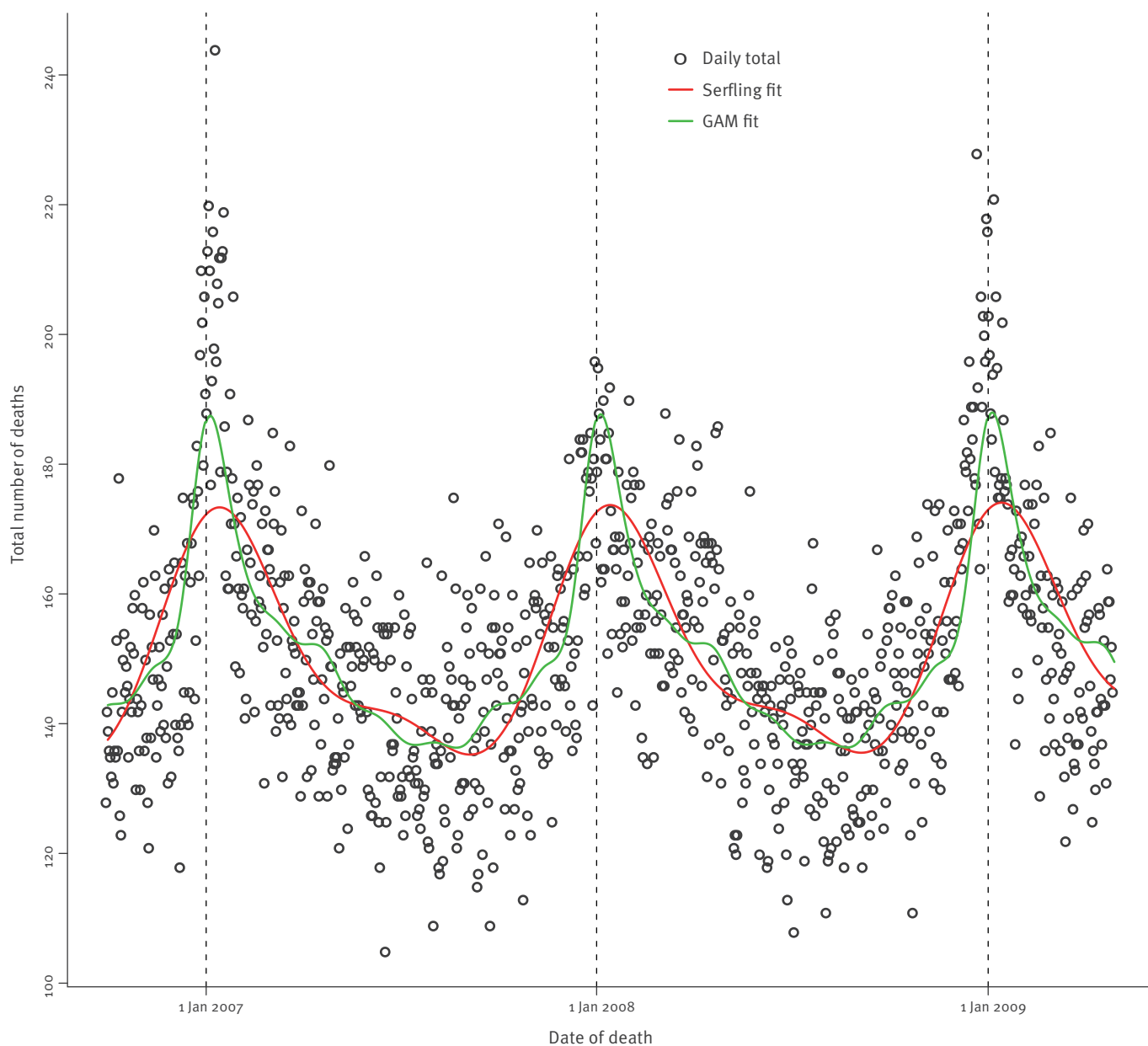
Methods

Collected data

Details of most deaths (≥95%) are transferred electronically to NRS from local Registrars' offices. Data is made available each weekday (Monday to Friday) to HPS on such deaths (currently, data is emailed but

FIGURE 1

The daily totals of deaths occurring in Scotland, 1 October 2006–29 April 2009



GAM: Generalised Additive Model.

The daily totals for all days of each year are shown. The dashed vertical black lines indicate 1 January for each year. The coloured lines give the fits of the two statistical models used to calculate expected numbers of deaths in the mortality surveillance system. The GAMs (green) follow the winter peaks, which occur at the beginning of each year, more closely, while the Serfling model (red) captures summer seasonality (corresponding with those times where the models are at their lowest values, the troughs between the winter peaks) more smoothly and consistently.

TABLE 1

Fit statistics of the Serfling models fitted to the daily totals of observed deaths occurring in Scotland, 1 October 2006–29 April 2009

Model description (unique elements, beyond those described in the caption)	Null deviance	Residual deviance	Explained deviance	% Deviance explained	Null degrees of freedom	Residual degrees of freedom	Used degrees of freedom	AIC
Interaction of age and first seasonal harmonic	105,993	12,277	93,716	88	11,303	11,277	26	54,059
Interaction of $Young_a$ and first seasonal harmonic	105,993	12,280	93,713	88	11,303	11,281	22	54,054
Second seasonal harmonic; second order interactions between both seasonal harmonics and each of age and $Young_a$	105,993	12,150	93,843	89	11,303	11,271	32	53,944
Model defined by Equation (1); as above, but no interaction between second seasonal harmonic and $Young_a$	105,993	12,152	93,841	89	11,303	11,277	26	53,934

AIC: Akaike Information Criterion.

All models include the first seasonal harmonic, factors for age group and sex, and second order interactions between: age group and sex; sex and the seasonal harmonics. Both by analyses of deviance and comparison of AICs (a lower AIC is favoured), the model defined by Equation (1) is the preferred model. It is not appropriate to use a goodness-of-fit test on this model, as it fits to very low counts in the youngest age groups, which violates the large sample assumptions of the test.

development is underway on automating the transfer). This arrangement is possible since all-cause mortality data has not been subject to the validity checking process whereby verification of the cause of death is established and ultimately published a number of months later as an official statistic. For each death, the sex, age at time of death, postcode of last known residence, date of death and date of registration are collected. Postcode sectors are used to link deaths with particular geographical regions in Scotland (typically, health boards, the 14 geographical subdivisions of the Scottish health service).

Data characteristics

On average, there are 152 deaths per day in Scotland, with little difference between days of the week [20]. Mortality reaches a peak around the turn of the year and is at its lowest during the summer. As would be expected in an industrialised country, there are relatively few deaths at young ages. We use the age groupings generally adopted by HPS: 0–14, 15–44, 45–64, 65–74, 75–84 and ≥85 years. There are different seasonalities in mortality among these groups and different levels between the sexes. The biggest differences are seen in the older groups (65–74, 75–84 and ≥85 year-olds), in both seasonal level and pattern.

Developing statistical models for the calculation of expected mortality

We fitted regression models to the daily totals of deaths aggregated by date of death (Figure 1). The following models were developed on data from 1 October 2006 to 14 May 2009, excluding the last two weeks to reduce the effect of unreported deaths.

Serfling model development

Serfling’s model is extended to the following, where a Poisson Generalised Linear Model (GLM) is fitted to the daily mean of observed deaths, μ_{tsa} , for each age group a and sex s :

$$\begin{aligned} \log(\mu_{tsa}) = & \beta_0 + Sex_s + Age.Gp_a + \beta_1 Trend_t \\ & + \beta_2 \sin(2\pi p_t) + \beta_3 \cos(2\pi p_t) \\ & + \beta_4 \sin(4\pi p_t) + \beta_5 \cos(4\pi p_t) \\ & + Sex_s: Age.Gp_a \\ & + Sex_s: \{ \sin(2\pi p_t) \\ & + \cos(2\pi p_t) + \sin(4\pi p_t) + \cos(4\pi p_t) \} \\ & + Young_a: \{ \sin(2\pi p_t) + \cos(2\pi p_t) \} \end{aligned} \quad (1)$$

where the β_k are coefficients; Sex_s is a factor with two levels; $Age.Gp_a$ is a factor with a level for each age group; $Trend_t$ is the number of day t as numbered from the first day ($Trend_t = t$); p_t gives the within-year time, which begins at zero on 1 January and increases

TABLE 2

Comparisons of Akaike Information Criterion between the Generalised Additive Models that adopt different approaches to modelling seasonality (separate models versus use of a factor)

Age in years	Separate GAM for each age and sex – Equation (2)			Sex factor – Equation (3)	Lowest AIC
	Females	Males	Sexes combined		
0–14	1,512.79	1,747.18	3,259.97	3,261.42	Equation (2)
15–44	3,263.70	4,140.83	7,404.53	7,411.06	Equation (2)
45–64	4,727.36	5,233.78	9,961.15	9,952.85	Equation (3)
65–74	5,084.46	5,426.35	10,510.81	10,500.31	Equation (3)
75–84	5,784.34	5,704.85	11,489.19	11,475.05	Equation (3)
≥85	5,944.66	5,211.47	11,156.13	11,158.30	Equation (2)

AIC: Akaike Information Criterion; GAM: Generalised Additive Model.

The GAM defined by Equation (2) has a separate spline for each age group and sex combination, while the GAM defined by Equation (3) uses the same spline for both sexes but utilises a factor to capture differences in seasonal levels. For the models defined by Equation (2), the AICs for each sex, within each age group, are summed ('Sexes combined') to allow for comparison with the models defined by Equation (3).

by 1/365 or 1/366 increments, depending on the numbers of days in that year; *Young_a* is similar to *Age.Gp_a*, except that the three youngest groups (0–14, 15–44 and 45–64) share the same level; *A : B* denotes the interaction of terms *A* and *B*. The inclusion of an interaction term allows two variables to affect the mean in a more complex way than simply additively; for example, *Sex_s : Age.Gp_a* allows a different level for each sex and age combination. The development of this model is outlined in Table 1. The introduction of *Young_a* allows for a parsimonious model while still allowing greater seasonal variation in the older groups.

Essentially, this model is very similar to those used in other mortality monitoring systems: there is a linear trend and seasonality is modelled with sine and cosine terms [3,4,7,9]. It differs from A-MOMO by being fitted to daily data, as in the approach by Cox et al. [7], and using data from the whole year, as described in the works of Cox et al. and Hardelid et al. [7,9]. In contrast with models from other systems, differences between age groups and sex are addressed in one model by using factors (*Age.Gp_a*, *Young_a* and *Sex_s*), while other systems fit separate regression models to appropriate data subsets. A standard Poisson model suffices for the Scottish data [20].

Generalised Additive Models development

Using GAMs, we fitted models that follow the winter peak more closely. A GAM allows a 'spline' to be fitted to data, a process by which the data range is split into separate sections, delineated by a series of 'knot points', within which a simple curve is fitted to the data contained therein (see [21] for an introduction). The Mixed GAM Computation Vehicle with GCV/AIC/REML

smoothness estimation (mgcv) package is used to fit the following GAMs [17].

We first determine if the daily means of deaths μ_t are best modelled with a separate spline for each age group *a* and sex *s* combination (n=12 models):

$$\log(\mu_t) = \beta_0 + \beta_1 Trend_t + f(p_t), \tag{2}$$

or, if the seasonality for both sexes in each age group can be modelled with the same spline, with an additive sex factor to address differences in level (n=6 models):

$$\log(\mu_{ts}) = \beta_0 + \beta_1 Trend_t + Sex_s + f(p_t). \tag{3}$$

In these Poisson GAMs, $f(p)$ is a cyclic cubic regression spline with 52 weekly knots, fitted to the within period time *p_t*. The use of a cyclic regression spline ensures a smooth seasonal pattern. We choose between these models by comparing their Akaike Information Criteria (AICs, [22]) as shown in Table 2. For each age group in the models defined by Equation (2), we sum the AICs of the models for each sex to allow for comparison with the models defined by Equation (3). The GAMs defined by Equation (3) are to be preferred in three age groups (45–64, 65–74 and 75–84 years), while in two groups there is little difference (0–14 and ≥85 years). Thus, we choose models defined by Equation (3).

Next, we investigate more parsimonious models resulting from using fewer knots. We choose from among sets of knots placed at regular intervals (weekly, fortnightly and monthly) and sets where knot locations vary from a higher density around the winter solstice, to a lower one around the summer solstice (Figure 2). The latter

TABLE 3

The Akaike Information Criteria of the Generalised Additive Models defined by Equation (3) with different sets of knot points

Age group in years	Sets of knots considered (n=number of knots)						
	Weekly (52)	Fortnightly (26)	Manually chosen-D (21)	Manually chosen-C (18)	Manually chosen-B (16)	Monthly (12)	Manually chosen-A (11)
0-14	3,261.42	3,261.72	3,263.55	3,263.21	3,258.56	3,258.56	3,258.56
15-44	7,411.06	7,411.07	7,411.08	7,411.13	7,411.14	7,411.12	7,411.22
45-64	9,952.85	9,953.32	9,952.22	9,954.31	9,951.37	9,961.44	9,962.50
65-74	10,500.31	10,503.66	10,503.53	10,501.76	10,506.11	10,506.23	10,506.44
75-84	11,475.05	11,475.76	11,480.02	11,482.72	11,485.77	11,485.98	11,485.85
≥85	11,158.30	11,158.43	11,157.82	11,158.72	11,158.65	11,164.53	11,162.58

Knots placed at regular intervals include those described as weekly, fortnightly and monthly. The position of the manually specified knot sets are shown in Figure 2.

sets of knots are motivated by noting an association between hours of daylight and levels of mortality during the winter [20].

The resulting AICs for fitting the models defined by Equation (3) to each age group and knot set are shown in Table 3. There is little change in AIC over this wide range of models, but the fit of the models in the older groups improves with the addition of more knots. The more clearly defined seasonal patterns in the older age groups, available from the greater number of deaths in these groups, can be more closely modelled by using a greater number of knots [20]. However, note that AICs cannot directly be used for choosing between knot sets. Instead, a pragmatic approach is adopted. For each age group, we choose the set of knots where the AIC levels off and is consistent with the fewest knots. However, to ease interpretation, we also want to use the same set of knots across all age groups. Given these constraints, we choose set B, containing 16 knots (Figure 2).

Reporting delay correction

There is a median reporting delay of one day. Thus, to increase the likelihood of detecting excesses more quickly, the empirical cumulative distributions of delays are used to inflate totals of reported deaths from recent days, which are likely to be underreported because of delayed reporting (similar methods are used by other authors [9,10]). As over 99% of deaths are reported within 14 days, only reporting totals from the last two weeks are corrected. Delays are measured in terms of working days, since negligible totals of deaths are reported at the weekend and on public holidays, when Registrars' offices are closed. Delay distributions are grouped into four day types: weekdays (Monday to Friday); Saturdays; Sundays; and public holidays. An example of these being used to correct national totals

is given in Supplementary Table 1 (<http://preview.tinyurl.com/c6ktyrr>).

The following formula is used to calculate the delay corrected reporting total C_{isat} for age a , sex s , for day $t - i$ ($0 < i < 13$) for use on day t :

$$C_{isat} = \frac{R_{sati}}{D_{d(t-i),g(a,s),w(i)}}$$

where R_{sati} is the total number of deaths occurring on day $t - i$ that have been reported by day t , and $D_{d(t-i),g(a,s),w(i)}$ is the proportion of deaths expected to be reported by $w(i)$ working days after day $t - i$, for the appropriate day type $d(t - i)$ (either weekday, Saturday, Sunday or public holiday), in group $g(a,s)$ (one of: 0-14, 45-64, 65-74, 75-84, ≥85 years age groups, but with separate groups for males and females aged 15-44 years). Currently, the $D_{d(t-i),g(a,s),w(i)}$ are calculated from the delays in the reporting of deaths occurring between 1 October 2006 and 29 April 2009. The groupings of $g(a,s)$ reflect that the delays for most age groups are similar across the sexes, except for the 15-44 year-old group, where reports for males tend to be delayed for longer, as they are frequently subject to autopsies following violent deaths.

Model for monitoring reporting levels

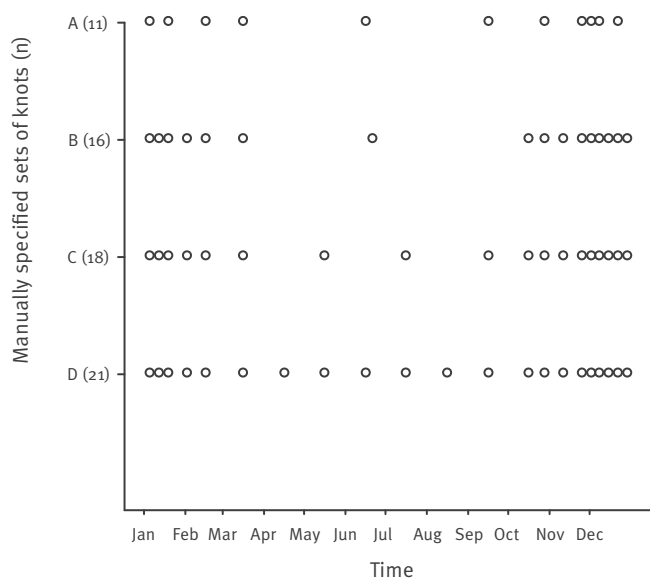
We fitted the following Poisson GAM to the mean daily total μ_{Rt} of reported deaths (that is, daily totals of reports aggregated by date of registration, rather than date of death as is used above):

$$\log(\mu_{Rt}) = \beta_0 + \beta_1 Trend_t + f(p_t) + Day_t + Holiday_t,$$

where: Day_t is a factor with a level for each week day (no deaths can be registered at the weekend); and $Holiday_t$ is a factor indicating if day t is a public holiday, modelling the lower levels of reporting on such days. Factor Day_t models the differences in reporting

FIGURE 2

The manually specified knot sets considered in the Generalised Additive Models



The letters A, B, C, D on the Y axis designate specific sets of knots. The number in parentheses next to each letter specifies the number of knots in the given set. Individual knots are depicted as small circles. In all sets, knots are placed more regularly around the winter solstice and then placed less frequently towards the summer solstice.

levels between days: for example, Mondays tend to have the highest level of reporting, since they are the first opportunity for weekend deaths to be reported. Linked with such deferred reporting, Dayt is set to Monday levels for any day following a public holiday. This model allows reporting levels to be monitored (Figure 3).

General system

The separate elements described thus far are brought together to produce a mortality surveillance system, as outlined in Figure 4. Details of the latest deaths are retrieved from NRS and used to update totals recorded within the system. The delay correction is used to inflate the reported totals for recent days to reduce the effect of delayed reporting. Expected mortality can be calculated from either the Serfling model or GAM, and then compared to the corrected totals. Plots and summaries of the expected and (delay corrected) observed values are made, and excesses determined. When an excess occurs, it can be investigated with the aid of the reporting level monitoring model, to determine whether an excess is genuine, or more likely to be an

artificial product of the delay correction and an unusual reporting pattern.

In the HPS system, an excess occurs when the corrected reporting total for any sex, within any age group, on any day, exceeds the upper limit of the 99% prediction interval for the expected mortality for that group on that day. The size of an excess is defined as the difference between the expected and corrected reporting total. The total excess deaths for any day are found by summing the sizes of any excesses occurring across the twelve age/sex combinations. Obviously, excesses will vary with the model chosen for calculating the expected numbers of deaths, particularly during the winter.

System use and outputs

In preparation for the influenza season, the statistical models are annually re-estimated at the end of September, using data from 2001 to the present period. Predictions of expected levels are then made from October onwards for the coming year. Similarly, from 2012, the reporting delay distributions will also be annually updated (however, they are relatively stable). The system and its data are audited as part of the seasonal influenza review.

Plots produced by the system show daily totals of deaths, ranges of expected values and any excesses, allowing easy interpretation of the current state of mortality (for example, see Figure 3). Recent data, subject to the reporting delay correction, is highlighted to emphasise how totals for this period can fluctuate, as delayed reports continue to accumulate. Besides a graph of national figures ($n=1$ plot), other plots are also produced for each sex ($n=2$), age group ($n=14$) and health board ($n=14$).

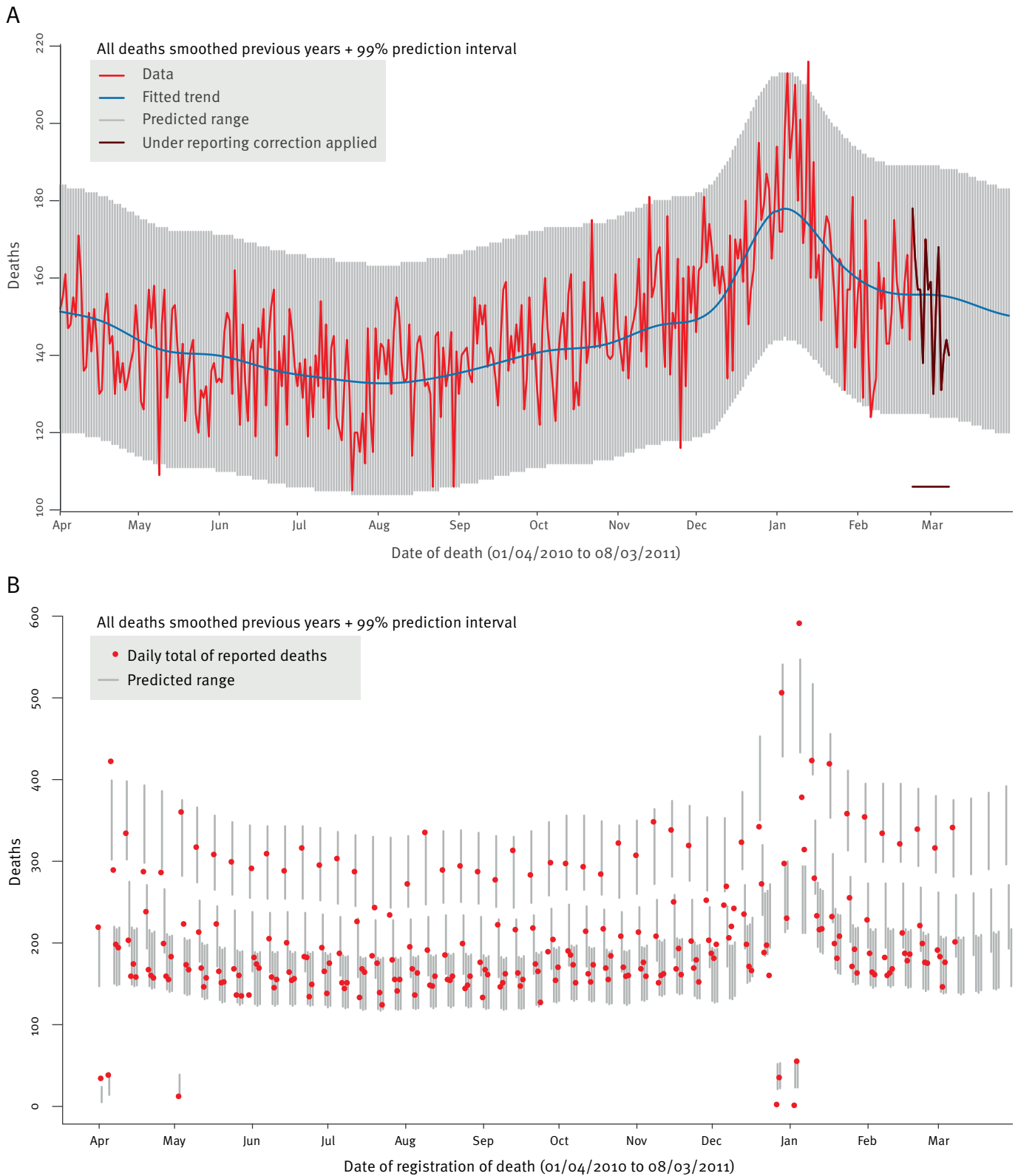
From 2011/12, the system is run daily at HPS (as it was during the influenza A(H1N1)pdm09 pandemic). Output from the system (primarily, the plots described above) is reviewed at least once a week by an epidemiologist on the Respiratory team (the team at HPS who take the lead on linking evidence of excess mortality to virus activity); during periods of intense monitoring, such as during the Olympics, output is reviewed daily (only Monday–Friday, as no new data is available at the weekend). Dependant on the review results, a protocol is followed to bring appropriate results to the attention of the consultant epidemiologist, who then decides if further investigation is warranted and if any excesses should be communicated to the NHS (for more detail see [20,23]).

Results

Output from running the system on a Tuesday is included in the weekly surveillance reports produced during the influenza season. For example, on 20 January 2011, HPS reported that an excess above the usual winter pattern had occurred during 3 to 9 January

FIGURE 3

Example output from the mortality surveillance system, Scotland, 01 April 2010–08 March 2011

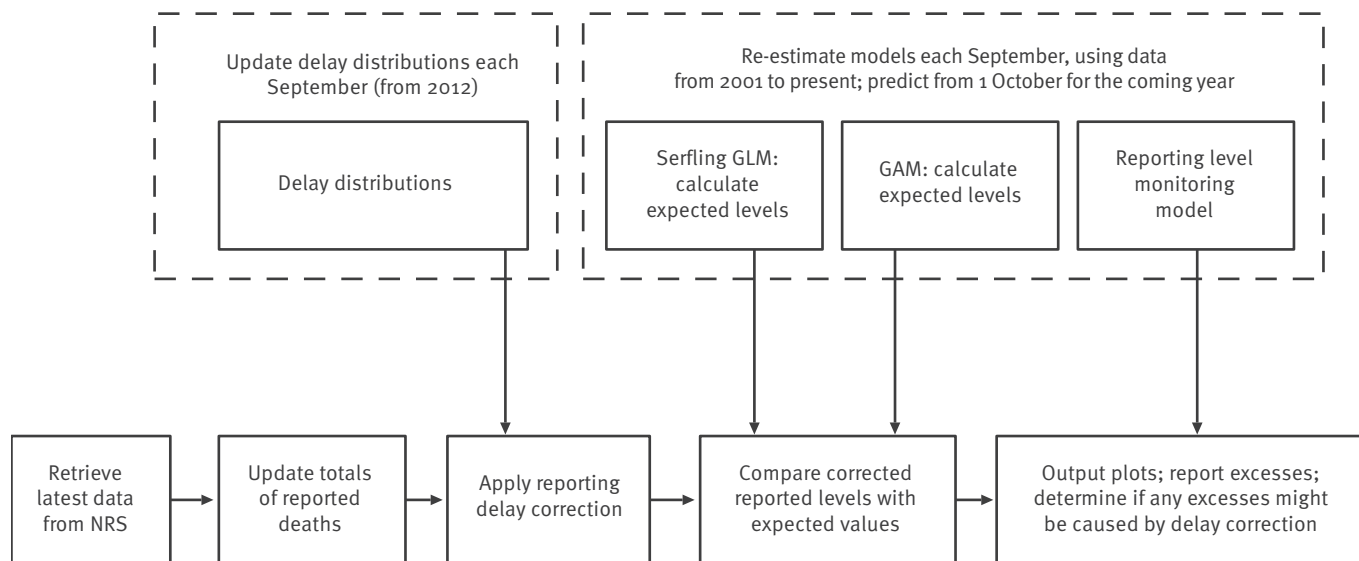


A. Daily national totals of deaths in Scotland, produced on 9 March 2011. The grey lines correspond to the 99% prediction intervals from the Generalised Additive Models; data points outside this, such as from the 3 to the 9 January, correspond to excesses of mortality. The inclusion of the horizontal line (in brown), indicating the period over which the delayed reporting correction is applied, helps to remind users that daily totals near to the present are subject to fluctuation, as delayed reports accumulate.

B. Output from the reporting level monitoring model. Red points are the daily totals of reported deaths and so there are only values for times corresponding to Monday through Friday, the only days on which deaths can be registered. The range of expected reporting levels are shown by the grey lines (99% prediction interval). The higher level of registration generally corresponds to report totals from Mondays.

FIGURE 4

A broad overview of the Scottish mortality surveillance system, from 2009



GAM: Generalised Additive Model; GLM: Generalised Linear Model; NRS: National Records of Scotland.

(shown in Figure 3, see [24] for details); no other excesses were reported in the winter of 2010/11.

Generally, the GAM and its associated output are preferred by the epidemiologists that use the system, due to its more direct and explicit interpretation. For example, the GAM based surveillance was a key element in demonstrating to the Scottish public that influenza A(H1N1)pdm09 had not significantly impacted upon general levels of mortality within Scotland (see, for example, [19]). However, later analysis has shown in other countries that influenza A(H1N1)pdm09 may have had an impact on mortality among the young [3,9]. Increases in the young (0–14 year-olds) are hard to detect, because even proportionately large increases result in only small increases in absolute totals.

The main challenge to using the system arises from the interpretation of the estimated daily deaths during the two week window in which the reporting delay correction is applied. Approximately 20% of deaths that occur on a working day are reported that same day, leading to a multiplication factor of five to convert the number of reported deaths to a ‘delay corrected’ estimate of the number of deaths that have actually occurred. Thus, during the most recent few days, when the delay correction has its biggest effect, there are frequent temporary excesses that disappear with the accumulation of further data on subsequent days. The model

monitoring levels of reporting can assist in determining genuine excesses, but this judgment generally requires a statistician.

As the system has only been running since winter 2009, it is too early to reliably comment on the sensitivity or specificity of the system (however, these are considered as part of the system’s annual review). In any case, calculating sensitivity and specificity for all-cause mortality systems is difficult as the ‘relevant events are hard to define’ and the ‘impact of known events is not certain, while other threats might not yet be known’ [7]. However, to date, the system has not missed anything which has had an impact on deaths, though nothing untoward has occurred. It has detected days and periods with an excessive numbers of deaths and these have been investigated by epidemiologists. Examples from 2012 include 10 and 11 May (173 and 174 deaths respectively) and 22 May (188 deaths); these excesses may be associated with the atypical and very changeable temperatures during May.

Discussion and conclusions

Elsewhere, all-cause mortality surveillance systems that utilise Serfling models have been widely demonstrated to be useful in monitoring seasonal influenza [1,4,7,13,25]. Specifically, in Scotland, we have previously noted the strong association between levels of seasonal influenza and levels of ‘standard’ excess mortality (deaths above the expected level as calculated

from a Serfling model) [14,16]. With the implementation of this system, Scotland now has an automated system for monitoring 'standard' excesses in different age and sex combinations, at both the national and regional level.

The use of two statistical models for calculating expected totals of deaths increases the flexibility and utility of the system. By including Serfling excesses in the system, we have a measure of excess mortality that is more directly comparable to that produced by the systems of other countries. Further, as the Serfling part of the system is similar to EuroMOMO's A-MOMO, barring the differences noted earlier, the Serfling model supports Europe wide monitoring, particularly with those countries that have adopted EuroMOMO methodology [3,6,10].

The Serfling model gives an estimated number of excess deaths relative to a model which predicts a smooth change in deaths from summer to winter. Thus, virtually every year, there will be an excess of deaths from the Serfling model during the sharp winter peak. In contrast, by fitting to the winter peak, the GAM only produces excess deaths when mortality levels around the peak are worse than in previous years. This allows users of the system to more easily and directly detect if seasonal influenza and other factors are significantly increasing mortality above expected winter levels.

The GAMs have separate trends and cyclical components within each age group. It would have been possible to try and develop a Generalised Linear Mixed Model (GLMM) with random effects to control the deviation from the general trend and cyclical effects along the lines of Durban et al. [26]; however, the deviations are not random effects, but rather systematic ones associated with age group – the seasonal cycle becomes progressively more pronounced with increasing age group. Therefore, deviations from the population average are better represented by fixed differences with age, rather than random differences. Effects due to sex could be random but we did not find substantial evidence of differing seasonal components for men and women.

Interpreting output from the system is made more challenging due to the reporting delay correction. However, it is generally accepted that such corrections are needed if mortality surveillance systems are going to detect excesses in a more timely fashion [2,7,9,10]. While the reporting level monitoring model helps determine if an excess is likely to be an artefact of unusual reporting patterns, outputs from the system would be more easily interpretable if they were automatically 'corrected' in some way to take account of the increased uncertainty arising from the use of the delay correction. We have considered some approaches for this, but further development is needed.

One approach would be to 'inflate' the prediction interval that is used to give the predicted range of mortality (the grey lines in Figure 3, panel A), to reflect the increased variability arising from the use of the delay correction. We would expect the adjusted prediction interval to have a funnel shape, being widest in more recent days and then tapering to have the same width as the unadjusted prediction interval, to reflect the diminishing contribution of the correction further away from the most recent data. Ideally, the width of the adjusted prediction interval would be determined by a statistical model that takes into account differences in delays between age groups and day type. This approach may benefit from the delay corrections being modelled by probability distributions, rather than working with the raw empirical values as we have. For example, A-MOMO models delays with a binomial distribution [10]. We have investigated using standard geometric and negative binomial distributions, but fits have been poor [20] and further work is needed.

Little attention has been given to developing statistical methods for addressing reporting delay in mortality monitoring, or other types of prospective surveillance systems [2,27]. For example, the publications by Heisterkamp et al. and Kanieff et al. [28,29] are among the few papers that focus on the topic. Given that delay in mortality reporting is a common issue across Europe, further work in this area would be of wide benefit [29].

Currently, when the regression models are fitted to previous mortality data, every observation is given equal weight, including those corresponding to periods of excess mortality, which may increase expected mortality inappropriately. Ideally, the fit of these models, and consequently the expected levels, should be most heavily influenced by periods with no excesses. There are two main approaches to achieving this. The first of these, the one adopted by A-MOMO and others, is to fit the models to a subset of observations which are unlikely to include excesses of mortality [10]. The second approach refits the models a number of times and weights observations by the reciprocals of earlier fits [7,9]. Observations from periods of excesses should have larger residuals, resulting in smaller reciprocal weights, and so, should have a smaller influence on fit. We would need to adopt the latter approach, as otherwise the GAM might not reliably capture the annual seasonal cycle.

The ability of the system to detect smaller, but sustained, shifts away from expected mortality may be improved through the use of cumulative sums (CUSUMs), as used in other systems [2,9,30].

An enduring criticism of all-cause mortality data is that the cause of the excesses detected may be imputed but not known with certainty. Imputation is usually investigated by considering the temporal, demographic and geographic nature of the data in relation to other data sets. All-cause mortality data may also mask deaths

attributed to rare conditions. Therefore, HPS are considering the use of provisional cause-specific data. Pilot work is looking at accessing deaths where cause is provisionally recorded as influenza related and gauging how timely such a system might be.

This manuscript has presented the deaths surveillance system that has been running in Scotland since 2009. The strengths of the system are: the automatic daily processing of data from NRS; the use of a reporting delay correction; using both a Serfling model (for comparison with other European countries) and a GAM (for easier comparison with what is expected in Scotland based upon the pattern from previous years). The system provides timely information to epidemiologists and can be used on a weekly or daily basis as required.

Acknowledgements

This research was supported by a CASE PhD studentship, funded by the EPSRC and Health Protection Scotland. Results were written up in association with the NIHR CLAHRC for Cambridgeshire and Peterborough.

Funding to enable the creation of the data transfer mechanism between NRS and HPS was provided by Robert Kirkwood (of the Scottish Government Management Team) and the Scottish Government Resilience team.

We would like to thank Frank Dixon, Gordon Munro and Imran Hussain of NRS for their work on providing us with data during the development and early use of the system.

References

1. Simonsen L, Clarke MJ, Stroup DF, Williamson GD, Arden NH, Cox NJ. A method for timely assessment of influenza-associated mortality in the United States. *Epidemiology*. 1997;8(4):390-5. <http://dx.doi.org/10.1097/00001648-199707000-00007> PMID:9209852
2. Unkel S, Farrington CP, Garthwaite PH, Robertson C, Andrews N. Statistical methods for the prospective detection of infectious disease outbreaks: a review. *J R Stat Soc Ser A Stat Soc*. 2012;175(1):49-82. <http://dx.doi.org/10.1111/j.1467-985X.2011.00714.x>
3. Mazick A, Gergonne B, Wuillaume F, Danis K, Vantarakis A, Uphoff H, et al. Higher all-cause mortality in children during autumn 2009 compared with the three previous years: pooled re-sults from eight European countries. *Euro Surveill*. 2010;15(5):pii=19480. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19480> PMID:20144446
4. Nogueira PJ, Machado A, Rodrigues E, Nunes B, Sousa L, Jacinto M, et al. The new automated daily mortality surveillance system in Portugal. *Euro Surveill*. 2010;15(13):pii=19529. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19529> PMID:20394709
5. Sartorius B, Jacobsen H, Törner A, Giesecke J. Description of a new all cause mortality surveillance system in Sweden as a warning system using threshold detection algorithms. *Eur J Epidemiol*. 2006;21(3):181-9. <http://dx.doi.org/10.1007/s10654-005-5923-6> PMID:16547832
6. EuroMOMO – European monitoring of excess mortality for public health action. [Accessed 22 Jul, 2011]. Available from: <http://www.euromomo.eu/>
7. Cox B, Wuillaume F, Van Oyen H, Maes S. Monitoring of all-cause mortality in Belgium (Be-MOMO): a new and automated system for the early detection and quantification of the mortality impact of public health events. *Int J Pub Health*. 2010;55(4):251-9. <http://dx.doi.org/10.1007/s00038-010-0135-6> PMID:20376690
8. Cox B, Wuillaume F, Maes S, Van Oyen H. Belgian Mortality Monitoring (Be-MOMO). Public health and surveillance. Brussels: Scientific Institute of Public Health, Epidemiology Unit. [Accessed 11 Apr 2013]. Available from: <https://www.wiv-isp.be/epidemiology/be-momo/>
9. Hardelid P, Andrews N, Pebody R. Excess mortality monitoring in England and Wales during the influenza A(H1N1) 2009 pandemic. *Epidemiol Infect*. 2011;139(9):1431-9. <http://dx.doi.org/10.1017/S0950268811000410> PMID:21439100
10. Mølbak K. European monitoring of excess mortality for public health action. [Accessed 20 Apr 2012]. Available from: http://ec.europa.eu/eahc/documents/news/technical_meetings/EuroMOMO.pdf
11. Health Protection Scotland (HPS). Health Protection Scotland – Protecting Scotland's Health. [Accessed 20 Apr 2012]. Glasgow: HPS. Available from: <http://www.documents.hps.scot.nhs.uk/about-hps/hps-brochure-2009.pdf>
12. General Register Office for Scotland. [Accessed 11 Apr 2013]. Available from: <http://www.nrscotland.gov.uk/>
13. Serfling RE. Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Rep*. 1963;78(6):494-506. <http://dx.doi.org/10.2307/4591848> PMID:19316455 PMID:1915276
14. General Register Office for Scotland. Increased Winter Mortality – Background. c2011. [Accessed 20 Apr 2012]. Available from: <http://www.gro-scotland.gov.uk/files2/stats/increased-winter-mortality/increased-winter-mortality-background-info.pdf>
15. Healy JD. Excess winter mortality in Europe: a cross-country analysis identifying key risk factors. *J Epidemiol Community Health*. 2003; 57(10):784-9. <http://dx.doi.org/10.1136/jech.57.10.784> PMID:14573581 PMID:1732295
16. Bowie N, Jackson G. The raised incidence of winter deaths. General Register Office for Scotland; 2002. [Accessed 11 Apr 2013]. Available from: <http://www.gro-scotland.gov.uk/files2/stats/increased-winter-mortality/occ-paper-winter-deaths.pdf>
17. Wood SN. Generalized Additive Models: An Introduction with R. London: Chapman & Hall/CRC; 2006.
18. Harvey A, Koopman SJ, Riani M. The Modeling and Seasonal Adjustment of Weekly Observations. *J Bus Econ Stat*. 1997;15(3):354-68.

19. Health Protection Scotland (HPS). Weekly Influenza Situation Report (Including H1N1v). Glasgow: HPS; 3 Feb 2010. Situation Report: 31. [Accessed 20 Apr 2012]. Available from: <http://www.documents.hps.scot.nhs.uk/respiratory/swine-influenza/situation-reports/weekly-influenza-sitrep-2010-02-04.pdf>
20. Wagner AP. Extending Scottish Exception Reporting Systems Spatially and Temporally [PhD thesis]. Glasgow: University of Strathclyde; 2010.
21. Zuur A, Ieno EN, Walker N, Saveliev AA, Smith GM. Mixed Effects Models and Extensions in Ecology with R. New York: Springer; 2009.
<http://dx.doi.org/10.1007/978-0-387-87458-6>
22. Sakamoto Y, Ishiquro M, Kitagawa G. Akaike Information Criterion Statistics. 2nd ed. London: Springer; 1986.
23. Kavanagh K, Robertson C, Murdoch H, Crooks G, McMenamin J. Syndromic surveillance of influenza-like illness in Scotland during the influenza A H1N1v pandemic and beyond. *J R Stat Soc Ser A Stat Soc.* 2012;175(4):939-58.
<http://dx.doi.org/10.1111/j.1467-985X.2012.01025.x>
24. Health Protection Scotland (HPS). Influenza Update Report (Week 2, 2010/11). Glasgow: HPS; 20 Jan 2011. [Accessed 20 Apr 2012]. Available from: <http://www.documents.hps.scot.nhs.uk/respiratory/seasonal-influenza/flu-update/2011-01-20.pdf>
25. World Health Organization (WHO). WHO statement: Comparing deaths from pandemic and seasonal influenza. Geneva: WHO. [Accessed 20 Apr 2012]. Available from: http://www.who.int/csr/disease/swineflu/notes/briefing_20091222/en/index.html
26. Durban M, Harezlak J, Wand MP, Carroll RJ. Simple fitting of subject-specific curves for longitudinal data. *Stat Med.* 2005;24(9):1153-67.
27. Farrington CP, Andrews NJ, Beale AD, Catchpole MA. A statistical algorithm for the early detection of outbreaks of infectious diseases. *J R Stat Soc Ser A Stat Soc.* 1996;159(3):547-63.
<http://dx.doi.org/10.2307/2983331>
28. Heisterkamp SH, Jager JC, Ruitenbergh EJ, van Druuten JA, Downs AM. Correcting reported AIDS incidence: a statistical approach. *Stat Med.* 1989;8(8):963-76.
<http://dx.doi.org/10.1002/sim.4780080807>
PMid:2799125
29. Kanieff M, Rago G, Minelli G, Lamagni T, Sadicova O, Šelb J, et al. The potential for a concerted system for the rapid monitoring of excess mortality throughout Europe. *Euro Surveill.* 2010;15(43):pii=19697. Available from: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19697>
PMid:21087579
30. Höhle M, Paul M. Count data regression charts for the monitoring of surveillance time series. *Comput Stat Data Anal.* 2008; 52(9):4357-68
<http://dx.doi.org/10.1016/j.csda.2008.02.015>