



**HAL**  
open science

## Joint assessment of temporal segmentation, time unit and detection algorithms in syndromic surveillance

Sophie Brilleaud, Benoit Durand, Yann Le Strat, Carole Sala

### ► To cite this version:

Sophie Brilleaud, Benoit Durand, Yann Le Strat, Carole Sala. Joint assessment of temporal segmentation, time unit and detection algorithms in syndromic surveillance. *Preventive Veterinary Medicine*, 2022, 203, pp.105619. 10.1016/j.prevetmed.2022.105619 . anses-03926751

**HAL Id: anses-03926751**

**<https://anses.hal.science/anses-03926751>**

Submitted on 22 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

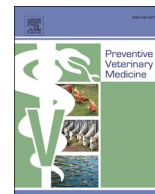
Copyright



Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# Preventive Veterinary Medicine

journal homepage: [www.elsevier.com/locate/prevetmed](http://www.elsevier.com/locate/prevetmed)



## Joint assessment of temporal segmentation, time unit and detection algorithms in syndromic surveillance

Sophie Brilleaud<sup>a,b,c,d,\*</sup>, Benoit Durand<sup>a</sup>, Yann Le Strat<sup>b</sup>, Carole Sala<sup>c</sup>

<sup>a</sup> Epidemiology Unit, French Agency for Food, Environmental and Occupational Health & Safety (ANSES), 14, rue Pierre et Marie Curie, 94706 Maisons-Alfort Cedex, France

<sup>b</sup> Santé publique France, French National Public Health Agency, 12 rue du Val d'Osne, 94415 Saint-Maurice Cedex, France

<sup>c</sup> French Agency for Food, Environmental and Occupational Health & Safety (ANSES), 31, avenue Tony Garnier, 69364 Lyon Cedex 07, France

<sup>d</sup> Paris-Est-Créteil University, France

### ARTICLE INFO

#### Keywords:

Modifiable temporal unit problem  
Syndromic surveillance  
Detection algorithm  
One health

### ABSTRACT

The choice of the aggregation that defines the temporal unit of epidemiological surveillance is part of the more theoretical framework of the modifiable temporal unit problem (MTUP). It has been demonstrated that this choice influences temporal cluster detection and may lead to false-positive results and poor estimation of regression model parameters. In syndromic surveillance (SyS), despite the choice of which temporal aggregation to use being crucial, it has not yet been addressed in the literature. In most SyS systems, this choice is driven by the frequency of the data collection and/or human resources available, although neither the temporal unit's influence on the performance of anomaly detection algorithms nor on the efficiency of the SyS are known. The main objective of our study was to analyze the influence of the temporal aggregation unit on the performances of SyS detection algorithms used routinely, according to the characteristics of specific syndromes and outbreaks. Simulating daily time series of various syndromes, we tested three different time series aggregation methods. For each of four anomaly detection algorithms and their variants, we calculated seven performance indicators and multi-criteria scores to guide epidemiologists in their choice of which temporal aggregation of surveillance to use. From 19,200 analyzed time series, we observed an effect of temporal aggregation on the performance of the detection algorithms tested. Results also showed that the time aggregation unit was linked to the detection algorithm used, and that strong aggregation-algorithm interactions need to be taken into account when deciding on which aggregation-algorithm pair to use. Using theoretical data, our study also showed that no one ideal aggregation-algorithm pair exists for all contexts when deciding on which temporal unit of surveillance to use, and that the choice depends on several parameters. Our results can help public health practitioners choose the most appropriate time series aggregation and algorithm according to their specific needs. Finally, the present work enabled us to develop recommendations for a One Health project where the same time aggregation type and detection method could be used for both human and animal syndromic surveillance data.

### 1. Introduction

Potential threats to human and veterinary public health are usually monitored by specific surveillance systems. Traditionally, these systems have been based on clinical case reports and programmed screening surveillance plans. Since the 1990s, syndromic surveillance (SyS) – which initially focused on the detection of bioterrorism events (Paterson and Durrheim, 2013) – has been added to this list of systems, in order to

improve the detection of unexpected/emergent health events (e.g., climatic events like heat waves), and to monitor endemic and seasonal diseases (e.g., gastroenteritis, influenza) (Henning, 2004). SyS is defined as the real-time or near real-time collection, analysis, interpretation and dissemination of health-related data to enable early identification of the impact (or absence of impact) of potential human or veterinary public health threats that require effective public health action (Henning, 2004; Triple, 2011). It is mostly based on data initially collected for

\* Corresponding author at: Epidemiology Unit, French Agency for Food, Environmental and Occupational Health & Safety (ANSES), 14, rue Pierre et Marie Curie, 94706 Maisons-Alfort Cedex, France.

E-mail addresses: [sophie.brilleaud@santepubliquefrance.fr](mailto:sophie.brilleaud@santepubliquefrance.fr) (S. Brilleaud), [benoit.durand@anses.fr](mailto:benoit.durand@anses.fr) (B. Durand), [yann.lestrat@santepubliquefrance.fr](mailto:yann.lestrat@santepubliquefrance.fr) (Y. Le Strat), [carolesala@yahoo.fr](mailto:carolesala@yahoo.fr) (C. Sala).

<https://doi.org/10.1016/j.prevetmed.2022.105619>

Received 29 July 2021; Received in revised form 15 March 2022; Accepted 17 March 2022

Available online 21 March 2022

0167-5877/© 2022 Elsevier B.V. All rights reserved.

purposes other than surveillance (Hulth et al., 2009; Brouwer et al., 2017). Many different SyS systems (SySS) have been implemented in several countries, in human (Caserio-Schönemann and Meynard, 2015; Heffernan et al., 2004; Salmon et al., 2016; Yan et al., 2013; Wu et al., 2008) and animal health (Sala et al., 2020; Burkom et al., 2019; Dórea et al., 2013) contexts. For example, in France, morbidity data recorded by hospital emergency departments for care management purposes, and mortality data recorded by the National Institute of Statistics and Economic Studies as a demographic indicator, are centralized daily and analyzed by the SySS SurSauD (Health surveillance of emergencies and deaths), which is managed by the National Public Health Agency (Santé publique France) (Caserio-Schönemann and Meynard, 2015). SurSauD was developed after the unexpectedly high excess human mortality due to the 2003 heatwave in France. More recently, OMAR, the French SySS for bovine mortality, analyses the weekly decrease in stock numbers at a national level (Sala et al., 2020). These two SySS were developed independently and are not interconnected. Using a One Health approach to connect these two human and animal SySS may not only increase detection of unusual events which affect both humans and animals (e.g., events originating from a shared environment), but may also help prevent the outbreaks of zoonotic diseases like West-Nile disease, salmonellosis and Q fever. However, combining the findings of the two systems is difficult because they do not use the same spatio-temporal resolution. To simplify, “temporal aggregation unit” is considered here as the method we used to “cut” the time series in this paper. While statistical methods for the temporal detection of anomalies (called “unusual events” by other authors) (Robertson et al., 2010), (Unkel et al., 2012; Noufaily et al., 2012; Bédubourg and Le Strat, 2017; Faverjon and Berezowski, 2018; Noufaily et al., 2019) have been extensively documented, the choice of the temporal unit of surveillance has not yet been addressed in the literature, despite it being a crucial element for SySS. In most SySS, this choice is driven by the frequency of data collection and/or the human resources available to interpret the results. However, the influence of the temporal aggregation unit is not generally analysed for time series monitored for SySS. The choice of which temporal unit to use is part of the more theoretical framework of the modifiable temporal unit problem (MTUP) raised by (Çöltekin et al., 2011) and described by (Cheng and Adepeju, 2014). The MTUP is the result of the influence which the choice of temporal unit has on statistical results when analyzing time series. (Cheng and Adepeju, 2014) propose considering the MTUP as having three components as follows: (i) temporal aggregation, which corresponds to the temporal unit (hour, day, week, etc.), (ii) temporal segmentation, which corresponds to the start of each temporal unit (for example, if the chosen time unit is the week, it could start on Sunday or Monday) and (iii) the boundary effect, which corresponds to the temporal frame of a time series: when does it start and when does it end? It has been shown that these three effects may lead to false-positive results, and to over-estimations or biased estimations of regression model parameters (Cheng and Adepeju, 2014; de Jong and de Bruin, 2012). To test the sensitivity of the effect of the MTUP, Pereira analyzed the number of jobs accessible to the general population when using public transportation according to different travel durations (Pereira, 2019). He found that a boundary effect existed which could lead to partial conclusions and misleading results. Furthermore, (de Jong and de Bruin, 2012) analyzed the effect of the MTUP on simulated time series of vegetation indices and showed variations in the regression results when analyzing vegetation cycles. To date, the MTUP has never been considered in SyS and we do not know what influence it has on the anomaly detection methods commonly used in SySS. The objective of our work was to explore the influence of the temporal aggregation unit on the performances of anomaly detection algorithms used in SyS, according to the algorithm used and the characteristics of the time series and health events in question specifically disease outbreaks. Using various simulated time series scenarii, we evaluated the performances of five families of algorithms according to three types of temporal aggregation. For realistic purposes, we used the simulation tool developed by

(Noufaily et al., 2019), which is based on real-world data collected by Public Health England (PHE). Their program helps create time series to mimic a wide range of infectious diseases (diarrhea, pneumonia, influenza-like-illnesses, etc.), non-infectious diseases (cardiac syndrome, heat stroke, etc.) and vector-borne diseases (from arthropod or insect bites). Moreover, the 16 scenarii proposed by Noufaily et al. combine different incidence levels (from 4 to 1000 daily cases), types of seasonality (annual, biannual, quarterly and no seasonality), a day-of-the-week effect (one or two week peaks), and trends (from no trend to a marked trend). We hypothesized that this diversity probably covers the majority of syndromes followed by SySS in human and animal health.

## 2. Material and methods

### 2.1. Time series simulation

We simulated baselines of daily counts (i.e., time series without outbreaks, representing routine variations) using a negative binomial distribution of mean  $\mu$  and variance  $\phi\mu$ , with a dispersion parameter  $\phi \geq 1$ . For each day  $t$ , the mean  $\mu(t)$  was obtained from the following model (Noufaily et al., 2019):

$$\begin{aligned} \mu(t) &= \exp \left\{ \theta + \beta(t+s) + \sum_{j=1}^{k_1} \left\{ \gamma_1 \cos\left(\frac{2\pi j(t+s)}{52 \times d}\right) + \gamma_2 \sin\left(\frac{2\pi j(t+s)}{52 \times d}\right) \right\} \right\} \\ &+ \sum_{j=1}^{k_2} \left\{ \gamma_3 \cos\left(\frac{2\pi j(t+s)}{d}\right) + \gamma_4 \sin\left(\frac{2\pi j(t+s)}{d}\right) \right\}, \end{aligned} \quad (1)$$

where  $\theta$  is the baseline frequency reports,  $\beta$  is the linear trend,  $d$  the day-week recording system ( $d = 5$  or  $7$ ),  $k_1$  is the seasonality ( $k_1 = 0$  corresponding to no seasonality,  $k_1 = 1$  and  $k_1 = 2$  to annual and biannual seasonality, respectively, and  $k_1 = 4$  to quarterly seasonality),  $k_2$  is the weekly pattern ( $k_2 = 0$  no specific pattern,  $k_2 = 1$  and  $k_2 = 2$  to one and two weekly peaks respectively),  $\gamma_1$  and  $\gamma_2$  are the seasonal effects,  $\gamma_3$  and  $\gamma_4$  represent different intensities of the day-of-the-week effect. The values of the parameters for each scenario are available in Table 1 in (Noufaily et al., 2019). We generated 100 randomly simulated 7-year (2548 days) time series for each scenario. Fig. 1 illustrates the diversity of the simulated baselines.

We added an outbreak event to each simulated time series, in the last

**Table 1**

Default parameters for the detection algorithms and the minimum length of historical period needed. All the algorithms are implemented in the R package surveillance (Höhle, 2007).

Detection algorithm	Default parameters	Minimum historical period
Farrington flexible	$w = 3, b = 5, \alpha = 0.01,$ trend=TRUE	5 years
EARS NB	c.ARL = 5	1 year
EARS C1	$\alpha = 0.01,$ method= “C1”	7 days/weeks
EARS C2	$\alpha = 0.01,$ method= “C2”	9 days/weeks
EARS C3	$\alpha = 0.01,$ method= “C3”	9 days/weeks
RKI 1a	$w = 6, “b=0”, \alpha = 0.01,$ actY=TRUE	6 weeks/42 days
RKI 2a	$w = 6, “b=1”, \alpha = 0.01,$ actY=TRUE	1 year
RKI 3a	$w = 4, “b=2”, \alpha = 0.01,$ actY=FALSE	2 years
Bayes 1a	$w = 6, “b=0”, \alpha = 0.05,$ actY=TRUE	6 weeks/42 days
Bayes 2a	$w = 6, “b=1”, \alpha = 0.05,$ actY=TRUE	1 year
Bayes 3a	$w = 4, “b=2”, \alpha = 0.05,$ actY=FALSE	2 years

<sup>a</sup> For RKI 1–3 and Bayes 1–3, we used  $w = 42$  for daily aggregation.

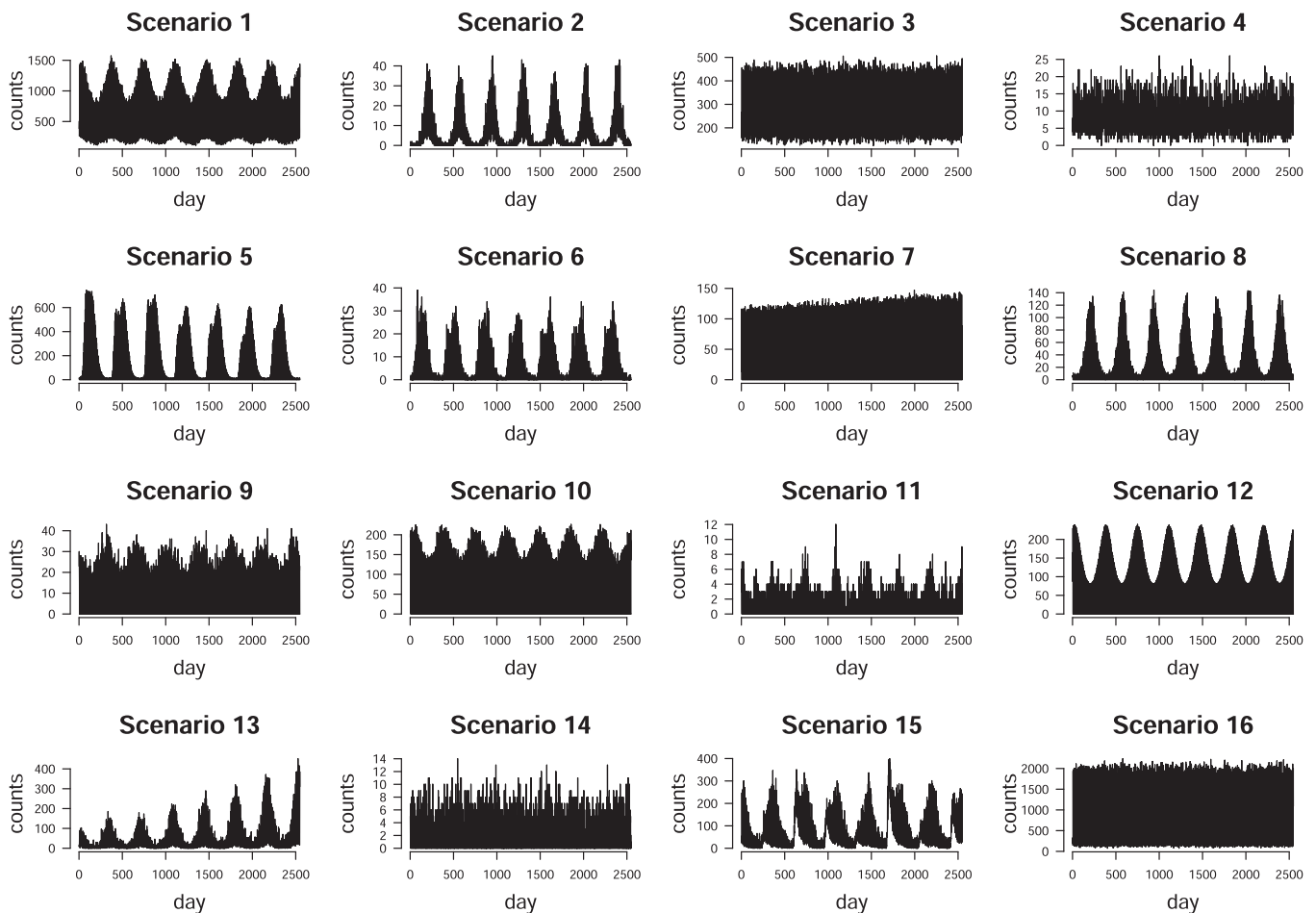


Fig. 1. Representation of one time series among the 100 simulated series for each scenario, in number of cases per day.

49 weeks. The size of each outbreak ( $m$ ) was generated by a Poisson distribution with a mean equal to  $m \times \sqrt{\mu_{(t)}\phi}$ ,  $t$  being the day when the outbreak started (randomly selected among the 49 last weeks of the time series). Four sizes of outbreak were simulated: very small ( $m = 2$ ), small ( $m = 3$ ), medium ( $m = 5$ ) and large ( $m = 10$ ). The mean duration of an outbreak depended on its size: from 10.5 days for very small outbreaks to 15 days for large ones. In order to create a real-world context, a supplementary day-of-the-week effect was added to each time series (Fig. 2). For the 7-day-week scenario (*i.e.*, when  $d = 7$ ), we doubled the number of cases during weekends, and for the 5-day-week scenario the number of daily cases was multiplied by 1.5 on Mondays and Fridays and left equal to zero during week-ends. Additionally, the simulations took into account the effect of public holidays similar to the 'weekend' effect, just mentioned: the number of cases were doubled during public holidays in the 7-day-week, scenario and multiplied by 1.5 before and after public holidays in the 5-day-week scenario.

To ensure comparability of the simulated data with real-world syndromic surveillance data, we corrected time series of scenarios 5, 7 and 12 by dividing the daily simulated number of cases (in the baseline and in the outbreak) by 10 or 100 depending on how unrealistically high they were compared with what was observed by PHE.

The time series were simulated using R 3.6.3 (R Core Team, 2021). The R code, largely based on that developed by (Noufaily et al., 2019), is available on GitHub at: [https://github.com/SophieBrd/SyS\\_data\\_simulation\\_MTUP\\_study](https://github.com/SophieBrd/SyS_data_simulation_MTUP_study).

## 2.2. Statistical analysis

### 2.2.1. Temporal aggregation units

We compared three types of temporal aggregation units corresponding to two surveillance frequencies:

- no aggregation, corresponding to a daily analysis of the gross daily count (Figs. 3a and 3b)
- a 7-day moving aggregation (daily moving segmentation) with a daily analysis of the sum of 7 days (*i.e.*, the 6 previous days plus the current day), so the value at day  $t$  corresponded to  $t$ :  $M_t = \sum_{i=t-6}^t x_i$  (where  $x_i$  is the number of cases at day  $i$ ), see Figs. 3c and 3d
- a weekly aggregation with a Monday-Sunday aggregation and an analysis each Monday, corresponding to the sum of seven daily counts from the preceding week (Figs. 3e and 3f).

### 2.2.2. Statistical detection algorithms

We tested five families of algorithms, with different tuning parameters, leading to 11 detection algorithms, all available in the R package 'surveillance' (Höhle, 2007):

- Farrington Flexible: expected values are estimated from a 5-year-long historical period, using the following quasi-Poisson regression model:

$$\log(\mu_i) = \theta + \beta t_i + \delta_{j(t_i)} \quad (2)$$

where  $j(t_i)$  is the seasonality parameters of the time unit  $t_i$  with  $j(t_0)$

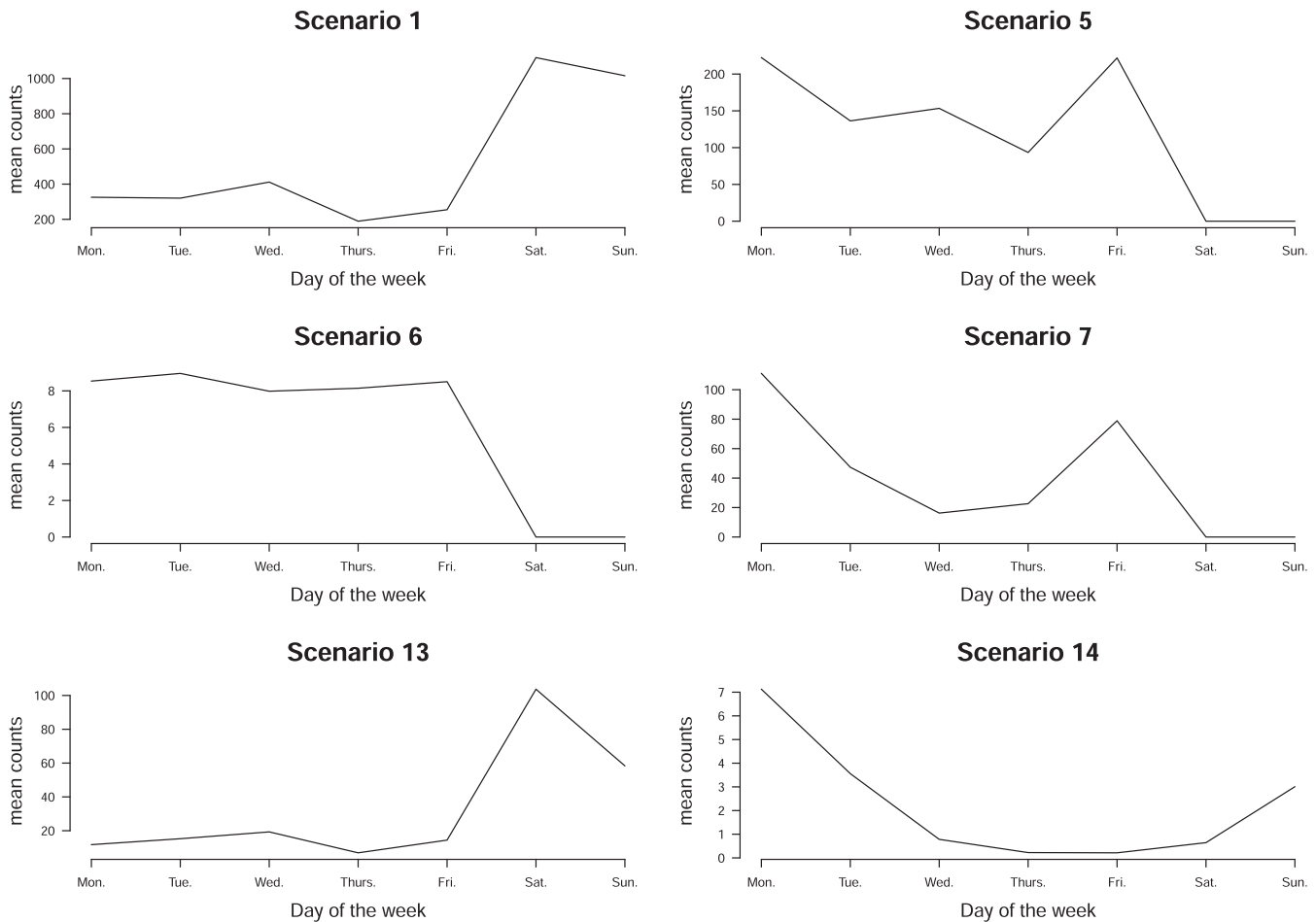


Fig. 2. From scenario 1, 5, 6, 7, 13 and 14, the daily mean counts were calculated for every week of the 100 simulations of each scenario.

= 0 and  $\delta_0 = 0$ . In this model a trend is always adjusted, even if it is not statistically significant. Then an excess score is calculated. An alarm is raised when the excess score  $X = \frac{y_0 - \hat{\mu}_0}{U - \mu_0}$  is greater or equal to

1, with  $y_0$  being the observed value,  $\hat{\mu}_0$  the expected value, and U the  $100(1 - \alpha)$  quantile of the normal distribution. For the present study, we chose  $\alpha = 0.01$ . To reduce the effect of past outbreaks in the baseline, a weight was assigned to each value of the baseline. The Farrington Flexible algorithm differs from the original Farrington algorithm in the way these weights are calculated (Noufaily et al., 2012; Farrington et al., 1996)

- Early aberration reporting system (EARS) C1, C2, C3: EARS is an anomaly detection method used when limited baseline data are available (Hutwagner et al., 2003). The search for an anomaly uses a predicted value, which depends on the EARS parameterization. We used 3 parameterizations: C1-Mild, C2-Medium and C3-Ultra (from C1 the least sensitive, to C3 the most sensitive). Under the null hypothesis (*i.e.*, when there is no outbreak), the three random variables C1, C2 and C3 (defined below) follow a normal distribution of mean 0 and standard deviation 1. An alarm is raised at time  $t$  when  $C1_t$ ,  $C2_t$  and  $C3_t$  are greater than a threshold value corresponding to the  $(1 - \alpha)$  quantile of a normal distribution of mean 0 and standard deviation 1. C1 is obtained with the following formula:

$$C1_t = \frac{x_t - \bar{x}_{1t}}{S1_t} \tag{3}$$

with  $x_t$  corresponding to the observed value of day  $t$ ,  $\bar{x}_{1t} = \frac{1}{7} \sum_{i=t-1}^{t-7} x_i$

the mean counts of the baseline and  $S1_t^2 = \frac{1}{6} \sum_{i=t-1}^{t-7} (x_i - \bar{x}_{1t})^2$  is the variance. C2 is calculated as:

$$C2_t = \frac{x_t - \bar{x}_{2t}}{S2_t} \tag{4}$$

with  $x_t$  corresponding to the observed value of day  $t$ ,  $\bar{x}_{2t} = \frac{1}{7} \sum_{i=t-3}^{t-9} x_i$

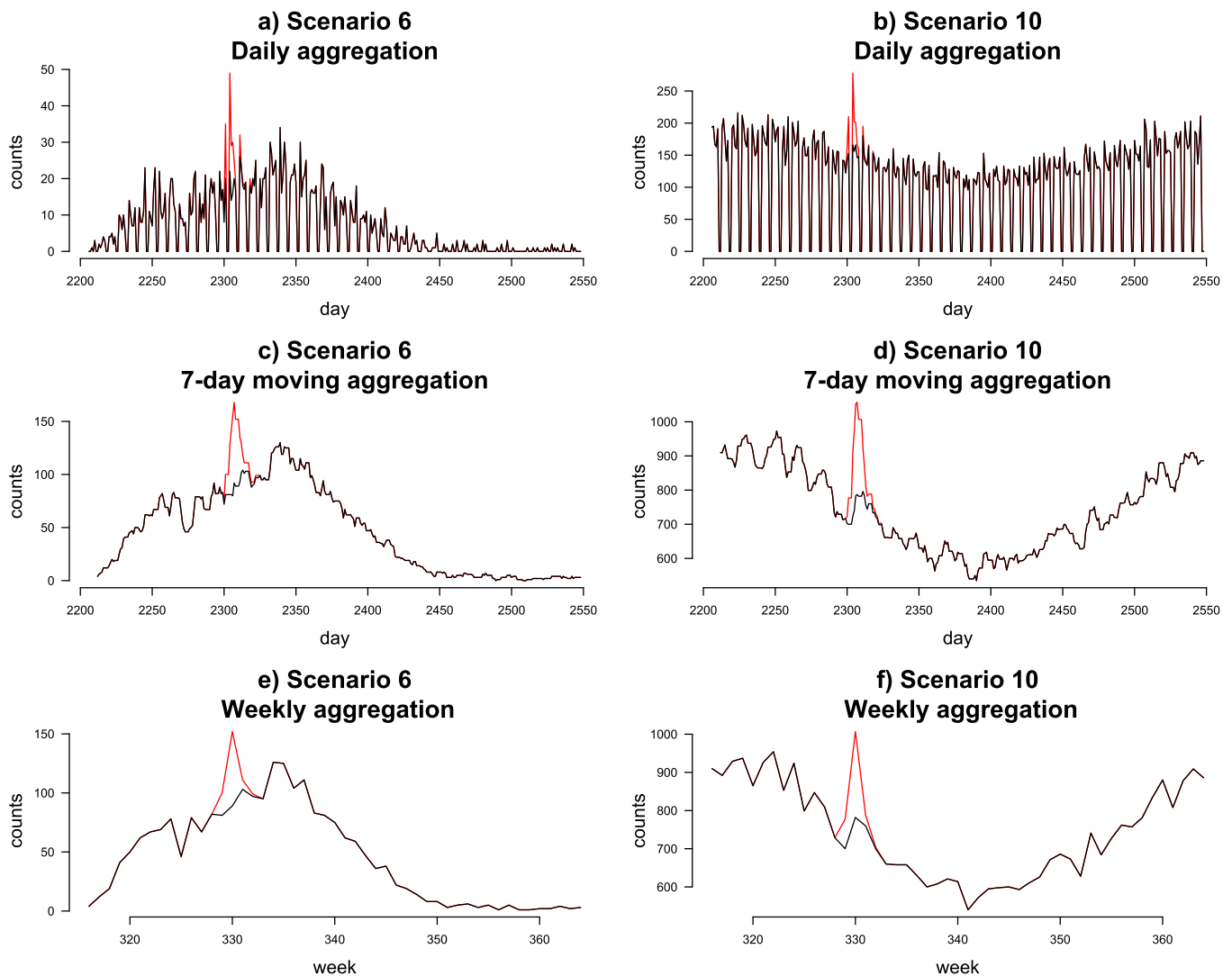
the baseline mean counts and  $S2_t^2 = \frac{1}{6} \sum_{i=t-3}^{t-9} (x_i - \bar{x}_{2t})^2$  is the variance.

C3 is obtained as:

$$C3_t = \sum_{i=t}^{t-2} \max\{0, C2_i - 1\} \tag{5}$$

EARS is usually used for weekly data and is known to have trouble handling seasonality effect (Fricker et al., 2008).

- EARS NB are count data charts combined with Poisson regression and binomial regression, based on the generalized likelihood ratio. In this method, time series are modeled with Poisson or negative binomial regression models. The detection problem is treated in a statistical process control (SPC) framework: the stopping rule used in CUSUM methods (which is usually based on the likelihood ratio) is in the generalized likelihood ratio (GLR) form. The GLR is estimated using Monte Carlo sampling. The EARS NB method is fully described in (Höhle and Paul, 2008).
- RKI 1, 2, 3: these algorithms were developed at the Robert Koch Institute. They evaluate time-points by comparing them with an



**Fig. 3.** Daily aggregation (top), 7-day-moving aggregation (middle) and weekly aggregation (bottom) for scenarii 6 and 10 and a medium outbreak size (red line).

upper limit (threshold) value. The latter is calculated from different reference values according to the RKI algorithm:

- RKI 1: the reference values correspond to the case numbers over the previous six weeks
- RKI 2: the reference values correspond to the case numbers over the six weeks preceding the current value and the case numbers for 13 weeks from the previous year, specifically 6 weeks before and 6 after the current week (*i.e.*, distributed symmetrically around the comparable time unit from the previous year of the current time-point)
- RKI 3: 18 reference values are needed: nine weeks from the year preceding the current time-point year and nine from two years preceding the current time-point year (also symmetrical around the comparable time unit of the current time-point)

The algorithm calculates the mean count  $\mu$  from the reference values. If  $\mu$  is greater than 20, the threshold is obtained as:  $upCI = \mu + 2\sigma$ , where  $\sigma^2$  is the variance of the reference values. If  $\mu$  is smaller than 20, the threshold value corresponds to an upper limit of a Poisson distribution confidence interval (Höhle, 2007). If the current time-point value is greater than the threshold value, an alarm is raised

- Bayes 1, 2, 3: these algorithms need reference values to calculate the  $(1 - \alpha)$  quantile of a predictive posterior negative binomial

distribution as a threshold. The binomial negative distribution of parameters  $r$  and  $p$  depends on:  $r$  the sum of observed values during the historical period and  $p$  the number of observed values. The way to use the historical period to calculate the reference value for each Bayes algorithm is the same as for the RKI algorithms: Bayes 1 uses the same way to calculate the reference value as RKI 1, Bayes 2's reference value is calculated from the same historical period as in RKI 2, and so on. For Bayes 1, 2 and 3, the  $\alpha$  value is fixed at 0.05 (Höhle, 2007).

For all of these algorithms, the  $\alpha$  value was intrinsic to the algorithm, fixed and constant. The default parameters associated with each algorithm in the R Surveillance package are described in Table 1.

### 2.3. Performance measures and analysis

We computed the following indicators for each time series: .

- the number of true positives (TP), *i.e.*, the number of time units (*i.e.*, a day or a week) for which an alarm was raised when an outbreak was simulated during the time unit;
- the number of false positives (FP), *i.e.*, the number of time units with an alarm when no outbreak was simulated in the unit,



- the number of true negatives (TN) *i.e.*, the number of time units for which no alarm was raised when no outbreak was simulated during the time unit,
- the number of false negatives (FN), *i.e.*, the number of time units without an alarm when an outbreak was simulated in the unit.

and calculated seven performance measures for each algorithm and time series:

- **Sensitivity (*Se*)** calculated as  $Se = TP / (TP + FN)$  and corresponding to the probability of an alarm being raised during an outbreak,
- **Specificity (*Sp*)** calculated as  $Sp = TN / (TN + FP)$  and corresponding to the probability of an alarm not being raised when there was no outbreak. We also calculated the False positive rate (FPR), equals to  $1 - Sp$ ,
- **Positive predictive value (PPV)** as  $PPV = TP / (TP + FP)$  and corresponding to the probability of an outbreak when an alarm was raised,
- **F1-measure** as  $F1 = 2 \times (Se \times PPV) / (Se + PPV)$ , which is the harmonic mean of PPV and sensitivity,
- **Probability of detection (POD)** as a binary measure equals 0 when an outbreak was not detected or 1 when an outbreak was detected for each time series,
- **Timeliness** as the proportion of a simulated outbreak that was missed before an alarm was raised:  $timeliness = (\rho - 1) / \delta$ , where  $\rho$  is the rank of the day when the first alarm was raised (from 1 for the first day of the outbreak to X, the last day of the outbreak), where  $\delta$  is the duration of an outbreak. The value of timeliness goes from 0 (*i.e.*, 0% of the outbreak has been missed (or equivalently an alarm was raised on the first day of an outbreak)) to 1 (*i.e.*, 100% of the outbreak was missed).

We used multivariate regressions (with a Poisson or negative binomial distribution in case of overdispersion), to evaluate the effect of the time aggregation unit on each of the seven performance measures, taking into account the algorithm, characteristics of the time series, and size of the outbreak (Bédubourg and Le Strat, 2017; Barboza et al., 2014; Buckeridge et al., 2008). The daily aggregation, Farrington Flexible algorithm, very small outbreaks, no seasonality, one weekly peak and no trend, were all used as the reference categories. We calculated the 95% confidence intervals using robust estimation of standard errors. We added an interaction between the temporal aggregation and detection algorithm. The regression analyses were performed using Stata 14.2 Anon (StataCorp, 2015).

### 2.3.1. Multi-criteria analysis through a performance score

SySS designers face multiple objectives that may conflict and involve trade-off decisions. We have therefore provided the following multi-criteria score to help resolve these trade-offs. SySS users generally want to receive timely and relevant alarms. For this purpose, we considered three major performance measures: sensitivity, PPV and timeliness. We calculated a score weighting these performance measures:

$$Score(w_{Se}, w_{PPV}, w_{time}, a, tau) = w_{Se} \times \overline{Se}(a, tau) + w_{PPV} \times \overline{PPV}(a, tau) + w_{time} (1 - \overline{Timeliness}(a, tau)), \quad (6)$$

where  $\overline{Se}(a, tau)$  was the average sensitivity for the detection algorithm  $a$  and the time aggregation unit  $tau$ ,  $\overline{PPV}(a, tau)$  the average PPV,  $\overline{Timeliness}(a, tau)$  the average timeliness,  $w_{Se}$  the weight allocated to the mean sensitivity,  $w_{PPV}$  the weight allocated to the mean PPV, and  $w_{Time}$  the weight allocated to the average timeliness. For each  $(w_{Se}, w_{PPV})$  pair, we selected the  $(a, tau)$  combination of an algorithm and a time aggregation unit pair for which the score was maximal. Timeliness was weighted at 1 while sensitivity and PPV had weights ranging from 0.1 to 10 on a log10 scale. Results were plotted, as represented in Fig. 4. This representation allowed us to show how the optimal choice of algorithm, according to the temporal aggregation unit, changes with the weighting of sensitivity, PPV and timeliness. Fig. 4 provides an algorithm comparison context for all 3-way inequality relationships among sensitivity, PPV and timeliness. As a result, the Fig. 4 can be divided in 6 different zones, described in the legend. Even though timeliness' weight is fixed to 1, it is possible to choose an algorithm that favors timeliness over sensitivity and PPV: when PPV and sensitivity weight are less than 1 (zone a and zone b for example in Fig. 4). We also represented how the choice of the temporal aggregation unit changed for each detection algorithm we used in this study.

### 3. Results

We simulated and analyzed a total of 19,200 time series (16 scenarii  $\times$  100 simulated time series  $\times$  4 sizes of outbreaks  $\times$  3 temporal aggregations). Table 2 provides the mean value of the 7 performance measures, according to the time aggregation unit and detection algorithm used. Sensitivity ranged from 9%, for EARS C1 with daily aggregation, to 73% for Bayes 3 with 7-day moving aggregation. Specificity ranged from 71%, for Bayes 2 in daily aggregation, to 99% for the Farrington Flexible algorithm in daily and weekly aggregation. PPV

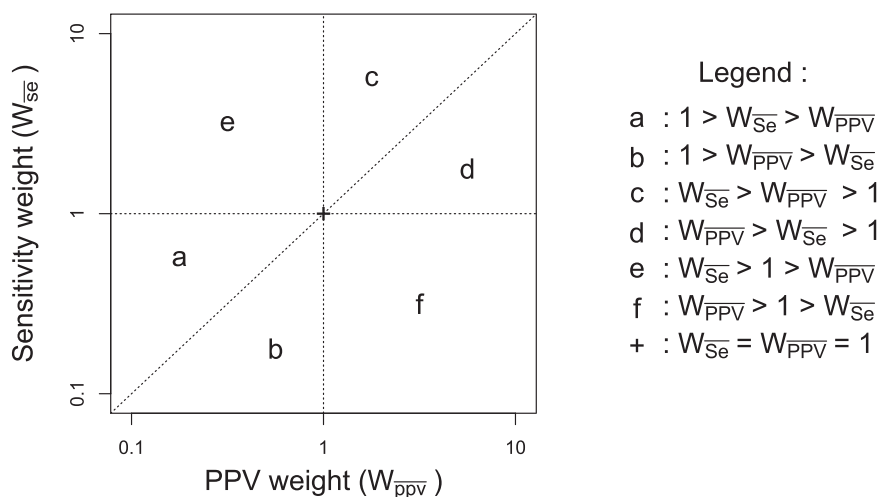


Fig. 4. Schematic representation of the multi-criteria choice of algorithm and temporal aggregation unit, based on the relative weights given to sensitivity, PPV and timeliness (fixed to 1). Each letter corresponds to a zone with a particular weight combination: for example zone a corresponds to a timeliness weight greater than the sensitivity weight which in turn is greater than the PPV weight.

**Table 2**  
Mean performance measures observed according to time aggregation type and anomaly detection algorithm.

Time aggregation	Detection algorithm	Performances measures						
		Sensitivity	Specificity	PPV	F1-Measure	FPR	POD	Timeliness
Daily aggregation	Farrington flexible	0.40	0.99	0.53	0.38	0.02	0.72	0.44
	EARS NB	0.28	0.89	0.40	0.14	0.11	0.96	0.14
	EARS C1	0.09	0.97	0.16	0.07	0.03	0.55	0.53
	EARS C2	0.13	0.97	0.22	0.10	0.03	0.58	0.51
	EARS C3	0.11	0.99	0.28	0.10	0.02	0.44	0.66
	RKI 1	0.20	0.87	0.05	0.08	0.13	0.88	0.25
	RKI 2	0.21	0.89	0.12	0.12	0.11	0.90	0.24
	RKI 3	0.34	0.90	0.18	0.15	0.10	0.89	0.25
	Bayes 1	0.43	0.72	0.05	0.09	0.28	0.99	0.09
	Bayes 2	0.53	0.71	0.07	0.11	0.29	0.99	0.08
	Bayes 3	0.60	0.72	0.08	0.13	0.28	0.99	0.08
	Farrington Flexible	0.37	0.98	0.48	0.34	0.02	0.68	0.56
	EARS NB	0.41	0.97	0.34	0.34	0.03	0.93	0.22
	EARS C1	0.26	0.92	0.10	0.14	0.08	0.93	0.18
7-day moving aggregation	EARS C2	0.44	0.83	0.08	0.13	0.17	0.95	0.16
	EARS C3	0.46	0.81	0.08	0.13	0.19	0.95	0.23
	RKI 1	0.29	0.90	0.11	0.15	0.10	0.92	0.19
	RKI 2	0.42	0.95	0.26	0.30	0.05	0.90	0.24
	RKI 3	0.66	0.88	0.24	0.32	0.12	0.92	0.22
	Bayes 1	0.43	0.91	0.19	0.25	0.09	0.89	0.23
	Bayes 2	0.63	0.86	0.19	0.28	0.14	0.95	0.02
	Bayes 3	0.73	0.83	0.18	0.27	0.17	0.96	0.17
	Farrington flexible	0.33	0.99	0.72	0.38	0.01	0.69	0.33
	EARS NB	0.23	0.93	0.30	0.22	0.07	0.75	0.27
	EARS C1	0.30	0.92	0.24	0.24	0.08	0.70	0.32
	EARS C2	0.42	0.84	0.22	0.26	0.16	0.68	0.33
	EARS C3	0.39	0.83	0.20	0.23	0.17	0.60	0.45
	Weekly aggregation	RKI 1	0.34	0.89	0.19	0.23	0.11	0.75
RKI 2		0.30	0.97	0.53	0.31	0.03	0.66	0.36
RKI 3		0.38	0.97	0.57	0.37	0.03	0.73	0.29
Bayes 1		0.48	0.77	0.16	0.22	0.23	0.85	0.17
Bayes 2		0.54	0.83	0.22	0.29	0.17	0.90	0.12
Bayes 3		0.51	0.83	0.28	0.35	0.17	0.93	0.09

ranged from 5%, for Bayes 1 in daily aggregation, to 72% for Farrington Flexible in weekly aggregation. F1-measure ranged from 7% (EARS C1 in daily aggregation) to 38% (Farrington Flexible in daily and weekly aggregations). POD ranged from 44%, for EARS C3 in daily aggregation, to 99% for Bayes 2 and 3 in daily aggregation. FPR ranged from 1% (Farrington Flexible with weekly aggregation) to 29% (Bayes 2 in daily aggregation). Timeliness ranged from 2%, for Bayes 2 in 7-day moving aggregation, to 66% for EARS C3 in daily aggregation. For a given detection algorithm, we observed differences in the mean value of performance measures according to the time aggregation unit. For example, for the Farrington Flexible algorithm, sensitivity ranged from 33% in weekly aggregation to 40% in daily aggregation. Timeliness ranged from 33% in weekly aggregation to 56% in 7-day moving aggregation. Mean PPV ranged from 5% in daily aggregation to 72% in weekly aggregation type. Fig. 5 represents the results of the multivariable regression model for sensitivity, PPV and timeliness. As these three characteristics are the most common performance measures used by epidemiologists, and with a view to simplifying the reading of this article, we chose to only show results for these three performance measures. The results for PPV, F1-measure, FPR and POD can be found in the appendix. We observed an effect of the time aggregation unit for all the performance indicators measured. Daily aggregation had higher sensitivity than 7-day moving aggregation (ratio = 0.97) and weekly aggregation (0.74) for a similar detection algorithm, scenario (or time series characteristics) and outbreak size. Furthermore, sensitivity varied with the detection algorithm: the use of Bayes 2 and Bayes 3 increased the sensitivity by 24% and 38%, respectively. Seasonality, strong weekly patterns (*i.e.*,  $k_2 = 2$  in Eq. (1)) and trends decreased sensitivity, as did dispersion and baseline frequency reports (*i.e.*,  $\theta$  in equation (1)). Sensitivity ratios increased with outbreak size: from 14% for small outbreaks to 40% for large outbreaks. PPV was also influenced by time aggregation unit. Indeed, weekly aggregated time series increased PPV of 34% and 7-day

moving aggregated time series decreased PPV. Compared with the Farrington Flexible algorithm, the use of other algorithms decreased PPV. Time series characteristics had influence on PPV; for example, a trend in a time series decreased PPV of 69%. As the outbreak size increased, PPV also increased: from 24% for small outbreaks to 85% for big outbreaks. With regard to timeliness (Fig. 5), in our data set the use of weekly aggregation decreased the proportion of missed outbreaks. The choice of the detection algorithm also influenced timeliness, with EARS C1, C2 and C3 being slower to detect an outbreak than the Bayes, RKI and EARS NB algorithms. Moreover, timeliness was more influenced by time series characteristics than specificity and sensitivity: timeliness decreased with increased dispersion, while it increased for all seasonality types (up to 93% for biannual seasonality (Table ATable 1)). Finally, timeliness decreased with the increase in outbreak size. Considering performance measures separately, we observed that to maximize sensitivity, the best match was the Bayes 3 algorithm paired with a 7-day moving aggregation. To maximize the PPV, the best match was the Farrington Flexible algorithm paired with weekly-segmented time series. The Bayes 2 and 3 algorithms coupled with daily counts and weekly aggregation minimized timeliness. An Excel file containing the regression model estimations for all performance measures is available in supplementary data.

With regard to the performance score for sensitivity, PPV and timeliness, results are presented in Fig. 6 for each type of time aggregation. For the daily aggregated time series, only 2 of the 11 detection algorithms are represented in the graph (Bayes 3 and Farrington Flexible), meaning that the other 9 algorithms were not optimal for any weighting combination. When the same weight was allocated to sensitivity, PPV and timeliness, Bayes 3 had the best score. When promoting timeliness (weight 10 times greater than sensitivity and PPV weights), Bayes 3 had the best score. When PPV and timeliness were both more important than sensitivity, Bayes 3 and Farrington Flexible had the best score. When



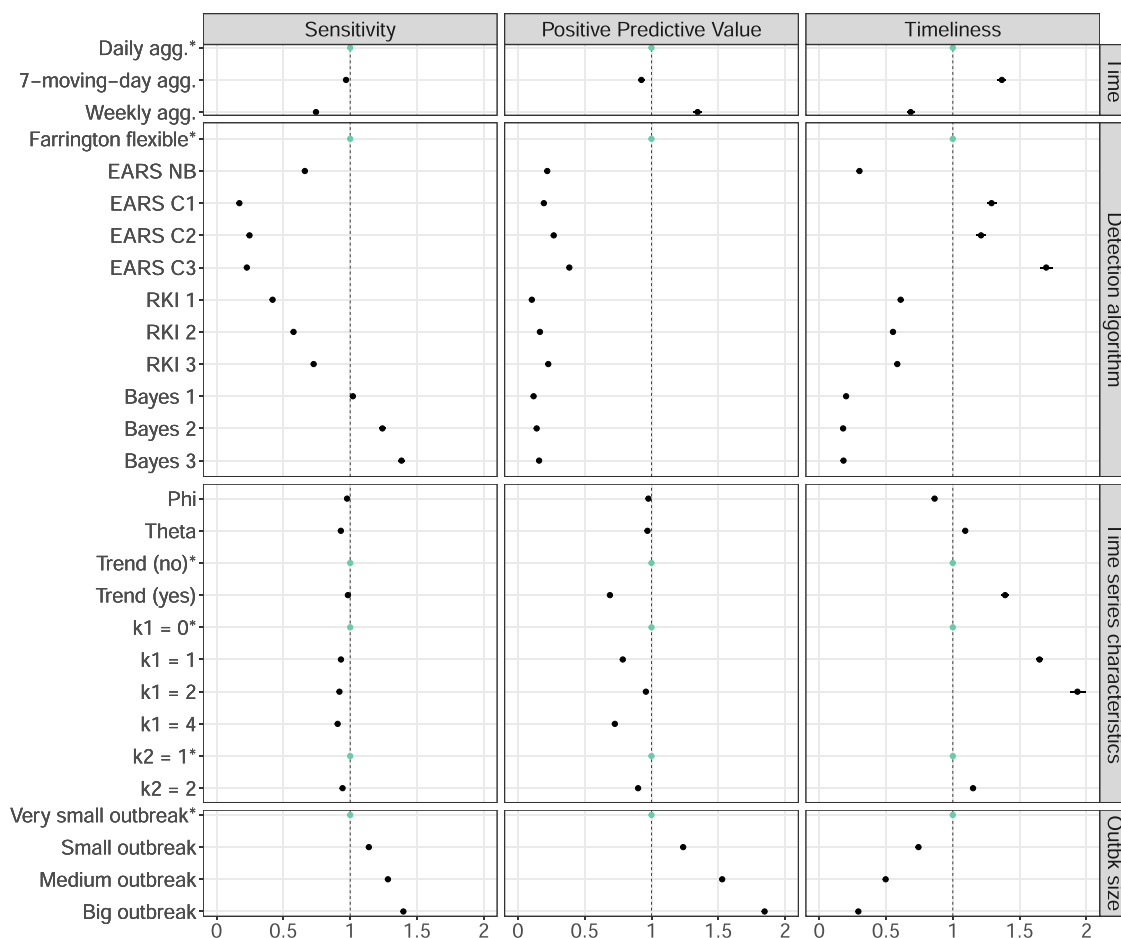


Fig. 5. Sensitivity, PPV and timeliness ratios for each covariate, taking into account interactions between time aggregation unit and detection algorithm. Green points correspond to covariates used as a reference.

PPV and sensitivity were weighted greater than timeliness, the use of the Farrington Flexible algorithm had the best score. Finally, when PPV was 1–10 times more important for a user than both sensitivity and timeliness, Farrington Flexible had the best score. Fig. 7 shows the score from a type of time aggregation perspective for each detection algorithm considered. For example, for the Bayes 1 algorithm, weekly aggregation prioritized sensitivity over PPV and timeliness. Conversely, 7-day moving aggregation was optimal when PPV was the priority. Finally, when timeliness was the most important performance measure, daily aggregation had the best score. Another representation of the multi-criteria score is available in supplementary materials (S5 and S6) where sensitivity and PPV weight range from 0.001 to 1000.

#### 4. Discussion

The main objective of our study was to explore the effects of the temporal segmentation and the time unit on the detection of temporal anomalies in syndromic surveillance. Our results showed that the mean values of the seven performance indicators tested depended both on the time aggregation unit and the detection algorithm used. In addition, for a given detection algorithm, substantial differences in the mean performances existed according to the time aggregation unit used. This study was designed to help public health practitioners implement statistical algorithms to detect anomalies using data from a syndromic surveillance system. One important question for epidemiologists is how best to aggregate a time series. This study shows that there is no one ideal aggregation unit/algorithm pairing that can be proposed, as the optimal pairing depends on several parameters, specifically the

characteristics of the time series and of the anomaly in question (e.g., outbreak/anomaly size) as well as the objective of the detection (do we want it to be sensitive, specific, or both?; do we want to obtain good timeliness?, etc.). The use of simulated times series helped us to control and explore parameters which potentially influence time series and outbreak characteristics. This approach is recognized as the best way to assess the performances of detection algorithms. While previous studies used simulated datasets (Noufaily et al., 2012; Enki et al., 2016; Noufaily et al., 2019; Bédubourg and Le Strat, 2017; Hutwagner et al., 2005; Fricker et al., 2008), or real data (Faverjon et al., 2019; Jackson et al., 2007) to create temporal series, all studies specifically looking at outbreaks used only simulated data in order to create a gold standard for calculating performance parameters of detection algorithms. The present simulation allowed us to explore the wide range of time series diversity encountered in real-world SySS. Although the parameters of the scenarios were estimated using human data, some also mimic animal data, as described in Perrin et al. (2010) and analyzed by OMAR (Sala et al., 2020). We included time series with or without trends associated with different weekly patterns (Fig. 2). The variations in the daily number of cases (from fewer than 10 cases a day in scenario 11 to approximately 1000 cases a day in scenario 1) and day-of-the-week effects, reflected the diversity of the different real-world temporal aggregations observed. These variations enabled us to evaluate the behavior of the detection algorithms according to different case numbers. The characteristics of the times series did not strongly affect the sensitivity of the detection algorithms, irrespective of the aggregation unit. On the contrary, PPV and more particularly timeliness was more influenced by trends and seasonality, the early detection of an

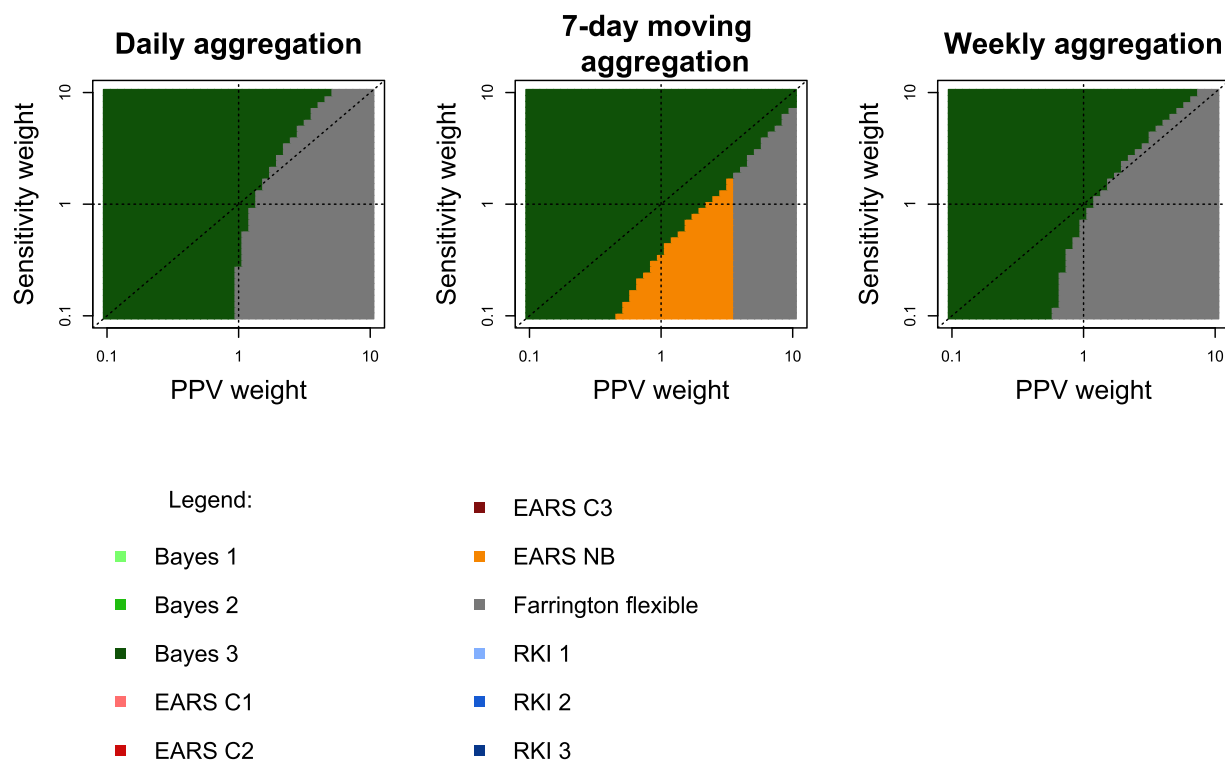


Fig. 6. Multi-criteria choice of the algorithm for each time aggregation unit, according to the relative weight given to sensitivity, PPV, and timeliness.

outbreak becoming increasing difficult when trends and complex seasonality were present. This latter finding is an important result, given that the objective of an SySS is early detection of anomalies and given that trend and seasonality are frequent in real-world time series (temporal variation of an at-risk population, increasing diagnostic capabilities, seasonality of many diseases, mortality, etc.). As expected, the increasing size (level and duration) of an outbreak positively influenced the sensitivity and timeliness of the detection algorithms assessed. Nevertheless, the simulated outbreaks were quite short in duration (maximum three weeks long). It would be interesting to examine different parameters to vary the duration and size of outbreaks, in particular to explore the efficiency of monitoring long-term outbreaks with the problem of baseline contamination (Shmueli and Burkom, 2010). The time aggregation unit influenced the performance of the detection algorithms for all time series and outbreak characteristics. However, the aggregation unit affected the seven performance measures in different ways (see Fig. 5 and supplementary material). This fact led to complex results and no single answer to the question of what the best time aggregation unit-detection algorithm pair to use is. Accordingly, users have to choose the performance measures that are most suitable for their particular SySS. Also, in this study we chose to calculate the F1-measure as it combines both sensitivity and PPV. It could be interesting to also consider a sensitivity analysis with  $F_{\beta}$  for multiple values of  $\beta$  in a later study. Indeed, the  $F_{\beta}$  function is a generalization of the F1-measure where a weight  $\beta$  is chosen such that sensitivity is considered  $\beta$  times as important as PPV. We examined sensitivity, PPV and timeliness, which are the three most commonly used performance indicators in public health surveillance; we encourage the reader to consult the results presented in the supplementary material. Each performance measure provides different information and they complement each other. Furthermore, in Figs. 6 and 7 we propose a visual aid which takes into account all three of these measures, to help epidemiologists and other SySS users choose the best temporal aggregation unit and a detection algorithm according to their context.

In this study, we chose to present the optimal combinations of time aggregation unit and detection algorithms with our simulated data set.

These optimal combinations we presented in Figs. 6 and 7 may be linked to the characteristics of the time series we simulated. Also, the multi-criteria score we calculated is based on the mean sensitivity, timeliness and PPV observed for the 16 scenarios. As a result, the optimal combination of time aggregation unit-detection algorithm is given for the 16 scenarios and may be different if we would be interested in only one scenario among the 16 we simulated or in a completely different kind of time series. The supplementary table S2 allows a reader to search for a combination according to the different time series characteristics. The choice of which statistical algorithm to use is also not an obvious one, as dozens of methods exist which can be tuned by modifying several parameters (such as the  $\alpha$ -value). We chose a selection of algorithms, with fixed parameters, from all those available in order to simplify the analysis. In Table 2, we found out that the F-measures for weekly aggregation and for the two daily aggregations are numerically comparable, but one should not conclude that weekly surveillance is adequate if daily results are available and public health response measures are feasible on a daily basis. Moreover, the sensitivity-PPV tradeoff between Farrington Flexible and Bayes 3 gives the impression that Bayes 3 is preferable if investigation resources are unlimited and sensitivity is paramount, but even then the poor timeliness for Bayes 3 may nullify the sensitivity advantage. There may be data environments where Bayes 3 should be adopted because timeliness is much better, but such choices require periodic system updates and available expertise. If the cost of false alarms is important, Farrington Flexible is clearly superior except in systems where the algorithm choice may be tailored to specific time series. Although Bayes 2 and 3 had the best sensitivity and timeliness, they could not manage the day-of-the-week effect. The reason for this shortcoming is that calculation of the threshold values in the Bayes algorithms is based on a simple sum of historical values to calculate the threshold value - which corresponds to  $(1-\alpha)$  quantile of a predictive posterior negative binomial distribution. For example, as scenario 1's day-of-the-week effect was very strong (almost doubling the daily cases every weekend), the Bayes 2 and 3 raised alarms every weekend in the daily aggregation for this scenario. The other types of time aggregation unit overcame the day-of-the-week effect while the 7-day moving counts

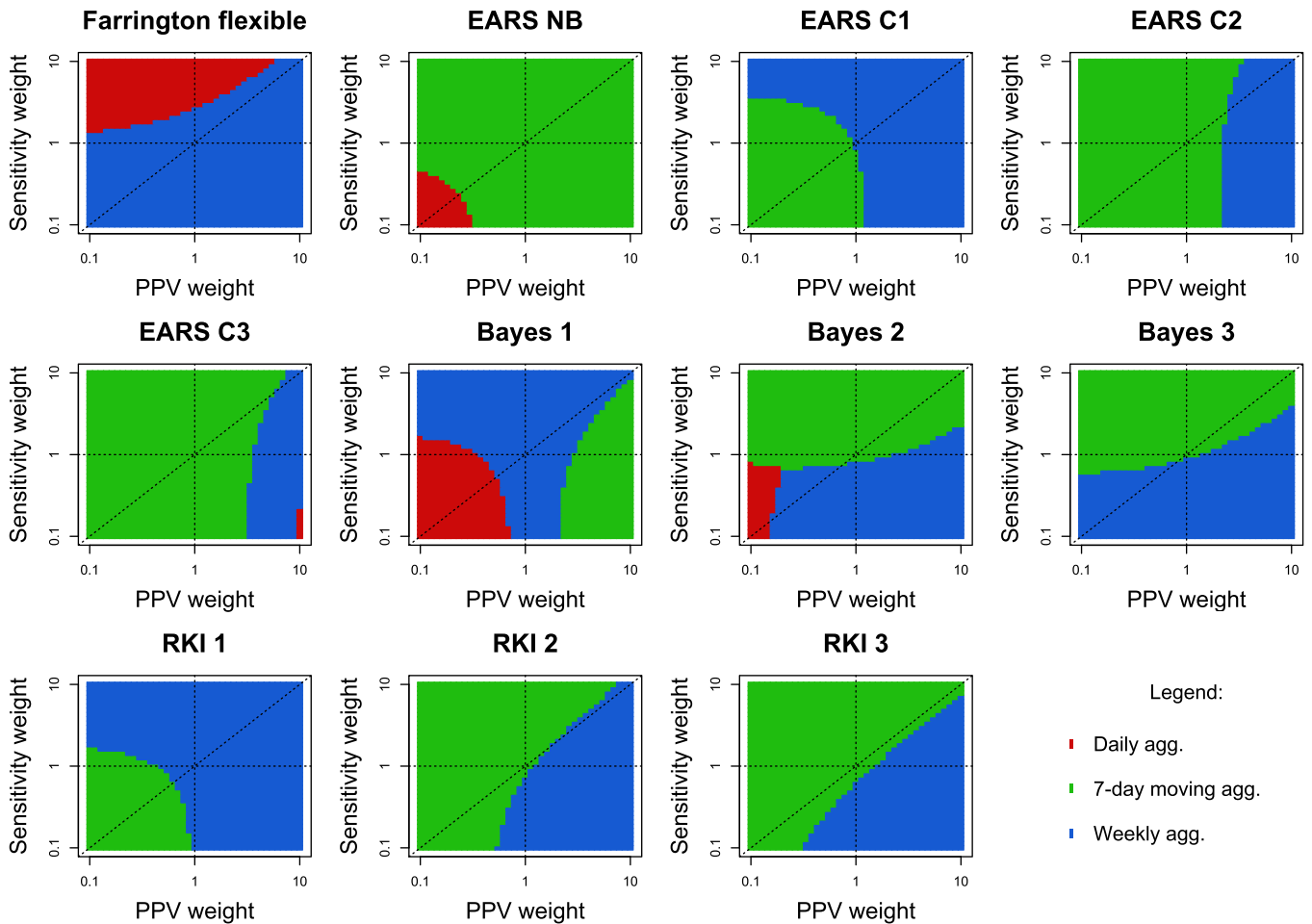


Fig. 7. Multi-criteria choice of the time aggregation unit for each algorithm, according to the relative weight given to sensitivity, PPV, and timeliness.

was a good compromise as it enabled data to be analyzed on a daily basis with a controlled day-of-the-week effect. We recommend associating Bayes 2 and 3 with 7-day moving aggregated time series to ensure good timeliness performance to control for the day-of-the-week effect, and to analyze data on a daily basis. Finally, it would be interesting to try different parameters to vary the duration and size of outbreak, and to use other algorithms to evaluate the effect of temporal aggregation unit on combinations of different algorithms.

Our study only focused on the temporal component of a surveillance system. We did not examine the spatial component, which is a very important element, as a spatial correlation between anomalies of interest often exists. Complementing the MTUP, the modifiable areal unit problem (MAUP) addresses spatial aggregation, which may also impact the statistical results (Openshaw, 1981). The continuation of this work on temporal data should study the effect of the MAUP on the detection of anomalies in Sys first of all using simulated data.

## References

Barboza, P., Vaillant, L., Le Strat, Y., Hartley, D.M., Nelson, N.P., Mawudeku, A.,

Madoff, L.C., Linge, J.P., Collier, N., Brownstein, J.S., Astagneau, P., 2014. Factors influencing performance of internet-based biosurveillance systems used in epidemic intelligence for early detection of infectious diseases outbreaks. *PLOS One* 9 (3), e90536. <https://dx.plos.org/10.1371/journal.pone.0090536>.

- Cheng, T., Adepeju, M., 2014. Modifiable temporal unit problem (MTUP) and its effect on space-time cluster detection. *PLOS One* 9 (6), 10.
- Salmon, M., Schumacher, D., Burmann, H., Frank, C., Claus, H., Höhle, M., 2016. A system for automated outbreak detection of communicable diseases in Germany. *Eurosurveillance* 21 (13). <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2016.21.13.30180>.
- Shmueli, G., Burkom, H., 2010. Statistical challenges facing early outbreak detection in biosurveillance. *Technometrics* 52 (1), 39–51. <http://www.tandfonline.com/doi/abs/10.1198/TECH.2010.06134>.
- Triple S Project, 2011. Assessment of syndromic surveillance in Europe. *Lancet* 378 (9806), 1833–1834. <https://linkinghub.elsevier.com/retrieve/pii/S0140673611608349>.
- H. Brouwer, A.D. Bont-Smolenaars, L.V. Wuijckhuise, G.V. Schaik, 2017. Syndromic surveillance on phone calls about cattle health problems for early detection of emerging diseases. In: *Proceedings of the 3rd International Conference on Animal Health Surveillance*, Rotorua, New Zealand.
- D.L. Buckeridge, A. Okhmatovskaia, S. Tu, C. Nyulas, M.A. Musen, 2008. Predicting Outbreak Detection in Public Health Surveillance: Quantitative Analysis to Enable Evidence-Based Method Selection. In: *Proceedings of the AMIA 2008 Symposium Proceedings*, 76–80, 2008.
- Bédubourg, G., Le Strat, Y., 2017B. Evaluation and comparison of statistical methods for early temporal detection of outbreaks: a simulation-based study. *PLOS One* 12 (7), e0181227. <https://dx.plos.org/10.1371/journal.pone.0181227>.
- Burkom, H., Estberg, L., Akkina, J., Elbert, Y., Zepeda, C., Baszler, T., 2019. Equine syndromic surveillance in Colorado using veterinary laboratory testing order data. *PLOS One* 14 (3), e0211335. <https://dx.plos.org/10.1371/journal.pone.0211335>.
- Caserio-Schönemann, C., Meynard, J.B., 2015. Ten years experience of syndromic surveillance for civil and military public health, France, 2004–2014. *Eurosurveillance* 20 (19). <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES2015.20.19.21126>.
- Dórea, F.C., McEwen, B.J., McNab, W.B., Sanchez, J., Revie, C.W., 2013. Syndromic surveillance using veterinary laboratory data: algorithm combination and customization of alerts. *PLOS One* 8 (12), e82183. <https://dx.plos.org/10.1371/journal.pone.0082183>.
- Enki, D.G., Garthwaite, P.H., Farrington, C.P., Noufaily, A., Andrews, N.J., Charlett, A., 2016. Comparison of statistical algorithms for the detection of infectious disease

- outbreaks in large multiple surveillance systems. *PLOS One* 11 (8), e0160759. <https://dx.plos.org/10.1371/journal.pone.0160759>.
- Farrington, C.P., Andrews, N.J., Beale, A.D., Catchpole, M.A., 1996. A statistical algorithm for the early detection of outbreaks of infectious disease. *J. R. Stat. Soc. A* 159 (3), 547–563.
- Faverjon, C., Berezowski, J., 2018. Choosing the best algorithm for event detection based on the intended application – a conceptual framework for syndromic surveillance. *J. Biomed. Inform.* 85, 126–135.
- Faverjon, C., Carmo, L.P., Berezowski, J., 2019. Multivariate syndromic surveillance for cattle diseases: epidemic simulation and algorithm performance evaluation. *Prev. Vet. Med.* 172, 104778. <https://linkinghub.elsevier.com/retrieve/pii/S0167587719301114>.
- Fricke, R.D., Hegler, B.L., Dunfee, D.A., 2008. Comparing syndromic surveillance detection methods: EARS' versus a CUSUM-based methodology. *Stat. Med.* 27 (17), 3407–3429 ([Online]. Available). <http://doi.wiley.com/10.1002/sim.3197>.
- Heffernan, R., Mostashari, F., Das, D., Besculides, M., Rodriguez, C., Greenko, J., Steiner-Sichel, L., Balter, S., Karpati, A., Thomas, P., Phillips, M., Ackelsberg, J., Lee, E., Leng, J., Hartman, J., Metzger, K., Rosselli, R., Weiss, D., 2004. New York city syndromic surveillance systems. *Morb. Mortal. Wkly. Rep. Suppl.* 53, 4.
- Höhle, M., 2007. Surveillance: an R package for the monitoring of infectious diseases. *Comput. Stat.* 22 (4), 571–582 ([Online]. Available). <http://link.springer.com/10.1007/s00180-007-0074-8>.
- Höhle, M., Paul, M., 2008. Count data regression charts for the monitoring of surveillance time series. *Comput. Stat. Data Anal.* 52, 4357–4368.
- Hulth, A., Rydevik, G., Linde, A., 2009. Web queries as a source for syndromic surveillance. *PLOS One* 4 (2), e4378.
- Hutwagner, L., Thompson, W., Seeman, G.M., Treadwell, T., 2003. The bioterrorism preparedness and response early aberration reporting system (EARS). *J. Urban Health: Bull. N. Y. Acad. Med.* 80 (2), 8.
- Hutwagner, L., Browne, T., Seeman, G.M., Fleischauer, A.T., 2005. Comparing aberration detection methods with simulated data. *Emerg. Infect. Dis.* 11 (2), 314–316. <http://wwwnc.cdc.gov/eid/article/11/2/04-0587article.htm>.
- Jackson, M.L., Baer, A., Painter, I., Duchin, J., 2007. A simulation study comparing aberration detection algorithms for syndromic surveillance. *BMC Med. Inform. Decis. Mak.* 7 (1), 6. <https://bmcmmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-7-6>.
- de Jong, R., de Bruin, S., 2012. Linear trends in seasonal vegetation time series and the modifiable temporal unit problem. *Biogeosciences* 9 (1), 71–77 ([Online]. Available). <https://bg.copernicus.org/articles/9/71/2012/>.
- Noufaily, A., Enki, D.G., Farrington, P., Garthwaite, P., Andrews, N., Charlett, A., 2012. An improved algorithm for outbreak detection in multiple surveillance systems. *Stat. Med.* 32 (7), 1206–1222 ([Online]. Available). <http://doi.wiley.com/10.1002/sim.5595>.
- Noufaily, A., Morbey, R.A., Elliot, A.J., Smith, G.E., Lake, I.R., McCarthy, N., 2019. Comparison of statistical algorithms for daily syndromic surveillance aberration detection. *Bioinformatics* 35 (17), 3110–3118.
- Openshaw, S., 1981. Le problème de l'agrégation spatiale en géographie. *Espace Géogr.* 10 (1), 15–24.
- Paterson, B.J., Durrheim, D.N., 2013. The remarkable adaptability of syndromic surveillance to meet public health needs. *J. Epidemiol. Glob. Health* 3 (1), 41–47. <https://www.atlantis-press.com/article/125905916>.
- Pereira, R.H.M., 2019. Future accessibility impacts of transport policy scenarios: equity and sensitivity to travel time thresholds for bus rapid transit expansion in Rio de Janeiro. *J. Transp. Geography* 74, 321–332. <https://linkinghub.elsevier.com/retrieve/pii/S09666692318302047>.
- Perrin, J.-B., Ducrot, C., Vinard, J.-L., Morignat, E., Gauffier, A., Calavas, D., Hendrikx, P., 2010. Using the national cattle register to estimate the excess mortality during an epidemic: application to an outbreak of Bluetongue serotype 8. *Epidemics* 2 (4), 207–214. <http://linkinghub.elsevier.com/retrieve/pii/S1755436510000617>.
- R Core Team, 2021. R: A language and environment for statistical computing. R. Found. Stat. Comput. (Vienna, Austria) <https://www.R-project.org/>.
- Robertson, C., Nelson, T.A., MacNab, Y.C., Lawson, A.B., 2010. Review of methods for space-time disease surveillance. *Spat. Spatio-Tempo Epidemiol.* 1 (2–3), 105–116 ([Online]. Available). <http://linkinghub.elsevier.com/retrieve/pii/S187758451000002X>.
- Sala, C., Vinard, J.-L., Pandolfi, F., Lambert, Y., Calavas, D., Dupuy, C., Garin, E., Touratier, A., 2020. Designing a syndromic bovine mortality surveillance system: lessons learned from the 1-year test of the french OMAR alert tool. *Front. Vet. Sci.* 6, 453. <https://www.frontiersin.org/article/10.3389/fvets.2019.00453/full>.
- StataCorp, 2015. StataCorp, Stata Statistical Software: Release 14. StataCorp LP, College Station, TX.
- A. Çöltekin, S.D. Sabbata, C. Willi, I. Vontobel, S. Pfister, M. Kuhn, and M. Lacayo, 2011. Modifiable temporal unit problem. In: Proceedings of the ISPRS/ICA Workshop Persistent Problems in Geographic Visualization - ICC2011, 2.
- K.J. Henning, 2004. What is Syndromic Surveillance? Morbidity and Mortality Weekly Report, vol. 53, Supplement: Syndromic Surveillance, no. Reports from a National Conference, 2003, 7–11, 2004.
- Unkel, S., Farrington, C.P., Garthwaite, P.H., Robertson, C., Andrews, N., 2012. Statistical methods for the prospective detection of infectious disease outbreaks: a review: detection of infectious disease outbreaks. *J. R. Stat. Soc.: Ser. A* 175 (1), 49–82. <http://doi.wiley.com/10.1111/j.1467-985X.2011.00714.x>.
- Wu, T.-S.J., Shih, F.-Y.F., Yen, M.-Y., Wu, J.-S.J., Lu, S.-W., Chang, K.C.-M., Hsiung, C., Chou, J.-H., Chu, Y.-T., Chang, H., Chiu, C.-H., Tsui, F.-C.R., Wagner, M.M., Su, I.-J., King, C.-C., 2008. Establishing a nationwide emergency department-based syndromic surveillance system for better public health responses in Taiwan. *BMC Public Health* 8 (1), 18. <http://bmcpubhealth.biomedcentral.com/articles/10.1186/1471-2458-8-18>.
- Yan, W., Palm, L., Lu, X., Nie, S., Xu, B., Zhao, Q., Tao, T., Cheng, L., Tan, L., Dong, H., Diwan, V.K., 2013. ISS-an electronic syndromic surveillance system for infectious disease in rural China. *PLOS One* 8 (4), e62749. <https://dx.plos.org/10.1371/journal.pone.0062749>.