

A bootstrap-based comparison of ILI intensity thresholds from the moving epidemic and WHO methods

Johannes Bracher, Jonas Littek

Karlsruhe Institute of Technology (KIT), Chair of Statistics and Econometrics

Abstract

The moving epidemic method (MEM) and the WHO method are widely used approaches to determine intensity levels for seasonal influenza and influenza-like illness (ILI). They are conceptually similar, but differ in two aspects. Firstly, the MEM involves a log-transformation of incidence data, while the WHO method typically operates on the original scale of the data. Secondly, the MEM method by default uses more than one observation from each past season, with a fixed total number of observations to include. The WHO method, on the other side, uses only the single highest value from each season to determine peak intensity thresholds. We provide statistical arguments and a simulation study to assess the impact of these choices on the resulting intensity thresholds and empirical exceedance proportions. The simulation study is based on a bootstrap approach using historical ILI data from France, Spain, Switzerland and the United States. We find that using data on the original scale leads to a too large proportion of seasons being classified as high or “very high intensity intensity, which can to some degree be mitigated by using a logarithmic transformation. When fixing the total number of observations to include as in the MEM, intensity thresholds tend to increase as more historical data becomes available. When data is only available for few seasons, there is a rather large chance of classifying a new season as high or very high intensity. We therefore advocate always using one observation per season with a log transformation, i.e. a hybrid between the default settings of the MEM and WHO methods. This adaptation is possible within an existing software implementation of the MEM approach and thus easy to apply.

1 Introduction

Following the 2009 influenza H1N1 pandemic, the need for a rapid assessment tool for influenza intensity was recognized. In a report, the Review Committee on the Functioning of the International Health Regulations and on Pandemic Influenza (H1N1) recommended that member states apply and evaluate their methods for severity assessment every year, thus improving their pandemic preparedness (WHO, 2011, p.118). In the subsequent WHO Pandemic Influenza Severity Assessment (PISA) guideline (2017), two specific statistical methods are recommended to determine influenza intensity thresholds. Firstly, the so-called *WHO method*, already described in a previous document (WHO, 2014) and in more detail in the PISA guide; and secondly, the *moving epidemic method* (MEM; Vega et al. 2013, 2015) which is also recommended by the European Centre for Disease Prevention and Control (e.g., ECDC 2017).

In particular the MEM approach has in the meantime been adopted by numerous national public health agencies (see references in Section 3) and seems to deliver practically useful thresholds. Its application has been facilitated by the availability of a user-friendly and well-documented R package (`mem`; Lozano 2020), which can even be used via a graphical user interface (Lozano and Vega, 2018). The statistical properties of the MEM and WHO methods, however, have not yet been studied in detail. We obtain some analytical results, complemented by simulation experiments. In order to ensure a close link to real influenza patterns, the latter are based on bootstrap re-sampling of historical surveillance data. This way, we provide an improved understanding and intuition for the consequences of different implementation choices.

The article is structured as follows. Section 2 provides concise definitions of the MEM and WHO methods, highlighting two important differences between these conceptually similar approaches. Section 3 consists of an overview of published applications of the MEM and WHO methods. This is followed by a brief examination of statistical properties of the two methods in Section 4. In Section 5 we describe

our simulation approach and the results based on historical data from France, Spain, Switzerland and the United States. Section 6 concludes with a discussion.

2 Definition of the moving epidemic and WHO methods

In the following we focus on the construction of thresholds for peak intensity, framing the MEM and WHO methods as different special cases of the same general approach. Details on the construction of baseline epidemic thresholds from pre-season observations are omitted. We assume that thresholds are based on data from m past seasons and applied to the $(m+1)$ -th season. Vega et al. (2015) recommend to use $5 \leq m \leq 10$ seasons for the MEM to ensure a recent data basis. Thresholds for a given intensity measure X are then obtained via the following steps:

1. Within each historical season $j = 1, \dots, m$ order all observations in decreasing order, denoting the i -th largest observation from season j by $X_i^{(j)}$.
2. Select the n largest observations from each of the m past seasons to construct a reference set $\mathcal{X} = \{X_i^{(j)}; j = 1, \dots, m; i = 1, \dots, n\}$.
3. Apply an (invertible) transformation $Y_i^{(j)} = f(X_i^{(j)})$ to all members of the reference set \mathcal{X} to obtain a reference set \mathcal{Y} of transformed historical observations.
4. Assume that the transformed values in \mathcal{Y} come from a normal distribution and compute estimates $\hat{\mu}, \hat{\sigma}$ of its mean and standard deviation.
5. Define intensity thresholds on the transformed scale as quantiles of the normal distribution $N(\hat{\mu}, \hat{\sigma}^2)$, i.e. compute

$$q_{Y,\alpha} = \hat{\mu} + z_\alpha \hat{\sigma},$$

where z_α is the α quantile of the standard normal distribution. A common choice for both methods (WHO, 2017), also implemented as the default in the R package `mem` (Lozano, 2020) is to use

- the 40th percentile $q_{Y,0.4} = \hat{\mu} - 0.25\hat{\sigma}$ as the threshold between low and medium intensity;
 - the 90th percentile $q_{Y,0.9} = \hat{\mu} + 1.28\hat{\sigma}$ as the threshold between medium and high intensity;
 - the 97.5th percentile $q_{Y,0.975} = \hat{\mu} + 1.96\hat{\sigma}$ as the threshold between high and very high intensity.
6. Obtain thresholds on the original scale by applying the inverse transformation, i.e. setting $q_{0.4} = f^{-1}(q_{Y,0.4})$, $q_{0.9} = f^{-1}(q_{Y,0.9})$, $q_{0.975} = f^{-1}(q_{Y,0.975})$.

The MEM and WHO methods are special cases of this procedure with the following specifications:

- In the WHO method, the standard procedure is to set f to the identity function (i.e. no transformation is applied). If peaks vary considerably, a log-transformation is recommended (WHO, 2017). Irrespective of the number m of historical seasons, $n = 1$ observation per season is used.
- In the MEM method, the default choice for f is the natural logarithm, while the number of observations per season is set to $n = 30/m$. (The exact specification used in the R package `mem` is $n = \max(1, \lfloor 30/m \rfloor)$, where $\lfloor \cdot \rfloor$ denotes rounding to the nearest integer.) The total number of historical observations is thus kept (approximately) fixed, ensuring a reasonably large number of reference observations

Note that the software implementation of the moving epidemic method in the R package `mem` is rather general and permits the user to choose n , m , and f (between the identity and logarithmic function). The term *moving epidemic method* could thus also be used as an umbrella term for the general procedure described above. We here use it in a more narrow sense to refer to the specific procedure described in Vega et al. (2015), WHO (2017) and reflected in the default parameters of the `mem` package.

The rationale behind the percentile-based approach is that “about 50–60% of the season peaks should be above the moderate threshold, $\pm 10\%$ above the high threshold and $\pm 2.5\%$ above the extraordinary threshold” (WHO, 2017). Note that in most previous descriptions (e.g., WHO 2014; Vega et al. 2015), the thresholds have been referred to as upper ends of one-sided confidence intervals associated with the arithmetic (WHO method) or geometric mean (MEM) of the historical reference observations in \mathcal{X} . This, however, is slightly imprecise terminology as in the computations the standard deviation $\hat{\sigma}$ rather than the standard error $\hat{\sigma}/\sqrt{nm}$ of $\hat{\mu}$ is used. If indeed confidence intervals were used, these would tend to get narrower the more historical seasons were available, and the thresholds for all levels would eventually converge to the arithmetic / geometric mean. This is obviously not the intended behaviour.

3 An overview of recent applications

To get a better understanding of the different settings in which the MEM and WHO methods are applied in practice we performed a literature search of articles published in English language and citing the papers Vega et al. (2015), WHO (2014) and WHO (2017) until December 2020 (identified via *CrossRef* and *Google Scholar*). The results are summarized in Table 1. As can be seen from the numerous entries from the years 2019 and 2020, the MEM has quickly become a standard approach in the determination of intensity thresholds for influenza and other respiratory diseases. Indeed, the contributions come from numerous countries and in many cases have been co-authored by representatives of national or regional public health agencies. In most analyses, the suggested threshold levels at the 40th, 90th and 97.5th percentile are used. Variability with respect to the number m of historical seasons included is considerable, with a range from three to 16 seasons. Consequently, the number n of observations included per season ranges from two to ten (none of the above papers indicated a modification of the default setting $n = 30/m$ of the moving epidemic method). We only found three published applications of the WHO method, one of them providing a comparison to the thresholds from the MEM method (Rguig et al., 2020).

Table 1: Applications of the MEM and WHO method to determine intensity thresholds for respiratory diseases. We do not include works where only baseline thresholds are computed. The number of seasons included to compute thresholds is denoted by m , the number of observations used per season by n . Abbreviations: SARI = severe acute respiratory infection; ILI = influenza-like illness; RSV = respiratory syncytial virus.

(a) Moving epidemic method						
region	disease	years covered	m	n	percentiles	authors
Australia	ILI/influenza	2012–2017	5	6	40, 90, 99	Vette et al. (2018)
Australia, Chile, New Zealand, South Africa	ILI/SARI	2013–2019	6	5	40, 90, 97.5	Sullivan et al. (2019)
Catalonia	ILI	2010–2016	5	6	?	Basile et al. (2018)
Catalonia	ILI	2005–2018	12	3	?	Basile et al. (2019)
Catalonia	ILI/influenza	2010–2017	7	4	?	Torner et al. (2019)
Egypt	SARI/ILI	2010–2017	6	5	40, 90, 97.5	AbdElGawad et al. (2020)
Egypt	SARI	2013–2015	3	10	?	Elhakim et al. (2019)
England	ILI	2010–2016	6	5	?	Wagner et al. (2018)
Finland	influenza	2011–2016	5	6	?	Pesälä et al. (2019)
Montenegro	ILI	2010–2018	7	4	40, 90, 97.5	Rakocevic et al. (2019)
Morocco	ILI	2005–2017	11	3	40, 90, 97.5	Rguig et al. (2020)
Netherlands	RSV	2005–2017	12	3	40, 90, 97.5	Vos et al. (2019)
Norway	influenza	2006–2015	9	3	?	Benedetti et al. (2019)
Pakistan	ILI, SARI	2008–2017	10	3	40, 90, 97.5	Nisar et al. (2020)
Portugal	ILI	2012–2017	5	6	40, 90, 97.5	Páscoa et al. (2018)
Scotland	influenza	2010–2018	7	4	?	Murray et al. (2018)
Scotland	influenza	2010–2019	7–8	4	40, 90, 97.5	Dickson et al. (2020)
Slovenia	RSV	2008–2018	10	3	40, 90, 97.5	Grilc et al. (2020)
Spain (17 regions)	ILI	2003–2015	4–10	3–8	40, 90, 97.5	Bangert et al. (2017)
Spain	ILI	2001–2018	16	2	40, 90, 97.5	Redondo-Bravo et al. (2020)
Tunisia	ILI	2009–2018	9	3	50, 90, 95	Bouguerra et al. (2020)
United Kingdom	ILI	2000–2013	10	3	40, 90, 97.5	Green et al. (2015)
United Kingdom	ILI/RSV	2011–2018	4–6	5–8	40, 90, 97.5	Harcourt et al. (2019)
USA	ILI/influenza	2003–2015	11	3	50, 90, 98	Biggerstaff et al. (2017)
USA	ILI	2010–2015	5	6	50, 90, 98	Dahlgren et al. (2018)
USA	influenza	2010–2016	6	5	50, 90, 98	Dahlgren et al. (2019)
(b) WHO method						
region	disease	years covered	m	n	percentiles	authors
Cambodia	ILI	2009–2015	7	1	mean, 90, 95	Ly et al. (2017)
Morocco	ILI	2005–2017	11	1	40, 90, 97.5	Rguig et al. (2020)
Philippines	ILI	2006–2012	7	1	90	Lucero et al. (2016)
Victoria/Australia	ILI	2002–2011	6–10	1	90, 95	Tay et al. (2013)

4 Analytical considerations on the choice of n

As detailed in the previous section, the MEM and WHO methods have been applied extensively in practice and overall seem to produce thresholds which are considered useful by domain experts. A more statistical assessment, however has so far been missing. Which choice of transformation function f is most reasonable is largely an empirical question. Little more can be said than that a logarithmic transformation can be helpful if the distribution of peak values is skewed. However, some statements can be made on how the value chosen for n impacts the expected results.

Denote by $\mathbf{Y}^{(j)}$ the vector of the n largest transformed observations from season j in decreasing order, i.e. $\mathbf{Y}^{(j)} = (Y_1^{(j)}, \dots, Y_n^{(j)})^\top$ with $Y_1^{(j)} \geq Y_2^{(j)} \geq \dots \geq Y_n^{(j)}$. Now assume that $\mathbf{Y}^{(1)}, \dots, \mathbf{Y}^{(m)}$ are independent and identically distributed with mean and covariance matrix

$$\mathbb{E}(\mathbf{Y}^{(j)}) = \boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \quad \text{and} \quad \text{Cov}(\mathbf{Y}^{(j)}) = \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{1,1} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n,1} & \cdots & \sigma_{nn} \end{pmatrix}, \quad (1)$$

respectively, where to make notation more intuitive we also write $\sigma_i^2 = \sigma_{i,i}$. It can then be shown that the expectations of $\hat{\mu}$ and $\hat{\sigma}^2$ are given by

$$\mathbb{E}(\hat{\mu}) = \frac{1}{n} \sum_{i=1}^n \mu_i, \quad (2)$$

$$\mathbb{E}(\hat{\sigma}^2) = \frac{m}{mn-1} \sum_{i=1}^n (\sigma_i^2 + \mu_i^2) - \frac{1}{n(mn-1)} \sum_{i=1}^n \sum_{i'=1}^n \sigma_{i,i'} - \frac{m}{n(mn-1)} \left(\sum_{i=1}^n \mu_i \right)^2, \quad (3)$$

respectively. The derivation is provided in Supplement **XXX**. For reasons detailed there, if the transformation function f is the identity or the natural logarithm,

$$\mathbb{E}(q_\alpha) \approx f^{-1} \left\{ \mathbb{E}(\hat{\mu}) + z_\alpha \sqrt{\mathbb{E}(\hat{\sigma}^2)} \right\} \quad (4)$$

usually holds in good approximation (with f^{-1} the identity or exponential function). This implies that for $n = 1$, we have

$$\mathbb{E}(q_\alpha) \approx f^{-1}(\mu_1 + z_\alpha \sigma_1).$$

This is a desirable property. Loosely speaking, if in addition the $Y_1^{(j)}$ come from a normal distribution, the nominal exceedance probabilities for the different thresholds would on average be achieved. If a different value $n > 1$ is chosen, this will generally no longer be the case. Equations (2)–(4) tell us by how much q_α can be expected to differ from the intended value $f^{-1}(\mu_1 + z_\alpha \sigma_1)$. By the definition of the μ_i (with μ_i the expectation of the i -th largest observation in a given season), $\mathbb{E}(\hat{\mu})$ decreases in n . In most (possibly all) settings this also translates to the q_α . As a consequence, when choosing $n > 1$, one must expect to classify a larger number of season peaks as high or very high intensity. When choosing $n = 30/m$ as suggested for the MEM, thresholds will then tend to increase the more years of data are used to compute thresholds.

5 Simulaton study

5.1 Goal

To complement the analytical results from the previous section we perform a simulation study to compare four versions of the general approach described in Section 2:

- (a) No transformation, $n = 1$. This corresponds to the WHO method.
- (b) No transformation, $n = 30/m$.
- (c) Logarithmic transformation, $n = 1$.
- (d) Logarithmic transformation, $n = 30/m$. This corresponds to the MEM approach.

Our main focus is on the resulting intensity thresholds and shares of seasons classified as low, medium, high and very high. We will put the analytical approximations from the previous section to the test and assess the impact of the chosen n and f on thresholds and exceedance rates empirically.

5.2 Bootstrap approach based on historical ILI data

A pure simulation study requires careful specification of a data-generating process mimicking the seasonal patterns of influenza. Given the considerable variability in timing and intensity of season peaks, this is not a trivial task. We therefore resort to a bootstrap approach based on historical data rather than generating synthetic data. Assume we have M years of historical data on a measure of influenza activity. We then repeat the following steps a total of 500 times:

- Sample an ordered sequence of 15 seasons from the M available seasons. As commonly done in bootstrap experiments, this is done with equal probability for each season and *with replacement*, meaning that the same season can appear more than once in the sampled sequence. This approach is also referred to as the *seasonal block bootstrap* (Politis, 2001).
- For each value $m = 5, \dots, 15$:
 - Restrict the generated sequence to the first m seasons.
 - Apply approaches (a)–(d) as described in the previous subsection. Compute and store intensity thresholds for medium (40th percentile), high (90th percentile) and very high intensity (97.5th percentile).
 - Compute which fraction of all M available historical seasons would be classified as low, moderate, high and very high using the obtained thresholds.

We can then assess how the thresholds and shares of the different intensity levels depend on the employed transformation function, the number of observations n used per season and the number m of seasons included. Here, the range $m = 5, \dots, 15$ is motivated by the range of values found in real-world applications of the moving epidemic method as listed in Section 3.

5.3 Data

We use publicly available influenza surveillance data from four countries. Data on the weekly incidence of influenza-like-illness in France, 1985–2019, were obtained from the Réseau Sentinelles web platform (INSERM/Sorbonne Université, <https://www.sentiweb.fr>, Flahault et al. 2006). Population data were obtained from Institut national de la statistique et des études économiques. Data on weekly confirmed influenza cases in Spain, 1998–2019, were extracted from graphs shown in the *Informe Anual* of the *Sistema de la vigilancia de gripe en España* (<https://vgripe.isciii.es/>; Sistema de Vigilancia de Gripe en España 2019) using the tool *WebPlotDigitizer* (<https://apps.automeris.io/wpd/>). Weekly ILI counts from Swiss routine surveillance, collected by the Swiss Federal Office of Public Health, 2000–2016, were obtained from the R package `HIDDA.forecasting` (Held and Meyer, 2019). Population data were obtained from the Swiss Federal Statistical Office. Weekly weighted ILI percentage rates (wILI) in the US from CDC FluView (Charbonneau and James, 2019), 1998–2017, were obtained from the CDC FluSight influenza forecasting platform (<https://github.com/FluSightNetwork/cdc-flusight-ensemble/>). These wILI values correspond to the percentage of general practitioner visits due to influenza-like symptoms.

All four time series are displayed in Figure 1. Note that we removed the pandemic 2009/2010 influenza season from each series as it is not comparable to the usual years with only seasonal influenza. We also show boxplots of the first through sixth largest observation per season (right column). In the notation introduced in Section 2, the first boxplot thus shows observations $X_1^{(j)}$, $j = 1, \dots, M$, the second $X_2^{(j)}$, $j = 1, \dots, M$ and so on. Little surprisingly (and by definition), the average values get smaller for increasing ranks. Interestingly, they also get less dispersed, meaning that the variability between e.g. the sixth largest observations of the different seasons is smaller than that of the largest observations.

5.4 Used software

All analyses were performed using the R language for statistical computing, version 4.0.3 (R Core Team, 2020), in particular using the package `mem`, version 2.16 (Lozano, 2020). Note that the function `mem::memmodel` provides sufficient flexibility to apply all four approaches (a)–(d) from Section 5.1.

5.5 Results

Results from our bootstrap simulation study for the four considered countries are shown in Figures 2 (France, Spain) and 3 (Switzerland and US). In the first and third column of each figure we also show the

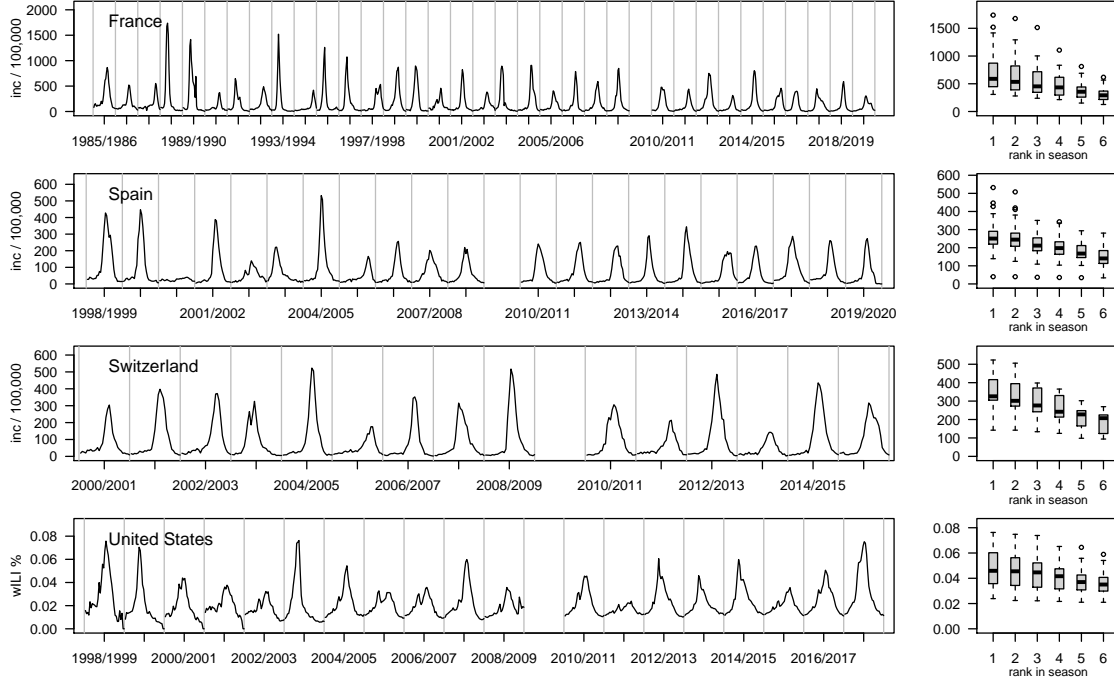


Figure 1: Time series of influenza activity measures in four countries: Weekly ILI cases per 100,000 population in France, 1985–2019; weekly number of confirmed influenza cases per 100,000 population in Spain, 1998–2019; weekly ILI cases per 100,000 in Switzerland, 2000–2016; weekly wILI percentages in the United States, 1998–2017. Off-season weeks are omitted in the plot, with grey lines delimiting the different seasons.

analytical approximations from Section 4 as small black dots. To compute these approximations, we just plugged the empirical mean vectors and covariance matrices into equations (2) – (4). As can be seen, the agreement with the bootstrap results is very good.

In all four countries, the thresholds obtained when using a log transformation as in the MEM are considerably higher than when no transformation is used. Especially in France and Spain, thresholds obtained without the log transformation are clearly too low, as can be seen from the large proportion of seasons classified as very high intensity (around 10% when using $n = 1$). Applying the log transformation mitigates this to some degree.

As expected, when letting the number of observations used per season depend on the number of available seasons via $n = 30/m$, average thresholds increase in m . As a particularly striking example, the average threshold for high intensity for France when using a log transformation increase from 834 for $m = 5, n = 6$ to 1011 for $m = 10, n = 3$ and 1080 for $m = 15, n = 2$. For $n = 1$, in which case the estimators can be interpreted as unbiased (Section 4), we obtain an average of 1147. This corresponds to an increase of roughly 35% relative to the thresholds obtained with $n = 6$. Including historical observations which are not from actual peak weeks thus leads to a considerable lowering of alarm thresholds and will necessarily increase the number of alerts for high and very high influenza activity. As an example, in Spain, applying the MEM with log transformation and $M = 5, n = 6$ leads to 25% of all seasons exceeding the high threshold and 11% exceeding the very high threshold. This is obviously rather far from the intended exceedance probabilities of 10% and 2.5% for the two thresholds.

The increase of thresholds as more historical data accumulates is not only a statistical tendency, but occurs with very high probability. Indeed, for all four countries, the probability that the threshold for high intensity increases when using the first ten rather than only the first five years of a sequence was around 90%. To illustrate this aspect further, we examine the following toy example: We repeat a sequence of five seasons of ILI incidence data from France (2014–2019) four times to obtain a time series of twenty seasons. We then apply the moving epidemic method with log transformation and $n = 30/m$ using the first $m = 5, 6, \dots, 20$ years of data, thus mimicking the development of thresholds as more historical data becomes available. As can be seen from Figure 4, this results in a quite steady increase of thresholds,

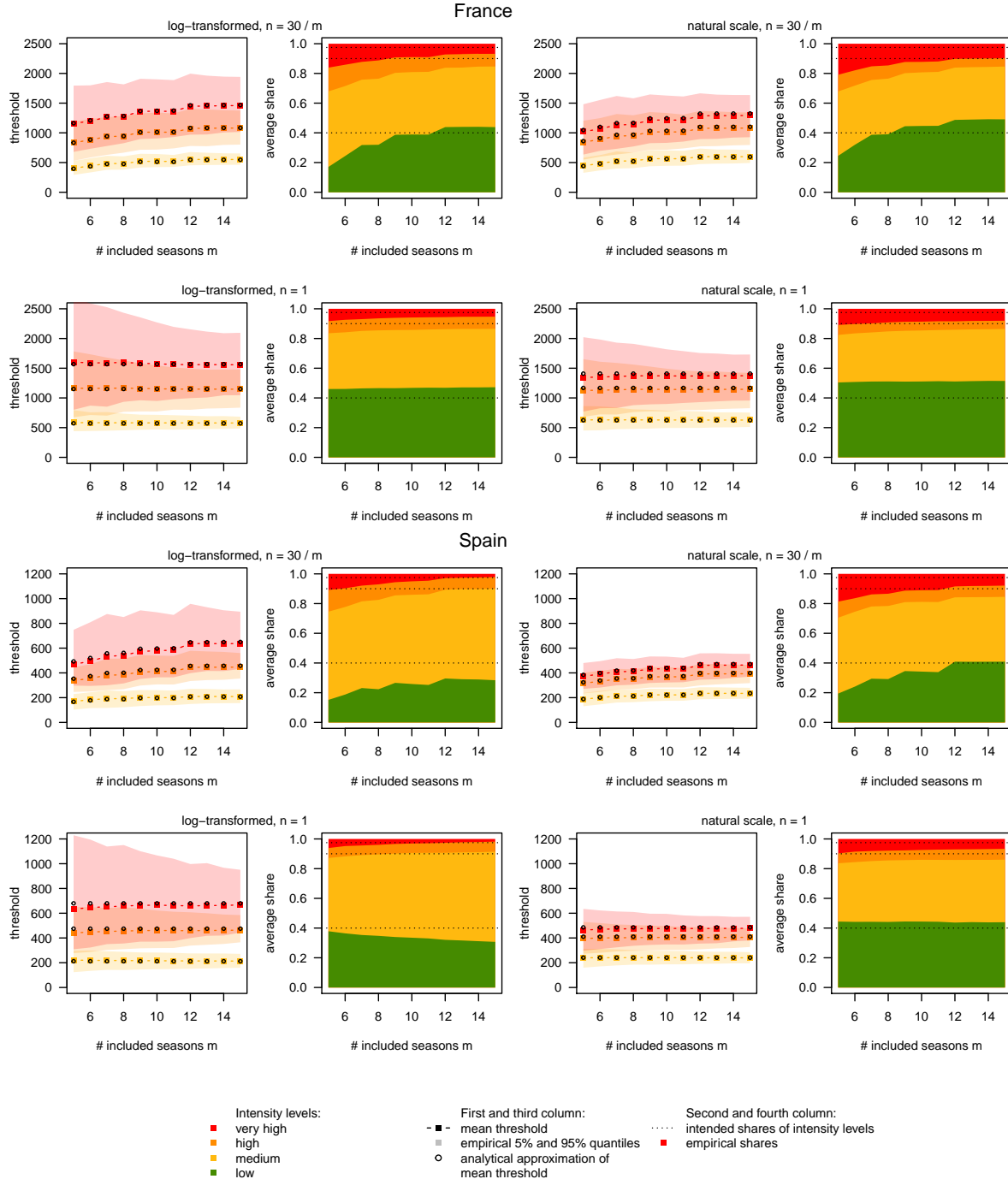


Figure 2: Results of bootstrap simulation study for France and Spain. First and third column: mean intensity thresholds along with bands delimited by the empirical 5% and 95% quantiles. The approximate mean threshold values obtained analytically are marked as little black dots. Second and fourth columns: resulting average shares of seasons classified as low, medium, high and very high intensity (shown as in a stacked barplot).

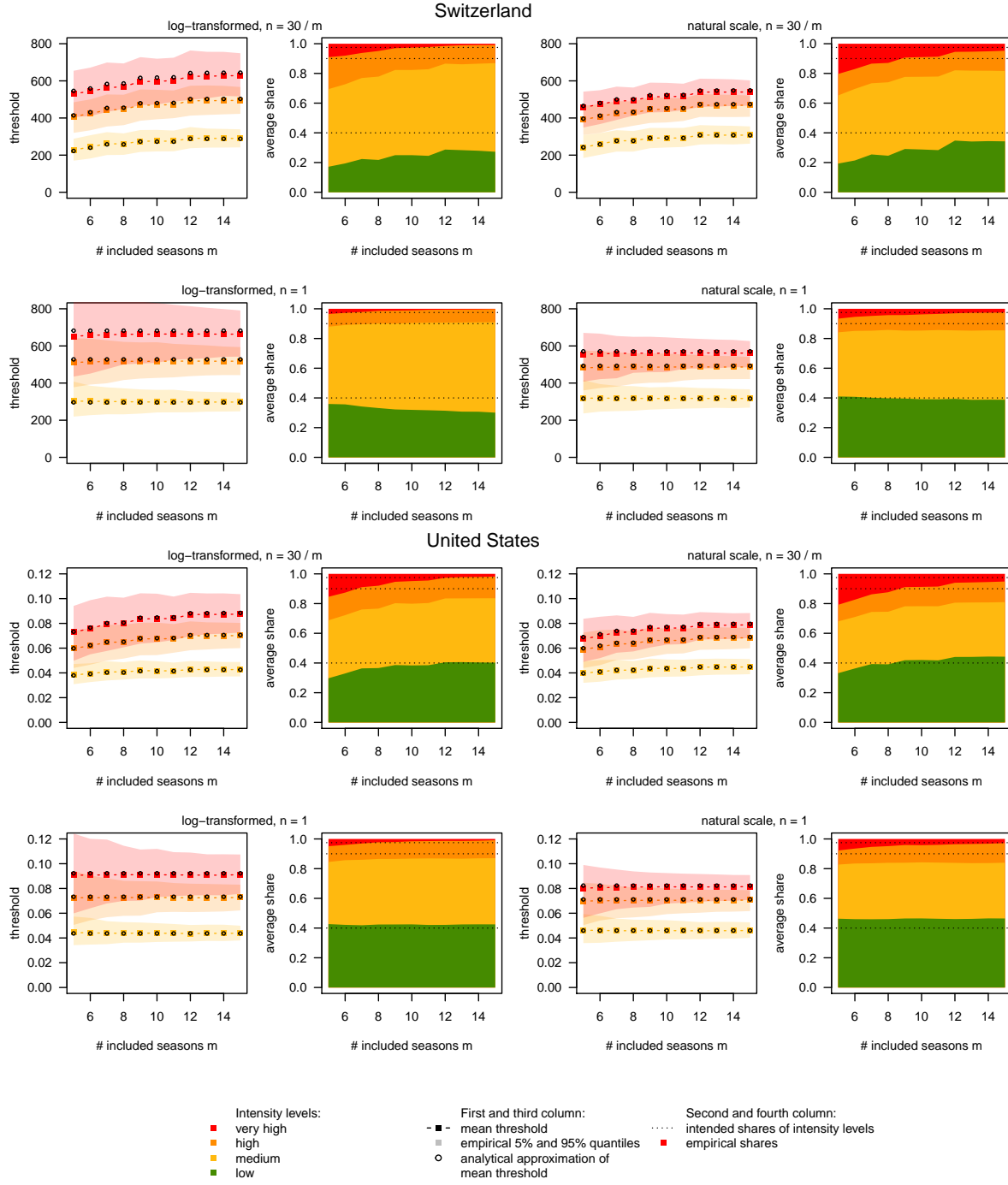


Figure 3: Results of bootstrap simulation study for Switzerland and the United States. First and third column: mean intensity thresholds along with bands delimited by the empirical 5% and 95% quantiles. The approximate mean threshold values obtained analytically are marked as little black dots. Second and fourth columns: resulting average shares of seasons classified as low, medium, high and very high intensity (shown as in a stacked barplot).

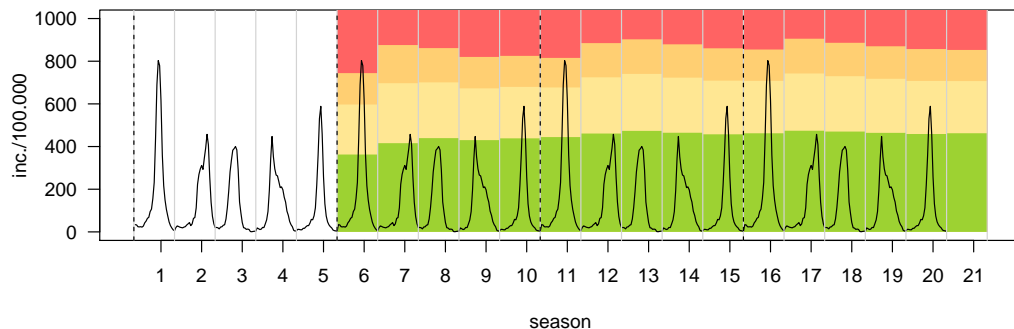


Figure 4: Toy example to illustrate increasing thresholds as historical data accumulates: Data from France, 2014/2015–2018/2019 are appended three times and thresholds are computed using $m = 5, \dots, 15$ years of training data. The thresholds based on years 1 through m are overlaid with the incidence time series from the $(m + 1)$ -th year.

despite the fact that the data in each five-year block behave exactly the same. For example, season 6 is classified as very high intensity based on seasons 1 through 5, but seasons 11 and 16 are not, even though they are identical to season 6 and the respective data used to construct thresholds follow exactly the same patterns. A similar discrepancy can be observed for season 7, which is classified as medium intensity, while seasons 12 and 17 are marked as low intensity.

When just using $n = 1$ irrespective of the number of historical seasons available, the average thresholds and shares of the different categories are more well-behaved also for small m . Of course, there is considerable variability in the estimated thresholds. This difficulty, however, is inherent in the problem of estimating a 90% or 97.5% quantile based on only 5–10 observations.

6 Discussion

We provided a statistical assessment of different implementation choices in a simple and widely used framework for the computation of influenza intensity thresholds. We found that using a log transformation to intensity measures prior to the computation of thresholds empirically led to better behaved thresholds with closer to nominal exceedance rates. Concerning the question of how many observations per historical season should be included in the computation of thresholds, we provided analytical arguments and empirical evidence that the common choice $n = 30/m$ leads to too low thresholds when few historical seasons are available. We therefore recommend to adopt $n = 1$ irrespective of the number M of available historical seasons. This is possible using the available implementation in the R package `mem` and only requires setting `i.n.max = 1` in the call of the `memmodel` to override the default value 30/ m .

We focused on the computation of influenza intensity thresholds, but the same general arguments apply to the computation of thresholds to determine the onset of a season (Vega et al., 2012). As variability among observations around the season onset may be lower than around the peak the problem may be less severe; we did not assess this empirically as not all our data sets fully covered the off-season.

More sophisticated methods to determine influenza intensity thresholds could be imagined, e.g. based on the broad literature on outbreak and anomaly detection (Unkel et al., 2012). However, we think that in practice a simple and interpretable tool with a robust software implementation like the MEM is a good choice. The emergence of a standard approach at the European level and beyond will contribute to improved comparability of results and facilitate the evaluation and refinement of the applied tool. With this work we hope to contribute a statistical perspective on this topic, which can complement public health practitioners’ experience from applied analyses.

References

AbdElGawad, B., Vega, T., El Houssinie, M., Mohsen, A., Fahim, M., Abu ElSood, H., Jabbour, J., Eid, A., and Refaey, S. (2020). Evaluating tools to define influenza baseline and threshold values using

- surveillance data, Egypt, season 2016/17. *Journal of Infection and Public Health*, 13(3):430 – 437.
- Bangert, M., Gil, H., Oliva, J., Delgado, C., Vega, T., De Mateo, S., and LArrauri, A. (2017). Pilot study to harmonize the reported influenza intensity levels within the Spanish influenza sentinel surveillance system (SISSS) using the Moving Epidemic Method (MEM). *Epidemiology and Infection*, 145(4):715–722.
- Basile, L., Oviedo de la Fuente, M., Torner, N., Martínez, A., and Jané, M. (2018). Real-time predictive seasonal influenza model in Catalonia, Spain. *PLOS ONE*, 13(3):1–15.
- Basile, L., Torner, N., Martínez, A., Mosquera, M., Marcos, M., and Jane, M. (2019). Seasonal influenza surveillance: Observational study on the 2017–2018 season with predominant B influenza virus circulation. *Vacunas*, 20(2):53 – 59.
- Benedetti, G., White, R. A., Pasquale, H. A., Stassijns, J., van den Boogaard, W., Owiti, P., and Van den Bergh, R. (2019). Identifying exceptional malaria occurrences in the absence of historical data in South Sudan: a method validation. *Public Health Action*.
- Biggerstaff, M., Kniss, K., Jernigan, D. B., Brammer, L., Bresee, J., Garg, S., Burns, E., and Reed, C. (2017). Systematic assessment of multiple routine and near real-time indicators to classify the severity of influenza seasons and pandemics in the United States, 2003–2004 through 2015–2016. *American Journal of Epidemiology*, 187(5):1040–1050.
- Bouguerra, H., Boutouria, E., Zorruga, M., Cherif, A., Yazidi, R., Abdeddaïem, N., Maazaoui, L., El-Moussi, A., Abid, S., Amine, S., Bouabid, L., Bougatef, S., Kouni Chahed, M., Ben Salah, A., Bettaieb, J., and Bouafif Ben Alaya, N. (2020). Applying the moving epidemic method to determine influenza epidemic and intensity thresholds using influenza-like illness surveillance data 2009–2018 in Tunisia. *Influenza and Other Respiratory Viruses*, 14(5):507–514.
- Charbonneau, D. H. and James, L. N. (2019). FluView and FluNet: Tools for influenza activity and surveillance. *Medical Reference Services Quarterly*, 38(4):358–368. PMID: 31687905.
- Dahlgren, F. S., Shay, D. K., Izurieta, H. S., Forshee, R. A., Wernecke, M., Chillarige, Y., Lu, Y., Kelman, J. A., and Reed, C. (2018). Evaluating oseltamivir prescriptions in centers for medicare and medicaid services medical claims records as an indicator of seasonal influenza in the United States. *Influenza and Other Respiratory Viruses*, 12(4):465–474.
- Dahlgren, F. S., Shay, D. K., Izurieta, H. S., Forshee, R. A., Wernecke, M., Chillarige, Y., Lu, Y., Kelman, J. A., and Reed, C. (2019). Patterns of seasonal influenza activity in U.S. core-based statistical areas, described using prescriptions of oseltamivir in Medicare claims data. *Epidemics*, 26:23 – 31.
- Dickson, E. M., Marques, D. F., Currie, S., Little, A., Mangin, K., Coyne, M., Reynolds, A., McMenamin, J., and Yirrell, D. (2020). The experience of point-of-care testing for influenza in scotland in 2017/18 and 2018/19 – no gain without pain. *Eurosurveillance*, 25(44).
- ECDC (2017). Risk assessment for seasonal influenza, EU/EEA, 2017–2018. Available online at https://www.ecdc.europa.eu/sites/default/files/documents/RRA%20seasonal%20influenza%20EU%20EEA%202017-2018-rev_0.pdf (last accessed 27 Dec 2020).
- Elhakim, M. M., Kandil, S. K., Abd Elaziz, K. M., and Anwar, W. A. (2019). Epidemiology of severe acute respiratory infection (SARI) cases at a sentinel site in Egypt, 2013–15. *Journal of Public Health*, 42(3):525–533.
- Flahault, A., Blanchon, T., Dorléans, Y., Toubiana, L., Vibert, J. F., and Valleron, A. J. (2006). Virtual surveillance of communicable diseases: a 20-year experience in France. *Statistical Methods in Medical Research*, 15(5):413–421. PMID: 17089946.
- Green, H. K., Charlett, A., Moran-Gilad, J., Fleming, D., Durnall, H., Thomas, D. R., Cottrell, S., Smyth, B., Kearns, C., Reynolds, A. J., and et al. (2015). Harmonizing influenza primary-care surveillance in the United Kingdom: piloting two methods to assess the timing and intensity of the seasonal epidemic across several general practice-based surveillance schemes. *Epidemiology and Infection*, 143(1):1–12.

- Grilc, E., Prosenc Trilar, K., Lajovic, J., and Sočan, M. (2020). Determining the seasonality of respiratory syncytial virus in Slovenia. *Influenza and Other Respiratory Viruses*, n/a(n/a).
- Harcourt, S. E., Morbey, R. A., Smith, G. E., Loveridge, P., Green, H. K., Pebody, R., Rutter, J., Yeates, F. A., Stuttard, G., Elliot, A. J., and et al. (2019). Developing influenza and respiratory syncytial virus activity thresholds for syndromic surveillance in England. *Epidemiology and Infection*, 147:e163.
- Held, L. and Meyer, S. (2019). Forecasting based on surveillance data. In Held, L., Hens, N., O’Neill, P. D., and Wallinga, J., editors, *Handbook of Infectious Disease Data Analysis*. Chapman and Hall/CRC.
- Lozano, J. (2020). mem: The moving epidemic method. R package, version 2.16 available via CRAN, <https://cran.r-project.org/web/packages/mem/index.html>.
- Lozano, J. and Vega, T. (2018). The moving epidemic method – the MEM web application: technical manual. available online at https://drive.google.com/file/d/0B0IUo_0NhT0oX29zc2p5RmlBUWc/view. R package memapp available at <https://cran.r-project.org/web/packages/memapp/>. Last accessed 17 Jan 2020.
- Lucero, M. G., Inobaya, M. T., Nillos, L. T., Tan, A. G., Arguelles, V. L. F., Dureza, C. J. C., Mercado, E. S., Bautista, A. N., Tallo, V. L., Barrientos, A. V., Rodriguez, T., and Olveda, R. M. (2016). National influenza surveillance in the Philippines from 2006 to 2012: seasonality and circulating strains. *BMC Infectious Diseases*, 16(1):762.
- Ly, S., Arashiro, T., Ieng, V., Tsuyuoka, R., Parry, A., Horwood, P., Heng, S., Hamid, S., Vandemaele, K., Chin, S., Sar, B., and Arima, Y. (2017). Establishing seasonal and alert influenza thresholds in Cambodia using the WHO method: implications for effective utilization of influenza surveillance in the tropics and subtropics. *Western Pacific Surveillance and Response Journal*, 8:22–32.
- Murray, J. L. K., Marques, D. F. P., Cameron, R. L., Potts, A., Bishop, J., von Wissmann, B., William, N., Reynolds, A. J., Robertson, C., and McMenamin, J. (2018). Moving Epidemic Method (MEM) applied to virology data as a novel real time tool to predict peak in seasonal influenza healthcare utilisation. the Scottish experience of the 2017/18 season to date. *Eurosurveillance*, 23(11):18–00079.
- Nisar, N., Aamir, U. B., Badar, N., Mahmood, M. R., Yaqoob, A., Tripathy, J. P., Laxmeshwar, C., Tenzin, K., Zaidi, S. S. Z., Salman, M., and Ikram, A. (2020). Epidemiology of influenza among patients with influenza-like illness and severe acute respiratory illness in Pakistan: A 10-year surveillance study 2008–17. *Journal of Medical Virology*, 92(12):3028–3037.
- Pesälä, S., Virtanen, M. J., Mukka, M., Ylilammi, K., Mustonen, P., Kaila, M., and Helve, O. (2019). Healthcare professionals’ queries on oseltamivir and influenza in Finland 2011–2016 – can we detect influenza epidemics with specific online searches? *Influenza and Other Respiratory Viruses*, 13(4):364–371.
- Politis, D. (2001). Resampling time series with seasonal components. *Frontiers in Data Mining and Bioinformatics: Proceed-ings of the 33rd Symposium on the Interface of Computing Science and Statistics: Orange County, California, June 13–17*, page 619–621.
- Páscoa, R., Rodrigues, A. P., Silva, S., Nunes, B., and Martins, C. (2018). Comparison between influenza coded primary care consultations and national influenza incidence obtained by the General Practitioners Sentinel Network in Portugal from 2012 to 2017. *PLOS ONE*, 13(2):1–10.
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rakocevic, B., Grgurevic, A., Trajkovic, G., Mugosa, B., Grujicic, S. S., Medenica, S., Bojovic, O., Alonso, J. E. L., and Vega, T. (2019). Influenza surveillance: determining the epidemic threshold for influenza by using the Moving Epidemic Method (MEM), Montenegro, 2010/11 to 2017/18 influenza seasons. *Eurosurveillance*, 24(12):1800042.
- Redondo-Bravo, L., Delgado-Sanz, C., Oliva, J., Vega, T., Lozano, J., Larrauri, A., and the Spanish Influenza Sentinel Surveillance System (2020). Transmissibility of influenza during the 21st-century epidemics, Spain, influenza seasons 2001/02 to 2017/18. *Eurosurveillance*, 25(21).

- Rguig, A., Cherkaoui, I., McCarron, M., Oumzil, H., Triki, S., Elmbarki, H., Bimouhen, A., Falaki, F. E., Regragui, Z., Ihazmad, H., Nejari, C., and Youbi, M. (2020). Establishing seasonal and alert influenza thresholds in Morocco. *BMC Public Health*.
- Sistema de Vigilancia de Gripe en España (2019). Guía de procedimientos para la vigilancia de gripe en España. Available online at https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Documents/GRIPE/GUIAS/Gu%C3%ADa%20de%20procedimientos%20para%20la%20vigilancia%20de%20gripe%20en%20Espa%C3%B1a_marzo%202019.pdf, last accessed 13 January 2020.
- Sullivan, S. G., Arriola, C. S., Bocacao, J., Burgos, P., Bustos, P., Carville, K. S., Cheng, A. C., Chilver, M. B., Cohen, C., Deng, Y.-M., El Omeiri, N., Fasce, R. A., Hellferscee, O., Huang, Q. S., Gonzalez, C., Jelley, L., Leung, V. K., Lopez, L., McAnerney, J. M., McNeill, A., Olivares, M. F., Peck, H., Sotomayor, V., Tempia, S., Vergara, N., von Gottberg, A., Walaza, S., and Wood, T. (2019). Heterogeneity in influenza seasonality and vaccine effectiveness in Australia, Chile, New Zealand and South Africa: early estimates of the 2019 influenza season. *Eurosurveillance*, 24(45).
- Tay, E. L., Grant, K., Kirk, M., Mounts, A., and Kelly, H. (2013). Exploring a proposed WHO method to determine thresholds for seasonal influenza surveillance. *PLOS ONE*, 8(10).
- Torner, N., Basile, L., Martínez, A., Rius, C., Godoy, P., Jané, M., Domínguez, A., Aizpurua, J., Alonso, J., Azemar, J., Aizpurua, P., Ardaya, P. M., Basas, M. D., Batalla, J., Biendicho, P., Bonet, M., Callado, M., Campos, S., Casanovas, J. M., Ciurana, E., Clapes, M., Cots, J. M., De la Rica, D., Domingo, I., Elizalde, G., Escapa, P., Fajardo, S., Fau, E., Fernandez, O., Fernandez, M., Ferrer, C., Forcada, A., Fos, E., Gadea, G., Garcia, J., Garcia, R., Gatiús, C., Gelado, M. J., Grau, M., Grivé, M., Guzman, M. C., Hernández, R., Jimenez, G., Juscafresa, A., LLussa, A. M., López, C., Kristensen, L., Macià, E., Mainou, A., Marco, E., Martínez, M., Martínez, J. G., Marulanda, K. V., Masa, R., Moncosí, X., Naranjo, M. A., Navarro, D., Ortolà, E., París, F., Pérez, M. M., Pozo, C., Pujol, R., Ribatallada, A., Ruiz, G., Sabaté, S., Sanchez, R., Sarrà, N., Tarragó, E., Teixidó, A. M., Torres, A., Valén, E., Van Esso, D., Van Tarjwick, C., Vink Schoenholzer, R., Zabala, E., Marcos, M. A., Mosquera, M. D. M., de Molina, P., Rubio, E., Isanta, R., Anton, A., Pumarola, T., Vilella, A., Gorrindo, P., Espejo, E., Andrés, M., Barcenilla, F., Navarro, G., Barrabeig, I., Pou, J., Alvarez, P., Plasencia, E., Rebull, J., Sala, M. R., Riera, M., Camps, N., Follia, N., Oller, A., Godoy, P., Bach, P., Rius, C., Hernández, R., Perez, R., Torra, R., Carol, M., Minguell, S., Marce, R., Garcia-Pardo, G., Olona, M., Alvarez, A., Ramon, J. M., Mòdol, J. M., Mena, G., Campins, M., Massuet, C., Tora, G., Ferràs, J., Ferrús, G., and on PIDIRAC Sentinel Surveillance of Catalonia, T. W. G. (2019). Assessment of two complementary influenza surveillance systems: sentinel primary care influenza-like illness versus severe hospitalized laboratory-confirmed influenza using the moving epidemic method. *BMC Public Health*, 19(1089).
- Unkel, S., Farrington, C. P., Garthwaite, P. H., Robertson, C., and Andrews, N. (2012). Statistical methods for the prospective detection of infectious disease outbreaks: a review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(1):49–82.
- Vega, T., Lozano, J., Meerhoff, T., and the Working Group on Influenza Qualitative Indicators (2012). Quantification of qualitative influenza indicators used in euroflu and tessy. Available online at <https://www.saludcastillayleon.es/profesionales/es/centinelas/informe-indicadores-cualitativos-actividad-gripal.ficheros/317117-DOC%201%20FINAL%20TECHNICAL%20REPORT%200N%202012%2011%2022.pdf>.
- Vega, T., Lozano, J. E., Meerhoff, T., Snacken, R., Beauté, J., Jorgensen, P., Ortiz de Lejarazu, R., Domegan, L., Mossong, J., Nielsen, J., Born, R., Larrauri, A., and Brown, C. (2015). Influenza surveillance in Europe: comparing intensity levels calculated using the Moving Epidemic Method. *Influenza and Other Respiratory Viruses*, 9(5):234–246.
- Vega, T., Lozano, J. E., Meerhoff, T., Snacken, R., Mott, J., Ortiz de Lejarazu, R., and Nunes, B. (2013). Influenza surveillance in Europe: establishing epidemic thresholds by the Moving Epidemic Method. *Influenza and Other Respiratory Viruses*, 7(4):546–558.
- Vette, K., Bareja, C., Clark, R., and Lal, A. (2018). Establishing thresholds and parameters for pandemic influenza severity assessment, Australia. *Bulletin of the World Health Organization*, 96:558–567.

- Vos, L. M., Teirlinck, A. C., Lozano, J. E., Vega, T., Donker, G. A., Hoepelman, A. I., Bont, L. J., Oosterheert, J. J., and Meijer, A. (2019). Use of the moving epidemic method (mem) to assess national surveillance data for respiratory syncytial virus (rsv) in the netherlands, 2005 to 2017. *Eurosurveillance*, 24(20):1800469.
- Wagner, M., Lamos, V., Cox, I. J., and Pebody, R. (2018). The added value of online user-generated content in traditional methods for influenza surveillance. *Scientific Reports*, 8(1):13963.
- WHO (2011). Strengthening response to pandemics and other public-health emergencies: Report of the review committee on the functioning of the international health regulations (2005) and on pandemic influenza (h1n1) 2009. Available online at https://www.who.int/ihr/publications/RC_report/en/ (accessed 26 Dec 2020)).
- WHO (2014). Who global epidemiological surveillance standards for influenza. Available online at https://www.who.int/influenza/resources/documents/influenza_surveillance_manual/en/ (last accessed 27 December 2020).
- WHO (2017). Pandemic influenza severity assessment (PISA). Available online at <https://apps.who.int/iris/handle/10665/259392> (accessed 27 Dec 2020).