Time-series modeling for the quantification of seasonality and forecasting antibiotic-resistant episodes: application to carbapenemase-producing Enterobacteriaceae episodes in France over 2010-2020

Jonathan ROUX[1], Narimane NEKKAB[2,3], Mélanie COLOMB-COTINAT[4], Pascal ASTAGNEAU[5,6], Pascal CREPEY*[,1]

[1] Univ Rennes, EHESP, REPERES (Pharmacoepidemiology and health services research) - EA 7449, Rennes, France

[2] Laboratoire MESuRS, Conservatoire National des Arts et Métiers, Paris, France

[3] Institut Pasteur, Cnam, Unité PACRI, Paris, France

[4] Santé publique France, Saint-Maurice, France

[5] Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Paris, France

[6] Centre de prévention des infections associées aux soins, Paris, France

Running title

Modeling of CPE in France over 2010-2020

Corresponding author

Pascal Crépey

EHESP – Département METIS

15, Avenue du Professeur Léon-Bernard

CS 74312

35043 RENNES Cedex

FRANCE

Mail: pascal.crepey@ehesp.fr

Phone number: +33 (0)2.99.02.28.07

Synopsis

Background Carbapenemase-producing Enterobacteriaceae (CPE) cause resistant healthcare-associated infections that jeopardize healthcare systems and patient safety worldwide. The number of CPE episodes has been increasing in France since 2009 but the dynamics are still poorly understood.

Objectives Use time-series modeling to describe the dynamics of CPE episodes from August 2010 to December 2016 and to forecast its evolution for the 2017-2020 period.

Methods We used time-series to analyze CPE episodes from August 2010 to November 2016 reported to the French national surveillance system. The impact of seasonality was quantified using seasonal-to-irregular ratios. Seven time-series models and three ensemble stacking models (average, convex and linear stacking) were assessed and compared to forecast CPE episodes during 2017-2020.

Results During 2010-2016, 3559 CPE episodes were observed in France. Compared to the average yearly trend, we observed a 30% increase in the number of CPE episodes in Autumn. We noticed a 1-month lagged seasonality of non-imported episodes compared to imported episodes. Average stacking gave the best forecasts and predicted an increase during 2017-2020 with a peak up to 345 CPE episodes (95% PI[124-1158], 80% PI[171-742]) in September 2020.

Conclusions The observed seasonality of CPE episodes sheds light on potential factors associated with the increased frequency of episodes which need further investigations. Our model predicts that the number of CPE episodes will continue to rise in the coming years in France, mainly due to local dissemination, associated with bacterial carriage of patients in the community which is becoming an immediate challenge to help controlling the outbreak.

Introduction

The increasing incidence of healthcare-associated infections caused by carbapenemase-producing Enterobacteriaceae (CPE) jeopardizes healthcare systems and patient safety in Europe,[1] the US, and many countries worldwide.[2] Because carbapenems remain the main antibiotic for the treatment of multi-resistant bacteria infections, CPE infections are associated with longer hospital stays and an excess risk of death.[3,4] The primary risk factors of CPE's spread involve consumption of broad-spectrum antibiotics (carbapenems, third and fourth generation cephalosporins and fluoroquinolones), cross-infection during hospital stays, and cross-border transfers of patients in healthcare settings.[5] In parallel, plasmidic gene transfers between bacteria contribute to the rising number of non-internationally imported cases.[6]

Despite reinforced national guidelines and strategies regarding the management and prevention of emerging CPE,[5,7] the number of episodes in France has risen steadily since 2009.[8,9] This phenomenon seems to be associated with inter-regional dissemination and importation of international cases as shown in several recent studies.[1,10–13] Moreover, in France CPE has advanced from its epidemiological stage 3 (regional spread) in 2013 to stage 4 (inter-regional spread) in the 2014-2015 period.[1,14] Despite a growing concern regarding this issue, the majority of studies conducted so far have been limited to the epidemiological description at the national level[8,9,13] and the reporting of specific local outbreaks,[10–12,15,16] with especially few information on the seasonality of episodes. Common methods used to understand and predict the dynamics of infectious diseases involve both compartmental and agent-based models.[17] These methods, however, either require a deep knowledge of the transmission pathways used by the pathogens, or rely on assumptions that are difficult to validate. Time-series analysis, on the other hand, appears to be easier to apply and has garnered much interest in the field, especially in modeling influenza dynamics[18–20] or antimicrobial resistance.[21] These methods rely on the identification of temporal patterns, with few assumptions modeling the mechanism of how CPE is spread. Time-series analysis can therefore be effectively used to describe and quantify the trend and seasonality of CPE episodes incidence, as currently little is known on the matter. In addition to providing deeper insights into the phenomenon, this approach also allows relatively simple forecasts of the number of cases, which could help public health authorities to better define and evaluate infection control guidelines.

Time-series analysis was thus used to firstly describe and quantify the dynamics (trend and seasonality) of CPE episodes from August 2010 to December 2016. The second objective was to identify a methodological process using time-series modeling to forecast the evolution of CPE episodes for the 2017-2020 period.

Materials and methods

Data sources

Surveillance data of CPE episodes notified between August 2010 and December 2018 were extracted from the French national Healthcare-Associated Infections Early Warning and Response System (HAI-EWRS).[22] An episode was defined as a single case or a cluster of cases carrying the same strain of CPE, and known to have been in contact with one another.[13] A case was defined as a positive CPE diagnosis from sample collected for infection or through systematic screening performed according to standard national recommendations.[7] Available characteristics on each episode were the occurrence date, the index case status (infected or carrier), and the importation status, i.e. the presence of a direct link with a foreign country for the potential index case (hereafter denominated as imported or in the case of no link, non-imported). Information on the mechanism of CPE resistance confirmed by the national reference center, and the involved bacteria species were also collected. CPE episodes were grouped according to month of occurrence and analyzed together and separately according to their importation status.

Since the HAI-EWRS collection system changed in January 2017 (international importation status of the index case not collected in the same way, missing data on the infection status of the index case), we decided not to include the period 2017-2018 in the main analysis but to use it as a control set for the prediction model (Figure 1B). In addition, during the descriptive phase of this work, the number of episodes in December 2016 appeared to be an outlier compared to previous months, possibly due to an “end-of-year” reporting issue before the new collection system. The main study period was thus set from August 2010 to November 2016. To build and validate the different models, data was split into two datasets: the training set from August 2010 to December 2014 and the test set from January 2015 to November 2016 (Figure 1A).

Figure 1 here

Statistical analysis

We considered three main time-series in this study: the first was composed of imported episodes, the second one of non-imported episodes and the third of all episodes, whatever their importation status. All considered time-series were multiplicative, a log transformation was thus used in each model built. To quantify the seasonality of episodes in the three time-series on the whole study period (August 2010-November 2016), we used seasonal-to-irregular (SI) ratios.[23] SI ratios were computed based on the de-trended time-series obtained using the X-11 seasonal adjustment method and corresponded to the product of the seasonal and the irregular parts of the multiplicative time series. Time-series were de-trended using moving average[24] and trends estimates using time-series linear model.[23] As X-11 seasonal adjustment method required non-zero values; one fictional episode was added each month over the study period for the non-imported time-series.

To forecast the evolution of the number of CPE episodes, we considered 10 methods issued from time-series modeling and ensemble methods. We built seven time-series models, hereafter called component models. These included: a seasonal autoregressive integrated moving average (SARIMA) model,[25] a time-series linear model (tslm),[23] a X-11 seasonal adjustment method,[24] an exponential smoothing state space (ETS) model,[26,27] a multiplicative Holt-Winters method with multiplicative errors,[26,27] a neural network autoregression (NNAR),[23] and a TBATS model (exponential smoothing state space model with Box-Cox transformation, ARMA (autoregressive-moving-average) errors, trend and seasonal components).[28] All these components models were fitted independently on the training set (August 2010-December 2014) and then used to forecast the number of CPE episodes over the test period (January 2015-November 2016). In addition, we implemented three ensemble stacking methods, which combined the fitted and forecasted values of the seven component models. First, an average stacking model was defined considering all combinations of component models. The retained combination was the one with the best predictive quality on the training set according to the mean absolute error (MAE) quality parameter.[29] Next, we used a convex stacking model composed of the best linear combination of the seven component models with non-negative coefficients summing to one.[29] Finally, we built a linear stacking model which estimated the best coefficients of the linear combination of all component models.[30] Both the convex and the linear stacking model were computed using the square loss as a performance criterion.[29] For each component model, 80% and 95% prediction intervals (PIs) were estimated based on the standard deviation of each step forecast.[23] For stacking methods, as no consensus exists to our knowledge, we carried out a linear combination of the PIs by using the stacking coefficients.

To select the best model to be used for forecasting, we assessed the quality of each of the 10 models compared to the observed data of training and test sets using MAE, mean absolute percentage error (MAPE), and root mean squared error (RMSE).[23] Forecasts being on the same scale, minimization of MAE was selected as the main selection criterion as suggested by Hyndman.[23] The model with the best forecasting quality on the test set was then trained on all available data (2010-November 2016) and used to predict CPE episodes from December 2016 to December 2020.

Analyses were performed using R (3.5.2)[31] and the packages forecast,[32] opera,[29] and seasonal.[24]

Results

Characteristics of CPE episodes

A total of 3559 CPE episodes – 1473 (41.4%) over the training period and 2086 (58.6%) over the test period – were reported in France between August 2010 and November 2016 (46 episodes occurring in December 2016 were excluded from the analyses). Out of these, 1624 (45.6%) episodes were internationally-imported cases and 1935 (54.4%) had no documented link with a foreign country (non-imported episodes). Among the 769 episodes with an infected index case, 280 (36.4%) were imported, whereas 1,350 (49.0%) were imported among the 2753 episodes with a carrier index case ($p<0.05$) (83 episodes had a missing index case status). The majority of CPE episodes had an OXA-48 (class D beta-lactamases with oxacillinase enzyme activity) mechanism of resistance with a total of 2684 (75.4%) episodes. NDM (New Delhi metallo-beta-lactamase) resistance corresponded to 541 (15.2%) episodes, KPC (Klebsiella pneumoniae Carbapenemase) to 168 (4.7%), and VIM (Verona integron-encoded metallo-beta-lactamase) to 145 (4.1%) episodes. The remaining 21 episodes (0.6%) had no reported resistance mechanism. *Klebsiella* *pneumoniae* and *Escherichia coli* were the two main bacteria species involved in the episodes (1915 episodes (53.8%) and 1335 episodes (37.5%), respectively).

Quantifying the seasonal effect

The SI ratios chart was obtained using the X-11 seasonal adjustment method, enabling to quantify the impact of seasonality on the number of CPE episodes during the 2010-2016 period (Figure 2). Across the years, the number of episodes reported was higher in Autumn and lower at the beginning of each year. The seasonal increase of the number of CPE episodes was 30% and 29% higher in September and October respectively, as compared to the average trend. In contrast, a 20% decrease in the number of CPE episodes was observed in February compared to the other months.

Results of the analysis by importation status using X-11 seasonal adjustment method are presented in Figure 3. The number of non-imported episodes appeared to grow faster than the imported ones. Indeed, we found trends equal to 0.59 and 0.88 for the imported and non-imported episodes, respectively, using the time-series linear model over the whole study period. When only considering data starting from 2012 (no fictional episode added in the non-imported time series), these trends were even higher and respectively equal to 0.67 and 1.01. A decrease in the trend for non-imported episodes was observed at the end of the period mainly due to the low number of episodes reported in November 2016. SI charts highlighted a 1-month lagged seasonal impact of the non-imported compared to imported episodes. Indeed, the number of imported CPE episodes increased by around 33% in August and September, while the peak was observed in October for non-imported episodes with an average rise of 33% compared to other months. We also looked at the seasonality of the two main episodes strains split according to their importation status (Supplementary figure 1). *K. pneumoniae* seemed to drive the seasonality of imported episodes with a 69% increase of the number of episodes in September. On the opposite, *E. coli* drove the non-imported episodes with a 75% increase of the number of episodes related to this strain in October.

Figure 2 here

Figure 3 here

Prediction of episodes evolution

Accuracy parameters over the training and test sets of the seven component models and the three stacking methods are presented in Table 1. Aside from the stacking techniques, TBATS appeared to be the best method to model the training set (MAE=3.35), but was less accurate than X-11 seasonal adjustment on the test set (MAE=13.08 versus 12.86 respectively). All three stacking techniques provided better fitted values on the training set, as compared to the component models. Linear stacking had the best adjustment on the training set (MAE=2.65). On the test period, average stacking method based on three component models (X-11 seasonal adjustment, multiplicative Holt-Winters method, and TBATS) produced the most accurate forecasts (MAE=12.65), performing better than X-11 seasonal adjustment alone.

Table 1 here

Based on its performance on the test set, we used the average stacking method to forecast the number of CPE episodes during 2017-2020. This method was trained on all episodes occurring over the August 2010-November 2016 period, and retained only two component models: X-11 seasonal adjustment method and ETS. The obtained forecasts and PI are presented in Figure 4. The number of CPE episodes was predicted to increase over the next 4 years in France with a peak up to 345 episodes in September 2020 (95% PI [124-1158], 80% PI [171-742]). In addition, the model predicted 177 episodes (95% PI [96-316], 80% PI [118-257]), 225 episodes (95% PI [106-480], 80% PI [136-365]), and 278 episodes (95% PI [1153-733], 80% PI [153-514]) in September 2017, September 2018, and September 2019, respectively. The accuracy parameters (RMSE, MAE, and MAPE) comparing the predictions over the 2017-2018 period and the observed values were equal to 25.38, 18.95 and 14.32, respectively. Moreover, the real values observed during control period (2017-2018) were included in the 80% prediction interval (Figure 4).

Figure 4 here

Discussion

This study is the first, to our knowledge, that investigates the dynamics of the spreading of CPE in France, and deploys time-series analysis for this purpose. Using the exhaustive database of notified cases from the beginning of the epidemic, we showed an ongoing increasing trend of CPE episodes and difference of seasonality according to case importation status. Despite wide prediction intervals, the number of CPE episodes is forecasted to continue to grow for the next 4 years.

OXA-48 was the dominant strain of carbapenemase in France, followed by NDM; which differs to what is observed in the US and other European countries.[1,33] Although France has one of the highest antibiotic consumption rates in Europe, the consumption of carbapenems is lower than average and has not significantly increased during the 2012-2016 period; therefore, antibiotic consumption may not explain the rise in the number of notified episodes we observed.[34] While the national public health agency ('Santé publique France') provides a general report on the national epidemiological situation,[8,9] our analyses enhances the results provided thus far by quantifying the seasonality in the number of episodes. Indeed, we revealed a 30% increase in the number of CPE episodes in Autumn, i.e. September and October, compared to other months. Moreover, when stratifying the dataset according to the origin of the episodes, two separate peaks were observed: a 33%-increase peak in August and September for the internationally-imported episodes and a 33%-increase peak in October for the non-imported ones. The peak of imported cases may be due to summer holidays occurring in July and August in France, and thus a consequence of hospitalizations and repatriations from foreign countries. We cannot, however, rule out the possibility that intra-national variations in patient flows linked to summer holidays tend to increase the population at-risk of infection or CPE carriage. The one-month delay in non-imported cases, causing a peak in October, could be linked to secondary cases occurring in healthcare facilities or communities. These hypotheses, however, require individual patient studies in order to be confirmed. It also appeared that there was a difference in seasonality between bacteria strains. This could be a consequence of the endemic status of *K. pneumoniae* in foreign European countries compared to *E. coli*,[35] but this hypothesis may need deeper analyses. In addition, it is more than probable according to national guidelines[7] that patients having experienced hospitalizations in a CPE-endemic foreign country will be more likely to be screened for CPE carriage and reported; a selection bias that could also explain a delay between imported and non-imported cases. Since

2014, the trend of non-imported episodes appeared to grow at a faster rate than that of imported cases. This could be explained by either an increase in community spreading of CPE, or the impossibility of reconstructing chains of episodes between healthcare facilities, which may be due to the inter-regional spread.[1,14]

Using average stacking methodology, our results suggest that the overall number of CPE episodes in France will continue to rise over 2017-2020 with peaks in September each year. This result is consistent with the epidemiological transition of France regarding CPE from stage 3 to 4 in 2014-2015 according to ECDC.[1,14] Despite the wide prediction intervals of the forecasts, the prediction model seems to be robust. Indeed, we showed that there was a high seasonality in the data and an ongoing trend that may not disappear without implementation of specific control measures or an external event.

Regarding statistical analysis, our study relies on a robust methodology aiming at finding the most effective method for prediction. Firstly, seven component models, which are the most frequently used in time-series analysis, were built. Then, three stacking models were implemented based on these component models. These ten methods were compared, to retain the one method that produced predictions, which best fitted the observed data of the test set. The wide variety of models deployed allowed us to consider different underlying generation process of time-series, such as the moving average with SARIMA or the exponential smoothing with ETS or Holt Winter's method. In addition, the use of ensemble techniques improved the quality of predictions, as suggested in the literature,[36] supported by the model's better performance on the quality of parameters, when compared to the single component models. The use of such ensemble methods, however, may suffer from an overfitting bias.[37] To limit this bias, we chose to split the entire dataset into training and test sets, and to select the best model according to the quality parameters on the test set only. The effective performance of our final model in predicting the final control set (period 2017-2018) indicates that such overfitting is limited. In addition, as shown in the results section, stacking methods rely on component models, which themselves rely on data. Different data may thus lead to a different combination of models retained by the stacking method. Therefore, we looked for a replicable methodological process rather than a model to be replicated in other similar studies.

Our study also relies on robust surveillance data. Indeed, due to mandatory reporting of CPE infection and colonization through various healthcare centers and surveillance networks we believe our data accurately reflects the situation of CPE episodes in France over the study period. In addition, the study period stretched over 76 months, leading to training and test sets 53-month and 23-month long, respectively. Because of the seasonality of the data, these lengths were considered long enough to obtain a good fitting of the different models on the data. Moreover, the predictions over 2017-2020 were based on a model trained on the whole dataset, i.e. 76 months. This hypothesis was also confirmed by the control data falling within the prediction intervals. The latter were wide, especially regarding the upper bound, which was mainly due to the ETS component model and the long-term prediction leading to high standard deviations at the end of 3-year forecasting period.

Another limit of our study arising from the data is the low number of non-imported episodes, especially at the beginning of the study period where zeroes were frequent. Because some time-series methods cannot deal with such repeated zeroes, we had to modify the non-imported time-series by adding a fictional episode to each month to ensure focus on the seasonality. This is unlikely to impact our findings regarding the seasonality and global trend of this time-series since the addition of an episode was constant through time. However, this low number may be a consequence of the criteria used to classify an episode as imported. Indeed, an episode was considered as internationally-imported if the first patient had been repatriated or hospitalized abroad in the past 12 months. These criteria may therefore underestimate the total number of non-imported episodes which are incorrectly considered as imported ones, but we assume that may correspond to only a few episodes.

In conclusion, time-series modeling appears to be a useful tool for the study of the spread of antibiotic resistance in both quantifying the seasonality and forecasting. The seasonality of CPE episodes highlighted in our study need to be further investigated, in order to better account for this phenomenon. Prevention and control efforts should be maintained to better control CPE epidemic including reinforcement of information to healthcare professionals to promptly detect CPE cases, especially in non-imported cases, and dedicated financial and human resources to healthcare facilities. In addition, probable spread of CPE in the community is becoming an immediate challenge to help controlling the outbreak. Infection prevention and control should thus be reinforced to avoid autochthonous cases, which would place undue strain on the French healthcare system.

Acknowledgements

The authors would like to thank Anne Berger-Carbonne for facilitating data access and providing comments, and Sam Raven for editing the manuscript.

Funding

This work was supported by the French national research agency (ANR), through the SPHINx (Spread of Pathogens on Healthcare Institutions Networks) project.

Transparency declarations

J.R. reports grants from French National Research Agency, during the conduct of the study. P.C. reports grants from French National Research Agency, during the conduct of the study. All other authors: none to declare.

References

1. European Centre for Disease Prevention and Control. *Rapid risk assessment: Carbapenem-resistant Enterobacteriaceae*. Stockholm: ECDC; 2016.

2. World Health Organization. *Antimicrobial resistance - Global Report on Surveillance*. 2014.

3. Morrill HJ, Pogue JM, Kaye KS, *et al.* Treatment Options for Carbapenem-Resistant Enterobacteriaceae Infections. *Open forum Infect Dis* 2015; **2**: ofv050.

4. Tischendorf J, de Avila RA, Safdar N. Risk of infection following colonization with carbapenem-resistant Enterobactericeae: A systematic review. *Am J Infect Control* 2016; **44**: 539–43.

5. European Centre for Disease Prevention and Control. *Risk assessment on the spread of carbapenemase-producing Enterobacteriaceae (CPE) through patient transfer between healthcare facilities, with special emphasis on cross-border transfer*. Stockholm; 2011.

6. Logan LK, Weinstein RA. The Epidemiology of Carbapenem-Resistant Enterobacteriaceae: The Impact and Evolution of a Global Menace. *J Infect Dis* 2017; **215**: S28–36.

7. Lepelletier D, Berthelot P, Lucet J-C, *et al.* French recommendations for the prevention of 'emerging extensively drug-resistant bacteria'' (eXDR) cross-transmission'. *J Hosp Infect* 2015; **90**: 186–95.

8. Pontiès V, Soing-Altrach S, Savitch Y, *et al. Episodes impliquant des entérobactéries productrices de carbapénèmases (EPC) en France, de 2004 à 2015. Bilan épidémiologique national au 31 décembre 2015*. 2017.

9. Dortet L, Cuzon G, Ponties V, *et al.* Trends in carbapenemase-producing Enterobacteriaceae, France, 2012 to 2014. *Eurosurveillance* 2017; **22**: 10–8.

10. Kassis-Chikhani N, Decré D, Gautier V, *et al.* First outbreak of multidrug-resistant Klebsiella pneumoniae carrying blaVIM-1 and blaSHV-5 in a French University Hospital. *J Antimicrob Chemother* 2006; **57**: 142–5.

11. Kassis-Chikhani N, Decré D, Ichai P, *et al.* Outbreak of Klebsiella pneumoniae producing KPC-2 and SHV-12 in a French hospital. *J Antimicrob Chemother* 2010; **65**: 1539–40.

12. Poirel L, Fortineau N, Nordmann P. International transfer of NDM-1-producing Klebsiella pneumoniae from Iraq to France. *Antimicrob Agents Chemother* 2011; **55**: 1821–2.

13. Vaux S, Carbonne A, Thiolet JM, *et al.* Emergence of carbapenemase-producing enterobacteriaceae in France, 2004 to 2011. *Eurosurveillance* 2011; **16**: 11–7.

14. Albiger B, Glasner C, Struelens MJ, *et al.* Carbapenemase-producing Enterobacteriaceae in Europe: assessment by national experts from 38 countries, May 2015. *Eurosurveillance* 2015; **20**: 17–34.

15. Legeay C, Hue R, Berton C, *et al.* Control strategy for carbapenemase-producing Enterobacteriaceae in nursing homes: perspectives inspired from three outbreaks. *J Hosp Infect* 2019; **101**: 183–7.

16. Semin-Pelletier B, Cazet L, Bourigault C, *et al.* Challenges of controlling a large outbreak of OXA-48 carbapenemase-producing Klebsiella pneumoniae in a French university hospital. *J Hosp Infect* 2015; **89**: 248–53.

17. Chretien JP, George D, Shaman J, *et al.* Influenza forecasting in human populations: A scoping review. *PLoS One* 2014; **9**: e94130.

18. Paul MJ, Dredze M, Broniatowski D. Twitter Improves Influenza Forecasting. *PLOS Curr Outbreaks* 2014.

19. Kandula S, Shaman J. Near-term Forecasts of Influenza-like Illness: An Evaluation of Autoregressive Time Series Approaches. *Epidemics* 2019; **27**: 41–51.

20. Ray EL, Reich NG. Prediction of infectious disease epidemics via weighted density ensembles. *PLoS Comput Biol* 2018; **14**: e1005910.

21. Peñalva G, Fernández-Urrusuno R, Turmo JM, *et al.* Long-term impact of an educational antimicrobial stewardship programme in primary care on infections caused by extended-spectrum β-lactamase-producing Escherichia coli in the community: an interrupted time-series analysis. *Lancet Infect Dis* 2020; **20**: 199–207.

22. RAISIN Working Group. RAISIN - a national programme for early warning, investigation and surveillance of healthcare-associated infection in France. *Euro Surveill* 2009; **14**: 5–12.

23. Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice, 2nd edition. *OTexts Melbourne, Aust* 2019. Available at: OTexts.com/fpp2.

24. Sax C, Eddelbuettel D. Seasonal Adjustment by X-13ARIMA-SEATS in R. *J Stat Softw* 2018; **87**: 1–17.

25. Box GE, Jenkins GM, Reinsel GC, *et al. Time Series Analysis: Forecasting and Control, 5th Edition*. Hoboken, New Jersey: John Wiley and Sons; 2015.

26. Hyndman RJ, Khandakar Y. Automatic Time Series Forecasting: The **forecast** Package for *R*. *J Stat Softw* 2008; **27**: 1–22.

27. Hyndman R, Koehler A, Ord K, *et al. Forecasting with Exponential Smoothing*. Berlin, Heidelberg:

Springer Berlin Heidelberg; 2008.

28. de Livera AM, Hyndman RJ, Snyder RD. Forecasting time series with complex seasonal patterns using exponential smoothing. *J Am Stat Assoc* 2011; **106**: 1513–27.

29. Gaillard P, Goude Y. opera: Online Prediction by Expert Aggregation [R package version 1.0]. 2016.

30. Cesa-Bianchi N, Lugosi G. *Prediction, Learning, and Games*. Cambridge University Press; 2006.

31. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria; 2020.

32. Hyndman R, Athanasopoulos G, Bergmeir C, *et al.* forecast: Forecasting functions for time series and linear models [R package version 8.5]. 2019.

33. Centers for Disease Control and Prevention. Tracking CRE. *CDC* 2017. Available at: https://www.cdc.gov/hai/organisms/cre/trackingcre.html.

34. European Centre for Disease Prevention and Control. *Summary of the latest data on antibiotic consumption in the European Union - ESAC-Net surveillance data November 2017*. Stockholm; 2017.

35. European Centre for Disease Prevention and Control. *Surveillance of antimicrobial resistance in Europe 2018*. Stockholm; 2019.

36. Clemen RT. Combining forecasts: A review and annotated bibliography. *Int J Forecast* 1989; **5**: 559–83.

37. Polikar R. Ensemble based systems in decision making. *IEEE Circuits Syst Mag* 2006; **6**: 21–44.

Table 1 Accuracy parameters of modeling methods of CPE episodes in France during training (2010-2014) and test (2015-2016) periods.

Model	Training set (2010-2014)			Test set (2015-2016)		
	RMSE	*MAE*	*MAPE*	*RMSE*	*MAE*	*MAPE*
SARIMA	5.91	4.77	19.23	17.90	13.32	14.99
Tslm	6.72	5.22	17.72	20.54	16.39	18.47
X-11 seasonal adjustment	4.95	4.41	18.08	17.40	12.86	14.06
ETS model	7.11	5.56	20.91	18.45	14.68	17.76
Holt-Winters model	8.09	6.53	21.46	25.16	17.60	20.25
NNAR	4.87	3.93	16.42	32.11	26.59	27.40
TBATS model	4.35	3.35	14.25	16.06	13.08	15.24
Average stacking*	3.72	2.81	11.14	15.92	12.65	14.39
Convex stacking	3.29	2.65	11.70	18.81	14.53	15.60
Linear stacking	3.13	2.55	11.24	20.22	16.08	17.09

RMSE: Root mean squared error. MAE: Mean absolute error. MAPE: Mean absolute percentage error.

* Average stacking based on the three following component models: X-11 seasonal adjustment, multiplicative Holt-Winters method, and TBATS.

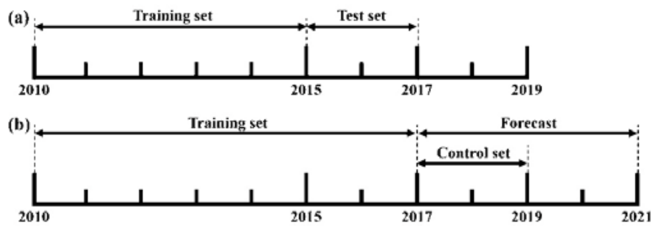

Figure 1 Summary of time periods used for (a) model building and (b) forecasting of the number of CPE episodes in France.

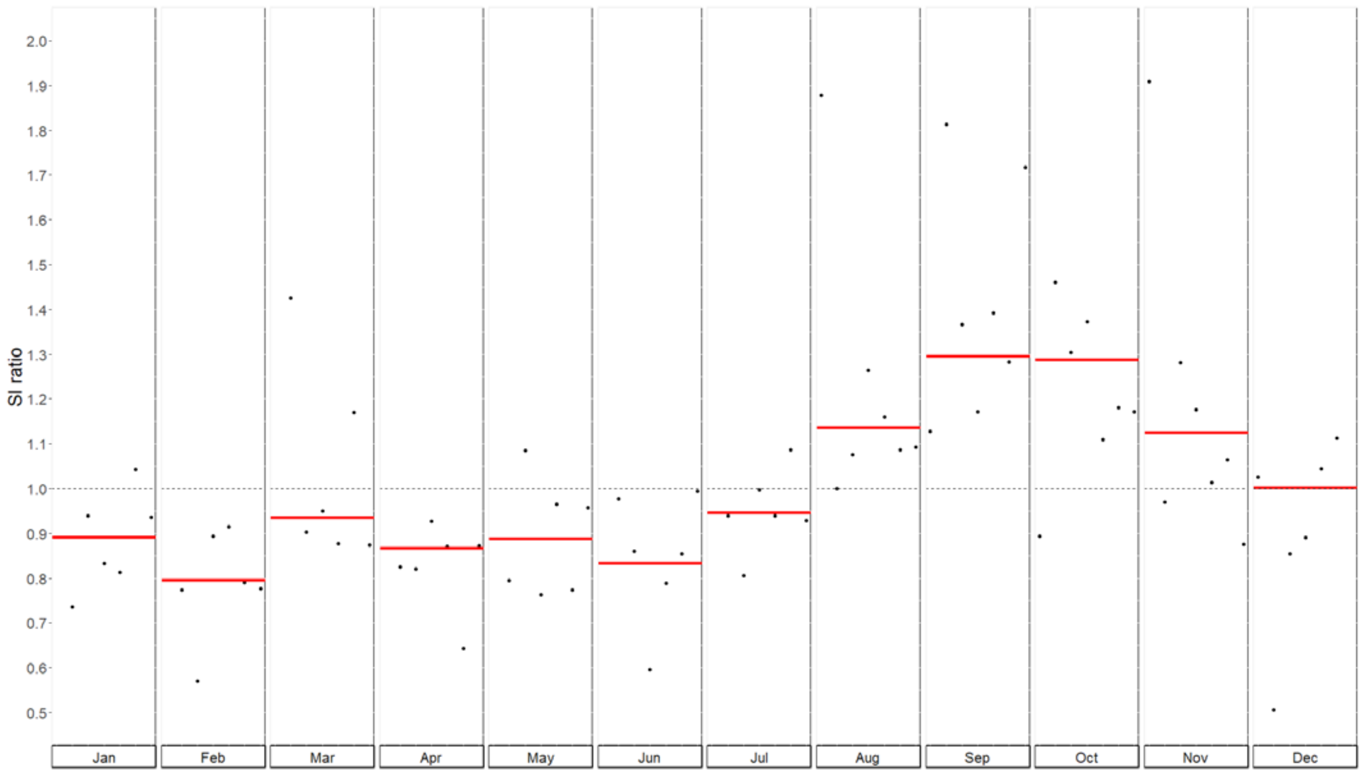

Figure 2 The SI ratios chart of CPE episodes using data on 3,559 episodes in France from August 2010 to November 2016. Average of seasonal factors per month is represented with lines and seasonal-to-irregular ratios by dots. This figure appears in colour in the online version of JAC and in black and white in the printed version of JAC.

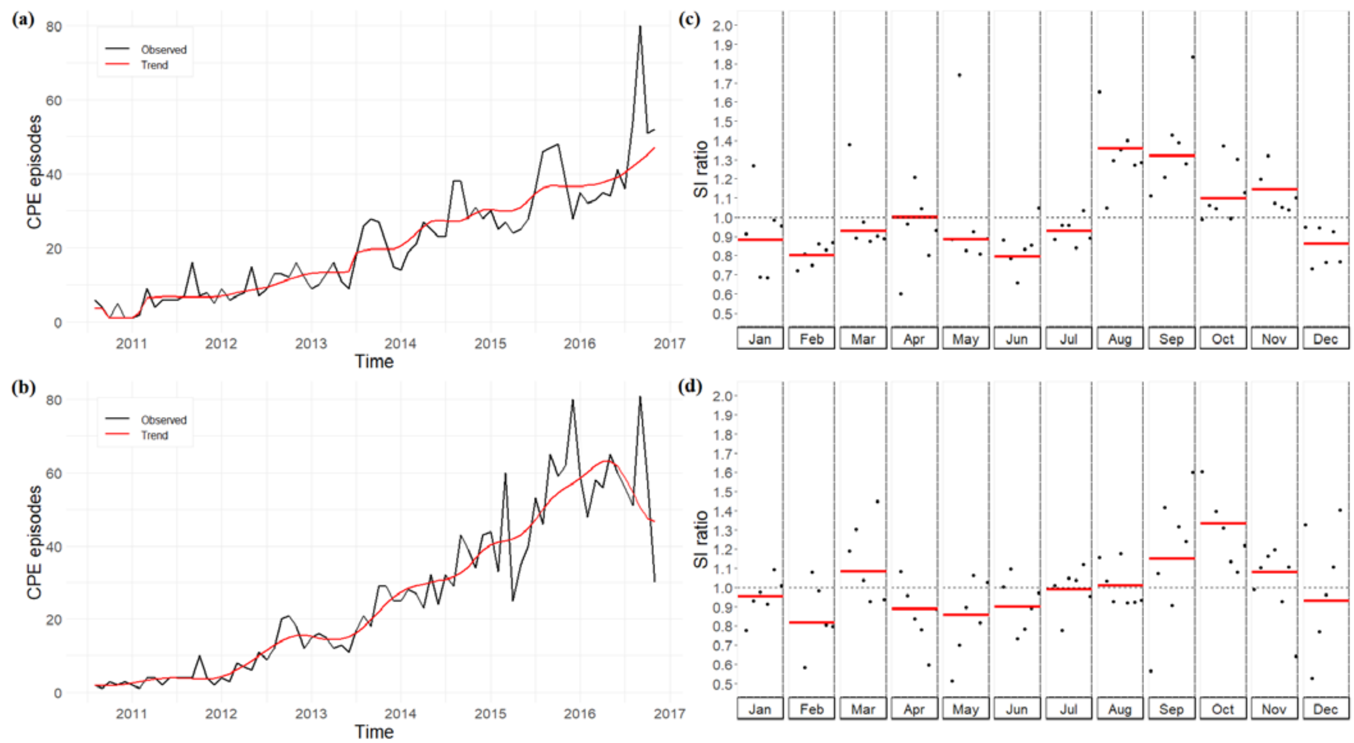

Figure 3 Representation of the number and corresponding trend of (a) imported episodes and (b) non-imported episodes, and SI ratios chart of (c) imported episodes and (d) non-imported episodes. One fictional episode was added each month over the study period for the non-imported time-series. Average of seasonal factors per month is represented with lines and seasonal-to-irregular ratios by dots. This figure appears in colour in the online version of JAC and in black and white in the printed version of JAC.

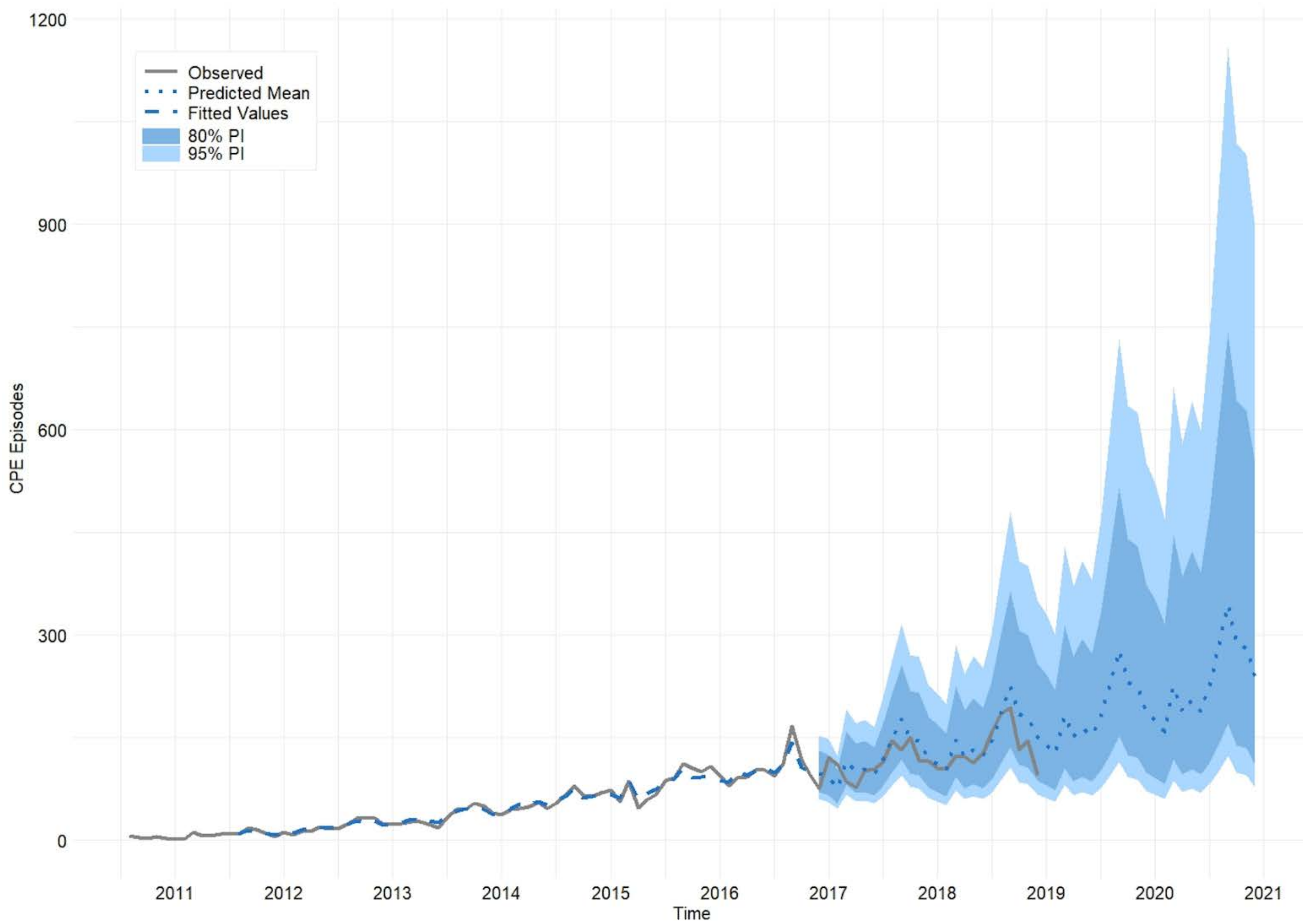

Figure 4 Prediction over a 4-year period of CPE episodes using data on 3559 episodes in France from August 2010 to November 2016 and observed number of CPE episodes from December 2016 to December 2018. This figure appears in colour in the online version of JAC and in black and white in the printed version of JAC.