BMJ Open

Time-series cohort study to forecast emergency department visits in the city of Milan and predict high demand: a 2-day warning system

Rossella Murtas, Sara Tunesi, Anita Andreano, Antonio Giampiero Russo

To cite: Murtas R, Tunesi S, Andreano A, *et al*. Time-series cohort study to forecast emergency department visits in the city of Milan and predict high demand: a 2-day warning system. *BMJ Open* 2022;**12**:e056017. doi:10.1136/bmjopen-2021-056017

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-056017).

Received 02 August 2021
Accepted 03 April 2022

Epidemiology Unit, Agency for Health Protection of Milan, Milan, Italy

Correspondence to
Antonio Giampiero Russo; agrusso@ats-milano.it

ABSTRACT

Objectives The emergency department (ED) is one of the most critical areas in any hospital. Recently, many countries have seen a rise in the number of ED visits, with an increase in length of stay and a detrimental effect on quality of care. Being able to forecast future demands would be a valuable support for hospitals to prevent high demand, particularly in a system with limited resources where use of ED services for non-urgent visits is an important issue.

Design Time-series cohort study.

Setting We collected all ED visits between January 2014 and December 2019 in the five larger hospitals in Milan. To predict daily volumes, we used a regression model with autoregressive integrated moving average errors. Predictors included were day of the week and year-round seasonality, meteorological and environmental variables, information on influenza epidemics and festivities. Accuracy of prediction was evaluated with the mean absolute percentage error (MAPE).

Primary outcome measures Daily all-cause EDs visits.

Results In the study period, we observed 2 223 479 visits. ED visits were most likely to occur on weekends for children and on Mondays for adults and seniors. Results confirmed the role of meteorological and environmental variables and the presence of day of the week and year-round seasonality effects. We found high correlation between observed and predicted values with a MAPE globally smaller than 8.1%.

Conclusions Results were used to establish an ED warning system based on past observations and indicators of high demand. This is important in any health system that regularly faces scarcity of resources, and it is crucial in a system where use of ED services for non-urgent visits is still high.

Strengths and limitations of this study

- This study is one of the few studies linking temporal periodicity, occurrence of festivities, local weather conditions and pollution to emergency department (ED) visits.
- We estimated an autoregressive integrated moving average model for each hospital, thus taking into consideration each specific characteristic and incorporating weekly and annual seasonality with Fourier terms.
- Results were used to establish an ED warning system based on past observations and indicators of high demand.
- We cannot exclude the possible presence of unmeasured variables that may better predict ED visits and overcrowding.

INTRODUCTION

The emergency department (ED) is the gateway (an open door) and the most critical area of a hospital, moving many activities and causing problems in the management of elective procedures when the number of patients who come knocking increases. In the last decade, many countries have seen a substantial rise in the number of ED visits, with an increase in length of stay[1] and associated detrimental effects on quality of care. ED visits are unavoidably subject to fluctuation, and several models to predict high demand have been developed in the last decade, aiming at effectively managing hospital beds and staff rosters.[2] In Italy, even though the number of ED visits has been decreasing since 2016, the mean waiting time in EDs was high, between 12 hours and 24 hours in 3.5% of cases in 2017, and over 24 hours in 2.1% of cases.[3] The definition of overcrowding[4] in the ED literature is not consistent, nor are the measures used to assess overcrowding, which vary from clinician perception of overcrowding to input measures (eg, waiting times and number of patients that arrived), throughput measures (eg, ED capacities and patient care time), output measures (eg, percentages of hospital admissions and hospital beds) or multidimensional indices such as the Emergency Department Work Index. This variety of measures corresponds to the different type

of factors studied as causes of ED crowding. We concentrate here on predicting the number of visits from input factors, that is, determinants and modalities of patient inflow, such as non-urgent visits and influenza season. In this case, it is better to speak of overflow.[5] We did not investigate throughput factors, describing organisational issues in the ED, such as inadequate staffing, nor output factors. The latter includes one of the major reasons for ED overcrowding, which is the shortage of acute care bed capacity.[6–10] Among the most investigated input factors are non-urgent visits, meaning 'patients who could have been assessed and treated in other facilities that treat less urgent cases'.[11] In Italy in 2017, only 23% of ED visits were classified as red or yellow at triage, while 13%[3] had a low level of priority, coded white triage in Italy. This use of ED services is a signal of lack of continuity of primary care and difficulty of access to both primary and specialist care. It is also not cost-effective and leads to an increase in waiting times in the EDs.[12 13]

Several factors potentially affect the daily number of ED visits. Among these are annual,[14 15] seasonal[16–19] and weekly[14–19] periodicity, as well as festivities.[14 16 20 21] The effect of local weather conditions and pollution on ED visit volumes is still being debated: while some studies confirmed a significant association with temperature,[15 17–19 22 23] precipitation,[17 19] humidity[22] and weather conditions,[23] other authors found these variables to be only mediocre predictors of the number of ED visits[16] and found air pollution mostly impacting cardiac and respiratory diseases.[22] An additional factor that has been studied in relation with ED visit volumes is influenza, with around 7% of total accesses attributable to influenza-like illness (ILI) during the epidemic season.[24] Murtas and colleagues[25] evaluated the hypothesis of the early presence of the COVID-19 epidemic in Italy by analysing data on trends of access to EDs using a Poisson regression model adjusted for seasonality and influenza outbreaks. In this work, they found that predicting ED visits by considering both seasonality and ILI rates, compared with a model taking into account only seasonality, notably increased the fitting of the model. Therefore, syndromic surveillance (such as ILI rates, which in Italy are provided weekly by the National Health Service Sentinel System) may be able to provide early warning of hospital bed capacity strain caused by seasonal respiratory disease.[26] To our knowledge, there is no study linking all this information together to ED visits.

The present study aimed to develop a model for forecasting ED arrivals using regression-based time-series analysis with autoregressive integrated moving average (ARIMA) errors, accounting simultaneously for the effect of meteorological and environmental variables, as well as information on influenza epidemics and festivities, on the number of ED visits in the city of Milan. The model is used to establish an innovative ED warning system (WS) providing a planning instrument for hospitals based on past observations and indicators of high demand.

METHODS

Study design

This is a retrospective study conducted in the area served by the Milan Agency for Health Protection using current healthcare databases of daily ED visits aggregated at hospital level. No individual-level data were used, and patients cannot be identified from aggregated data which do not contain low counts (ie, cells with ≤5 counts).[27]

Study setting and population

We collected all ED visits, including patients registered at triage that voluntarily left the ED premises before being evaluated by a physician, between 1 January 2014 and 31 December 2019 in the five largest hospitals located in the city of Milan (figure 1). All five hospitals are public hospitals and received 49% of all emergency room access of the city of Milan, which has a total of 17 EDs, with a mean number of daily ED visits during 2014–2019 ranging from 124 for hospital C to 247 for hospital E.

Study protocol

Aggregated data on daily ED visit volumes, by age and gender, were extracted from the regional health database. Meteorological and environmental information was extracted from the Regional Environmental Protection Agency (ARPA).[28] Daily mean temperature, relative humidity (RH), cumulative precipitation, nitrogen dioxide (NO_2) and particulate matter with a diameter of ≤10 µm (PM10) were collected from two monitoring stations (one measuring meteorological indicators and one measuring air pollution) located in the centre of Milan (figure 1). For the sensitivity analysis, we also investigated the effect of minimum, maximum and apparent temperatures on daily ED visits.[29] Missing values on a specific day were imputed with the average of the measure in that specific year. Weekly data on ILI notifications were taken from the National Health Service Sentinel System (InfluNet).[30] Weekly incidence rates of ILI were expressed as the number of cases per 1000 inhabitants per week. All available information was linked to daily ED visit volumes for each of the five hospitals included in the study. Datasets were divided into training (from 1 January 2014 to 31 December 2018) and validation sets (from 1 January 2019 to 31 December 2019). For each hospital, we first estimated model parameters on the training dataset and evaluated postsample accuracy in the validation set. We included, in each model, only factors that significantly influenced the number of ED visits. Multicollinearity was evaluated calculating Pearson pairwise correlation between variables and variance inflation criterion (VIF).[31]

Patient and public involvement

Patients were not involved in this research.

Data analyses

Development of the predictive model

To predict the daily volume of visits in each ED, we used a time-series approach consisting of a regression model with ARIMA errors.[32] The statistical units were days, 1826 days

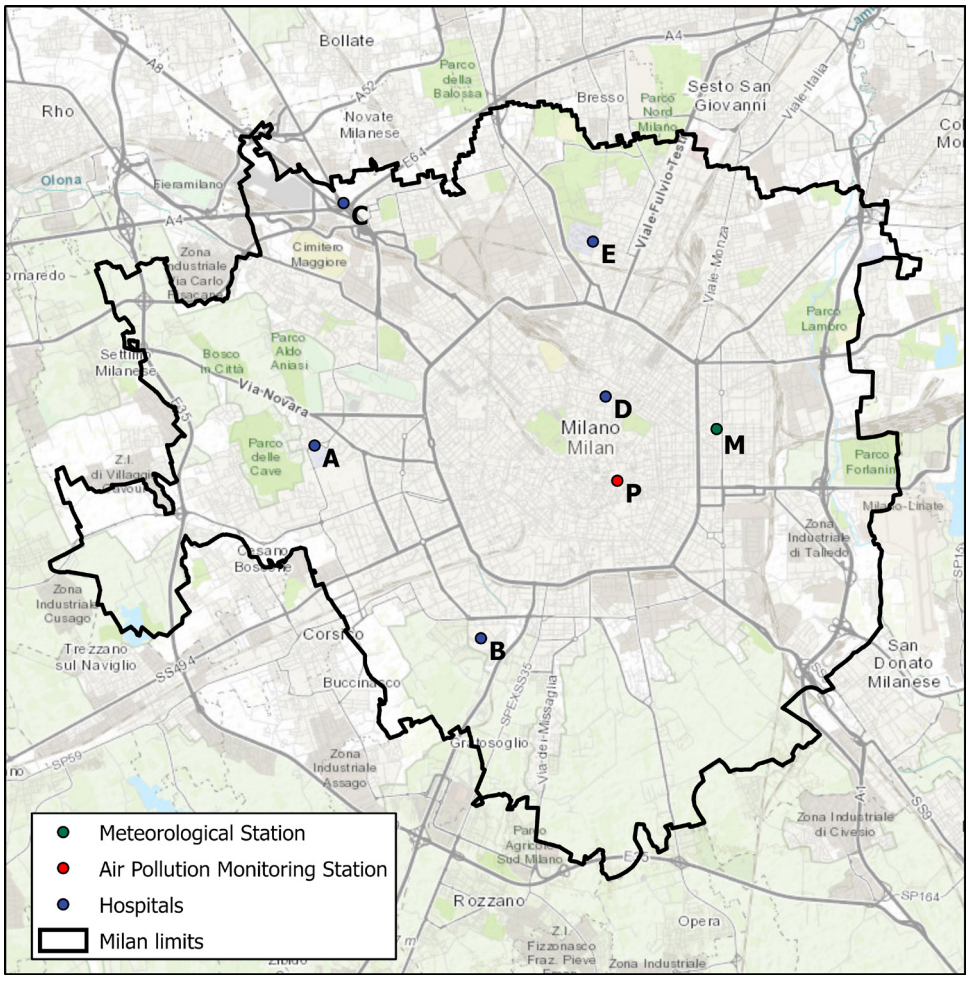

Figure 1 Location of the five participating hospitals and of meteorological and air pollution monitoring stations in the city of Milan.

in the training set and 365 days in the validation set. This model is able to combine two powerful statistical methods: linear regression and ARIMA. Linear regression of Y on X is usually described by the equation $Y_t = \alpha + \beta x_t + \epsilon_t$, where Y_t and x_t are the values of Y and X at day t; α and β are the intercept and the slope of the regression line; and ϵ_t is the error of the model at day t (the deviations from the fitted line to the observed values) assumed to be independent from other values. The ARIMA model deals with autocorrelation between errors through two components: the autoregressive and the moving average (MA) process. The autoregressive component assumes that previous observations are good predictors for future values, while the MA component allows the model to update the predictions if the level of a constant time-series changes. ARIMA specification is described by three parameters (p, d and q), where p is the order of autoregression that is the number of time lags; d is the degree of differencing (the number of times the data have had past values subtracted to make the time-series stationary); and q is the order of the MA process. For each hospital, these parameters were identified examining total autocorrelation function (ACF) and partial autocorrelation function (PACF), as well as statistical significance (p value <0.05), and minimal Akaike information criteria (AIC). Day of the week and year-round seasonality were controlled for by including Fourier terms, a series of sine–cosine functions capable of approximating periodicity.[20 32] The number of Fourier terms was chosen to minimise the AIC for each seasonal period (up to seven for day of the week seasonality and up to 365 for year-round seasonality). Each seasonal component can be written in the model equation as

$$\sum_{j=1}^{n}\left[\alpha_j \sin\left(\frac{2\pi jt}{m}\right) + \beta_j \cos\left(\frac{2\pi jt}{m}\right)\right],$$

where n is the number of Fourier terms chosen to minimise the AIC (up to seven for day of the week seasonality and up to 365 for year-round seasonality) and m is the seasonal period (seven for day of the week and 365 for year-round seasonality).

Therefore, meteorological and environmental variables, as well as information on influenza epidemics and festivities, were retained in the final model only if statistically significant. As festivities, we considered Italian public holidays with school and office closures: New Year's Day, Epiphany, Easter Sunday and Monday, Italian Liberation Day, Labour Day, Foundation of the Italian Republic,

assumption day, All Saints' Day, Saint Ambrose's Day (local patron saint), Feast of the Immaculate Conception, Christmas Day, Saint Stephen's Day and New Year's Eve. In addition, we created dummies for specific festivities that were responsible for a significant variation in the number of ED visits: New Year's Eve and Assumption Day (15 August). Diagnostics of the finally selected models included the Jarque-Bera test of normality, and correlation among the residuals was obtained according to the Ljung-Box test. Variables and tests were considered statistically significant if the p value was <0.05.

The ARIMA model was compared with a simple regression model (M1) including only meteorological, environmental and festivity covariates and with a generalised linear model (M2) also including the Fourier terms to control for seasonality. P values were calculated by comparing the full model (ARIMA) to M1 and M2 using the likelihood ratio test.

Forecasting accuracy

Predicted values on validation sets were estimated using one-step forecast.[32] We estimated parameters only on training sets. However, we calculated forecasts on validation sets using all of the data preceding each observation. The accuracy of predictions was evaluated with the mean absolute percentage error (MAPE), which expresses, as percentages, a unit-free measure of performance:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|y_t-\hat{y}_t|}{y_t} * 100,$$

with y_t and $\hat{y}_t$, respectively, as the observed and predicted numbers of visits at day t, and n the number of days in the validation set (n=365 in this study).

High demand definition

We proposed a definition of high ED demand as days where the number of visits exceeded the median of the preceding 31 days. The days were defined as green (level 1) if the number of visits exceeded the median by less than 5%, yellow (level 2) if between 5% and 10%, red (level 3) if higher than or equal to 10%. High demand was calculated on the observed and predicted ED visits in validation sets; we thus calculated the proportion of observed high ED demand that is correctly classified by predicted high ED demand (called sensitivity or recall metrics for multiclass classification problems).[33] In addition, we calculated the accuracy of predictions as the number of correct classifications over the total number of observations. All statistical analyses were performed with R V.3.6.3.[34]

To evaluate the proposed definition, we further calculated high demand as the number of visits exceeding the median of the preceding 7, 14 and 21 days and the number of visits exceeding the mean of the preceding 7, 14, 21 and 31 days, defining green, yellow, and red levels of high demand as previously discussed. We chose 7, 14 and 21 lag days in order to adjust for weekly variation in the number of ED visits by design. We further calculated high demand as defined by the Lombardy Region[35]: when the number of visits exceeded the 91st percentile of the previous year time series. Low demand days were defined as those with a number of visits smaller than the 25th percentile, medium demand days as those with a number of visits between the 25th percentile and the 75th percentile, high demand days if between the 75th percentile and the 90th percentile, and finally very high demand days if over the 91st percentile.

ED warning system

In the month of January 2020, we established an ED WS, which was used by the selected hospitals in Milan as a planning instrument for EDs and consists in a transmission of daily reports. This WS continued until February when the COVID-19 outbreak started in Italy. According to the model choices highlighted by the aforementioned methodology (validation and calibration of the model were performed with data from 2014 to 2019), parameters were updated weekly and used to establish the WS, which operated in January 2020. A hypothetical daily report received from a hospital on 5 January 2020 can be found in figure 2. The report included forecasts of the number of visits for the following 2 days, with 95% margin errors and a high demand indicator (green, yellow or red). The forecasts were made incorporating in the model past meteorological and environmental information via an application programming interface where 2-day future forecasts of meteorological and environmental information were provided by ARPA Lombardia. Weekly information on ILI was downloaded every week from InfluNet and included in the predictive models. Daily reports were constructed and dispatched automatically using R and R Markdown. During the WS campaign, we established a monitoring service capable of estimating daily sensitivity, accuracy of predictions and MAPE separately for prediction 1 and 2 days ahead.

All analyses were performed with R software V.4.0.2 (R Core Team, Vienna, Austria); models and Fourier terms were estimated, respectively, using the Arima and the Fourier functions in the R package forecast[36] using the parameter xreg for covariate specification. VIF was calculated using the VIF function in the car package.[37]

RESULTS

ED visit volumes

Between 1 January 2014 and 31 December 2019 (training set of 1826 and validation set of 365 days), we observed 2 223 479 visits, 370 633 on average every year. Daily mean number of visits by hospital, temporal, meteorological and patient characteristics in the training sets are summarised in table 1. Missingness, over the whole period of 2014–2019, in meteorological and environmental variables was found in 8 days for temperature, 7 days for precipitation and 37 days for humidity. Description of training and validation sets, and plots of each hospital's time series are summarised in online supplemental table 1 online supplemental file 1. The Pearson correlation between

EMERGENCY ADMISSION 2-DAY WARNING SYSTEM

Agency for Health Protection of Milan

Hospital A. Prediction for the day:

06 Nov 2019	07 Nov 2019
Livello 1	Livello 2

Predicted ED visits ± 95% Margin Errors	227 ± 12	254 ± 14
Covariates		
NO2	46 $\mu g/m^3$	38.4 $\mu g/m^3$
PM10	6 $\mu g/m^3$	3.9 $\mu g/m^3$
TEMPERATURE	5°	6°
UMIDITY	65%	57%
PRECIPITATION	0 mm	0 mm
ILI RATE	10.87	10.87

Methodology

Level 1	Number of visits exceeded the median by less than 5%.
Level 2	Number of visits exceeded the median between 5% and 10%.
Level 3	Number of visits exceeded the median by more than 10%.
Prediction	Prediction based on a regression model with ARIMA errors (Hyndman 2018).
Environment	Meteo and Pollution Forecast from ARPA Lombardia.
ILI rates	Influence Like Illness rate, number of cases per 1,000 inhabitants per week (InfluNet).

Sistema Socio Sanitario
Regione Lombardia
ATS Milano
Città Metropolitana

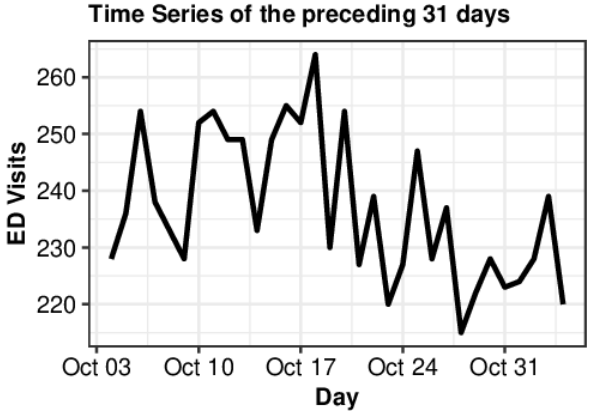

Figure 2 Hypothetical daily report received from a hospital on 5 January 2020. ED, emergency department.

Table 1 Total number of visits and mean number of daily visits by hospital, temporal and meteorological factors, and patient characteristics between 1 January 2014 and 31 December 2019 in five emergency departments of the city of Milan, Italy

	n (%)*	Mean (min–max)†	Cumulative precipitation (mm)	n (%)‡	Mean (min–max)§
Hospitals					
A	421 741 (19)	192 (107–301)	≤0.6	1 678 953 (75.5)	1018 (563–1295)
B	457 021 (20.6)	209 (65–302)	0.7+	544 526 (24.5)	1005 (627–1392)
C	272 308 (12.2)	124 (61–197)	NO_2 (µg/m^3)		
D	530 519 (23.9)	242 (125–337)	≤32	564 957 (25.4)	974 (563–1295)
E	541 890 (24.4)	247 (133–346)	33–44	570 284 (25.6)	1022 (723–1292)
Total	2 223 479	1015 (563–1392)	45–57	536 484 (24.1)	1032 (698–1392)
Gender			58+	551 754 (24.8)	1035 (693–1272)
F	1 113 405 (50.6)	508 (277–782)	PM10 (µg/m^3)		
M	1 087 903 (49.4)	497 (277–661)	≤20	571 339 (25.7)	990 (563–1261)
Age			21–29	575 530 (25.9)	1017 (688–1295)
≤14	360 600 (16.4)	165 (55–443)	30–44	544 147 (24.5)	1023 (710–1392)
15–65	1 307 139 (59.4)	597 (317–860)	45+	532 463 (23.9)	1032 (693–1272)
66+	533 569 (24.2)	244 (141–385)	ILI (number of weekly new cases per 1000 inhabitants)		
	N (%)‡	Mean (min–max)§			
Temperature (°C)			≤1.2	1 310 096 (58.9)	1001 (563–1295)
≤9.2	564 744 (25.4)	1021 (693–1392)	1.3–2.5	303 072 (13.6)	1031 (799–1256)
9.3–15.6	563 764 (25.4)	1033 (813–1261)	2.6–5.6	303 102 (13.6)	1031 (698–1261)
15.7–22.3	563 757 (25.4)	1025 (656–1295)	5.7+	307 209 (13.8)	1045 (693–1392)
22.4+	531 214 (23.9)	980 (563–1292)	Day before/after festivity		
Relative humidity (%)			No	2 096 838 (94.3)	1012 (563–1392)
≤50	560 870 (25.2)	1018 (563–1295)	Yes	126 641 (5.7)	1055 (688–1295)
51–62	552 865 (24.9)	1009 (637–1292)	Festivity		
63–76	554 041 (24.9)	1017 (627–1392)	No	2 144 726 (96.5)	1018 (677–1392)
77+	555 703 (25)	1016 (786–1278)	Yes	78 753 (3.5)	938 (563–1253)

*Total number of visits by hospital, gender and age. The percentages of the number of visits out of the total (2 223 479 total number of visits; 2 201 308 with information on age and gender) are in parentheses.
†Mean, minimum and maximum number of daily visits by hospital, gender and age.
‡Total number of visits by temporal and meteorological factors (ie, total number of visits in days with a particular value of temperature, humidity, etc). The percentages of the number of visits of the total (2 223 479 total number of visits) are in parentheses.
§Mean, minimum and maximum number of daily visits by temporal and meteorological factors (ie, mean number of daily visits in the days with a particular value of temperature, humidity, etc).
F, female; ILI, influenza-like illness; M, male; NO_2, nitrogen dioxide; PM10, particulate matter with a diameter of ≤10 µm.

predictors varied from weak (absolute correlation <0.3) to moderate (absolute correlation between 0.3 and 0.7), with a maximum of −0.67 between temperature and ILI and 0.61 between NO_2 and PM10. VIF was smaller than 5 for all variables, with a maximum of 2.8 for temperature and 1.9 for ILI. We therefore included all the variables in the models, selecting the final model according to the statistical significance of predictors and minimal AIC.

Model specification and ARIMA results

All models showed a very strong day of the week and year-round seasonality effect, according to ACF and PACF plots. To normalise residuals, outliers (in the training sets only) were replaced by the mean of the observations of the same day in the other years; consequently, all models showed residual normally distributed according to the Jarque-Bera test (number of replaced outliers are presented in online supplemental table 2). All models showed a lack of fitting on New Year's Eve and/or 15 August; for this reason, we chose to define a specific dichotomous variable ('1' for the peculiar festivity and '0' for the other days) capable of detecting this extra variation. Table 2 displays the ARIMA parameters fitted for each model

Table 2 ARIMA specifications and covariate effects on the number of ED visits between 1 January 2014 and 31 December 2018 (training sets)

		Hospitals				
		A	B	C	D	E
Model specification	ARIMA parameters (p, d and q)	(0, 1, 2)	(1, 1, 1)	(1, 1, 2)	(1, 1, 1)	(1, 1, 1)
	Fourier terms*	3, 13	3, 14	3, 13	3, 16	3, 15
Covariate effects (SE)†	Temperature (°C)	1.29 (0.15)	1.23 (0.14)	0.68 (0.11)	1.16 (0.18)	1.84 (0.18)
	Humidity (%)				−0.08 (0.04)	
	Precipitation (mm)	−0.2 (0.05)	−0.12 (0.05)		−0.13 (0.07)	−0.31 (0.06)
	NO_2 (µg/m^3)	−0.08 (0.03)			−0.09 (0.04)	
	PM10 (µg/m^3)			0.03 (0.02)		
	ILI (weekly new cases per 1000 inhabitants)	1.74 (0.41)	1.05 (0.37)	0.73 (0.29)		0.97 (0.46)
	Festivity		−28.23 (1.98)	−12.96 (1.45)	−25.42 (2.23)	−14.56 (2.39)
	Special festivity‡	−43.16 (6.31)	−57.64 (6.36), −62.61 (6.29)	−42.06 (4.92)	−59.86 (7.58)	−63.24 (7.92)
	Day before/after festivity	7.14 (1.5)	9.06 (1.58)	3.75 (1.22)		13.89 (1.96)

*Number of sine and cosine terms used to approximate day of the week and year-round seasonality.
†Parameter estimates and SEs in parentheses. Predictors were retained in the final model only if statistically significant (p value <0.05).
‡New Year's Eve for hospitals A, C and D, and New Year's Eve and 15 August for hospital B.
.ARIMA, autoregressive integrated moving average; ILI, influenza-like illness; NO_2, nitric oxide; PM10, particulate matter with a diameter of ≤10 µm.

and the number of Fourier terms that minimised AIC. All models were non-stationary in mean and needed one differencing to make the time-series stationary (d=1). ARIMA parameters and Fourier terms were different across hospitals, showing that each time series needed different model specification. table 2 also displays, for each hospital, the factors that significantly influenced the number of ED visits and that were included in the models. High temperatures were always associated with a statistically significant increase in ED visit volumes, with a maximum increase of 1.84 daily visits every 1°C increase (hospital E, SE 0.18). RH was significantly associated with a limited decrease of total ED visits (−0.08, SE 0.04) for a 1% increment of RH only at hospital D. High levels of cumulative precipitation were associated (except for hospital C) with a statistically significant decrease in ED visits, with a maximum decrease of 0.31 daily visits every 1 mm of precipitation (hospital E, SE 0.06). Concerning air pollution, we found an opposite effect of NO_2 and PM10 on ED visits, with a mild significant negative effect for NO_2 in two hospitals (−0.08 and −0.09) and an even milder positive association with PM10 in one (0.03). Except for hospital C, the effect of ILI was always associated with the number of ED visits, showing an increase of daily visits between 0.73 and 1.74 (SE 0.29 and 0.41, respectively) at every unit increase in weekly ILI rates. Festivities were associated with a decrease in ED visits of between 13 and 28 (SE 1.45 and 1.98), while special festivities were associated with the greatest decrease of at least 42 ED visits (SE 4.94). ACF and correlation among residuals according to the Ljung-Box test by hospital and up to 30 and 366 lags can be found in online supplemental figure 1. ACF plots of residuals were overall in significance limits and the Ljung-Box test showed overall no significant correlation between residuals at different lags, except Hospital E, which showed residual autocorrelation up to lag 366.

Forecasting accuracy and high demand definition

The accuracy of predictions (MAPE) in the validation sets, sensitivity and accuracy between observed and predicted high ED demand are displayed in table 3. Model performance was good, with small MAPEs in validation sets, ranging from a minimum of 5.5% for hospital D to a maximum of 8.1% for hospital C. The models showed high sensitivity on days with green-level high demand; almost 90% of days with predicted green-level high demand were confirmed from observed values. On days with yellow-level high demand, sensitivity between predicted and observed demand was scarce, ranging from 0.04 for hospital B to 0.28 for hospital A. Sensitivity of red-level high demand varied between hospitals, with a minimum of 0.25 for hospital A to a maximum of 0.57 for hospital D. Observing table 3, we can suggest that, for each hospital, at least 54% of the observed red-level high demand days were classified, from predictions, as being at least yellow-level. Accuracy was high, with at least 67%

Table 3 Indicators of performance of the developed models: accuracy of predictions (MAPE) in the validation sets, and accuracy and sensitivity of high demand classification

				Predicted high ED demand (%, Sensitivity)		
Hospital	MAPE	Accuracy (%)	Observed high ED demand	Green	Yellow	Red
A	5.9	72	Green	93	6	1
			Yellow	64	28	8
			Red	46	29	25
B	5.7	72	Green	92	8	0
			Yellow	85	4	11
			Red	35	15	50
C	8.1	67	Green	88	8	4
			Yellow	78	10	12
			Red	45	20	35
D	5.5	76	Green	91	6	3
			Yellow	65	27	8
			Red	35	9	56
E	6.1	74	Green	90	8	2
			Yellow	59	24	17
			Red	34	28	38

.ED, emergency department; MAPE, mean absolute percentage error.

of the days with exactly the same predicted and observed high demand level (green, yellow or red).

All ARIMA models fitted the data significantly better than a simple regression model (M1) and a generalised linear model (M2), with MAPE for M1 and M2 above 13.5% and 9.8%, respectively (online supplemental Table 3). Observed and predicted ED visits in the validation sets (from 1 January 2019 to 31 December 2019) by date and hospital can be found in online supplemental figure 2. In online supplemental table 2, we compared ARIMA results for different temperature specifications: mean, minimum, maximum and apparent temperature. The greatest effect on ED visits was attributed to mean temperature, while indicators of performance and AIC where generally superior for mean temperature compared with minimum, maximum and apparent temperature. In online supplemental table 2, we also calculated, only for outlier days, the relative error mean of observed versus predicted values in order to evaluate if extreme temperatures were better outlier predictors than mean temperature. Number of outliers replaced ranged from two for hospital A to seven for hospital D; results suggested an overall better fit of outliers using minimum temperature (three out five hospitals with smaller relative errors).

In online supplemental table 4, we compared the high demand definition used in the ED WS with similar definitions. There was slight improvement in percentage accuracy between the definition used and the other algorithms, and there was no favourite algorithm for all hospitals: hospital B had a maximum improvement of 4% using the mean of the preceding 31 days or the median of the proceeding 21 days; hospitals A and C had an improvement of 2% using the mean of the preceding 31 days; hospital D had an improvement of 2% using the mean of the preceding 21 days; and finally hospital E had an improvement of 1% using the mean of the preceding 21 or 31 days. Using the high demand definition used by the Lombardy Region, we did not find any improvement in accuracy, with an overall percentage of matched classification between 50% and 64%. High demand was always predicted less well compared with the definition used in our ED WS. However, results showed good prediction of very high demand days with a sensitivity between 38% and 67%.

ED warning system

In table 4A,B we provided the accuracy of predictions (MAPE), sensitivity and accuracy between observed and predicted high ED demands in January (the operating period of the WS) for horizons of 1 and 2 days. Errors of prediction (MAPE) were slightly higher than in the validation set, with MAPE for 1-day horizon always smaller than MAPE for 2 days horizons. Accuracy between observed and predicted high ED demand was never smaller than 0.45 and generally smaller than in the validation set.

DISCUSSION

In this work we proposed and implemented in daily practice, a system to predict the number of ED visits in five hospitals of the city of Milan. The system is based on regression models with ARIMA errors, where ARIMA parameters were allowed to vary between hospitals, according to their specific characteristics, and it provides

Table 4 Accuracy of predictions (MAPE), sensitivity and accuracy between observed and predicted high ED demand in January 2020 (the operating period of the Ws) with a 1-day (A) and 2-day (B) horizon

	MAPE	Accuracy (%)	Observed high ED demand	Predicted high ED demand (%, sensitivity)		
				Green	Yellow	Red
(A)						
Hospital A	7.8	52	Green	94	6	0
			Yellow	100	0	0
			Red	71	29	0
Hospital B	7.8	81	Green	87	13	0
			Yellow	0	100	0
			Red	17	17	67
Hospital C	8.6	52	Green	100	0	0
			Yellow	67	33	0
			Red	73	27	0
Hospital D	6.6	45	Green	55	36	9
			Yellow	0	33	67
			Red	50	33	17
Hospital E	11	45	Green	100	0	0
			Yellow	100	0	0
			Red	92	8	0
(B)						
Hospital A	8.1	55	Green	100	0	0
			Yellow	100	0	0
			Red	71	29	0
Hospital B	8.6	71	Green	73	27	0
			Yellow	0	100	0
			Red	25	17	58
Hospital C	9	45	Green	93	7	0
			Yellow	83	17	0
			Red	82	18	0
Hospital D	7.6	48	Green	50	18	32
			Yellow	0	0	100
			Red	33	0	67
Hospital E	11.2	45	Green	100	0	0
			Yellow	100	0	0
			Red	92	8	0

.ED, emergency department; MAPE, mean absolute percentage error.

daily reports on the number of visits predicted for the two subsequent days at the five hospitals participating in the study. The models showed a good overall performance with the MAPEs always smaller than 5.5% and 8.1%. Our results are slightly better than other studies: Marcilio and colleagues[20] forecast daily ED visits with Generalised Linear Models, finding MAPEs between 5.4% and 11.5%, according to different forecasting horizons and controlling for temperature effect.

Jones and colleagues,[16] using similar models, found MAPEs that varied between 8.5% and 15.5%. However, Duwalage *et al*[15] using a generalised additive model found MAPEs consistently lower than 5% for 14-day forecasts, which significantly improved including temperature in the model. Although the number of predicted ED visits was close to the observed values, and there was good sensitivity in predicting mild (green) high demand, there was moderate sensitivity in predicting the spike of ED visit volumes (red-level high demand) for some hospitals and acceptable sensitivity for hospital D. This is particularly important for the scope of this study, which aimed to forecast ED visits in order to develop a 2-day WS. For

this reason, a better predictive performance of the red-level forecast would be desired. In fact, one of the major reasons for ED overcrowding is the shortage of acute care bed capacity compared with the huge number of visiting patients. Comparing our definition with similar definitions, we found a slight improvement in percentage accuracy, around 1% and 4%, but there was no a favourite algorithm for all hospitals. Furthermore, using the definition of very high demand for ED visits defined by the Lombardy Region, we found sensitivity was better compared with our models, and we plan to implement this in further evolutions of our WS. However, we found good sensitivity in classifying observed red-level demand as at least yellow from predictions, and accuracy among observed and predicted high demand levels was always close to 70%. The definition of high ED demand is not straightforward as it relies on the specific hospital's characteristics. It is one of the main causes of ED overcrowding, which is the most problematic issue in EDs, thus deserving the effort in trying to predict it. In this study, we proposed a definition based on percentage increases compared with the median of the preceding month to warn EDs of requests rising over the levels they managed in the preceding month.

During the operating period of the WS, January 2020, we found a worse adaptation of the models than in the validation year 2019. This could be due to the ongoing outbreak of COVID-19, as ED visits for non-critical problems were discouraged.[38]

Concerning potential predictors, we found a strong day of the week and year-round seasonality effect, adequately captured by the terms used to approximate periodicity (Fourier terms). Even if the aim of this work was to develop a forecasting model and not an explanatory model, here we found statistically significant effects of meteorological factors on ED visits. Temperature was always positively associated with outcome, with an increase in the number of visits for each 1° increase in temperature across hospitals, in accordance with previous results.[17 19 23] As reported in another study,[39] high temperatures are associated with ED visits, especially for the most susceptible population, as persons with diabetes or cancer, so it is important for public health officials to implement adaptation measures to manage the impact of high temperatures on population health. Here we found a slightly better fit for outliers using minimum temperature instead of mean temperature. Nonetheless, we decided to include mean temperature in the ED WS because it showed the greatest effect on ED visits. Further work has to be done in order to investigate the role of extreme temperature on ED visit fluctuations. The role of precipitations has not yet been well established. To our knowledge, only one study measured an indirect effect in reducing ED visit volumes.[19] In accordance with these results, rainy days were found to be mildly associated with reduced numbers of ED visits. NO_2 and PM10 had a mild significant effect only in two hospitals and in one hospital, respectively, and were discordant, with a negative effect of NO_2 and a positive effect of PM10 on the number of ED visits. This may be explained considering that the effect of pollution on ED visits is generally exerted and measured on respiratory conditions, especially asthma, and/or cardiac rather than with total visits, and it may be diluted when analysing all ED visits. Only a few studies found a positive association of total suspended particles with all visits but trauma, going in the same direction as the small significant increase in the number of visits related to PM10 we found.[40] In addition, pollution estimated from the monitoring station (classified as from urban traffic) used in the analysis might be of a greater magnitude than that really observed in each hospital. However, even though the hospitals were mostly located on the outskirts of the city of Milan, they are all located in urban areas characterised by a similar air pollution pattern. ILIs were found to significantly increase the number of ED visits, as found by other researchers.[41]

This study indicated a moderate to good sensitivity in predicting high demand, showing some difficulties in anticipating the exact red-level days. In the future, we aim to investigate models capable of directly predicting ED peaks instead of predicting the number of ED visits such as copulas used for detecting spikes in signal processing in brain circuits[42] or machine learning models. Finally, when interpreting these results, it is necessary to be aware of the possible multicollinearity problem between variables, which may alter the magnitude and statistical significance of coefficients. However, according to Vatcheva *et al*,[31] only high correlations between variables would result in a change of sign of the coefficients, and furthermore, VIFs were always smaller than 5. Correlated factors were the pollution variables (NO_2 and PM10), which were never considered in the same model together. Given that the highest correlation was found among temperature and ILI, the effect of these variables on the number of ED visits may potentially be biased due to multicollinearity. However, we included both terms in the models, given the fact that they have an effect on ED visits independently from one another.

Another limitation is the choice of the hospitals considered for this work, that is, major hospitals located in the city of Milan. This methodology might not be feasible for use by small hospitals as they might have low counts or even no visits at all on particular days. A solution can be provided by implementing different statistical models, for example, negative binomial or zero-inflated Poisson models, and would be one of our aims in the next years.

High-demand ED forecasting has a dual nature that should be addressed: first, knowing in advance the number of expected visits would allow a more reasoned choice of the hospital to which request assistance, and second, forecasts should be followed immediately by an evaluation of the available beds and of the staff needed to accommodate these expected visits. These two problems were not addressed in this work, given that this study was intended to estimate ED demand only and does not include information on hospital capacity but

are fundamental ingredients that should be considered in the future.

In conclusion, we proposed a hospital-specific ED WS based on predictive models developed on previous attendances that can be used as a planning instrument in hospitals to increase resources, and to prevent high patient demand when a higher number of attendances is expected. This is important in any health system that usually deals with scarcity of resources, and it is crucial in a system where use of ED services for non-urgent visits is still high.

Contributors RM and AGR conceptualised the study and defined the methodology, and accessed and verified the data. RM analysed the data set. ST and AA contributed to the literature search, data interpretation and writing of the manuscript and made substantial contributions to the revision of the paper. AGR, author acting as guarantor, supervised and managed the project.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Map disclaimer The depiction of boundaries on this map does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. This map is provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval In accordance with Italian legislation, this study was not submitted for ethics approval.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

ORCID iDs
Rossella Murtas http://orcid.org/0000-0002-8989-7605
Sara Tunesi http://orcid.org/0000-0002-8576-942X
Anita Andreano http://orcid.org/0000-0002-9667-7010
Antonio Giampiero Russo http://orcid.org/0000-0002-5681-5861

REFERENCES

1 Pines JM, Hilton JA, Weber EJ, *et al*. International perspectives on emergency department crowding. *Acad Emerg Med* 2011;18:1358–70.
2 Wargon M, Guidet B, Hoang TD, *et al*. A systematic review of models for forecasting the number of emergency department visits. *Emerg Med J* 2009;26:395–9.
3 Programma Nazionale Esiti - PNE. Home, 2018. Available: https://pne.agenas.it/index.php [Accessed Feb 18, 2020].
4 Hwang U, McCarthy ML, Aronsky D, *et al*. Measures of crowding in the emergency department: a systematic review. *Acad Emerg Med* 2011;18:527–38.
5 Asplin BR. Measuring crowding: time for a paradigm shift. *Acad Emerg Med* 2006;13:459–61.
6 Wang Z, Xiong X, Wang S, *et al*. Causes of emergency department overcrowding and blockage of access to critical services in Beijing: a 2-year study. *J Emerg Med* 2018;54:665–73.
7 Jeyaraman MM, Copstein L, Al-Yousif N, *et al*. Interventions and strategies involving primary healthcare professionals to manage emergency department overcrowding: a scoping review. *BMJ Open* 2021;11:e048613.
8 Trzeciak S, Rivers EP. Emergency department overcrowding in the United States: an emerging threat to patient safety and public health. *Emerg Med J* 2003;20:402–5.
9 McKenna P, Heslin SM, Viccellio P, *et al*. Emergency department and hospital crowding: causes, consequences, and cures. *Clin Exp Emerg Med* 2019;6:189–95.
10 Hoot NR, Aronsky D. Systematic review of emergency department crowding: causes, effects, and solutions. *Ann Emerg Med* 2008;52:126–36.
11 Howard M, Davis B, Anderson C, *et al*. Patients' perspective on choosing the emergency department for nonurgent medical care: a qualitative study exploring one reason for overcrowding. *J Emerg Nurs.* 2005.
12 Angelillo IF, Ricciardi G, Nante N, *et al*. Appropriateness of hospital utilisation in Italy. *Public Health* 2000;114:9–14.
13 Hakkoymaz H. How appropriate are the emergency department admissions? A study from a tertiary hospital emergency department. *EJMI* 2020.
14 Wargon M, Casalino E, Guidet B. From model to forecasting: a multicenter study in emergency departments. *Acad Emerg Med* 2010;17:970–8.
15 Duwalage KI, Burkett E, White G, *et al*. Forecasting daily counts of patient presentations in Australian emergency departments using statistical models with time-varying predictors. *Emerg Med Australas* 2020;32:618–25.
16 Jones SS, Thomas A, Evans RS, *et al*. Forecasting daily patient volumes in the emergency department. *Acad Emerg Med* 2008;15:159–70.
17 Rising WR, O'Daniel JA, Roberts CS. Correlating weather and trauma admissions at a level I trauma center. *J Trauma* 2006;60:1096–100.
18 Holleman DR, Bowling RL, Gathy C. Predicting daily visits to a walk-in clinic and emergency department using calendar and weather data. *J Gen Intern Med* 1996;11:237–9.
19 Diehl AK, Morris MD, Mannis SA. Use of calendar and weather data to predict walk-in attendance. *South Med J* 1981;74:709–12.
20 Marcilio I, Hajat S, Gouveia N. Forecasting daily emergency department visits using calendar variables and ambient temperature readings. *Acad Emerg Med* 2013;20:769–77.
21 McCarthy ML, Zeger SL, Ding R, *et al*. The challenge of predicting demand for emergency department services. *Acad Emerg Med* 2008;15:337–46.
22 Díaz J, Alberdi JC, Pajares MS, *et al*. A model for forecasting emergency hospital admissions: effect of environmental variables. *J Environ Health* 2001;64:9–15.
23 Macgregor DM. Effect of weather on attendance with injury at a paediatric emergency department. *Emerg Med J* 2003;20:204–5.
24 Wargon M, Brun-Ney D, Beaujouan L, *et al*. No more winter crisis? forecasting daily bed requirements for emergency department admissions to hospital. *Eur J Emerg Med* 2018;25:250–6.
25 Murtas R, Decarli A, Russo AG. Trend of pneumonia diagnosis in emergency departments as a COVID-19 surveillance system: a time series study. *BMJ Open* 2021;11:e044388.
26 Morbey RA, Charlett A, Lake I, *et al*. Can syndromic surveillance help forecast winter Hospital bed pressures in England? *PLoS One* 2020;15:e0228804.
27 Italia. Provvedimento del Garante N. 2 del 16 giugno 2004. *Codice di deontologia e di buona condotta per i trattamenti di dati personali per scopi statistici e scientifici*.
28 Agenzia Regionale per la Protezione dell'Ambiente della Lombardia. ARPA Lombardia. Available: https://www.arpalombardia.it:443/Pages/ARPA_Home_Page.aspx [Accessed March 10, 2020].
29 Steadman RG. Norms of apparent temperature in Australia. *Aust Met Mag* 1994;43:1–16.
30 Ministero della Salute. Influenza. Available: http://www.salute.gov.it/portale/influenza/homeInfluenza.jsp [Accessed March 10, 2020].
31 Vatcheva KP, Lee M, McCormick JB, *et al*. Multicollinearity in regression analyses conducted in epidemiologic studies. *Epidemiology* 2016;6. doi:10.4172/2161-1165.1000227. [Epub ahead of print: 07 03 2016].

32 Hyndman RJ, Athanasopoulos G. *Forecasting: principles and practice*. OTexts, 2018.
33 M H, M.N S. A review on evaluation metrics for data classification evaluations. *IJDKP* 2015;5:01–11.
34 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020. https://www.R-project.org/
35 Salute M della. Pronto soccorso, via libera DELLA Conferenza Stato-Regioni alle nuove linee di indirizzo. Available: http://www.salute.gov.it/portale/news/p3_2_1_1_1.jsp?lingua=italiano&menu=notizie&p=dalministero&id=3849 [Accessed April 24, 2020].
36 Hyndman [aut R, cre, cph. Forecast: Forecasting functions for time series and linear models, 2020. Available: https://CRAN.R-project.org/package=forecast [Accessed Nov 14, 2020].
37 Fox J, Weisberg S. *An R companion to applied regression*. SAGE Publications, 2018.
38 Jeffery MM, D'Onofrio G, Paek H, *et al*. Trends in emergency department visits and hospital admissions in health care systems in 5 states in the first months of the COVID-19 pandemic in the US. *JAMA Intern Med* 2020;180:1328.
39 Lavigne E, Gasparrini A, Wang X, *et al*. Extreme ambient temperatures and cardiorespiratory emergency room visits: assessing risk by comorbid health conditions in a time series study. *Environ Health* 2014;13:5.
40 Samet JM, Bishop Y, Speizer FE, *et al*. The relationship between air pollution and emergency room visits in an industrial community. *J Air Pollut Control Assoc* 1981;31:236–40.
41 Schanzer DL, Schwartz B. Impact of seasonal and pandemic influenza on emergency department visits, 2003-2010, Ontario, Canada. *Acad Emerg Med* 2013;20:388–97.
42 Hu M, Li M, Li W, *et al*. Joint analysis of spikes and local field potentials using copula. *Neuroimage* 2016;133:457–67.