Demographic and epidemiological characteristics of scorpion envenomation and daily forecasting of scorpion sting counts: the case of Touggourt in Algeria

Original article

Kaouthar Boubekeur [1], Mohamed L'Hadj [2] and Schehrazad Selmane [1,*]

[1]L'IFORCE, Faculty of Mathematics, University of Sciences and Technology Houari Boumediene, Bab Ezzouar, Algiers 16111, Algeria

[2]Health and Hospital Reform Services, Ministry of Health, Population and Hospital Reform, Algiers, Algeria

Funding: This research received no funding.

Corresponding author:

Schehrazad Selmane

[1]L'IFORCE, Faculty of Mathematics, University of Sciences and Technology Houari Boumediene, Bab Ezzouar, Algiers 16111, Algeria

Tel: +213-21247907, Fax: +213-21247907, E-mail: cselmane@usthb.dz

Running title: Epidemiology and forecasting of scorpion stings

CONFLICT OF INTEREST: The authors have no conflicts of interest to declare for this study.

ORCID

Schehrazad Selmane https://orcid.org/0000-0002-3035-8172

Mohamed L'Hadj https://orcid.org/0000-0001-5589-6148

Kaouthar Boubekeur https://orcid.org/0000-0003-4938-2022

ABSTRACT

OBJECTIVES: This study intended to provide a better insight into the demographic and epidemiological characteristics of scorpion envenomation in an endemic area in Algeria and to identify the model that best predicts daily scorpion sting counts.

METHODS: Daily sting data from January 1, 2013 to August 31, 2016 enrolled in this study were extracted from the designed questionnaires for scorpion stings in the two emergency services of Touggourt. Count regression models were applied to daily sting data.

RESULTS: A total of 4712 scorpion sting cases were notified. Of these, 70% were aged between 10 and 49 years old. The male-to-female ratio was 1.3. The upper and lower limbs were the human body parts the most subject to scorpion stings (90.4% of cases). In most instances, stings have been mild (92.8%). The percent of stung people inside dwellings was 68.8%. The hourly distribution of stings showed a peak between 10:00 a.m. and 11:00 a.m. The daily number of stings ranged from 0 to 24. The occurrence of stings was highest on Sundays. The incidence of scorpion stings increases sharply in the summer. The mean annual incidence rate was 542 cases/100000 inhabitants. The fitted count regression models showed that the negative binomial Hurdle model was appropriate to forecast daily stings in terms of temperature and relative humidity and the fitted data agreed considerably with the actual data.

CONCLUSIONS: This study shows that daily scorpion sting data provides more meaningful insights; moreover, the negative binomial Hurdle model should be preferred for predicting daily scorpion sting counts.

KEY WORDS: Algeria, Count regression models, Epidemiology, Scorpion stings, Touggourt

INTRODUCTION

Scorpion stings constitute an actual public health concern in many arid, semi-arid or Saharan regions throughout the world within the stripe of 50 ° latitude, both south and north. The populations living in scorpion stings risk areas are estimated at approximately two billion people. Each year, an estimated 1.2 million are victims of scorpion stings worldwide. Second only to snakebites in terms of venomous animal-related human fatalities, scorpions are responsible for an estimated 3000 deaths each year. These estimates concern only the few countries having a reliable system for scorpion sting epidemiological surveillance [1].

Owing to its climatic and geographic conditions and ecological characteristics, Algeria houses a diverse scorpion fauna [2]. The scorpion envenomations represent a major public health issue with nearly three-quarters of the country's population at risk of scorpion stings [3]. A national program for scorpion sting control has been running since 1986 [4,5]. The yearly number of scorpion stings fluctuates around 50 000 cases, thus imposing a heavy burden on the nation's health care expenditures [2,6]. For the past ten years, the yearly number of deaths has been around 50 deaths; most, if not all, fatalities were caused by the Androctonus australis stings [7].

Several epidemiological surveys were conducted in several affected regions by scorpionism [8,9,10]. Most of the mathematical approaches intended to analyse the collected data are based on descriptive statistics [11,12]. The association of the scorpion sting incidence with climate variables, using monthly data, in many affected regions, was previously performed using multiple linear regressions [13,14,15]. Until recently, other statistical approaches, using monthly data, such times series analysis and count data, are now taking over [16-19].

In this study, we scrutinised the demographic and epidemiological characteristics of scorpion stings in Touggourt region and we presented a best fit model to forecast the daily scorpion sting counts in that region using climate variables. Even though the used approach was applied in several studies related to other health issues [20-25], however, it is the first time it is applied to daily scorpion stings.

MATERIALS AND METHODS

Study area

Touggourt lies in the Sahara, in the Righ valley region, north-eastern Algeria, along with sand dunes and salt lakes to the north and south and small hills to the west. It is made up of eight districts. The region is characterised by a harsh winter and a hot and dry summer.

Scorpion sting data

Daily scorpion sting data from January 1, 2013 to August 31, 2016 were drawn from scorpion sting designed questionnaires from the two emergency services of the study region, namely, public hospital establishment of Touggourt and nearby public health establishment of

Touggourt. The questionnaires have been anonymised. The extracted data included, inter alia, the following information: gender, age, date of sting, address, anatomical sting site, sting time, time of the 1st health medical examination, grade on first clinical examination, location (inside/outside the dwelling), and treatment administered. Data was analysed and graphs were generated to recognize trends and epidemiological and demographic features of scorpion stings using the software application Epi Info version 7.2.1.0, freely available at http://www.cdc.gov/epiinfo/.

Weather data

Daily data on the average temperature (T) and relative humidity (RH), accumulated rainfall (P), wind speed (W), and evaporation value (E) from January 1, 2013 to August 31, 2016, was supplied by Touggourt meteorological station.

Modelling method

Count data occur in many fields, among others, in public health and in epidemiology. Among the various models put forward to fit count data are charted the Poisson, zero-inflated Poisson (ZIP), Poisson hurdle (PH), negative binomial (NB), zero-inflated negative binomial (ZINB), and negative binomial hurdle (NBH). The Poisson model is the most widely used model when it comes to count events arising within a specific period. One key feature of Poisson model is the equality between the mean and the variance; an equality that does not always occur when dealing with real life data. Indeed, life data are often characterised by overdispersion, that is, the variance exceeds the mean. When overdispersion is present, Poisson regression is usually given up in favour of NB regression that adjusts the overdispersion. However, the NB model can only take into account overdispersion referring to unobserved heterogeneity in data but not to overdispersion arising from excess of zeros. To handle overdispersion resulting only from excess of zeros, zero-inflated Poisson and hurdle Poisson models are used. To account for overdispersion from excess of zeros and unobserved heterogeneity, zero-inflated NB or hurdle NB models are preferred given their flexibility [26-29]. We got away the technicalities to enable readers to follow the presented material; a detailed description of each involved model as well as the various statistical tests to assess overdispersion, to compare the regarded models and to select the best model is given in a supplementary content. We briefly review only the model that was appropriate to forecast the daily stings in terms of climate variables.

The hurdle model is a two-component model; a hurdle component used to model large zero counts, and a truncated count component used to model only the positive counts. The hurdle model can be expressed as follows:

$$P(Y_i = y_i | x_i, z_i, \beta, \gamma) = \begin{cases} f_{zero}(0; z_i, \gamma) & if\ y_i = 0 \\ \dfrac{1 - f_{zero}(0; z_i, \gamma)}{1 - f_{count}(0; x_i, \beta)} f_{count}(y_i; x_i, \beta) = \Phi\ f_{count}(y_i; x_i, \beta) & if\ y_i > 0 \end{cases}$$

where y is the dependent variable, x is a vector of covariates for positive counts and z is a vector of covariates in the zero part. The model parameters β, γ are related to x and z respectively, and are estimated by maximum likelihood. f_{zero} is a probability density function at least on {0, 1} or {0, 1, 2, …}, and f_{count} is a probability density function on {0, 1, 2, …}. The f_{zero} part (where $y_i = 0$) is modelled using a binary logit (logistic regression) model, where all positive counts are given a value of 1. The probability of $y_i = 0$ is given by

$$f_{zero}(0; z_i, \gamma) = \frac{1}{1 + e^{z_i \gamma}}$$

Evidently, the probability of a nonzero count is given by $1 - f_{zero}(0; z_i, \gamma)$. Regarding $f_{count}(y_i; x_i, \beta)$, it is modelled with a left-truncated ($y_i > 0$) Poisson model or negative binomial model in case of overdispersion with log link. The corresponding mean regression relationship is given by

$$\log(\mu_i) = x_i^T \beta + \log\left(1 - f_{zero}(0; z_i, \gamma)\right) - log\left(1 - f_{count}(0; x_i, \beta)\right).$$

Some climate variables were shown to affect the count of monthly scorpion stings [3,13,14], so we considered incorporating the available climate variables at Touggourt meteorological station into the models. In order to take into account other factors which may contribute to scorpion sting accidents, a trend variable (Tr) has been incorporated into the considered models. The scatter plot, along with the Pearson product-moment correlation coefficient between the scorpion sting variable, S, and each one of climate variables are used as a guide in selecting the appropriate climate covariates.

The six regression models outlined above were fitted to the data. The Akaike's information criterion (AIC) and Bayesian information criterion (BIC) were estimated to assess the goodness of fit of the six models. The like better model is the one with the minimum AIC and BIC values [26]. To compare and test the goodness of fit between model pairs, likelihood ratio tests and Vuong statistics were performed. And ultimately to evaluate the six models, the metrics mean absolute error (MAE) and root mean squared error (RMSE) were estimated in order to measure the closeness of the observed values to the predicted values [28].

The RStudio software, version 1.1.383 © 2009-2017 was used to perform all analysis and modelling part [30-32]. The R code for mapping the incidence and the R code for fitting and forecasting the daily scorpion sting data are provided in supplementary content.

RESULTS

Spatial distribution of scorpion sting incidence in different districts of Touggourt

Figure 1 depicts the spatial repartition per district of scorpion sting incidence per 100 000 population for the years 2013, 2014 and 2015. An increase of scorpion sting incidence per year is observed in Blidet Amor district, a decrease is observed in Sidi Slimane district, and slight fluctuations prevail in the remaining six districts. The annual incidence rate was 561 cases/100 000 inhabitants in 2013, 541 cases/100 000 inhabitants in 2014, 523 cases/ 100 000 inhabitants in 2015. At district level, the incidence ranged between 341 and 977 per 100 000 inhabitants; exceeding largely the mean national incidence rate estimated at 125/100 000 inhabitants [6]. The correlation between the number of scorpion stings per district and the size of the population was high; the Pearson product-moment correlation coefficient was 0.91 in 2013, 0.81 in 2014 and 0.75 in 2015.

Epidemiological features of scorpion envenomations

From January 1st, 2013 to August 31, 2016, a total of 4712 scorpion sting victims of which seven scorpion-related human fatalities were notified at the two emergency services in Touggourt. Table 1 gives a breakdown of notified cases by gender, by age groups, by anatomical sting site, by grade on first clinical examination, by sting time and by location. Victims were predominantly male (56.4%). Children under the age of 15 account for about one-fifth of stung people (19.9%) and the population in active age (18-64 years) was as expected the most stung with 68.7% of whom 57.2% males and 42.8% females and the elderly aged 65 and more represented 7.5% of cases. The mean ± standard deviation age was 32 ± 19 years (95% CI = 31.47, 32.57). The limbs were the most exposed body parts to scorpion stings gathering 91.3% of cases of which 52.8% (40.5% females and 59.5% males) affected the lower limbs and 38.5% (47.7% females and 52.3% males) the upper limbs. The percent of stung people inside dwellings was 70.7% (51.4% of females and 48.6% males) with the highest recorded frequency between 10 a.m. and 11 a.m. Outside dwellings the percent of stung people was 29.3% with a male predominance (75.8% males and 24.2% of females) with the highest recorded frequency between 10 a.m. and 11 a.m. Medical care was performed in a timely manner; 92.7% of victims received medical assistance within 2 hours following the sting accident with 54% in less than one hour.

At the first medical examination, 4371 (92.8%) cases were classified as mild with male-to-female ratio 1.3, 187 (4%) cases as moderate with male-to-female ratio 1.4, seven case as severe of whom 5 females, and for 147 (3.1%) cases the information have not been filled in. Antivenom has been used in 4455 (94.5%) patients. Seven scorpion-related human fatalities were notified (3 males and 4 females). At the first medical exam, five cases were classified as severe. The victims were aged respectively 4, 15, 26, 30, 39, 80, and 88. All of them were stung inside dwelling and on the limbs (of which 5 on lower limbs).

Stings occur year-round as depicted in Figure 2. Over the 1339-day covering the study period, no stings were recorded during 320 days in the emergency services. For the years 2013-2015, more than half of the sting cases were recorded during the summer, followed by the spring, then the autumn and the winter (Figure 3A). August tops the monthly frequency of events followed by September then July (Figure 3A). The hourly distribution of stings, displayed in Figure 3B, shows a peak between 9 a.m. and 11 a.m. (17.6% of cases) and off-peak at 5 a.m. and at 4 p.m. The maximum daily stings occurred on September 29, 2013 with 24 sting cases of which 20 stings occurred inside dwellings, followed by 21 sting cases of which 19 stings inside dwellings on September 1, 2015, and 19 sting cases of which 16 stings inside dwellings on August 18, 2015 (Figure 2). It should be noted that the daily peaks mentioned were all mild.

Statistical modelling output

The dataset includes 1339 observations and was divided into a training set including 1095 (82%) observations from January 1, 2013 to December 31, 2015 to estimate the parameters of the best model, and a testing set including 244 (18%) observations from January 1, 2016 to August 31, 2016 to evaluate the model.

Examination of the relationship between monthly scorpion sting data and the climate variables showed a strong correlation with monthly average temperature (r = 0.91) with an apparent quadratic relationship ($S = -0.0004\ T^2 + 0.19\ T + 9.73$; $R^2 = 0.95$) and high correlation with relative humidity (r = -0.70) and a quadratic relationship ($S = 0.0007\ RH^2 - 0.29\ RH + 61.74$; $R^2 = 0.72$); which is consistent with obtained results in other geographical regions affected by scorpionism [16,17]. At daily level, the relationship is not readily apparent, the correlation is high with T (r = 0.69) and significant with RH (r = - 0.45), however, these climates variables were statistically significant in the modelling process.

We started by fitting the Poisson model with S, the daily scorpion sting count, as dependent variable and all climate variables and trend as covariates. Drawing on the likelihood ratio test only average temperature, relative humidity, and trend had been withheld in the final fitted standard Poisson regression model.

As the variance (14.78) is 4 times higher than the mean (3.52) which refers to overdispersion and close to quarter of the observations are zeros (23.9%), regression models accounting for both overdispersion and excess of zeros have been considered. Hence, the ZIP, PH, NB, ZINB, and NBH models were fitted to daily scorpion sting count using the same covariates as for Poisson model. In Table 2 are displayed the statistical tests used to compare and test the goodness-of-fit among the six models. NB-type models (NB, ZINB and NBH) have smaller AIC and BIC values than the Poisson-type models (Poisson, ZIP and PH). Likewise LR $\chi 2$ is highly significant for all models and NB-type models have smaller values than the Poisson-type models. Moreover, the NBH model fitted the daily scorpion sting data better than all the other models. NB-type models are preferred to Poisson-type models to handle the overdispersion of the daily scorpion sting counts and the statistical tests are indicative of the fact that overdispersion is due to both heterogeneity and excess zeros. Moreover, Vuong tests indicate that both the zero-inflated models and Hurdle models were better and effective at handling the excessive zero counts than the Poisson and NB models. Finally, the testing set was used to evaluate the models. The metrics RMSE and MAE were estimated. The evaluation of the models showed that the metric MAE is smaller while predicting the frequency of stings from January to March using Poisson model and from April to August using NBH model. These findings therefore suggest that the negative binomial Hurdle model should be preferred when modelling daily scorpion sting count using considered criteria.

For brevity, only the NBH model estimates are presented. The estimated parameters, standard errors, and associated p-values are displayed in Table 3. The NBH model consists two parts:

The zero part contains information about variables that the nested logistic regression model has used to estimate the probability ϕ of observing a zero count

$$logit(S = 0) = -3.1503 + 0.1556 * T + 0.0022 * RH + 2.4777 * Tr$$

Notice that the logistic regression model did not find RH variable useful for estimating ϕ. Indeed, the regression coefficient was found to be not statistically significant at the 95% confidence level, as indicated by the p value (0.8071) which is greater than 0.05. The two variables that the logistic regression model determined as useful for estimating the probability of observing a zero count were T and Tr.

The Count Part contains information about variables that the NBH model used to estimate scorpion sting count on the condition that S > 0.

$$\log(S) = -2.4052 + 0.1129 * T + 0.0148 * RH + 0.1242Tr$$

We can see that the coefficients for T and RH are statistically significant, as evidenced by their respective p values. The coefficient for T and RH are positive meaning that as T and RH goes up, the number scorpion sting goes up.

By aggregating the estimated daily data for the validation period into weekly data and monthly data, the predicted data were strongly correlated with the actual data (Pearson product-moment correlation coefficient: r = 0.94 for weekly data and r = 0.98 for monthly data) as shown in Figure 4, confirming the appropriateness and effectiveness of the NBH model to predict the scorpion stings in terms of climate variables with very high accuracy.

DISCUSSION

The purpose of this paper was, first, to provide a better insight into the demographic and epidemiological characteristics of scorpion stings in an endemic region in Algeria, and second, to model the daily scorpion sting counts as function of climate factors using six count data models and to compare the performance of these models.

The epidemiological analysis carried out showed a male predominance following thus similar pattern to that at the national level [6]. Similar findings have been documented in Morocco and Iran [9,33]. However, that is not a hard rule; a study carried out in the Rio Grande do Norte State in Brazil shows a female predominance [34]. The drawn findings must be viewed cautiously as it concerns the stratification by gender of the number of scorpion stings and not the stratification by gender of the incidence rate. Body parts mostly prone to stings are legs and hands. Several studies carried out in other affected provinces of the country and also in other affected countries by this issue are consistent with these findings [8,9,16,17,33]. The most likely reason that moving body parts are at greater risk of scorpion stings is the fact that many victims do not protect themselves to ovoid the stings in most of their activities; they bear a significant share of responsibility for that accidents either through ignorance or negligence. With respect to location, a significant variation according to the gender of the patient is observed, and almost three-quarters of sting accidents occurred inside dwellings. The percentage of people stung inside dwellings differs from province to province;

in Biskra province, situated in central-eastern area, 46% of sting cases occurred inside dwellings, whereas in El Bayadh province, western of Algeria, 65% of sting cases arose inside dwellings [16,17]. The high frequency of accidents inside dwelling may be related to a great exposure to scorpion accidents in home environment while accomplishing domestic activities. Moreover, due to high temperature in that Saharan region, people spend the majority of their waking hours inside dwellings. Ninety percent of sting accidents occurred between April to October, corresponding to the hottest months (July and August) that records the highest frequency of accidents, followed the harvest date season (September-October) and the planting season (April-May) where scorpion sting accidents are common on farms. The high incidence in summer is perceived in various regions affected by scorpionism [9,33]. Several studies in different geographical regions have documented varied age-class distributions of stings corroborating thus the geographical variation within epidemiological indicators as noted in the global appraisal on the epidemiology of scorpionism by Chippaux & Goiffon [1]. Individuals aged 20 to 29 years old made up the largest number of cases, accounting for 24% of all cases; this was also reported in the Rio Grande do Norte State [34]. Sting cases were predominantly with mild severity and progressing towards recovery. Similar results were reported in other regions [8,9,10]. The time prior to first medical care was for most victims less than 3 hours after sting, revealing that the population becomes aware of the need to seek care right after the sting and revealing also the failure of public awareness campaigns to avoid the sting itself. Due to the seasonal pattern of scorpion stings, launching regular public awareness-raising campaigns in high risk regions during the peak months can significantly reduce the number of stings incidents.

To our references knowledge, modelling of scorpion stings was performed with monthly data using statistical approaches such multiple linear regression, time series analysis, and count data [13-17]. No modelling of daily observations has been conducted to date. Aside from delivering relevant insights, a modelling at the daily data basis could produce high accuracy predictions provided the availability of accurate forecasts of the climate variables.

In this study and for the first time, scorpion sting data are exploited and evaluated without recourse to monthly aggregation as it has been done in the literature to date. The fitted count regression models to daily scorpion sting count as dependent variable, and temperature, relative humidity and the trend as independent variables showed that the negative binomial Hurdle model yielded adequate fits. It is shown that T and RH were effective factors for sting accidents and that T had the highest effect on that accident. Additionally, daily predictions by means of NBH model provide highly accurate monthly forecasts. These predictions could assist public health decision-makers in initiating appropriate, rapid and effective measures and also in containing any an unusual situation.

What is emerging from this study is the improvement in forecasting accuracy of scorpion sting cases in terms of weather condition. However, in view of the resulting impact on either financial or human terms, more research is needed to define alternative policies to avoid sting accidents rather than accepting them as inevitable. It is felt that this first study on daily scorpion sting count in terms of the used statistical approach and the achieved results is a contribution both towards literature and the modelling of scorpion stings and it is thought to encourage further investigations and studies about scorpionism issue.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare for this study.

SUPPLEMENTARY MATERIAL

Supplementary material is available at http://www.e-epih.org/.

REFERENCES

1. Chippaux JP, Goyffon M. Epidemiology of scorpionism: a global appraisal. Acta trop 2008;107(2):71-9.
2. Vachon M. Etudes sur les scorpions; 1952 [cited 2020 Apr 24]. Available from: http://www.ntnu.no/ub/scorpion-files/vachon.php (French).
3. Selmane S. Scorpion Envenomations and Climate Conditions: the case of Naama Province in Algeria. Inter J Math Models Methods Appl Sci. 2015;9:261-8.
4. Ministry of Health, Population and Hospital Reform (2009). National committee for the fight against the scorpion envenomment. Support for scorpion envenomation. (French)
5. Ministry of Health, Population and Hospital Reform, Prevention Department. Instruction no 04 / MSPRH / DP of March 4, 2012 on the new plan of notification and management of scorpionic envenomation cases. Available from: http://www.sante.dz/inst_env_scorp.pdf (French).
6. Laïd Y, Boutekdjiret L, Oudjehane R, Bachiri K. Scorpion envenomation: the annual report on the epidemiological situation in Algeria. Health and Environment Unit, National Institute of Public Health 2000-2014 [cited 2020 Apr 24]. Available from: http://www.insp.dz/index.php/envenimation-scorpionique (French).
7. National Institute of Public Health - Epidemiological characteristics of deaths from scorpionic envenomation in Algeria. 2011. http://insp.dz/index.php/Non-categorise/envenimation-scorpionique.html (French).
8. Selmane S, Benferhat L, L'Hadj M, Zhu H. Scorpionism in Sidi Okba, Algeria: A cross-sectional study of 2016 stung patients between 2014 and 2015. Tropical biomedicine. 2017;34(2):425-32.

9. El-Yamani R, Soulaymani A, Serhier Z. Epidemiology of Scorpion Envenoming in the Prefecture of Figuig, Morocco. . Int J Med Toxicol Forensic Med. 2019;9:17-24.
10. Ozkan O, Uzun R, Adiguzel S, Cesaretli Y, Ertek M. Evaluation of scorpion sting incidence in turkey. J Venom Anim Toxins Incl Trop Dis. 2008;14:128-40.
11. Nejati J, Saghafipour A, Rafinejad J, Mozaffari E, Keyhani A, Abolhasani A, et al. Scorpion composition and scorpionism in a high-risk area, the southwest of Iran. Electron Physician. 2018;10(7):7138-45.
12. El hidan Ma, Touloun O, El Oufir R, Boumezzough A. Epidemiological and spatial analysis of scorpion stings in two regions of Morocco: Marrakesh-Tensift-Al Haouz and Souss-Massa-Draa. J. Coast. Life Med. 2016;4:299-304.
13. Chowell G, Hyman JM, Díaz-Dueñas P, Hengartner NW. Predicting scorpion sting incidence in an endemic region using climatological variables. Int J Environ Health Res. 2005;15(6):425-35.
14. Selmane S, L'Hadj M. Regression Analysis on Scorpion Envenomation and Climate Variables in M'Sila Province, Algeria from 2001 to 2010. Int J Math Trends Technol (IJMTT). 2014;13:1-9.
15. Selmane S, Hadj H, Benferhat L, editors. The Impact of Climate Variables on the Incidence of Scorpion Stings in Humans in M'Sila's Province in Algeria. Proceedings of the World Congress on Engineering; 2014; 638-641.
16. Selmane S, L'Hadj M. Forecasting and prediction of scorpion sting cases in Biskra province, Algeria, using a seasonal autoregressive integrated moving average model. Epidemiology and Health. 2016;38.
17. Selmane S, Benferhat L, L'Hadj M, Zhu H. Modelling the scorpion stings using surveillance data in El Bayadh Province, Algeria. Asian Pac J Trop Dis. 2016;6(12):961-968.
18. Ebrahimi V, Hamdami E, Moemenbellah-Fard MD, Ezzatzadegan Jahromi S. Predictive determinants of scorpion stings in a tropical zone of south Iran: use of mixed seasonal autoregressive moving average model. J Venom Anim Toxins Incl Trop Dis. 2017;23:39.
19. Kassiri H, Shemshad K, Kassiri A, Shemshad A-A, Teimori A. Epidemiological and Climatological Factors Influencing on Scorpion Envenoming in Baghmalek County, Iran. Acad J Entomol. 2013;6:47-54.
20. Chipeta MG, Ngwira BM, Simoonga C, Kazembe LN. Zero adjusted models with applications to analysing helminths count data. BMC Research Notes. 2014;7(1):856.
21. Cheung YB. Zero-inflated models for regression analysis of count data: a study of growth and development. Stat Med. 2002;21(10):1461-9.
22. Lee JH, Han G, Fulp WJ, Giuliano AR. Analysis of overdispersed count data: application to the Human Papillomavirus Infection in Men (HIM) Study. Epidemiol Infect. 2012;140(6):1087-94.
23. Rose CE, Martin SW, Wannemuehler KA, Plikaytis BD. On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. J Biopharm Stat. 2006;16(4):463-81.
24. Tüzen MF, Erbaş S. A comparison of count data models with an application to daily cigarette consumption of young persons. Communications in Statistics - Theory and Methods. 2018;47(23):5825-44.
25. F, Moghimbeigi A, Mahjub H, Nasehi M, Khodadost M. Factors associated with mortality from tuberculosis in Iran: an application of a generalized estimating

equation-based zero-inflated negative binomial model to national registry data National. Epidemiol Health. 2019;41:e2019032.
26. Cameron AC, Trivedi PK. Regression Analysis of Count Data. 2 ed. Cambridge: Cambridge University Press; 2013.
27. Hilbe JM. Negative Binomial Regression. 2 ed. Cambridge: Cambridge University Press; 2011.
28. Hilbe JM. Modeling Count Data. Cambridge: Cambridge University Press; 2014.
29. McCullagh P, Nelder JA. Generalized Linear Models, Second Edition: Chapman & Hall; 1989.
30. Beaujean A, Morgan G. Tutorial on using regression models with count outcomes using R. Pract assess res eval. 2016;21.
31. Friendly M, Meyer D, Zeileis A. Discrete data analysis with R: Visualization and modeling techniques for categorical and count data. New York: Chapman and Hall/CRC; 2015. 1-525 p.
32. Zeileis A, Kleiber C, Jackman S. Regression Models for Count Data in R. Journal of Statistical Software. 2008;27(8):25.
33. Ghorbani E, Mohammadi Bavani M, Jafarzadeh S, Saghafipour A, Jesri N, Moradiasl E, et al. Spatial Distribution of Scorpionism in Ardabil Province, North West of Iran. Int J Pediatr. 2018;6(9):8241-51.
34. Araújo KAM, Tavares AV, Marques MRV, Vieira AA, Leite RS. Epidemiological study of scorpion stings in the Rio Grande do Norte State, Northeastern Brazil. Rev Inst Med Trop São Paulo. 2017;59:e58.

Figure legends

Figure 1. Spatial distribution map of the scorpion sting incidence per 100 000 inhabitants in Touggourt, for the years 2013, 2014 and 2015.

Figure 2. Frequencies of daily scorpion sting in Touggourt from January 1, 2013 to August 31, 2016.

Figure 3. (A) Seasonal average and monthly average of recorded scorpion sting cases in Touggourt in the period 2013-2015 and (B) The hourly breakdown of scorpion sting cases.

Figure 4. Scatter plots of predicted versus actual of scorpion sting cases from January to August 2016. A. weekly data, B. Monthly data.

Table 1. Demographic and epidemiological characteristics of patients stung by scorpions from January 1, 2013 to August 31, 2016

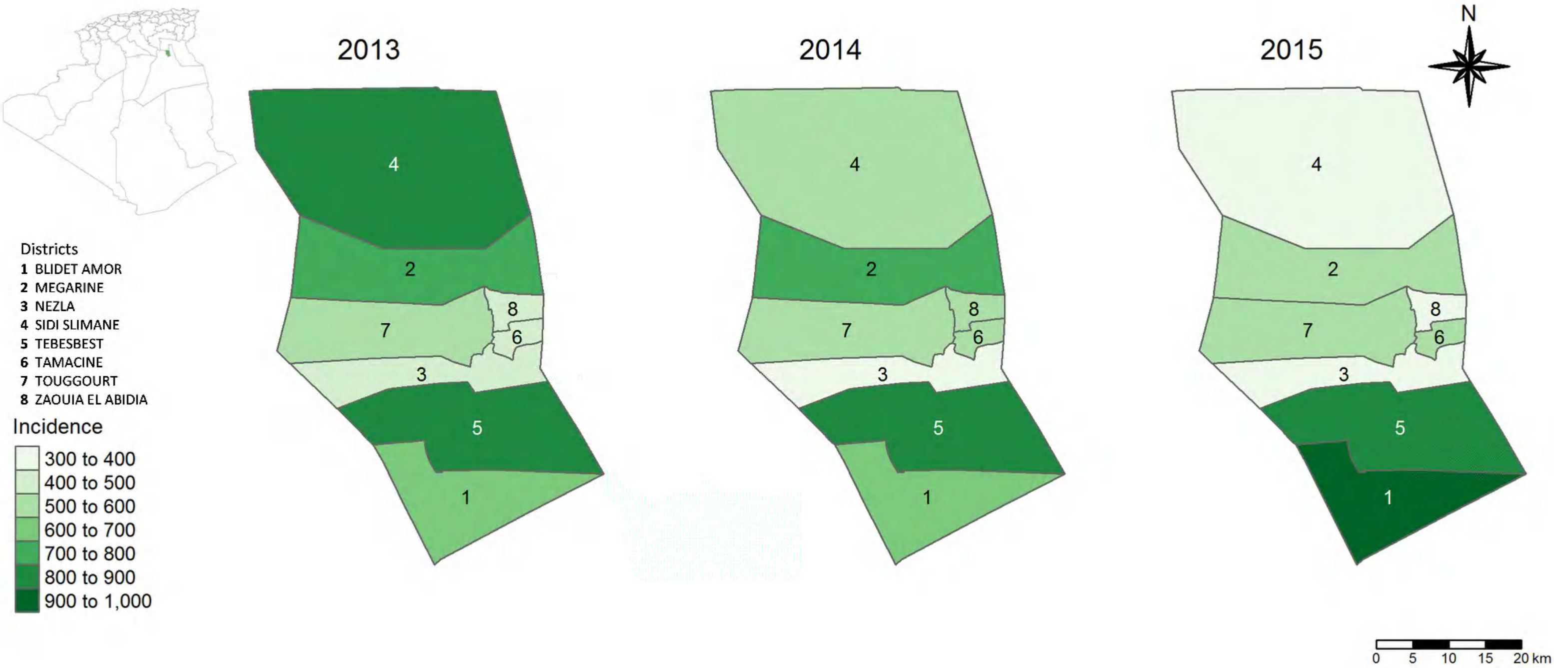
Gender	F	M	M/F	Total
Age groups				
0 - 9	200	218	1.1	418
10 - 19	370	511	1.4	881
20 - 29	437	677	1.5	1114
30 - 39	355	431	1.2	786
40 - 49	227	272	1.2	499
50 - 59	198	197	1.0	395
60 - 69	120	158	1.3	278
70 - 79	57	98	1.7	155
> 80	27	40	1.5	67
Unknown	65	54	0.8	119
Anatomical sting site				
Lower Limbs	998	1467	1.5	2465
Upper Limbs	855	939	1.1	1794
Head/Neck	120	113	0.9	233
Trunk	65	108	1.7	173
Unknown	18	29	1.6	47
Grade on first clinical examination				
Mild	1910	2461	1.3	4371
Moderate	77	110	1.4	187
Severe	5	2	0.4	7
Unknown	64	83	1.3	147
Sting time				
0 - 5	414	531	1.3	945
6 - 11	683	854	1.3	1537
12 - 17	453	548	1.2	1001
18 - 23	453	668	1.5	1121
Unknown	53	55	1.0	108
Location				
Inside dwellings	1667	1575	0.9	3242
Outside dwellings	325	1020	3.1	1345
Unknown	64	61	1.0	125

Table 2. AIC, BIC, Likelihood Ratio Chi-Square results of the models. LRT and Vuong test results of models.

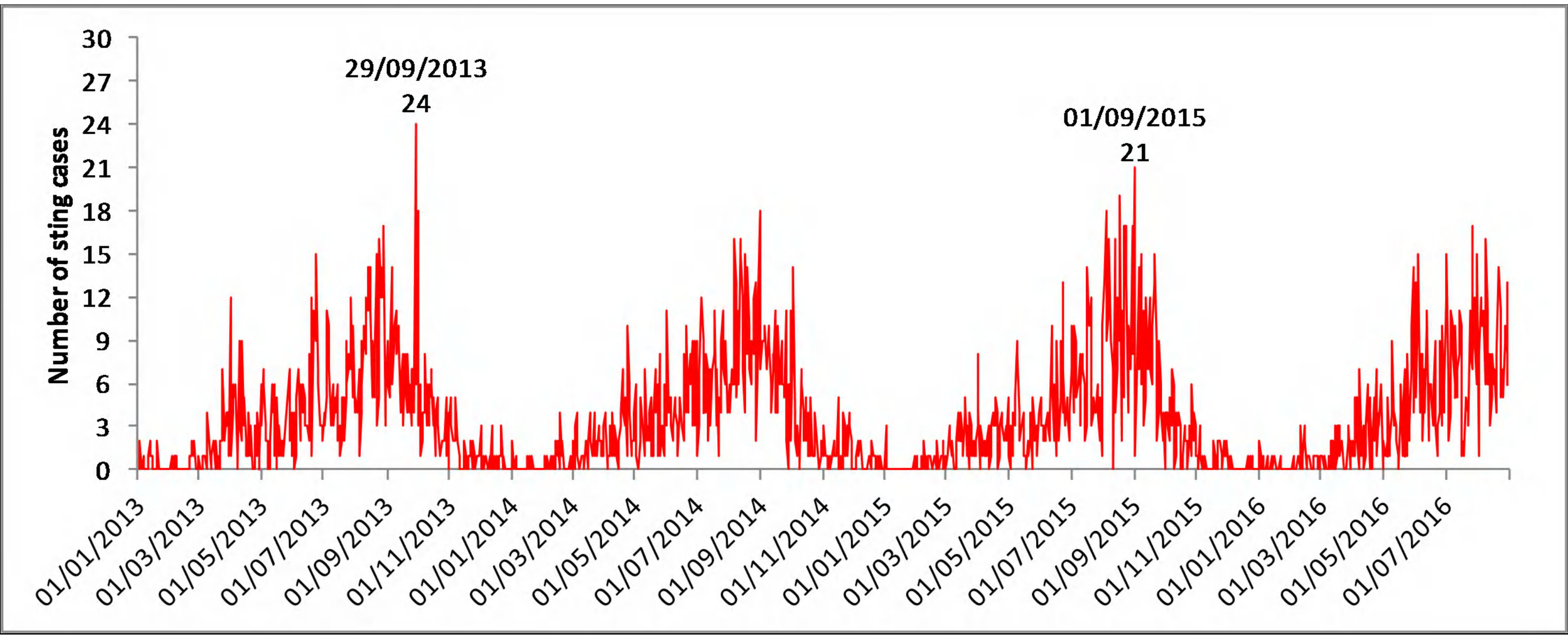
	Poisson	NB	ZIP	ZINB	PH	NBH
AIC	4472.9	4240.1	4393.2	4216.8	4392.1	4208.9
BIC	4492.9	4265.1	4433.2	4261.8	4432.1	4253.9
LR χ^2	4464.9	4230.1	4377.2	4198.8	4376.1	4190.9
χ^2 p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
LRT	NB-P	ZINB-ZIP	NBH-PH	P-ZIP	NB-ZINB	
	234.83 ($p < 0.001$)	178.44 ($p < 0.001$)	185.23 ($p < 0.001$)	87.68 ($p < 0.001$)	31.28 ($p < 0.001$)	
Vuong	ZIP-P	ZINB -NB	PH-P	NBH-NB	PH-ZIP	NBH-ZINB
	4.07	2.79	4.07	3.05	0.28	1.2
	$p < 0.001$	$p = 0.0026$	$p < 0.001$	$p = 0.0012$	$p = 0.3879$	$p = 0.1145$

Table 3. Estimated parameters, standard errors (SE), and associated p-values for the NBH model.

	Zero Part			Count Part		
Factor	Estimate	SE	*p*-value	Estimate	SE	*p*-value
Intercept	-3.1503	0.7252	< 0.001	-2.4052	0.2533	< 0.001
T	0.1556	0.0271	< 0.001	0.1129	0.0073	< 0.001
RH	0.0022	0.0090	0.8071	0.0148	0.0028	< 0.001
Tr	2.4777	0.6397	0.0001	0.1242	0.2196	0.572



2013
2014
2015
N
Districts
1 BLIDET AMOR
2 MEGARINE
3 NEZLA
4 SIDI SLIMANE
5 TEBESBEST
6 TAMACINE
7 TOUGGOURT
8 ZAOUIA EL ABIDIA
Incidence
300 to 400
400 to 500
500 to 600
600 to 700
700 to 800
800 to 900
900 to 1,000
0 5 10 15 20 km



Number of sting cases
30
27
24
21
18
15
12
9
6
3
0
29/09/2013
24
01/09/2015
21
01/01/2013
01/03/2013
01/05/2013
01/07/2013
01/09/2013
01/11/2013
01/01/2014
01/03/2014
01/05/2014
01/07/2014
01/09/2014
01/11/2014
01/01/2015
01/03/2015
01/05/2015
01/07/2015
01/09/2015
01/11/2015
01/01/2016
01/03/2016
01/05/2016
01/07/2016

Figure 3. (A) Seasonal average and monthl

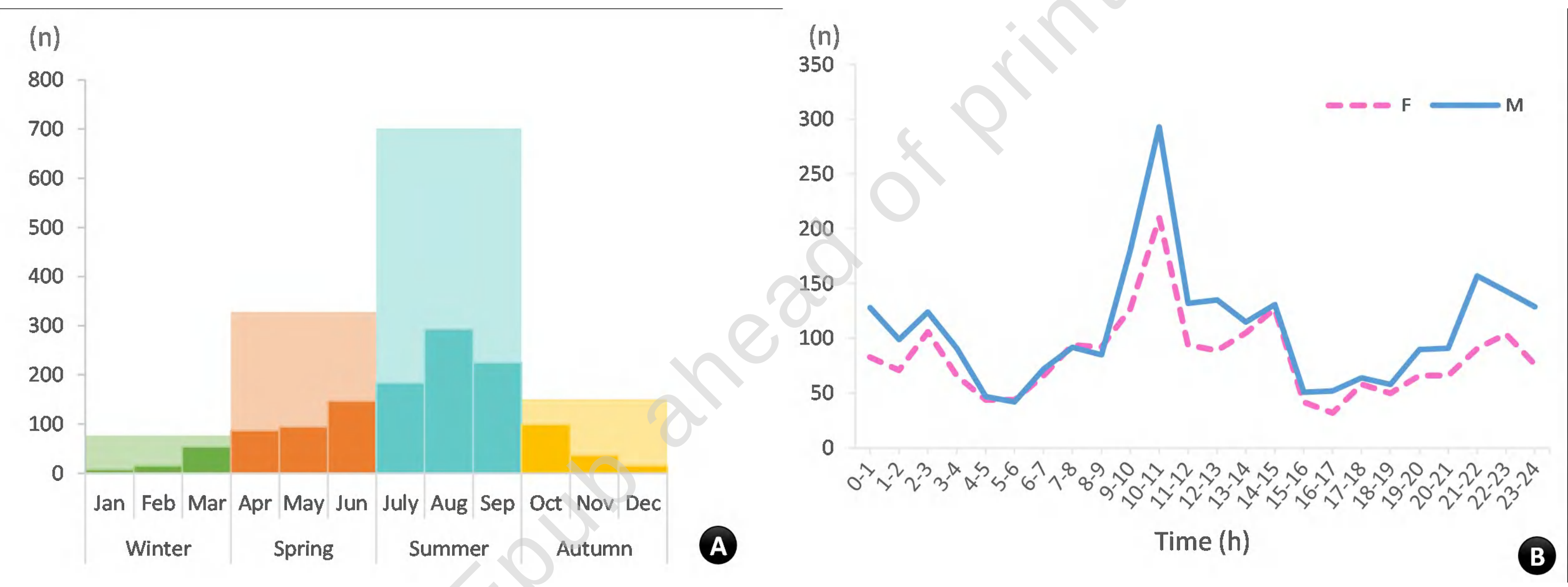


Figure 4. Scatter plots of predicted versus