Bayesian Zero- Inflated Poisson model for prognosis of demographic factors associated with using crystal meth in Tehran population

Asma Pourhoseingholi1,2, Ahmad Reza Baghestani1, Erfan Ghasemi1, Alireza Akbarzadeh Baghban3*, Mariet Ghazarian4

Received: 16 Aug 2017                Published: 19 Mar 2018

Abstract
Background: Use of methamphetamine (MA) and other stimulants has increased steadily over the past 10 years. Risk factor evaluation to reduce the problem in the community is one solution to protect people from addiction. This study aimed at using Bayesian zero-inflated Poisson (ZIP) model to investigate the relationship between the number of using crystal meth and some demographic factors in Tehran population.

Methods: A cross-sectional study was conducted to investigate crystal meth abuse in Tehran, the capital of Iran, in 2012. Stratified sampling method was used to select samples from 22 urban areas of Tehran. Trained researchers referred to the public places, such as streets, parks, squares, and libraries, to perform face-to-face interviews with the randomly selected samples. Bayesian ZIP model was used to perform the analysis, and SAS 9.3 program was used for data analysis.

Results: A total of 993 individuals were studied. According to Bayesian ZIP model, sex (mean= -0.27, 95%CI (-0.485, -0.061)), age (mean= 0.03, 95%CI (0.018, 0.043)), high school level education (mean= 1.276, 95%CI (0.699, 1.69)), diploma level education (mean= 10.4, 95%CI (0.511, 1.69)), and university level education (mean= 0.69, 95%CI (0.142, 1.33)) were all found to have significant associations with crystal meth usage, being the dependent variable.

Conclusion: Males, those with higher education levels, and older people in Tehran population are more likely to use crystal meth. This demographic information may be useful in designing preventive programs. Moreover, it is better to analyze count data with excessive zeroes using Bayesian zero-inflated model instead of the usual count models.

Keywords: Bayesian analysis, Crystal meth, Zero-inflated Poisson, Tehran population

Introduction
Opiates use has a long history in Iran (1). Especially, opium and hashish have been used since 300 years ago (2). On the other hand, Iranians have used opium and its pharmacological and psychotropic effects to manage some medical conditions related to pain (3, 4). However, in this century, many different abuse drugs exist, which have been surging in Iran (5, 6). These different drugs encompass heroin, high-grade crystal heroin (Kerack), crystal methamphetamine, buprenorphine, cocaine, tramadol, and ecstasy (6, 7). Among these drugs, the use of methamphetamine (MA) has increased, especially among young Iranians (8). Recent studies have shown a rapid increase in MA usage...
Bayesian zero inflated model in using crystal meth in Tehran population

among young adults (8). For example, a study found that the mean drug abuse initiation age of 25% of substance abusers, who reported using MA, was 18 years (9). Another study revealed that 16% of addicts, who referred to addiction treatment centers, had tried MA (10). Thus, it seems that using MA and other stimulants has increased steadily over the past 10 years.

Smoking, sniffing, injection, and ingestion are the main routes of methamphetamine usage (11). However, smoking is the most common route of methamphetamine administration in Iran. The smokable form of methamphetamine hydrochloride is called "crystal meth" and is more likely to cause addiction than other forms of methamphetamine (12, 13). Many social, legal, and emotional problems are caused as the result of the increasing increase in the usage of this drug. Furthermore, the destructive effect of this drug on health makes it an important health concern in any society. Memory loss, malnutrition, aggression, psychopathic behavior, and severe dental problems are the most common health problems associated with using crystal meth (14, 15). Transmission of infectious diseases, such as hepatitis and HIV, is another major problem associated with crystal meth usage.

It seems that we can find numerous reasons for prevalence of crystal meth due to geographical location of Iran and some other reasons such as easy access, ease of production, lack of awareness of negative consequences, and lack of serious preventive programs (14). Also, there is a wrong perception about this drug in the society; for example, its effect on losing weight or helping increase wakefulness and performance (16, 17). Despite all that, there is limited information about this drug. As a result, it is of paramount importance to study the risk factors and other prognosis of this drug to provide more information for policymakers and set preventive programs.

It seems that demographic factors have an important role as risk factors in addiction. In medical studies, statistical modeling is used as an approach for risk factor analysis, and in this approach, selecting an appropriate model is more important. When the response variable is counted, several models could be used for data analysis. Sometimes, count data have an overdispersion problem because of having large number of zeros. This phenomenon is called zero inflation. Using usual count models in zero-inflated data causes misleading results. Lambert (18) proposed zero-inflated Poisson (ZIP) regression model for independent count data. Using these models has recently increased in the application areas of medical and health care fields (19-24).

Usually, statistical inference is based on classical approaches, such as expectation-maximization (see, eg, (25)) on maximum likelihood. However, Bayesian inference is another approach for estimating the parameters. Use of Bayesian approach has recently been increased in applied statistical methods (26). Developing computational Bayesian statistics is the key point to make major changes in statistics. Bayesian estimate of zero-inflated models have been introduced just recently. Using these models instead of classical models has also been increased in application field. Accordingly, this study aimed at applying Bayesian method to estimate the parameters in a zero-inflated Poisson model. In this model, the relationship between the number of crystal meth use and some demographic factors was investigated.

Methods

Participant

The data of this study were part of a cross-sectional study that investigated crystal meth abuse in Tehran, Iran, in 2012. The sample size was selected according to previous prevalence of crystal meth in Iran (prevalence of almost 5%, and α= 0.05 for calculating the sample size). The sample size formula was used according to prevalence. Those aged 14 to 55 years were selected as a target population because a previous study found this age range to be a high-risk group (9).

Stratified sampling method was conducted to select samples from 22 urban areas of Tehran. The sample size in each area was selected proportionately to the population size. Trained researchers referred to the public places, such as streets, parks, squares, and libraries, to perform face-to-face interviews with the randomly selected samples. A valid questionnaire with 62 items was used to collect data. The Cronbach alpha of this questionnaire for this study was 87%, which is at an acceptable range, as it is greater than the threshold level of 70%. The present study used the following demographic information: sex, age, education level (illiterate, high school, diploma, and university level education), and marital status (single, married, and widowed) as covariates and the count of using the drug as a dependent variable. Other data were collected using the same questionnaire; however, it was decided to only use the data about the count of using crystal meth and some demographic aspects of participants for this study. Informed consent was obtained from all study participants before the project began.

Statistical analysis

Descriptive statistics and frequency distribution, such as mean, standard deviation, and percentage, were calculated according to standard methods. Non-parametric tests including Mann-Whitney and Kruskal Wallis were used for univariate analysis due to non-normality of the data.

The outcome variable was the count of using crystal meth, and count models were needed to perform the analysis (27). Bayesian ZIP model was used to analyze the count data. It has been established that prior probability for all favorite parameters is needed in all Bayesian methods. In ZIP model, in addition to covariate parameters, there is a mixture parameter, for which there is a need to assume prior distribution. In this study, the mixture proportion had a uniform (0, 1) prior distribution, and all covariates parameters had diffuse normal (0, σ2 = 1000) prior distribution, expressing the lack of knowledge about the regression parameters. P-values less than 5% were considered as significant. For categorical variables, p-values were compared with reference group. Because of using Bayesian approach in the analysis, the credit interval was used to make a decision about the significant variables. When the credit interval included zero, the parameter was considered not significant.
Markov chain Monte Carlo (MCMC) algorithm was employed for fitting the Bayesian model. The SAS 9.3 university edition (free statistical software) program was used to compute the results.

**Results**

A total of 993 individuals were investigated in this study, among whom 448 (45.1%) were female. The mean (SD) age was 25.98 (5.96) years, ranging from 14 to 52 years. Table 1 demonstrates the distribution of different characteristics of the participants.

Of the entire participants in this study, 74 persons (7.5%) have reported at least one time use of crystal meth in their life. Distribution of the number of crystal usage among the studied participants is presented in Table 2. The number of using crystal meth was found to be statistically different between male and female participants (p<0.001). Also, educational level and marital status were found to be significant predictors of using crystal (p<0.001). The mean of using crystal in males was higher than in females; this mean was higher in the widowed and those with high school level education (Table 3).

In the next step of this study analysis, Bayesian ZIP was used to investigate the effect of all considered factors on the number of using crystal meth simultaneously. Table 4 demonstrates the result of this model. ZIP model revealed that sex, age, high school level education, diploma level education, and university level education were associated with using crystal meth. According to these results, females (mean= -0.27; 95%CI (-0.485, -0.061)), age (mean= 0.03; 95%CI (0.018, 0.043), high school level education (mean=...
Bayesian zero inflated model in using crystal meth in Tehran population

1.276; 95%CI (0.699, 1.90)), diploma level education (mean = 1.04; 95%CI (0.511, 1.69)), and university level education level of education (mean= 0.69; 95%CI (0.142, 1.330)) were found to have significant association with crystal usage. The expected mean of the frequency of using crystal meth in females, younger people, and illiterates were less than in males, older people, and those with higher education, respectively.

Discussion

According to this study, the prevalence of at least one time use of crystal meth in Iranian population was 7.5%. The information regarding this drug among the population in general is limited. According to Drug Situation Analysis Report of the United Nation Office of Drug and Crime, methamphetamine usage has increased from none in 2004 to 3.5% in 2007 in Iran (28), so it seems there is a dramatic change in this prevalence during 5 years. On the other hand, due to lack of data in many countries, there is no report on its exact worldwide prevalence. However, there are several reports from other countries who reported the prevalence of meth including Philippines (10%), Guam (7%), Estonia (5.5%), Nicaragua and New Zealand (4%), Poland (3.7%), Denmark (3.1%), Germany (2.4%), and France (1.8%). It seems that the prevalence may have increased in recent years (29).

In this study, sex, age, and education had a statistically significant relationship with crystal meth usage in the studied population in Tehran. According to the results of this study, the expected mean of the frequency of using crystal meth is almost 25% less in females. Also, the expected mean among people with high school level education, diploma level education, and university level education was 3.58, 2.82, and 1.99 times more than illiterate people, respectively. In addition, older people were more likely to use crystal meth, and the expected mean was 1.03 time of the younger people. In another study done in Hamedan (an Iranian province in the west), being single and having higher education level had a significant relationship with using crystal meth (8). Another study on Iranian heroin kerack-dependent females revealed that younger and low-educated people in that population were more likely to use methamphetamine (30). This study and also the Hamedan study both indicated that highly educated people were a group at-risk. However, another study on heroin kerack-dependent females had completely different results, showing that the population of the study is important.

As mentioned earlier, there are limited studies on crystal meth usage, so comparing other results with this study is not possible. In contrast, this can show the importance of this study being among the first of its kind. Lack of studies and information about using this drug can affect preventive strategies. Furthermore, a long questionnaire that we used for collecting information in this study may have affected the participants and could be one weak point of this study.

The main feature of the data of this study was count. The first count model that researchers referred to was Poisson regression (PR) model. However, because of the related problem of this model (overdispersion), sometimes this model is not suitable and its results cannot be trusted. According to the source of overdispersion, using remedial techniques is needed. Sometimes, the main reason for this overdispersion is due to many zeroes in the data set. In this case, there are different methods for excessive zero data, such as zero-inflated models, which were used in this paper and in Hurdle models (31). Many studies have applied these models recently (32-37). There are 2 different approaches to estimate the parameters in statistical modeling there are two different approaches: classical and Bayesian approach. Classical methods include different methods, such as maximum likelihood estimation (MLE), which has some drawbacks. Uncertainty in the parameter estimates is one of the problems. Also, these methods are usually based on the asymptotic normality assumption, which is often not tenable for small sample (sparse) data (38, 39). However, Bayesian approach corrects these drawbacks of classical estimation; specially, the estimation according to Bayesian method are reliable regardless of the smallness of the data (40). On the other hand, recent advances in simulation method, such as Markov chain Monte Carlo calculation, now allow for fitting models of this type, which have a complex structure (41). In recent years, several authors have proposed Bayesian alternatives to fitting zero-inflated models (42, 43). In the application field, this method also has been used by different authors in different research areas.

Jang et al. have considered zero-inflated Poisson and zero-inflated negative binomial regression models to analyze discrete count road safety data. They estimated regression coefficients associated with several safety countermeasures using Bayesian method. Their empirical results showed that zero-inflated models perform better than the classical approach (44).

In a study investigating spatiotemporal HIV/TB child mortality, Bayesian analysis of zero inflation was used by Musenge et al. They used a computationally efficient alternative to Markov chain Monte Carlo (MCMC) called ‘Integrated Nested Laplace Approximation’ to estimate the parameters (45).

In another HIV mortality study, Musal and Aktekin introduced parameter estimation issues of Bayesian zero-inflated Poisson models and discussed MCMC method implications (46).

Swallow et al. have considered ecological data corresponding to an annual average of 26 weekly maximum counts of birds. They used Bayesian hierarchical Tweedie regression model that can directly accommodate the excess number of zeroes common to this type of data. Implementation of the model is conducted in a Markov chain Monte Carlo (MCMC) framework (47).

Conclusion

Thus, it seems that using these Bayesian models have increased in recent years. In this study, Bayesian ZIP model was performed to examine related demographic factors with count of using crystal meth. In conclusion, using Bayesian approach because of its benefits, compared to classical approach, is a good alternative for statistical modeling in different fields of studies. Another important issue
of this study was that to analyze count data with zeroes, needs arise for conducting the analysis adopting special zero-inflated modeling approach instead of the usual count modeling. Therefore, it is essential to apply the appropriate count model (zero-inflated in this case) to have meaningful results.

**Ethical considerations**

Ethical issues (plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely adhered to by the authors of this study.

**Acknowledgement**

This paper was part of a research for a PhD degree at School of Paramedical Sciences, Shahid Beheshti University of Medical Sciences (BUMS) and was financially supported by the mentioned university.

**Conflict of Interests**

The authors declare that they have no competing interests.

**References**

Bayesian zero inflated model in using crystal meth in Tehran population

47. Swallow B, Buckland ST, King R, Toms MP. Bayesian hierarchical modelling of continuous non-negative longitudinal data with a spike at zero: An application to a study of birds visiting gardens in winter. Biom J. 2015.

http://mjiri.iums.ac.ir