Study Protocol

Application of Spatio-temporal Model to Estimate Burden of Diseases, Injuries and Risk Factors in Iran 1990 – 2013

Mahboubeh Parsaeian PhD Candidate[1,2], Farshad Farzadfar MD MPH DSc[2,3], Hojjat Zeraati PhD[1], Mahmood Mahmoudi PhD[1], Gelareh Rahimighazikalayeh MSc[4], Iman Navidi BSc[2], Sharareh R. Niakan Kalhori PhD[2,5], Kazem Mohammad PhD•[1], Majid Jafari Khaledi PhD•[6]

Abstract

Background: Identifying the burden of disease and its inequality between geographical regions is an important issue to study health priorities. Estimating burden of diseases using statistical models is inevitable especially in the context of rare data availability. To this purpose, the spatio-temporal model can provide a statistically sound approach for explaining the response variable observed over a region and various times. However, there are some methodological challenges in analysis of these complex data. Our primary objective is to provide some remedies to overcome these challenges.

Method: Data from nationally representative surveys and systematic reviews have been gathered across contiguous areal units over a period of more than 20 years (1990 – 2013). Generally, observations of areal units are spatially and temporally correlated in such a way that observations closer in space and time tend to be more correlated than observations farther away. It is critical to determine the correlation structure in space-time process which has been observed over a set of irregular regions. Moreover, these data sets are subject to high percentage of missing, including misaligned areal units, areas with small sample size, and may have nonlinear trends over space and time. Furthermore, the Gaussian assumption might be overly restrictive to represent the data. In this setting, the traditional statistical techniques are not appropriate and more flexible and comprehensive methodology is required. Particularly, we focus on approaches that allow extending spatio-temporal models proposed previously in the literature.

Since statistical models include both continuous and categorical outcomes, we assume a latent variable framework for describing the underlying structure in mixed outcomes and use a conditionally autoregressive (CAR) prior for the random effects. In addition, we will employ misalignment modeling to combine incompatible areal units between data sources and/or over the years to obtain a unified clear picture of population health status over this period. In order to take parameter uncertainties into account, we pursue a Bayesian sampling-based inference. Hence, a hierarchical Bayes approach is constructed to model the data. The hierarchical structure enables us to "borrow information" from neighboring areal units to improve estimates for areas with missing values and small number of observations. For their general applicability and ease of implementation, the MCMC methods are the most adapted tool to perform Bayesian inference.

Conclusion: This study aims to combine different available data sources and produce precise and reliable evidences for Iranian burden of diseases and risk factors and their disparities among geographical regions over time. Providing appropriate statistical methods and models for analyzing the data is undoubtedly crucial to circumvent the problems and obtain satisfactory estimates of model parameters and reach accurate assessment.

Keywords: Burden of diseases, Iran, misalignment, spatio-temporal models, study profile

Cite this article as: Parsaeian M, Farzadfar F, Zeraati H, Mahmoudi M, Rahimighazikalayeh G, Navidi I, et al. Application of spatio-temporal model to estimate burden of disease, injuries and risk factors in Iran 1990 – 2013. *Arch Iran Med.* 2014; **17(1):** 28 – 32.

Introduction

Evaluating the burden of diseases and related risk factors is essential to identify key health priorities. Some studies have investigated the global and national pattern of diseases worldwide.[1–7] However, there is little information about Iran since there is only one national burden of diseases study, which was conducted in 2003 by the Ministry of Health and Medical Education.[8] This study was conducted at the national level and included six preselected provinces and the results showed a significant variability among these provinces.

Providing information about the variations between provinces can help policy makers to allocate resources more efficiently and specify inequalities. Thus, it is essential to model and measure subnational variations to reveal epidemiological similarities and disparities among provinces.

Usually observations from areal units (provinces or districts) close to each other are more similar and exhibit spatial correlation. Although including important area level covariates in the model can reduce high proportion of this correlation, identifying and measuring this covariates is not always possible. Thus, unknown or unmeasured spatially correlated cofounders cause spatially correlated residuals; on the other hand, traditional regressions ignore such a correlation and lead to biased and inefficient estimates.[9] In

Authors' affiliations: [1]Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran, [2]Non-Communicable Disease Research Center, Endocrinology and Metabolism Population Science Institute, Tehran University of Medical Sciences, Tehran, Iran, [3]Endocrinology and Metabolism Research Center, Endocrinology and Metabolism Researsh Institute, Tehran University of Medical Sciences, Tehran, Iran, [4]Department of Epidemiology and Biostatistics, School of Public Health, University of South Carolina, Columbia, South Carolina , USA, [5]Social Determinants of Health Research Centre, School of Health, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran, [6]Department of Statistics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran, Iran.
•Corresponding author and reprints: Kazem Mohammad PhD, Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran. Majid Jafari Khaledi PhD, Department of Statistics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran, Iran.
E-mail: mohamadk@ tums.ac.ir, jafari-m@modares.ac.ir.
Accepted for publication: 3 December 2013

addition to spatial correlation, the data have been collected over time. Consequently, observations closer in time tend to be more correlated than observations farther away. The mentioned advantages of spatio-temporal correlation between areal units make it possible to impute the estimates that are rare but needed in the burden of diseases studies.

The present article aims to discuss the statistical challenges in analysis of disease burden and propose novel models in spatio-temporal framework and appropriate approaches to handle these challenges.

Method

Data Sources

Data from nationally representative surveys and systematic reviews have been collected over contiguous areal units (provinces, districts and census tracts) through a period of more than 20 years (1990 – 2013).[10] These nationally representative surveys include Non-Communicable Disease Surveillance Surveys (NCDSS), National Health Surveys (NHS), Demographic Health Survey (DHS), Census information and Household Expenditure Surveys, Hospital Data Survey,[11] Outpatient Data and other national health surveys. Table 1 presents an overview of available data sources, their areal unit, and measurement time. As shown, the areal units and the time spans are different between surveys and it is not an easy task to combine these data to produce a unified result.

Statistical challenges and possible solutions

In the following section, we will discuss practical challenges that we will encounter in the data analysis step and we will briefly explain how statistical methods can be extended to overcome these limitations.

Modeling Spatio-temporal correlation structure

Spatial correlation of response variables across studied region violates assumption of independence in ordinary regressions. One way to overcome these drawbacks is the extension of ordinary regression to include spatial random effects. In this way, model can easily capture any over-dispersion or spatial autocorrelation that remains after accounting for available covariates. A well-known framework used in this context is Bayesian hierarchical model with conditional autoregressive (CAR) prior for these random effects. In this setting, it is essential to define an adjacency matrix for areal units that characterize the neighboring structure and importance weight of each neighbor. Typical approaches for defining these weights are based on the distance between two area centroids and sharing a common boundary.[9] Accordingly, model estimate at a given location is associated with nearby estimates, so it reflects some kind of autocorrelation.[12] This modeling framework is flexible and can be implemented via Markov chain Monte Carlo (MCMC) simulations. For their general applicability and ease of implementation, the MCMC methods are the most adapted tool to perform Bayesian inference. However, in our study these methods face several limitations. When model involves a large number of highly correlated latent variables, the conventional MCMC algorithm may converge slowly or even fail to converge. To deal with this obstacle and alleviate the convergence problem through MCMC methods, we will use and assess some advanced techniques which have been suggested in recent years (e. g. Inverse Bayes Formula[13–16] and Integrated Nested Laplace Approximation[17] as alternatives to MCMC). Using these techniques, the statistical inference and model fitting can be carried out more efficiently.

Another important feature of this study is to obtain trends of outcomes during this period. Since observations of areal units over time are correlated, special strategy for modeling trends is required. Note that observations are simultaneously correlated across space and time, so statistical methodology for space or time separately is not reasonable. Hence, we will use spatio-temporal models which can simultaneously consider spatially correlated residual and time trends. Note this model can consider both linear and non-linear trends that may happen in real phenomena. Therefore, it can overcome the limitations of the traditional regression models.

Misalignment

Linking several data sets from different sources such as national health surveys, censuses or systematic reviews in which data were recorded based on different areal units require special methodology.[18] For example, consider we want to model fasting plasma glucose that has been measured by NCDSS and NHS. These measurements are recorded based on different areal units. In fact, while NCDSS areal unit is recorded at districts level, NHS is recorded at province level. A problem arises here is how to combine data from two different levels and especially how to define adjacency structure of areal units that are not compatible for these two data sets. Two remedies are possible: We can omit the district level data and use only province level data for both these data sources. This ad-hoc method does not provide satisfactory estimates of the model parameters. In fact, the two sources of information are valuable in different ways. A more sophisticated solution is to design a modeling framework that can consider different level of data aggregation. This strategy, which has been popular in recent years, is based on utilizing misalignment techniques to combine data from different sources.

In addition to spatial misalignment, the number of cities and provinces has changed during this period. There were a total of 24 provinces in 1990 which has increased to 31 provinces in 2013. It should be noted, the province boundaries has changed after division. So, administrative divisions produce incompatible areal units over the course of study. This problem become more serious at districts level since the number of districts is nearly doubled in this period.

Thus, incompatible areal units between data sources and/or over the years necessitate modeling spatio-temporally misaligned data.[19,20]

Dealing with data scarcity

Data scarcity is a major shortcoming of this study and it should be considered thoroughly. The target responses are not measured for all districts and/or in all years. For example, NCDSS has been measured only in 6 years and 49 districts have been studied in 2011. Indeed, available data points are far fewer than what is necessary to produce reliable subnational inference during this period.

The most common remedy for this problem is to consider unmeasured responses as missing data. Latent class model provides a flexible framework for dealing with missing data. In this framework missing data along with model parameters will be considered unknown and will be estimated via the proposed Bayesian

Table1. Available data sources and their time span

Data Source	Areal unit	Time span
NCDSS	District, Province	2005, 2006, 2007, 2008, 2009, 2011
NHS	Province	1991, 2000
DHS	Province	2000
Census	Census tract	1986,1996, 2006, 2011
Household Expenditure	District, Province	1984 – 2011
Hospital Data Survey	Hospital	1996 – 2013
Outpatient Data	Medical Science University	1998 – 2012
Systematic Review	District, Province	1990 – 2013

hierarchical model. Accordingly, information from nearby areal units will be used to impute missing values.

Small area estimation

In addition to the presence of missing data, the number of observations in some areal units is too small to produce reliable estimates. For example, in areas with small number of death, the death rate estimates are unstable and it is better to use other neighboring data to improve this estimate. To overcome this problem, we will use Bayesian spatial models such as Besage, York and Mollie (BYM) to produce more precise estimates of both mean and variance.[9] In this framework, model-based estimates are a combination of direct estimates and neighboring estimates. Indeed, these estimates borrow information from nearby neighbors to improve the naive unstable direct estimates. It has been shown that small area models improve estimates even for relatively large sample sizes.[21]

Different distribution of spatial data

This study includes different types of spatio-temporal data including continuous (normal and possibly non-normal) and discrete (Poisson, Binomial and ordinal and categorical). Although there are several methods of spatial model for normal and Poisson distributions, statistical models for non-normal, ordinal and categorical data are less developed and are not routinely programmed in many statistical packages.

More specifically, the continuous responses could be highly non-Gaussian and may show features like heavier tails or skewness. Furthermore, ordered categorical spatial data can be regarded as the result of clipping an underlying latent continuous random process.[21]The customary approach for such models relies on the assumption that the latent random variables are normally distributed. Although, the use of Gaussian latent process for the analysis of categorical and continuous outcomes facilitates inferences and prediction, this assumption might be overly restrictive to represent the data. In fact, the violation of those assumptions results in degradation of model performance.

Therefore, new statistical model and methodology are required for this kind of responses. All required programs are written through R and WinBUGS statistical packages, which are two commonly used programming languages especially in the Bayesian context.

Model selection criteria

There are a wide variety of models that can be used to estimate trends of diseases and risk factors. These models may be different in terms of covariates, statistical methodology or even the form of response variable. So model selection criteria are required to choose the best model among all possible ones. Usually we want to generalize the results of the models, and estimate how accurately a predictive model performs in practice. So the best model is selected based on better out of sample performance. Hence the data is divided into two parts, performing the analysis on one part and validating the results on the other part.

Three types of criteria were used to select the best model. They include absolute or relative error in prediction, error in estimating accurate trends and coverage of uncertainty intervals. One model may perform better in one criterion and worse in the other. So it is sensible to combine the results of single models that are best based on these different criteria. Indeed the final ensemble model is a weighted average of these single models. There are many methods including Bayesian Model Averaging which can be used to develop weights for these single models.[23]

Missing data imputation for older age group

Another challenge is how to deal with important groups that had not been measured in none of the surveys. For example, metabolic risk factors had not been measured for age groups above 65 years in NCDSS surveys. Hence, we don't have any information about these groups and need to extrapolate. The statistical strategy for this issue is somewhat different from modeling missing data in which interpolating rather than extrapolating is required.

To overcome this problem for example in metabolic risk factors, we will use 47 surveys from other countries that have measured metabolic risk factors for all age ranges. Detailed description of this methodology is explained elsewhere.[24] However, in Brief:

I. First, we select individuals aged 30 – 60 years in both NCDSS and 47 international surveys and estimate the linear slope of metabolic risk factors versus age separately by provinces and surveys, respectively. Then we divide both set of the slopes into three groups based on the tertile values of international slopes.

II. For each tertile, we will estimate overall and province specific intercept via a fixed effects model that contains the effects of 5-year age groups and provinces on NCDSS data set for men and women separately.

III. For each tertile, we fit the same fixed effect model in 47 international surveys. Note that this time data includes older age groups. After removing survey specific effect, beta coefficient of older age groups can be obtained.

IV. For older age groups, the province specific mean and variance estimates of each metabolic risk factor can be obtained through adding the overall intercept, province specific intercept of NCDSS and older age group coefficient from relevant tertile of international surveys.

Conclusion

In the present study, we assess a wide variety of diseases, injuries and risk factors along with powerful statistical methods to characterize the health status of Iranian people and related changes

during a period of over 20 years. To our knowledge, there is only one Iranian national burden of disease study that was conducted in 2003 by the Ministry of Health and Medical Education.[8] This study was conducted at the national level and included six preselected provinces. Hence, the results just render an overall picture of the society and more detailed studies are required to provide comprehensive evidences for the whole country.

Some studies have investigated the national and subnational pattern of disease worldwide.[25,26] In a study conducted in Mexico, burden of disease was investigated at the national and subnational levels, and subnational disparities between states were identified to set health priorities.[26]

Recently conducted burden of disease studies, such as Global Burden of Disease 2010, pay more attention to quantifying the variations among areal units and they explicitly incorporates these variations in the modeling framework. In the present study, we have developed a Bayesian hierarchical model with spatial and temporal correlation. This modeling strategy is comparable to Global Burden of Disease 2010 and provides a flexible framework that enables us to "borrow information" from neighboring areal units and nearby time periods to obtain more accurate and efficient estimates. Using this additional information for imputing missing values, this model outperforms other alternative imputation methods such as Amelia II package.[27] Also additional prior information can be incorporated in the model and it is also possible to estimate via small number of observations.

Since in the Burden of disease study we primarily use existing data sets rather than gathering new information, the problem of different areal units is prevalent. Also administrative divisions result in increased number of provinces and districts, creating many incompatible areal units especially at district level. This problem is less pronounced in Global Burden of Disease since the areal units such as countries, almost remains constant. We have used spatio-temporal misalignment model to combine the results from disparate data sources and to produce a unified clear picture at province or district level as well as national level.

Two commonly used models that can be employed in this setting are multilevel and spatial models. Both models are well functioning to quantify this heterogeneity, however extra information about distances between observations will be used in the latter. In this study we mainly discussed spatial models, but both spatial and multilevel models[28] will be used in the analysis of burden of disease to compare their predictive ability and reduce any model dependency in the final results. In addition Ensemble model which is the weighted average of the best spatial and multilevel models will be used as a final model. It has been shown that ensemble models have outperform any single model and are less sensitive to model specification bias and produce smaller prediction error. These models incorporate both uncertainty for any single model and uncertainty due to different model specifications.[23]

As mentioned, data scarcity is one of the major shortcomings of this study. Available data points are far fewer than what is necessary to produce reliable inference. Therefore, an attempt is made to use all available information and a sophisticated modeling strategy to compensate for the lack of sufficient and accurate databases. In the presence of more reliable data for each province-year, less complicated models are required. Nevertheless, as expected modeling just can solve part of this problem and estimates with wide uncertainty intervals is obtained for some areas with small number of observations. In this study we have tried to provide an accurate picture of health status in Iran based on available data sources and the results can inform policy makers about the current and future health status of society and reveal possible gaps across different geographic regions.

Author's Contribution

General designing of paper: *Farshad Farzadfar, Majid Jafari Khaledi, Kazem Mohammd, Hojjat Zeraati, Mahmood Mahmoudi, Mahboubeh Parsaein*

Designing of models: *Majid Jafari Khaledi, Farshad Farzadfar, Kazem Mohammad, Hojjat Zeraati, Mahmood Mahmoudi, Mahboubeh Parsaeian, Sharareh R. NiakanKalhori*

Writing primary draft: *Mahboubeh Parsaeian, Farshad Farzadfar, Kazem Mohammad, Majid Khaledi Jafari, Gelareh Rahimighazikalayeh*

Manuscript revision: *Farshad Farzadfar, Majid Khaledi Jafari, Kazem Mohammad, Mahboubeh Parsaeian, Hojjat Zeraati, Mahmood Mahmoudi, Iman Navidi*

Rule of funding Sources

Our research was funded by Iran's Ministry of Health and Medi cal Education of Islamic Republic of Iran and Setad-e-Ejraie Farmane Imam. The spatio-temporal analysis of these date sets is the PhD thesis of Mahboubeh Parsaeian in Biostatistic.

Acknowledgments

The authors thank our colleagues in Non Communicable Disease Research Center and appreciate the collaboration of organizations and researchers that provided necessary data for this project. We would also like to express thanks to Dr.Masoud Moradi for his precise editing of the text and Dr. Sirin Djalalinia and Ms Rosa Hagh Shenas for thier efforts on managing coordinative and administrative processes.

References

1. Wang H, Dwyer-Lindgren L, Lofgren KT, Rajaratnam JK, Marcus JR, Levin-Rector A, et al. Age-specific and sex-specific mortality in 187 countries, 1970 – 2010: a systematic analysis for the Global Burden of Disease Study. *The Lancet*. 2010; **380(9859):** 2071 – 2094.
2. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study. *The Lancet*. 2010; **380(9859):** 2129 – 2143.
3. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study. *The Lancet*. 2010; **380(9859):** 2224 – 2260.
4. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study. *The Lancet*. 2010; **380(9859):** 2095 – 2128.
5. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990 – 2010: a systematic analysis for the Global Burden of Disease Study. *The Lancet*. 2010; **380(9859):** 2197 – 2223.
6. Salomon JA, Wang H, Freeman MK, Vos T, Flaxman AD, Lopez AD, et al. Healthy life expectancy for 187 countries, 1990-2010: a systematic analysis for the Global Burden Disease Study. *The Lancet*. 2010; **380(9859):** 2144 – 2162.
7. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et

al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990 – 2010: a systematic analysis for the Global Burden of Disease Study. *The Lancet*. 2010; **380(9859):** 2163 – 2196.
8. Naghavi M, Abolhassani F, Pourmalek F, Lakeh M, Jafari N, Vaseghi S, et al. The burden of disease and injury in Iran 2003. *Popul Health Metr*. 2009; **7:** 9.
9. Earnest A, Morgan G, Mengersen K, Ryan L, Summerhayes R, Beard J. Evaluating the effect of neighbourhood weight matrices on smoothing properties of Conditional Autoregressive (CAR) models. *Int J Health Geogr*. 2007; **6(54):** 1 – 12.
10. Farzadfar F, Delavari A, Malekzadeh R, Mesdaghinia A, Jamshidi HR, Sayyari A, et al. NASBOD 2013: Design, definitions, and metrics. *Arch Iran Med*. 2014; **17(1):** 7 – 15.
11. Niakan Kalhori SR, Tayefi B, Noori A, Mearaji M, Rahimzadeh S, Zandian. E, et al. Inpatient data, inevitable need for policy making at national and sub-national levels: A lesson learned from NASBOD. *Arch Iran Med*. 2014; **17(1):** 16 – 21.
12. Montero JM, Mínguez R, Durbán M. SAR models with nonparametric spatial trends. A P-spline approach. *Estadística Española*. 2012; **54(177):** 89 – 111.
13. Yu J-w, Tian G-l. Efficient algorithms for generating truncated multivariate normal distributions. *Acta Math Appl Sin Engl Ser*. 2011; **27(4):** 601 – 612.
14. Tan M, Tian G-L, Ng K. A Noninterative Sampling Method for Computing Posteriors in the Structure of EM-type Algorithms. *Statistica Sinica*. 2003;**13:** 625 – 639.
15. Tan M, Tian G-L, Ng KW. Hierarchical models for repeated binary data using the IBF sampler. *Comput Stat Data Anal*. 2006; **50(5):** 1272 – 1286.
16. Wang WL, Fan TH. Bayesian analysis of multivariate t linear mixed models using a combination of IBF and Gibbs samplers. *Journal of Multivariate Analysis*. 2012; **105(1):** 300 – 310.
17. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2009; **71(2):** 319 – 392.
18. Young L, Gotway C. Linking spatial data from different sources: the effects of change of support. *Stochastic Environmental Research and Risk Assessment*. 2007; **21(5):** 589 – 600.
19. Zhu L, Carlin BP. Comparing hierarchical models for spatio-temporally misaligned data using the deviance information criterion. *Stat Med*. 2000; **19(17–18):** 2265 – 2278.
20. Hund L, Chen JT, Krieger N, Coull BA. A Geostatistical Approach to Large-Scale Disease Mapping with Temporal Misalignment. *Biometrics*. 2012; **68:** 849 – 858.
21. You Y, Zhou QM. Hierarchical Bayes small area estimation under a spatial model with application to health survey data. *Survey Methodology*. 2011; **37(1):** 25 – 37.
22. Higgs MD, Hoeting JA. A clipped latent variable model for spatially correlated ordered categorical data. *Computational Statistics & Data Analysis*. 2010; **54(8):** 1999 – 2011.
23. Foreman KJ, Lozano R, Lopez AD, Murray CJ. Modeling causes of death: an integrated approach using CODEm. *Popul Health Metr*. 2012; **10(1):** 1 – 10.
24. Farzadfar F, Danaei G, Namdaritabar H, Rajaratnam JK, Marcus JR, Khosravi A, et al. National and subnational mortality effects of metabolic risk factors and smoking in Iran: a comparative risk assessment. *Popul Health Metr*. 2011; **9(1):** 55.
25. Begg SJ, Vos T, Barker B, Stanley L, Lopez AD. Burden of disease and injury in Australia in the new millennium: measuring health loss from diseases, injuries and risk factors. *The Medical Journal of Astralia*. 2008; **188(1):** 36 – 40.
26. Stevens G, Dias RH, Thomas KJA, Rivera JA, Carvalho N, Barquera S, et al. Characterizing the Epidemiological Transition in Mexico: National and Subnational Burden of Diseases, Injuries, and Risk Factors. *PLoS Med*. 2008; **5(6):** e.125.
27. Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. *Journal of Statistical Software*. 2011; **45(7):** 1 – 47.
28. Kasaeian A, Eshraghian MR, Rahimi Foroushani A, Niakan Kalhori SR, Mohammad K, Farzadfar F. Bayesian autoregressive multilevel modeling of burden of diseases, injuries and risk factors in Iran 1990 – 2013. *Arch Iran Med*. 2014; **17(1):** 22 – 27.