Work-related accidents among the Iranian population: a time series analysis, 2000–2011

Masoud Karimlou¹, Masoud Salehi², Mehdi Imani³, Agha-Fatemeh Hosseini⁴, Afsaneh Dehnad⁵, Nasim Vahabi⁶, Mahmood Bakhtiyari⁷

¹Department of Biostatistics and Computer, University of Social Welfare & Rehabilitation Sciences, Tehran, Iran, ²Health Management and Economics Research Center, Department of Biostatistics, Faculty of Public Health, Iran University of Medical Sciences, Tehran, Iran, ³Department of Statistics and Mathematics, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran, ⁴Department of Biostatistics, Faculty of Public Health, Iran University of Medical Sciences, Tehran, Iran, ⁵Department of Foreign Languages, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran, ⁶Department of Biostatistics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran, ⁷School of Public Health & Health Research Institute, Tehran University of Medical Sciences, Tehran, Iran

Background: Work-related accidents result in human suffering and economic losses and are considered as a major health problem worldwide, especially in the economically developing world.

Objectives: To introduce seasonal autoregressive moving average (ARIMA) models for time series analysis of work-related accident data for workers insured by the Iranian Social Security Organization (ISSO) between 2000 and 2011.

Methods: In this retrospective study, all insured people experiencing at least one work-related accident during a 10-year period were included in the analyses. We used Box–Jenkins modeling to develop a time series model of the total number of accidents.

Results: There was an average of 1476 accidents per month (1476.05 ± 458.77 , mean $\pm$ SD). The final ARIMA (p, d, q) (P, D, Q)_s model for fitting to data was: ARIMA(1, 1, 1) $\times$ (0, 1, 1)₁₂ consisting of the first ordering of the autoregressive, moving average and seasonal moving average parameters with 20.942 mean absolute percentage error (MAPE).

Conclusions: The final model showed that time series analysis of ARIMA models was useful for forecasting the number of work-related accidents in Iran. In addition, the forecasted number of work-related accidents for 2011 explained the stability of occurrence of these accidents in recent years, indicating a need for preventive occupational health and safety policies such as safety inspection.

Keywords: Work-related accidents, Time series analysis, ARIMA modeling, Iran

Introduction

Work-related accidents are common and a worldwide public health concern. According to the International Labor Organization, 5000 workers die daily from work-related accidents. Approximately 270 million annual occupational accidents result in 335 000 fatalities and 1250 billion lost dollars each year.¹ In Iran, the rate of work-related accidents among employees insured by Social Security Organization of Iran (ISSO) was 3.3 per 1000 between 2001 and 2005. In 2007, the rates of work-related accidents were between two to eight per 100 000 in European countries.² This difference between Iran, a developing

country, and Europe may be in part due to limitations in the quality of information, including access to official data sets.^{3–5} However, the prevalence of work-related injuries is likely a result of different working conditions, safety precautions, and safety policies.

ISSO has defined work-related accidents as accidents among insured employees at work or as a consequence of job insecurity. The purpose of this retrospective study was to investigate trends and seasonal factors in work-related accidents over a 10-year period in Iran. We use a time series forecast of work-related accidents to make inferences about the number of work-related accidents for the year 2011. Findings from this study can be used to design preventative policies and guide the implementation of work-related interventions in Iran.

Correspondence to: M. Salehi, Faculty of Public Health, Iran University of Medical Sciences, No 6, Rashid Yasemi st., Vali-e-Asr Ave., Tehran, Iran. Email: salehi74@yhoo.com

Materials and Methods

Study population

The study population included insured Iranians who experienced work-related accidents in which the beneficiaries filed a compensation claim with ISSO between March 2000 and March 2011. Data were obtained from the Office of Statistics and Economic–Social Calculations of ISSO.

Statistical methods

Box–Jenkins modeling was used to find the best seasonal autoregressive moving average (ARIMA) model for the aggregated monthly number of accidents. Statistical forecasting using the fitted model was used to predict future work-related accidents.

Box–Jenkins modeling

We used Box–Jenkins modeling to develop a time series model of the total number of work-related accidents. The following processes were considered for data series generation: moving average, autoregressive, and a combination of both. The series with characteristics of both processes are referred to as ARMA Process and the formula model derived from this process is:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \quad (1)$$

where the average and autoregressive processes are of the second order.

Wangdi proposed a seasonal ARIMA model in which parameters are determined by (p,d,q) and $(P,D,Q)s$. In this model, p and P stand for the autoregressive and seasonal autoregressive parameters, d and D represent the non-seasonal differences and seasonal differencing parameters, q and Q represent the moving average parameters and seasonal moving average parameters, and s stand for the length of the seasonal period.⁶

Box–Jenkins modeling includes three stages: identification, estimation, and forecasting. Since the Box–Jenkins model falls within the range of stationary panel time series data, in the identification step, the data are examined for the series stationarity. Generalized Dickey–Fuller and seasonal Dickey–Fuller tests are used to examine the series. There are five factors that result in non-stationary states: outliers, random walk, drift, trend, and changes in variance. Each requires its own transformation to become ineffective in the series. A seasonal non-stationary state is another factor to examine before the aforementioned factors.^{7,8}

After examining the stationarity of the series, the parameters of the model were estimated to select a suitable process to fit time series, using auto-correlation function (ACF) and partial auto-correlation function (PACF). The last stage concerns time series forecast for a certain period of time at the end of data interval.^{6,9,10} For this stage, the expected

values for the number of monthly accidents occurring in 2011–2012 were predicted.

Forecasting

Statistical forecasting predicts seasonal or cyclical trends and patterns. An optimal decision on time (t) depends on predicted values of the random variable (Y_{t+h}). The lead time h is the number of forecasted time points. A forecast is the selection process of the closest values to the real value at time $t+h$. Decreasing the difference between predicted points and real values is the aim of forecasting. This difference e_t is the forecast error at time t for which several forecast error measures are recommended to identify and determine the best fitted model. We used mean absolute percentage error (MAPE) and mean square forecast error as the common forecast accuracy criteria, explained as the average of differences between the true and forecasted values.^{11–15} Akaike information criteria (AIC) were used to model diagnostics in which the lowest AIC value was considered a good model.^{12–15}

We used SAS software (version 9.2) for statistical modeling of time series including ARIMA models and related tests. The Ethical Committee of University of Social Welfare and Rehabilitation Sciences in Tehran, Iran approved this study.

Results

The monthly time series (time space between each two consequent months considered as a lag) of work-related accidents over the 10-year period (Fig. 1) shows a trend with seasonal changes, suggesting strong autocorrelation.

At the first step of the identification stage, monthly mean and standard deviation of the series were 1476.05 and 458.77, respectively. The autocorrelation and ACF graph with 95% confidence intervals of autocorrelation coefficients for different lags supported the application of the Box–Jenkins model (Fig. 2).

Results indicated that the rank of moving average in the ARIMA model with one-stage differencing was 2, while the rank of autoregressive was 1. The reason for one-stage regular differencing and seasonal differencing was specified through generalized Dickey–Fuller test. Table 1 shows the results of the augmented Dickey–Fuller and seasonal augmented Dickey–Fuller tests before and after series differentiation.

The work-related accident data time series was non-stationary due to the presence of random walk and drift (Table 1). According to this table, after one stage of regular and seasonal differencing, the series reached the regular and seasonal stability, indicating that values of d and D parameters were equal to one.

The ACF and PACF diagrams showed that the ACF in Fig. 2 had noticeable damping, while PACF did not, confirming the convergence of the series. In

Figure 1 Time series values for total number of work-related accidents in Iran, 2000–2011.

the next stage, parameters of the model and their equality to zero were tested (Table 2).

Findings (Table 2) confirmed that the first ordering of the autoregressive, moving average, and seasonal moving average parameters were all statistically significant ($P < 0.05$). Therefore, the final ARIMA (p, d, q) (P, D, Q)_s model for fitting to data was

$$\text{ARIMA}(1,1,1) \times (0,1,1)_{12}$$

Table 3 shows the goodness of fit results for the final model. The AIC and MAPE in the final model by were equal to 1601.257 and 20.942, respectively.

The final model was:

$$\begin{aligned} & (1-B)(1-B^{12})_{\text{Total}} \\ &= -6.24392 + (100.84007B)(1-0.74182B^{12}) \\ &/(1-0.57835B)a_t \end{aligned}$$

After determining the optimal model, the series forecast for the monthly number of work-related accidents in 2010–2011 was calculated (Fig. 3).

As illustrated in Fig. 3, the expected incidence of accidents in the first month of spring (March) is higher than in the final month of the winter (February). The lowest expected incidence of accidents is in February 2002.

Discussion

Economic loss and human suffering are the social consequences of morbidity and mortality caused by work-related accidents. Statistical modeling can be used to help explain the phenomena of the accidents.¹⁶ We found that these accidents were mainly independent and random in nature, reflecting an uncertainty and fluctuation in the occurrence requiring statistical modeling due to their inherent nature.¹⁷

Work-related accidents are the result of multiple factors including individual behaviors, environmental factors, and workplace conditions.^{18,19} Low educational attainment, lack of prior work experience, boldness, and recklessness contribute to work-related accidents for younger populations.²⁰ Education and prevention are suggested to pave the ground for effective ways to ultimately reduce the number of work-related accidents and their consequences in Iran.²¹

The results of this type of analysis can be used to guide governmental decisions related to the programmatic interventions and policies to improve work life.^{22,23} While multiple statistical modeling methods exist for analyzing work-related accident data, we selected time series modeling and forecasting as the most appropriate method to show the trend work-related accidents over a 10-year period.

Figure 2 Trend and correlation analysis for time series of work-related accident data in Iran.

We found that $ARIMA(1,1,1) \times (0,1,1)_{12}$ is a reliable model for work-related accident data in Iran. The strength of the ARIMA technique for modeling and forecasting have been reported in previous studies.^{6,17,24} Our findings indicated that work-related accidents occur more frequently at different times of the year.

Our findings indicated that regardless of preventive measures, accidents would have followed the same trend in the subsequent years. Therefore, it is necessary to improve preventative approaches to control and reduce the rate of occupational accidents. Although our model showed stability in the later years, the forecasted values for subsequent years were

still high (2010–2011). Further studies are needed to identify the hazardous factors of work-related accident in order to effectively reduce the number accidents. Although some studies focused on risk factors exist, there is a need for more research in this area.²⁵ In addition, activities such as safety training

Table 2 Conditional least squares estimation

Parameter	Estimate	Standard error	t value	Approx. Pr> t	Lag
MU	-6.24392	2.51855	-2.48	0.0146	0
MA1,1	0.84007	0.10209	8.23	<0.0001	1
MA2,1	0.74182	0.06834	10.86	<0.0001	12
AR1,1	0.57835	0.15217	3.80	0.0002	1

Table 1 Augmented Dickey–Fuller unit root tests and seasonal augmented Dickey–Fuller unit root tests, before and after differencing

	Type	Lags	Rho	Pr<Rho	Tau	Pr<Tau	F	Pr>F
Augmented Dickey–Fuller unit root tests before differencing	Zero mean	0	-0.6905	0.5289	-0.43	0.5261		
	Single mean	0	-19.2850	0.0112	-3.74	0.0045	7.24	0.0010
	Trend	0	-32.8398	0.0028	-4.54	0.0019	10.74	0.0010
Seasonal augmented Dickey–Fuller unit root tests before differencing	Zero mean	0	3.1305	0.7763	1.13	0.8910		
	Single mean	0	-70.2594	0.0013	-12.65	<0.0001		
Augmented Dickey–Fuller unit root tests after differencing	Zero mean	0	-141.454	0.0001	-13.35	<0.0001		
	Single mean	0	-141.535	0.0001	-13.31	<0.0001	88.54	0.0010
	Trend	0	-141.767	0.0001	-13.29	<0.0001	88.37	0.0010
Seasonal augmented Dickey–Fuller unit root tests after differencing	Zero mean	0	-158.714	0.0001	-17.51	<0.0001		
	Single mean	0	-158.738	0.0001	-17.43	<0.0001

Figure 3 Forecasts for time series of total number of work-related accidents in Iran, 2011–2012.

and orientation for new employees and education for managers and supervisors would likely help prevent occupational accidents. The most important limitation of this study is that in Iran, there are different types of insurance; therefore, we accessed to people who registered officially in the ISSO data set. Also, due to the time-consuming process of claim, some accident victims choose not to receive compensation, which results in an underestimation of accident rates in the ISSO data set.

Results from this study revealed an autocorrelation and trend between lags. Considering these autocorrelations and using Box–Jenkins model, ARIMA $(1,1,1) \times (0,1,1)_{12}$ was fitted well to the data. We also found an ascending trend of work-related accidents, indicating a need for preventive occupational health and safety policies such as safety inspection and workplace health promotion.

Table 3 Goodness-of-fit indices of the ARIMA $(1,1,1) \times (0,1,1)_{12}$

Constant estimate	−2.63274
Variance estimate	39 538.42
Stand error estimate	198.8427
MAPE	20.942
AIC	1601.257
SBC	1612.373
Number of residuals	119

Disclaimer Statements

Contributors MS suggested the initial concept and provided administrative and technical support for conducting the study. MK and MI contributed to the study design and did critical revision of the manuscript for important intellectual content. AFH, AD, and MB contributed to the study design by preparing the final report and drafting the manuscript. MS and MI performed the statistical analysis and interpretation of data. NV provided the technical and material support for the study and supervised the data collection. All authors read and approved the final manuscript.

Funding We thank all the research deputy of University of Social Welfare & Rehabilitation Sciences who funded and supported this study and Mostafa Arji, member of Social Security Organization of Iran for preparation of the study data.

Conflicts of interest None.

Ethics approval None.

References

- 1 Benavides FG. Occupational injury, a public health priority. *Rev Esp Salud Publica*. 2006;80:553–65.
- 2 European Commission, Eurostat, Directorate-General for Employment, Social Affairs and Inclusion. Health and safety

- at work in Europe (1999–2007): a statistical portrait. Luxembourg: Publications Office of the European Union; 2010.
- 3 Smith G. Public health approaches to occupational injury prevention: do they work? *Injury Prev.* 2001;7:3–10.
- 4 Mosayebzadeh S. Occupational accidents and how to prevent. Tehran: Tamin Ejtemaei Publications; 2002.
- 5 Fabiano B. Port safety and the container revolution: a statistical study on human factor and occupational accidents over the long period. *Saf Sci.* 2010;48:980–90.
- 6 Wangdi K, Singhasivanon P, Silawan T. Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: a case study in endemic districts of Bhutan. *Malar J.* 2010;9:251–9.
- 7 SAS Institute Inc. SAS/STAT 9.2 user's guide. Cary, NC: SAS Institute Inc.; 2009.
- 8 Yaffee R, McGee M. Introduction to time series analysis and forecasting with applications of SAS and SPSS. New York: Academic Press Inc.; 2000.
- 9 Brockwell PJ, Davis RA. Time series: theory and methods. New York: Springer; 2009.
- 10 Fan J, Yao Q. Nonlinear time series: nonparametric and parametric methods. New York: Springer; 2005.
- 11 Maharaj EA. Comparison of non-stationary time series in the frequency domain. *Comput Stat Data Anal.* 2002;40:131–41.
- 12 Chatfield C. The analysis of time series: theory and practice. New York: Chapman and Hall/CRC; 1975.
- 13 Chatfield C. The analysis of time series: an introduction. New York: Chapman and Hall/CRC; 2003.
- 14 Freivalds A, Johnson AB. Time-series analysis of industrial accident data. *J Occup Accid.* 1990;13:179–93.
- 15 Montgomery DC, Jennings CL, Kulahci M. Introduction to time series analysis and forecasting. New York: John Wiley & Sons Inc; 2008.
- 16 Hovden J, Albrechtsen E, Herrera IA. Is there a need for new theories, models and approaches to occupational accident prevention? *Saf Sci.* 2010;48:950–6.
- 17 Sari M, Selcuk AS, Karpuz C, Duzgun HSB. Stochastic modeling of accident risks associated with an underground coal mine in Turkey. *Saf Sci.* 2009;47:78–87.
- 18 Ale BJM, Bellamy LJ. Accidents in the construction industry in the Netherlands: an analysis of accident reports using story-builder. *Saf Sci.* 2008;93:1523–33.
- 19 Attwood D, Khan F, Veitch B. Occupational accident models — where have we been and where are we going? *J Loss Prev Process Ind.* 2006;19:664–82.
- 20 Ciarapica FE, Giacchetta G. Classification and prediction of occupational injury risk using soft computing techniques: an Italian study. *Saf Sci.* 2009;47:36–49.
- 21 Zahedi SH. Industrial relationships, work relationships rules (based on work-related accidents). Tehran: Markaz-e Nashr-e Daneshgahi; 2004.
- 22 Jorgensen K. Audits, inspections and investigations: reporting and compiling accident statistics. In: ILO encyclopedia of occupational health and safety. Geneva: ILO; 1998. p. 1–31. Available from: <http://www.ilocis.org/en/contilo8.html>
- 23 Rantanen J, Kauppinen T, Toikkanen J. Accidents at work. In: Work and health country profiles: country profiles and national surveillance indicators in occupational health and safety. Helsinki: Finnish Institute of Occupational Health; 2001. p. 42–8.
- 24 Cadenas E, Jaramillo OA, Riverab W. Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method. *J Renew Energy.* 2010;35:925–30.
- 25 Bakhtiyari M, Delpisheh A, Riahi SM, Latifi A, Zayeri F, Salehi M, et al. Epidemiology of occupational accidents among Iranian insured workers. *Saf Sci.* 2012;50(7):1480–4.