Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia
Vol. 10, 2022, page 42 - 50

Logistic Regression Analysis on Factors Affecting Mortality of Trauma Patients

Ayu Nur Athirah Mohd Jamri, Norhaiza Ahmad*
Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia
*Corresponding author: norhaiza@utm.my

Abstract
The purpose of this study is to investigate the factors affecting mortality of trauma patients admitted to hospital emergency departments using logistic regression analysis. In this study, a logistic regression analysis is carried out on trauma patient data to identify the factors affecting mortality of trauma patient by gender. The dataset consists of many variables such as age, gender, Systolic Blood Pressure (SBP) , Injury Severity Score (ISS) , Glasgow Comma Scale (GCS), mechanism of injury, location of injury, Revised Trauma Scale (RTS) and mortality . Here, mortality is considered as the response variable for the analysis. Mortality has two outcome which is alive and dead. Univariate logistic regression analysis are performed between all the independent variable against mortality regardless of gender. From the univariate analysis, all the independent variables which are age, gender, SBP, ISS, GCS, mechanism of injury, location of injury and RTS are statistically significant (p- value <0.05) with the mortality. From the results, multivariate logistic regression is performed between the independent variable which are age, SBP, ISS, GCS, mechanism of injury, location of injury and RTS against mortality for male and female respectively. By doing the backward regression, all the tested variables are statistically significant (p- value <0.05) with mortality for both male and female respectively. From the findings, the odds ratio (OR) are reported as follows; age (OR 1.0386, 95% CI (0.03,0.04)), RTS (OR 0.6659, 95% CI (-0.52,-0.30)), ISS (OR 0.2026, 95% CI (-1.79,-1.40)) and GCS (OR 0.7900, 95% CI (-0.28,-0.19)) statistically significant (p- value <0.05) with mortality for male while age (OR 1.0458, 95% CI (0.04,0.05)), RTS (OR 0.6843, 95% CI (-0.56,-0.21)), ISS (OR 0.2038, 95% CI (-1.90,-1.27)), GCS (OR 0.7800, 95% CI (-0.31,-0.18)), SBP (OR 0.9933, 95% CI (-0.01,-0.0002)) and location of injury (OR 0.6768, 95% CI (-0.68,-0.10)) statistically significant (p- value <0.05) with mortality for female. This study shows that female patients with older age, lower RTS, ISS, GCS, SBP and also has only one location of injury are more likely to be dead while male patients with older age, lower RTS, ISS, GCS are more likely to be dead.

Keywords Trauma patient; Logistic Regression Analysis; Mortality; Gender.

1. Introduction

An emergency department in a hospital is responsible for the provision of medical and surgical care to patients arriving at the hospital in need of immediate care.

Trauma patient at hospital emergency department is the interest of this study as we want to know the variable that affect the mortality of trauma patient by gender. The trauma patient dataset consists of many variables such as age, somatic comorbidities, ISS, GSC and intubation status. In this study, we are interested in how variables such as age, somatic comorbidities, ISS, GSC and intubation status affect the mortality of trauma patient by gender. Our interest, mortality is considered as the response variable for the analysis. Variable mortality has two outcomes which are alive and dead. Thus, we can treated the mortality as categorical data with alive is 0 and dead is 1. Therefore, logistic regression model will be use to analyze the data.

In logistic regression model, we need regressor and predictor. Mortality acts as the regressor while the other variable as the predictor. Logistic regression model is used in R programming to interpret the results. Some model evaluation such as the AIC, null deviance and residual deviance will be used to evaluate the model. The aim of this study is to identify the factors affecting mortality of trauma patients by gender.

2. Literature Review

Observational study at five emergency medical services (EMS) agencies and 11 hospitals representing all 9-1-1 transfers within a country [1]. To evaluate the association of trauma center transport and functional outcomes, multivariate ordinal logistic regression analyses was conducted on multiple imputed data for all patients with TBI and patients with traumatic intracranial hemorrhage. The severity of injury in a patient with multiple traumas with a single number is difficult at best; therefore, multiple alternative scoring systems have been proposed, each with its own problems and limitations[2]. Traumatic brain injury (TBI) is a common healthcare problem related to disability. An easy-to-use trauma scoring system informs physicians about the severity of trauma and helps to decide the course of management. The purpose of the study is to use the combination of both physiological and anatomical assessment tools that predict the outcome and develop a new modified prognostic scoring system in TBIs[3].

Experimental research suggests that females have a higher survival rate after trauma. This study sought to determine the role of sex on mortality among trauma patients in China[4]. Multiple logistic regression was used to analyze the association between sex and post-injury complications and mortality. In this case study, the example regressor used are age, admission, ISS, Glascow Coma Scale (GCS), presence of hypotension and intubation status. The variable that acts as regressor are anticoagulation therapy, psychiatric comorbidity, penetrating injury, nocturnal admission, ISS, GCS, shock on arrival and mortality[5]. It is well known that males and females have different life expectations, comorbidities, and distinct differences in acute and chronic disease manifestation[6]. It has been shown that females have a more efficient immune system with subsequent improved clinical outcomes after severe infections and circulatory shock.The variable that the study used are age, ISS, prehospital shock index, comorbidities and expected mortality. Currently,two models of artificial neural network(ANN) and logistic regression(LR) are known as models that extensively used in sciences [7]. ANN has better performance than Logistic regression in predicting the terminal outcomes of traumatic patients in both the AUC and accuracy rate. Using an ANN to predict the final implications of trauma patients can provide more accurate clinical decisions.

Traumatic brain injury (TBI) is among the leading causes of mortality and long-term disability. Prognosis assessment is a primary factor of clinical decision-making by emergency physicians[8]. This study aimed to investigate the prognostic factors of TBI in the patients admitted to a typical emergency department. The analysis is conducted using simple binary logistic regression and multivariate logistic regression.

Previous studies have identified that the reverse shock index multiplied by the Glasgow Coma Scale score (rSIG) is a good predictor of mortality in trauma patients[9]. The outcome measures were the predictive ability of rSIG for coagulopathy, in-hospital mortality, and 24-h mortality. The comparison was done between the prognostic performance of rSIG with the shock index, age shock index, and quick Sequential Organ Failure Assessment. Mortality of trauma patient based on gender are analyze using multiple logistic regression[10,11,12,13].

3. Methodology

3.1. Logistic Regression Analysis Model

Logistic regression is the appropriate regression analysis to conduct when the dependent variable is dichotomous (binary). The logistic regression is a predictive analysis. Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables. It is a suitable model for this case study as gender is a binary variable (female and male).

The simple logistic regression model is given by

$$y = \log(\frac{p}{1-p}) = \beta_0 + \beta_1 x \tag{1}$$

Where *p* is a binomial proportion and *x* is the explanatory variable.The parameter of the logistic regression are β_0 and β_1. This also called simple binary logistic regression.

The form of multiple logistic regression follows the simple binary logistic regression as in equation

$$y = \log(\frac{p}{1-p}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \tag{2}$$

This multiple logistic regression is also called logit odds ratio. To get the odds ratio, take the exponential for both side

$$\frac{p}{1-p} = \exp^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n} \tag{3}$$

Thus, we will used the odds ratio to intepret the model of logistic regression model. Also, to get the true value of parameter estimate β is obtained as:

$$\beta \pm z_{\alpha/2}(\text{square of error } \hat{\beta}) \tag{4}$$

With 95% as the confidence interval.

The assumptions of the Logistic Regression Model are:

-The outcome is binary.

-There is a linear relationship between logit outcome and each independent variable.

-No extreme values or outliers in the continuous predictors.

-No high intercorrelations (multicollinearity) among the predictors.

3.2 Wald Test

Wald Test is used to test the statistical significance of logistic regression parameters, β logistic regression model. The formulation of the Wald test statistic relies on the parameter estimate and its variance estimate.

$$\text{Wald} = \frac{\beta}{\text{Standard Error of Estimate (s.e.}(\beta))} \tag{5}$$

The hypotheses for wald test are

$H_0: \beta = 0$

$H_1: \beta \neq 0$

This wald test is also called z-test. The significant test can be interpret using p-value. If p-value less than 0.05 with 95% confidence interval, the null hypothesis will be rejected. Thus, it shows that the variable are statistically significance.

3.3 Univariate and Multivariate Logistic Regression Analysis

Univariate analysis is the simplest form of analyzing data. It considers only one variable in the analysis. Each exploratory variable in the database is modeled in logistic regression once at the time to the response variable, y. This analysis looks at the relationship between each predictor to y without considering other variables. Therefore, the equation follow the simple binary logistic regression model as in Equation (1) . The univariate logistic regression is the preprocess for multivariate logistic regression. Unlike univariate analysis, the multivariate analysis combined all predictors to study relation with the responding variable, y. It is defined as the statistical study of data where multiple measurements are made on each experimental unit and where the relationships among multivariate measurements and their structure are important. It is to study the correlation between all predictors with y. The equation follows the multiple logistic regression model in Equation (2). After the multivariate logistic regression done, backward regression analysis is carried out. Based the results, only the significant variables will be considered in the backward logistic regression. The variables that are not significant will be eliminate from the backward regression analysis. Backward regression is done until there are only the significant variables left in the model. The model are evaluated using AIC , null deviance and residual deviance.

4. Results and discussion

4.1. Trauma patients data

For this research, trauma patients data is used. The variables contain in the data are age, gender, location of injury, mechanism of injury, Revised Trauma Score (RTS), Injury Severity Score (ISS), Systolic Blood Pressure (SBP), Glasgow Coma Scale (GCS), and mortality.

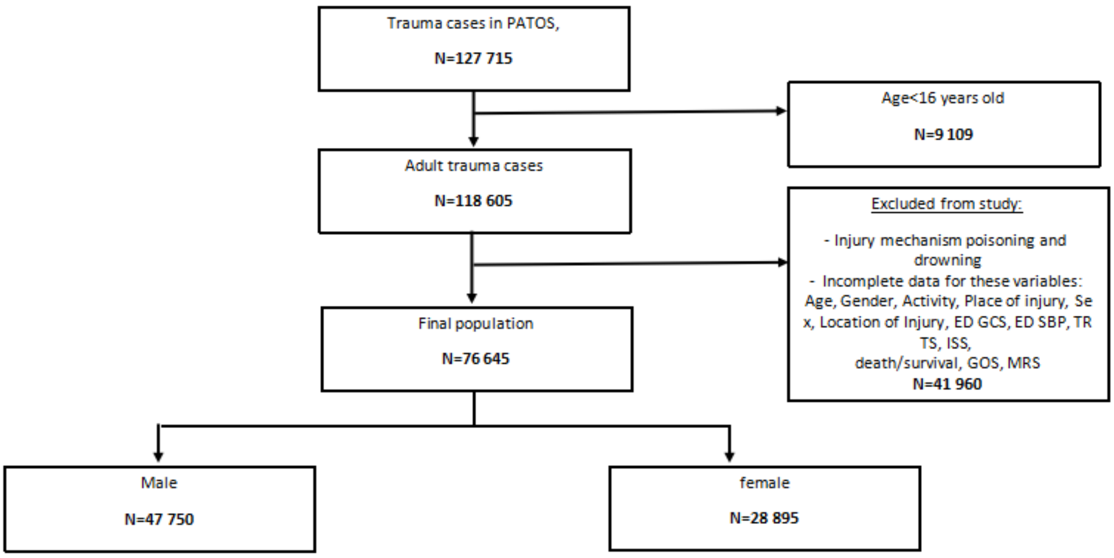

Figure 4.1 : Flowchart for selection of data

4.2. Univariate Logistic Regression

Univariate logistic regression is used to analyze the dependents variable significant with the independent variable mortality. The analysis is done between mortality and age, mortality and gender, mortality and mechanism of injury, mortality with TRTS, mortality and ISS, mortality with GCS, mortality with SBP, and mortality with location of injury. This is to find the relationship between the variables with mortality. This is the preprocess step before doing the multivariate logistic regression.

Table 4.1 : Univariate Analysis for the data Regardless of Gender

	Coefficient (CI)	P-value	Odds ratio
Age	0.0279 (0.0249,0.0309)	<0.05	1.0283
Gender	-0.4162 (-0.55, -0.28)	<0.05	0.6596
Mechanism of injury			
Traffic injury	0.7874 (0.55,1.05)	<0.05	2.197
Fall/Slip down	0.6027 (0.35,0.87)	<0.05	1.827
Assault			
Penetrating Injury	-0.8969	<0.05	
Others	0.2556 (-.151,-0.35)		0.4078
RTS	-1.1489 (-.1.189, -1.1.09)	<0.05	0.3170
ISS	-2.9823 (-3.11, -2.84)		
Minor (< 15)		<0.05	0.0507
Major (>15)			
GCS	-0.4625 (-0.47,-0.45)	<0.05	0.6297
ED Systolic BP (mmHg)	-0.0221 (-0.024, -0.20)	<0.05	0.978
Location	-0.4162 (-1.068 , -0.824)		
One		<0.05	0.3885
Many			

From the results, all the variables have p-value less than 0.05 and the null hypothesis can be rejected. Thus, the tested variables are statistically significant against mortality. Hence, the variables affect the mortality of the trauma patients.

The interest of study is the mortality of trauma patients by gender, univariate logistic regression analysis is done again one by one the variable against the mortality by the gender of trauma patient. The variable age, mechanism of injury, TRTS, ISS, GCS, SBP, and location of injury against mortality but this time the gender is separated. Thus, there are two analysis, for male and female. The results is in Table 4.2.

Table 4.2 : Univariate Logistic Regression Analysis based on Gender

UNIVARIATE	*Female*	*P-value*	*Male*	*P-value*
Count(%)	*28,895 (37.9)*		*47,750 (62)*	
Age,mean	*55*	*<0.05*	*47*	*<0.05*
Age,median(IQR)	*57*		*46*	
Mechanism of injury *Traffic injury* *Fall/Slip down* *Assault* *Penetrating Injury* *Others*	 *10,793 (37.4)* *12,754 (44.1)* *2,655 (9.2)* *1,297 (4.5)* *1,396 (4.8)*	 *<0.05* *<0.05* 	 *21,740 (45.5)* *14,816 (31.0)* *6,214 (13.0)* *3,032 (6.3)* *1,948 (4.1)*	 *<0.05* *<0.05*
RTS *5 - 12* *≤ 4*	 *28,868 (99.9)* *27 (0.1)*	*<0.05*	 *47,702 (99.9)* *48 (0.1)*	*<0.05*
ISS *Minor (<8)* *Moderate (9-15)* *Severe (16-75)*	 *21,660 (75.0)* *5,230 (18.1)* *2,005 (6.9)*	*<0.05*	 *34,957 (73.2)* *7,415 (15.5)* *5,378 (11.3)*	*<0.05*
GCS *Mild (13-15)* *Moderate (12-9)* *Severe (3-8)*	 *28,088 (97.2)* *403 (1.4)* *404 (1.4)*	*<0.05*	 *45,384 (95.0)* *1,082 (2.3)* *1,284 (2.7)*	*<0.05*
ED Systolic BP *(mmHg)* *>89* *76-89* *50-75* *<50*	 *28,448 (98.5)* *222 (0.8)* *179 (0.6)* *46 (0.2)*	*<0.05*	 *46,600 (97.6)* *566 (1.2)* *505 (1.1)* *79 (0.7)*	*<0.05*
Mortality *Alive* *Dead*	 *28,588 (98.9)* *307 (1.1)*		 *46,985 (98.4)* *765 (1.6)*	
Location *One* *Many*	 *20,373 (65.1)* *8,522(34.8)*	*<0.05*	 *29,581* *18,169*	*<0.05*

From the results, all the variables are statistically significant against mortality since the null hypothesis rejected (p-value<0.05). Hence,the variable affects the mortality of the trauma patients. Based on this, next analysis can continue to do the multiple logistic regression using all the variables.

4.3. Multivariate Logistic Regression Analysis

Multivariate Logistic Regression analysis used between mortality and other independent variables which are age, mechanism of injury, TRTS, ISS, GCS, SBP, and location of injury. The analysis is separated between gender, male and female to see the difference of significance in both male and female.

4.3.1. Multiple Logistic Regression for Female

A multiple logistic regression analysis is done for female. From the analysis, a backward logistic regression is done only considering the variables that significant at the previous analysis. From the previous analysis, age, Revised Trauma Score (RTS), Injury Severity Score (ISS), Glascow Coma Scale (GCS), Systolic Blood Pressure (SBP) and location of injury. Thus only this variables is considered in the backward logistic regression analysis. The final result for the model is as in Table 4.3.

Table 4.3 : Backward Regression of Multivariate Logistic Regression of Mortality of Trauma Patient Based on Gender (Female)

	Coefficient (CI)	P-value	Odds Ratio
Age,mean	-0.0447 (0.04,0.05)	<0.05	1.0458
RTS	-0.3794 (-0.56,-0.21)	<0.05	0.6843
ISS Minor (<15) Major (>15)	-1.5906 (-1.90,-1.27)	<0.05	 0.2038
GCS	-0.2485 (-0.31,-0.18)	<0.05	0.7800
ED Systolic BP (mmHg)	-0.0067 (-0.01,-0.002)	<0.05	0.9933
Location One Many	-0.3905 (-0.68,-0.1)	<0.05	 0.6768

From the analysis, age, RTS, ISS, GCS, SBP and location of injury has the p-value less than 0.05. So, the null hypothesis is rejected. This indicated that the variables significance with mortality. The odd ratios can be interpreted that older patients have 1.0458 more tendencies than older patients to be dead. Patients with lower RTS have 0.6843 more tendencies to be dead than patients with higher RTS. Patients with minor ISS have 0.2038 more tendencies to be dead than patients with major ISS. Patients with lower GCS have 0.7800 more tendencies to be dead than patients with higher GCS. Patients with lower SBP have 0.9933 more tendencies to be dead than patients with higher SBP. Patients that have one location of injury have 0.6768 more tendencies to be dead rather than patients that have more location of injury.

4.3.2. Multiple Logistic Regression for Male

A multiple logistic regression analysis is done for male. From the analysis, a backward logistic regression is done only considering the variables that significant at the previous analysis. From the previous analysis, age, Revised Trauma Score (RTS), Injury Severity Score (ISS) and Glascow Coma Scale (GCS). Thus only this variables is considered in the backward logistic regression analysis. The final result for the model is as in Table 4.4.

Table 4.4 : Backward Multivariate Logistic Regression of Mortality of Trauma Patient Based on Gender (Male)

	Coefficient	P-value	OR (CI)
Age,mean	0.0378 (0.03,0.04)	<0.05	1.0386
RTS 5 - 12 ≤ 4	-0.4162 (-0.52,-0.30)	<0.05	0.6659
ISS Minor (<8) Moderate (9-15) Severe (16-75)	-1.6023 (-1.79,-1.40)	<0.05	0.2026
GCS Mild (13-15) Moderate (12-9) Severe (3-8)	-0.2324 (-0.28,-0.19)	<0.05	0.7900

From the analysis, age, RTS, ISS, GCS, and SBP has the p-value less than 0.05. The null hypothesis is rejected. This indicate that the variables significance with mortality. The odd ratios can be interpreted that older patients have 1.0386 more tendencies than older patients to be dead. Patients with lower RTS have 0.6659 more tendencies to be dead than patients with higher RTS. Patients with minor ISS have 0.2026 more tendencies to be dead than patients with major ISS. Patients with lower GCS have 0.7900 more tendencies to be dead than patients with higher GCS.

The AIC, null deviance and residual deviance are used to evaluate the model and the backward regression model is the better model for both male and female.

Conclusion

Factors affecting mortality of trauma patient for male and female were analyze using logistic regression analysis. From the results, we can see that the factors affecting mortality of trauma patient for female is more than male. Female has 6 independent variables that affect the mortality which are age, Revised Trauma Scale (RTS), Injury Severity Score (ISS), Glascow Coma Scale (GCS), Systolic Blood Pressure (SBP) and location of injury while male only has 4 independent variables which are age, Revised Trauma Scale (RTS), Injury Severity Score (ISS) and Glascow Coma Scale (GCS). For female, patients with older age, lower Revised Trauma Scale (RTS), Injury Severity Score (ISS), Glasgow Coma Scale (GCS), Systolic Blood Pressure (SBP) and also has only one location of injury are more likely to be dead. For male, patients with older age, lower Revised Trauma Scale (RTS), Injury Severity Score (ISS) and Glasgow Coma Scale (GCS) are more likely to be dead.

Acknowledgement

The researcher would like to thank all people who have supported the research. Thank you to my supervisor, Dr Norhaiza Ahmad for guiding throughout this study. Thank you to my parents that always supporting me. Last but not least, thank you to everyone that involved in this study directly or indirectly.

References

[1] Nishijima, D. K., Gaona, S. D., Faul, M., Tancredi, D. J., Waechter, T., Maloney, R., Bair, T., Blitz, A., Elms, A. R., Farrales, R. D., Howard, C., Montoya, J., Garzon, H., & Holmes, J. F. (2020). The Association of Trauma Center Transport and long-term functional outcomes in head-injured older adults transported by Emergency Medical Services. Academic Emergency Medicine, 27(3), 207–216.

[2] Sethuraman, K. N., Marcolini, E. G., McCunn, M., Hansoti, B., Vaca, F. E., & Napolitano, L. M. (2014). Gender-specific issues in traumatic injury and resuscitation: Consensus-based recommendations for Future Research. Academic Emergency Medicine, 21(12), 1386–1394.

[3] Mahadewa, T. G. B., Golden, N., Saputra, A., & Ryalino, C. (2018). Modified revised trauma–marshall score as a proposed tool in predicting the outcome of moderate and severe traumatic brain injury. Open Access Emergency Medicine, Volume 10, 135–139.

[4] Zhu, Z., Shang, X., Qi, P., & Ma, S. (2017). Sex-based differences in outcomes after severe injury: An analysis of blunt trauma patients in China. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, 25(1).

[5] Larsson, E., Lindström, A.-C., Eriksson, M., & Oldner, A. (2019). Impact of gender on post-traumatic intensive care and outcomes. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, 27(1).

[6] Taha, P. H., & Sijbrandij, M. (2021). Gender differences in traumatic experiences, PTSD, and relevant symptoms among the Iraqi internally displaced persons. International Journal of Environmental Research and Public Health, 18(18), 9779.

[7] Hassanipour, S., Ghaem, H., Arab-Zozani, M., Seif, M., Fararouei, M., Abdzadeh, E., Sabetian, G., & Paydar, S. (2019). Comparison of artificial neural network and logistic regression models for prediction of outcomes in trauma patients: A systematic review and meta-analysis. Injury, 50(2),pp. 244–250.

[8] Bolvardi, E., Ahmadzadeh, M., Feiz Disfani, H., Rezaeitalab, F., Ravanshad, Y., & Foroughian, M. (2019). Predicting outcomes 3 months after traumatic brain injury in patients admitted to emergency department. Universa Medicina, 38(3), 147–155.

[9] Lee, Y. T., Bae, B. K., Cho, Y. M., Park, S. C., Jeon, C. H., Huh, U., Lee, D.-S., Ko, S.-H., Ryu, D.-M., & Wang, I. J. (2021). Reverse shock index multiplied by Glasgow Coma Scale as a predictor of massive transfusion in trauma. The American Journal of Emergency Medicine, 46, 404–409.

[10] Püsküllüoğlu, Sema & Açikalin, Ayça & Ay, Mehmet & Kozaci, Nalan & Avci, Akkan & Gulen, Muge & Icme, Ferhat & Satar, Salim. (2015). Analysis of Adult Trauma Patients Admitted to Emergency Department. Cukurova Medical Journal. 40. 563-573.

[11] Magnotti, L. J., Fischer, P. E., Zarzaur, B. L., Fabian, T. C., & Croce, M. A. (2008). Impact of gender on outcomes after blunt injury: A definitive analysis of more than 36,000 trauma patients. Journal of the American College of Surgeons, 206(5),984–991.

[12] Pape, M., Giannakópoulos, G. F., Zuidema, W. P., de Lange-Klerk, E. S., Toor, E. J., Edwards, M. J., Verhofstad, M. H., Tromp, T. N., van Lieshout, E. M., Bloemers, F. W., & Geeraedts, L. M. (2019). Is there an association between female gender and outcome in severe trauma? A multi-center analysis in the Netherlands. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, 27(1).

[13] Chalya, P.L., Gilyoma, J.M., Dass, R.M. et al. Trauma admissions to the Intensive care unit at a reference hospital in Northwestern Tanzania. Scand J Trauma Resusc Emerg Med 19, 61 (2011).