

## RESEARCH ARTICLE

# Survival analysis of heart failure patients: A case study

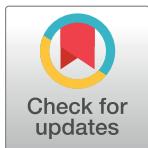
Tanvir Ahmad, Assia Munir, Sajjad Haider Bhatti\*, Muhammad Aftab, Muhammad Ali Raza

Department of Statistics, Government College University, Faisalabad, Pakistan

\* [sajjad.haider@gcuf.edu.pk](mailto:sajjad.haider@gcuf.edu.pk)

## Abstract

This study was focused on survival analysis of heart failure patients who were admitted to Institute of Cardiology and Allied hospital Faisalabad-Pakistan during April-December (2015). All the patients were aged 40 years or above, having left ventricular systolic dysfunction, belonging to NYHA class III and IV. Cox regression was used to model mortality considering age, ejection fraction, serum creatinine, serum sodium, anemia, platelets, creatinine phosphokinase, blood pressure, gender, diabetes and smoking status as potentially contributing for mortality. Kaplan Meier plot was used to study the general pattern of survival which showed high intensity of mortality in the initial days and then a gradual increase up to the end of study. Martingale residuals were used to assess functional form of variables. Results were validated computing calibration slope and discrimination ability of model via bootstrapping. For graphical prediction of survival probability, a nomogram was constructed. Age, renal dysfunction, blood pressure, ejection fraction and anemia were found as significant risk factors for mortality among heart failure patients.



## OPEN ACCESS

**Citation:** Ahmad T, Munir A, Bhatti SH, Aftab M, Raza MA (2017) Survival analysis of heart failure patients: A case study. PLoS ONE 12(7): e0181001. <https://doi.org/10.1371/journal.pone.0181001>

**Editor:** Chiara Lazzeri, Azienda Ospedaliero Universitaria Careggi, ITALY

**Received:** February 26, 2017

**Accepted:** June 23, 2017

**Published:** July 20, 2017

**Copyright:** © 2017 Ahmad et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Heart failure is the state in which muscles in the heart wall get fade and enlarge, limiting heart pumping of blood. The ventricles of heart can get inflexible and do not fill properly between beats. With the passage of time heart fails in fulfilling the proper demand of blood in body and as a consequence person starts feeling difficulty in breathing.

The main reason behind heart failure include coronary heart disease, diabetes, high blood pressure and other diseases like HIV, alcohol abuse or cocaine, thyroid disorders, excess of vitamin E in body, radiation or chemotherapy, etc. As stated by WHO [1] Cardiovascular Heart Disease (CHD) is now top reason causing 31% of deaths globally. Pakistan is also included in the list of countries where prevalence of CHD is increasing significantly. According to report by Al-Shifa hospital [2], 33% of Pakistani population above 45 has hypertension, 25% of patients over 45 years suffer diabetes mellitus, and CHD deaths in Pakistan has reached about 200,000 per year i.e. 410/100,000 of the population). All this results in increased prevalence of heart failure. Rate of heart failure patients in Pakistan is estimated to be 110 per million [1]. Rising stress of economic and social issues in the modern era, greasy food with little exercise results towards increased prevalence of heart failure in Pakistan.

Despite of this alarming situation, Pillai and Ganapathi [3] have reported that there are no reliable estimates of heart failure incidence and prevalence in this region while they are required due to poor and oily diet, lack of exercise and poor health care policies in this region. There are some projections based on prevalence data only from western countries.

In addition to relative scarcity of studies focusing on heart failure in this region, the present study has specific importance in the Pakistani context, as diet patterns in Pakistan are different with other the countries of South Asia like India, Bangladesh, Nepal and Sri Lanka.

The main objective of this study is to estimate death rates due to heart failure and to investigate its link with some major risk factors by choosing Faisalabad (third most populous city of Pakistan) as study area.

## Methods

### Detail of data

Current study is based on 299 patients of heart failure comprising of 105 women and 194 men. All the patients were more than 40 years old, having left ventricular systolic dysfunction and falling in NYHA class III and IV. Follow up time was 4–285 days with an average of 130 days. Disease was diagnosed by cardiac echo report or notes written by physician. Age, serum sodium, serum creatinine, gender, smoking, Blood Pressure (BP), Ejection Fraction (EF), anemia, platelets, Creatinine Phosphokinase (CPK) and diabetes were considered as potential variables explaining mortality caused by CHD. Age, serum sodium and CPK are continuous variables whereas EF, serum creatinine and platelets were taken as categorical variables. EF was divided into three levels (i.e.  $EF \leq 30$ ,  $30 < EF \leq 45$  and  $EF > 45$ ) and platelets was also divided into three level on the basis of quartiles. Serum creatinine greater than its normal level (1.5) is an indicator of renal dysfunction. Its effect on mortality was studied as creatinine  $> 1.5$  vs  $\leq 1.5$ . Anemia in patients was assessed by their haematocrit level. Following McClellan et al. [4] the patients with haematocrit less than 36 (minimum normal level of haematocrit) were taken as anemic. The information related to risk factors were taken from blood reports while smoking status and blood pressure were taken from physician's notes.

The study was approved by Institutional Review Board of Government College University, Faisalabad-Pakistan and the principles of Helsinki Declaration were followed. Informed consent was taken by the patients from whom the information on required characteristics were collected/accessed.

### Statistical techniques

Due to the presence of censored data, survival analysis was used to estimate the survival and mortality rates. Kaplan & Meier [5] product limit estimator was used to make comparisons between survival rates at different levels explanatory variables. Cox regression as presented by Collett [6] was used to develop a model that can link the hazard of death for an individual with one or more explanatory variables and test the significance of these variables.

Let hazard of death depends on  $p$  explanatory variables  $X_1, X_2, \dots, X_p$  then the hazard function for  $i^{th}$  individual can be defined by Cox model as

$$h_i(t) = h_0(t) e^{\beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi}}$$

For determining the functional form of any particular independent variable following Fitrianto & Jin [7] and Gillespie [8], plot of Martingale residuals versus different values (or levels) of a variable were used. The functional form of CPK was not linear therefore it was log transformed.

**Table 1.** Baseline characteristics for dead and censored patients.

Continuous Variables			Categorical Variables		
Variable	Dead (96)	Censored (203)	Variable	Dead (96)	Censored (203)
Creatinine	1.83	1.18	Male	62 (64%)	132 (65%)
Sodium	135.39	137.22	Smoking	30 (31%)	66 (32%)
CPK	670	540	Diabetes	40 (42%)	85 (42%)
Age	65.21	58.76	BP	40 (42%)	66 (32%)
Platelets	256381	266657	Anemia	54 (56%)	83 (40%)
EF	33.46	40.267			

<https://doi.org/10.1371/journal.pone.0181001.t001>

Following Pavlou et al. [9] model validation was assessed by bootstrapping [10–12] with 200 bootstrap replications. Internal validation of model was further checked by calculating calibration slope [13] for the average linear predictor. The calibration slope helped in estimating the ability of model for survival probability prediction. Discriminating ability of model was assessed by ROC curve [14]. A nomogram [15] was also built to predict the survival probabilities graphically.

## Results

Up to end of follow-up period, 96 (32%) patients died due to CHD. Table 1, presents different baseline characteristics of dead and censored patients at the end of follow up period.

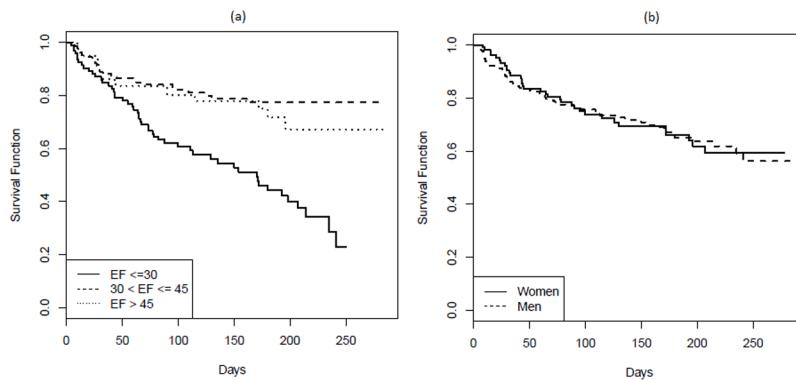
The results of Cox regression model are presented in Table 2. As Cox regression is semi parametric model, hence estimate of intercept (baseline hazard) was not provided by model fitting. According to Cox model, age was most significant variable.

Coefficient concerning age indicated that chances of death due to CHD increase with growing age. Hazard of death due to CHD increases by 4% for every additional year of age. EF was another significant factor, hazard rate among patients with EF  $\leq$ 30 was 67% and 59% higher as compared to the patients with 30 $<$ EF $\leq$ 45 and EF $\geq$ 45 respectively. In Fig 1(a), Kaplan Meier survival curve was constructed for each level of EF. It is obvious that survival for EF  $\leq$ 30 was lower than other two levels. Moreover, relatively small difference between the survival of patients with 30 $<$ EF $\leq$ 45 and EF $\geq$ 45 can be observed. Serum creatinine was significant with p-value = 0.0026. It means death hazard gets more than double for unit increase in Serum

**Table 2.** Significance of variables under Cox regression.

Variable	$\beta$ -coefficient	HR	Z-value	P-value
Age	0.0462	1.0473	4.81	0.0000
Gender	-0.1978	0.8205	-0.80	0.4239
EF $\leq$ 30 vs 30 $<$ EF $\leq$ 45	-1.1068	0.3306	-4.36	0.0000
EF $\leq$ 30 vs EF $\geq$ 45	-0.8894	0.4109	-3.18	0.0015
Smoking	0.1928	1.2127	0.77	0.4431
Diabetes	0.0992	1.1043	0.45	0.6553
Blood pressure	0.5043	1.6558	2.33	0.0195
Serum creatinine	0.8051	2.2368	3.01	0.0026
Serum sodium	-0.0658	0.9363	-2.79	0.0052
log(CPK)	0.1444	1.1554	1.41	0.1596
Anemia	0.5709	1.7699	2.59	0.0096
Platelets( $\leq Q_1$ )	0.4054	1.4999	1.62	0.1042
Platelets( $\geq Q_3$ )	0.3926	1.4800	1.46	0.1446

<https://doi.org/10.1371/journal.pone.0181001.t002>



**Fig 1.** (a) Kaplan Meier curves were used to study the survival at different levels of EF and in part (b), survival curves of male and female.

<https://doi.org/10.1371/journal.pone.0181001.g001>

creatinine. Serum sodium was significant with p-value = 0.0052 and its one unit (meq/L) increase decreases the hazard by 6%. Anemia was significant variable with p-value = 0.0096 and an anemic patient had 76% more chances of death as compared to non-anemic patient. According to results in Table 2, gender, smoking, diabetes, CPK and platelets were found to be non-significant.

Ejection fraction is an important measurement of how well one's heart is pumping and is used to help classify heart failure and guide treatment. The EF is also found to be significant correlate of deaths among heart failure patients from Cox regression for present sample. Keeping its importance in view, EF is further analyzed through baseline characteristics (Table 3) and Kaplan Meier curves (Fig 1(a)) which shows similar pattern as presented in Cox regression results.

In Fig 1(b), Kaplan Meier survival curves were constructed for both genders showed almost identical survival pattern.

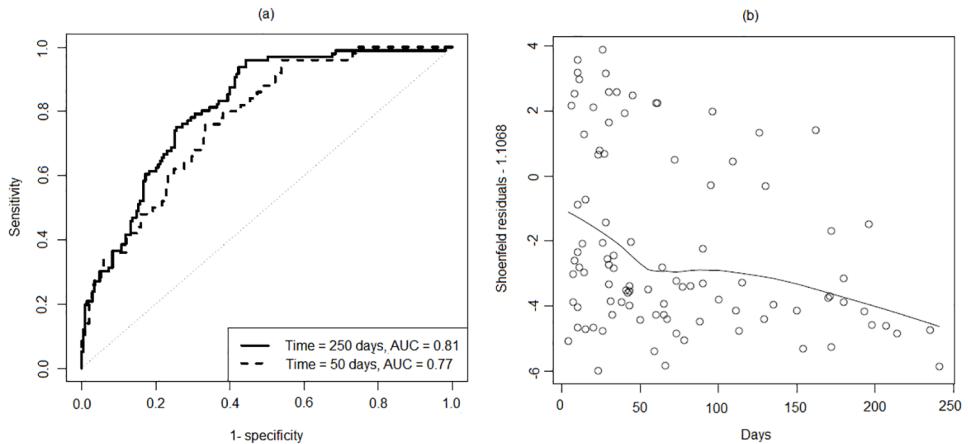
### Model validation

For model validation, calibration slope and ROC curve are developed from 200 bootstrapped samples. Calibration slope was equal to 0.96, which showed that model was not over fitted and predictions made by this model would neither be overestimated nor under estimated.

**Table 3.** Baseline characteristics with respect to EF levels.

	Variable	Dead (96)			Censored (203)		
		EF≤30 (51)	30<EF<45 (26)	EF≥45 (19)	EF≤30 (42)	30<EF<45 (100)	EF≥45 (61)
<b>Continuous Variables</b>	Creatinine	1.5447	2.225	2.084	1.345	1.122	1.179
	Sodium	134.76	135.73	136.53	136.6	137.23	137.59
	CPK	601	923	511	568	614.8	398.6
	Age	61.84	68.92	69.21	56.95	58.59	60.30
	Platelets	258577	260863	244353	256937	261280	282166
<b>Categorical Variables</b>	Male	35 (69%)	18 (69%)	09 (47%)	30 (71%)	63 (63%)	39 (64%)
	Smoking	17 (33%)	09 (35%)	04 (21%)	15 (36%)	32 (32%)	19 (31%)
	Diabetes	20 (39%)	20 (38%)	10 (53%)	18 (43%)	47 (47%)	20 (33%)
	BP	20 (39%)	12 (46%)	07 (37%)	14 (33%)	26 (26%)	26 (43%)
	Anemia	25 (49%)	11 (42%)	10 (53%)	18 (43%)	38 (38%)	27 (44%)

<https://doi.org/10.1371/journal.pone.0181001.t003>



**Fig 2. (a) ROC curves for discriminant ability w.r.t time (b) Effect of EF w.r.t time.**

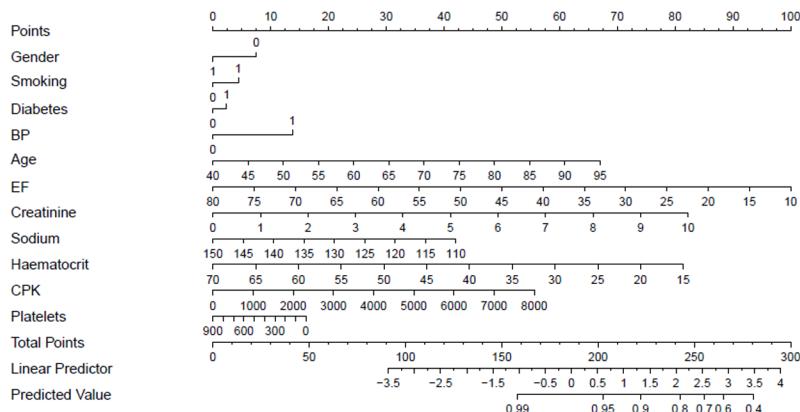
<https://doi.org/10.1371/journal.pone.0181001.g002>

Discrimination ability was checked by ROC curve in Fig 2(a). Area under the curve (AUC) was 0.81 at time of 250 days and 0.77 at time of 50 days thus it can be interpreted that the model was able to correctly recognize the event of death for 81% and 77% patients within 250 and 50 days respectively. It shows that discrimination ability of Cox model is higher at longer follow up time. The reason of this difference may be due to the violation of constant effect assumption of EF which is evident in Fig 2(b) which displays that effect of EF increases with the passage of time. As EF is highly significant for mortality (see Table 2), hence with passage of time model's discrimination ability increases.

### Nomogram for prediction

Calibration slope and discrimination ability suggested that Cox model is able to predict probability of survival and hazard sufficiently. On the basis of these results, nomogram is presented in Fig 3 to provide the graphical predictions of probability after assigning different points to each independent variable with respect to their significance. Sum of these points provides an estimate of probability of survival.

For example, an 80 year old non-smoker female diabetic patient with high blood pressure, EF = 40, haematocrit = 35, sodium = 140, creatinine = 5.2, platelets = 300 thousands and



**Fig 3. Nomogram for survival probability.**

<https://doi.org/10.1371/journal.pone.0181001.g003>

CPK = 3000 have points equal to  $50+0+8+3+14+56+50+12+40+10+20 = 263$  and probability of her survival is 0.60. The Cox model used for constructing this nomogram was fitted on original values of variables.

## Discussion

The statistical analysis identified age, EF, creatinine, sodium, anemia and BP as the significant variables affecting the likelihood of mortality among heart failure patients. Most of studies [16–17] supported the male gender as predictor of mortality among heart failure patients. However, like Román et al. [18] in this study odd ratio of men/women is not significant. With respect to significance and importance of variables the findings of the present study are more in lines with Rahimi et al. [19]. The results are found to be similar with other international studies like [20–23].

The findings that seem surprising are non-significance of smoking and diabetes. However, similar results concerning diabetes and smoking have been reported in other studies [24–25] as well. The reason behind may be smoking and diabetes are basically causes of heart problem at initial stages. We were only concerned with patients of NYHA class III and IV which are advanced stages of heart failure. Up to these stages, these factors (diabetes and smoking) may probably be controlled by medications and hence these factors do not have significant effect on deaths due to heart failure in class III and IV.

Performance of model was checked using calibration slope and ROC curve. Both concluded in adequacy of model for prediction. ROC curves were also used to discuss the goodness of model with respect to time. Nomogram was used to find the probability of survival by graphical method. It was observed that fall of survival probability was almost same for Kaplan Meier plot and nomogram.

## Conclusion

It can be concluded that growing age, renal dysfunction (having serum creatinine greater than its normal level 1.5), high BP (higher than normal range), higher level of anaemia and lower values of ejection fraction (EF) are the key factors contributing towards increased risk of mortality among heart failure patients. Increased level of serum sodium can reduce the odds of death. No significant differences were found due to smoking status, diabetes and gender of patients.

## Supporting information

S1 Data. DATA\_MINIMAL.

(CSV)

## Author Contributions

**Conceptualization:** Tanvir Ahmad, Sajjad Haider Bhatti.

**Data curation:** Assia Munir.

**Formal analysis:** Assia Munir, Muhammad Aftab.

**Investigation:** Muhammad Aftab.

**Methodology:** Tanvir Ahmad, Assia Munir, Sajjad Haider Bhatti, Muhammad Aftab, Muhammad Ali Raza.

**Project administration:** Tanvir Ahmad.

**Software:** Muhammad Ali Raza.

**Supervision:** Tanvir Ahmad.

**Validation:** Muhammad Aftab.

**Visualization:** Muhammad Ali Raza.

**Writing – original draft:** Tanvir Ahmad, Assia Munir, Sajjad Haider Bhatti.

**Writing – review & editing:** Tanvir Ahmad, Sajjad Haider Bhatti, Muhammad Ali Raza.

## References

1. WHO. Fact sheet on CVDs. Global Hearts. World Health Organization. 2016.
2. Al-Shifa IH. Cardiac Diseases in Pakistan [Internet]. 2016 [cited 15 Jun 2017]. <http://www.shifa.com.pk/chronic-disease-pakistan/>
3. Pillai Harikrishnan Sivadasan Ganapathi S. Heart failure in South Asia. *Curr Cardiol Rev.* 2013; 9: 102–111. <https://doi.org/10.2174/1573403X11309020003> PMID: 23597297
4. McClellan WM, Flanders WD, Langston RD, Jurkowitz C, Presley R. Anemia and renal insufficiency are independent risk factors for death among patients with congestive heart failure admitted to community hospitals: a population based study. *J Am Soc Nephrol.* 2002; 13: 1928–1936. PMID: 12089390
5. Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. *J Am Stat Assoc.* 1958; 53: 457–481.
6. Collett D. Modelling Survival Data in Medical Research. 2nd ed. Taylor & Francis; 2003.
7. Fitrianto A, Jiin RLT. Several Types of Residuals in Cox Regression Model: An Empirical Study. *Int J Math Anal.* 2013; 7: 2645–2654.
8. Gillespie B. Checking Assumptions in the Cox Proportional Hazards Regression Model. Midwest SAS Users Group (MWSUG). 2006.
9. Pavlou M, Ambler G, Seaman SR, Guttmann O, Elliott P, King M, et al. How to develop a more accurate risk prediction model when there are few events. *BMJ.* 2015; 351: h3868. <https://doi.org/10.1136/bmj.h3868> PMID: 26264962
10. Efron B. Bootstrap methods: another look at the jackknife. *Ann Stat.* 1979; 7: 1–26.
11. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. Chapman and Hall, New York; 1993.
12. Cribari-Neto F, Zarkos SG. Leverage-adjusted heteroskedastic bootstrap methods. *J Stat Comput Simul.* 2004; 74: 215–232. <https://doi.org/10.1080/0094965031000115411>
13. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: Seven steps for development and an ABCD for validation. *Eur Heart J.* 2014; 35: 1925–1931. <https://doi.org/10.1093/euroheartj/ehu207> PMID: 24898551
14. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics.* 2005; 61: 92–105. <https://doi.org/10.1111/j.0006-341X.2005.030814.x> PMID: 15737082
15. Yang D. Build Prognostic Nomograms for Risk Assessment Using SAS. Proceedings of SAS Global Forum 2013. 2013.
16. Barlera S, Tavazzi L, Franzosi MG, Marchioli R, Raimondi E, Masson S, et al. Predictors of Mortality in 6975 Patients With Chronic Heart Failure in the Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico-Heart Failure Trial Proposal for a Nomogram. *Circ Hear Fail.* 2013; 6: 31–39. <https://doi.org/10.1161/CIRCHEARTFAILURE.112.967828> PMID: 23152490
17. Tavazzi L, Senni M, Metra M, Gorini M, Cacciato G, Chinaglia A, et al. Multicenter prospective observational study on acute and chronic heart failure one-year follow-up results of in-hf (italian network on heart failure) outcome registry. *Circ Hear Fail.* 2013; 6: 473–481. <https://doi.org/10.1161/CIRCHEARTFAILURE.112.000161> PMID: 23476054
18. Román AV, Shamagian LG, Diéguez MAB, Veloso PR, González-Juanatey JR. Influence of sex on mortality in hospitalized patients with congestive heart failure and preserved or depressed systolic function. *Rev Española Cardiol.* 2005; 58: 1171–1180.
19. Rahimi K, Bennett D, Conrad N, Williams TM, Basu J, Dwight J, et al. Risk prediction in patients with heart failure: A systematic review and analysis. *JACC Hear Fail.* 2014; 2: 440–446. <https://doi.org/10.1016/j.jchf.2014.04.008> PMID: 25194291
20. Klein L, O'Connor CM, Leimberger JD, Gattis-Stough W, Piña IL, Felker GM, et al. Lower Serum Sodium Is Associated With Increased Short-Term Mortality in Hospitalized Patients With Worsening

- Heart Failure. Circulation. 2005; 111: 2454–2460. <https://doi.org/10.1161/01.CIR.0000165065.82609.3D> PMID: 15867182
- 21. Gustafsson F, Torp-Pederse C, Seibæk M, Burchardt H, Køber L. Effect of age on short and long-term mortality in patients admitted to hospital with congestive heart failure. Eur Heart J. 2004; 25: 1711–1717. <https://doi.org/10.1016/j.ehj.2004.07.007> PMID: 15451149
  - 22. Redondo-Bermejo B, Pascual-Figal DA, Hurtado-Martínez JA, Peñafiel-Verdú P, Saura-Espín D, Garrido-Bravo IP, et al. Influence of Gender on the Clinical Characteristics and Prognosis of Patients Hospitalized for Heart Failure. Rev Española Cardiol (English Ed. 2007; 60: 1135–1143.
  - 23. Varela-Roman A, Shamagian LG, Caballero EB, Ramos PM, Veloso PR, Gonzalez-Juanatey JR. Influence of diabetes on the survival of patients hospitalized with heart failure: A 12-year study. Eur J Heart Fail. 2005; 7: 859–864. <https://doi.org/10.1016/j.ejheart.2005.01.017> PMID: 15923138
  - 24. Otero-Raviña F, Grigorian-Shamagian L, Fransi-Galiana L, Názara-Otero C, Fernández-Villaverde JM, del Álamo-Alonso A, et al. Morbidity and mortality among heart failure patients in Galicia, N.W. Spain: the GALICAP study. Int J Cardiol. 2009; 136: 56–63. <https://doi.org/10.1016/j.ijcard.2008.04.025> PMID: 18617282
  - 25. Cheraghi M, Sadeghi M, Sarrafzadegan N, Pourmoghadas A, Ramezani MA. Prognostic Factors for Survival at 6-Month Follow-up of Hospitalized Patients with Decompensated Congestive Heart Failure. ARYA Atheroscler 2010. 2010; 6: 112–117.