

Analysis of Length of Stay (LOS) Data from the Medical Records of Tertiary Care Hospital in Saudi Arabia for Five Diagnosis Related Groups: Application of Cox Prediction Model

Sara AL-Gahtani¹, Mohamed M. Shoukri^{2*}

¹Department of Internal Medicine, King Fahd Medical City, Riyadh, Saudi Arabia

²Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada

Email: s_algahtani@live.com, *mshoukr@uwo.ca, *Shoukri.mohamed@gmail.com

How to cite this paper: AL-Gahtani, S. and Shoukri, M.M. (2021) Analysis of Length of Stay (LOS) Data from the Medical Records of Tertiary Care Hospital in Saudi Arabia for Five Diagnosis Related Groups: Application of Cox Prediction Model. *Open Journal of Statistics*, 11, 99-112.
<https://doi.org/10.4236/ojs.2021.111005>

Received: December 28, 2020

Accepted: January 24, 2021

Published: January 27, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Background: One of the main objectives of hospital managements is to control the length of stay (LOS). Successful control of LOS of inpatients will result in reduction in the cost of care, decrease in nosocomial infections, medication side effects, and better management of the limited number of available patients' beds. The length of stay (LOS) is an important indicator of the efficiency of hospital management by improving the quality of treatment, and increased hospital profit with more efficient bed management. The purpose of this study was to model the distribution of LOS as a function of patient's age, and the Diagnosis Related Groups (DRG), based on electronic medical records of a large tertiary care hospital. **Materials and Methods:** Information related to the research subjects were retrieved from a database of patients admitted to King Faisal Specialist Hospital and Research Center hospital in Riyadh, Saudi Arabia between January 2014 and December 2016. Subjects' confidential information was masked from the investigators. The data analyses were reported visually, descriptively, and analytically using Cox proportional hazard regression model to predict the risk of long-stay when patients' age and the DRG are considered as antecedent risk factors. **Results:** Predicting the risk of long stay depends significantly on the age at admission, and the DRG to which a patient belongs to. We demonstrated the validity of the Cox regression model for the available data as the proportionality assumption is shown to be satisfied. Two examples were presented to demonstrate the utility of the Cox model in this regard.

Keywords

Diagnostic Related Groups, K-Means Clustering, In Hospital Length of Stay, Cox Proportional Hazard Models, Relative Risk Estimation

1. Introduction

1.1. Background

Length of hospital stay (LOS) is an important indicator to assess the quality of care and the efficient use of medical resources. Decreased LOS has been associated with decreased risks of opportunistic infections and side effects of medication, and with improvements in treatment outcome and lower mortality rates. Furthermore, shorter hospital stays reduce the burden of medical fees and increase the bed turnover rate, which in turn increases the profit margin of hospitals, while lowering the overall the administrative costs [1] [2].

LOS among patients with the same disease may vary owing to a variety of factors, some of them are attributed to patients and others attributed to the hospital management practices participating in the study. Therefore, in order to understand how the risk factors measured at both levels of hierarchy and their possible interactions contribute to the variability in LOS, we should carefully select these factors to maximize our ability to predict LOS [3] for each patient at the time of admission to the care facility.

1.2. Ethical Data Acquisition

In this study, we used the electronic medical records (EMR) of the King Faisal Specialist Hospital and Research Center (KFSHRC), after obtaining the *Institutional Review Board* to analyze the LOS defined as the time between admission and discharge of patients forming the study sample.

The main contribution of this research is the use of the Cox Proportional Hazard Regression model to replace the traditional regression modeling techniques. The proposed approach overcomes the inadequacy of traditional regression models which require the response variable to belong to the well-known Gaussian family. This requirement is in fact the corner stone assumption needed for most regression analyses. We noted that the LOS is reported as the number of days elapsed from admission to discharge. As such it was recorded as an integer value with rightly skewed histogram.

In section 2, we provide summary of the basic features of the data, with emphasis on the DRG's selected for the study. In section 3, we use descriptive analyses of the data with the LOS being the primary outcome of interest. In section 4, we use Cox regression model to predict the risk of long stay. We provide a general discussion in section 5.

2. Study Data

The data set that we used have three complete observation for each patient, the LOS, Age at Admission, and the Diagnostic Related Group. The concept of Diagnostic Related Groups (DRGs) was first developed at Yale University in 1975. The main objective was to group patients with similar treatments and conditions for comparative studies. The DRGs were designed to be homogeneous units of hospital activity to which binding prices could be attached. A central theme in the advocacy of DRGs was that the reimbursement system would oblige hospital administrators to alter the behavior of the physicians and surgeons comprising their medical staff. Hospitals were forced to leave the “nearly risk-free” world of cost reimbursement and face the uncertain financial consequences associated with the provision of health care.

Krumholz *et al.* [4] discuss several factors that should be considered when assessing hospital performance. These relate to differences in the chronic and clinical acuity of patients at hospital presentation, the numbers of patients treated at a hospital, the frequency of the outcome studied, the extent to which the outcome reflects a hospital quality signal, and the form of the performance metric used to assess hospital quality. However, issues related to DRG have not been considered as factors of importance. Since the outcome of interest is LOS, any attempt to predict this variable that does not take into account the relative importance of DRG will produce biased findings.

We searched the electronic medical at KFSHRC between (January 2014) and (December 2016) for patients with complete information regarding their DRG, age, and LOS. We were able to obtain such information for five DRG groups as listed below, and we included the ICD-10 in brackets:

- 1) Acute Leukemia (R60B)
- 2) Lymphoma (R61B)
- 3) Endocrine metabolic diseases (K64B)
- 4) Kidney diseases (L04C)
- 5) Diseases of the respiratory systems (E62B)

In what follows we provide short literature review regarding the above five DRGs.

Acute Leukemia

Leukemia is a malignant neoplasm of hematopoietic origin, characterized by diffuse replacement of bone marrow and peripheral blood with neoplastic cells [5]. Although, many subtypes of leukemia were known, four main subtypes were frequently seen in diagnosis such as: Acute Myeloid Leukemia (AML), Chronic Myeloid Leukemia (CML), Acute Lymphoblastic Leukemia (ALL) and Chronic Lymphocytic Leukemia (CLL). Globally, between 1990 to 2018, the number of leukemia cases markedly increased from 297,000 to 437,033 [6]. Thus, according to GLOBOCAN report in 2018, leukemia was ranked the 13th among cancers worldwide, while leukemia deaths increased by 16.5% in the same year. According to the reported data from the GLOBOCAN for region of Middle-East and

Northern Africa (MENA), the estimated crude incidence is 5.3 per 100,000 among male population and 4.0 per 100,000 females [7]. Moreover, Arab Gulf Cooperation Council report on cancer, ranked leukemia as the 4th among the most common cancers in the area. The national health survey reported that increased prevalence of leukemia lesions among Saudi population is alarming for the healthcare service. This is because of serious complications of leukemia [7]. In 2016, the Saudi Cancer Registry, stated that leukemia was ranked 5th among cancers in both genders of all ages in the Saudi population. The overall prevalence of leukemia was 7.6% in males and 4.4% in females in Saudi population [8]. When looking at the age group of older than 14 years of age, leukemia ranked in the top seventh (3.7%), while it ranked the first (38.8%) among Saudi children of less than 14 years of age, with higher rates in males compared to females (59.6% vs. 40.9%). In this study, we aim to define the burden of leukemia with respect to LOS [9].

Lymphoma

In 2008, Non-Hodgkin Lymphoma (NHL) was one of the most prevalent types of cancer in Saudi Arabia, and ranked second in cancer incidence among the male population, with a ratio of 122:100 for men to women [10]. The International Agency for Research on Cancer (IARC) estimated that the age-standardized incidence rate (ASIR) for NHL was 6.5 per 100,000 men in 2012, and the age-standardized mortality rate (ASMR) was 4.3 per 100,000 men. Furthermore, the registry of King Faisal Specialist Hospital and Research Centre in Saudi Arabia recorded 5493 cases (7.6%) of NHL with admission to the hospital from 1975 to 2011. In Saudi Arabia, the ASIR of NHL is higher than that in the other Arabian Gulf countries [10].

Kidney diseases

Chronic Kidney Diseases (CKD) are a group of illnesses that constitute serious problem worldwide. However, data on the burden of CKD in the Arab world remains poorly understood [11]. The kingdom of Saudi Arabia which is the largest country in the Arabian Peninsula in Southwest Asia has an estimated population of 32 million, including approximately 5.5 million non-nationals. Data available on the exact incidence and prevalence of chronic kidney disease is limited to patients with end-stage renal disease. In the annual report of the Saudi Center for Organ Transplantation (SCOT) the incidence of dialysis in the Kingdom of Saudi Arabia was 136 new cases per million population (PMP). This compares to 360 PMP in the United States, 4585 PMP in Europe and to 163 PMP in India. The SEEK-Saudi study (Screening and Early Evaluation of Kidney Disease) is aimed at evaluating the burden of CKD and its predictors in the kingdom of Saudi Arabia using standardized GFR prediction equations. We shall investigate the non-epidemiologic burden of CKD using the available data and attempt to predict the LOS after adjusting for age and the effect of other DRG's.

Endocrine metabolic diseases

Although Saudi Arabia reports one of the highest prevalence levels of obesity

and diabetes, a very limited number of epidemiological studies have examined the prevalence of metabolic syndrome. In a recent study about the metabolic syndrome in Saudi Arabia [12], a total of 12,126 Saudi subjects were randomly recruited from the Kingdom's 13 administrative regions and were evaluated for metabolic syndrome and its risk factors. The prevalence of metabolic syndrome in Saudi Arabia was found to be 39.8% (34.4% in men and 29.2% in women) and 31.6% (45.0% in men and 35.4% in women), according to the NCEP ATP III and IDF criteria, respectively. Metabolic syndrome was also observed to be more prevalent among men and older subjects. The most frequently observed component of metabolic syndrome was found to be low levels of high-density lipoprotein (HDL), followed by abdominal obesity. The most significant risk factors in the studied cohort included age ≥ 45 , smoking history, low educational level, and living in urban areas. As can be seen, age is a very important risk factor for the metabolic syndrome and we shall demonstrate that it significantly predicts the risk of long stay.

Diseases of the respiratory system

According to the World Health Organization (WHO), chronic obstructive pulmonary disease (COPD) and asthma are among the most common respiratory diseases affecting people worldwide [13]. Respiratory diseases are associated with excess mortality, reduced quality of life for patients, and high health-care costs [14] [15]. In 2014, respiratory diseases combined represented the fifth leading cause of death in Saudi Arabia, according to the Kingdom's Ministry of Health (MOH) [16]. Approximately, 3388 people with respiratory diseases died in 2014, compared to 1892 in 2010. The WHO has estimated that > 65 million people have moderate-to-severe COPD worldwide [17].

3. Data Analyses and Results

LOS has consistently been measured as an indicator of health care quality due to its availability, objective nature [18] and close association with outcomes [19]. Previous research has linked decreased LOS to worse patient outcomes, such as higher rates of hospital readmissions in a wide variety of patient populations [19] [20]. We used SPSS version 26 to analyze the data.

Our study sample (N = 5894) contains the complete records for the LOS, DRG, and the Age at Admission. The data were extracted from the Electronic Medical Records of the largest tertiary hospital in Saudi Arabia. The fundamental objective is to analyze the relationship between diseases diagnoses and hospital LOS, controlling for age as a known confounder. Overall, 26% (1543) of inpatients were discharged within 9 days. The Acute Leukemia patients group had the highest average LOS at 17 days. The results demonstrate significant differences in hospital LOS by disease diagnosis. Previous studies linked the insurance status of inpatients to the treatment outcome. Uninsured patients are known to have worse outcomes, including mortality and decreased access to health care resources [21] [22]. These studies also indicated that publicly insured

patients tend to fall on the other end of the LOS spectrum, experiencing extended stays that are potentially both dangerous and costly. Shorter LOS has been associated with hospitals and physicians that maintain better quality of care ratings, evidenced by improved patient satisfaction and decreased mortality [23]. Research has also demonstrated that extended hospitalization carries a number of risks, particularly as patients age.

Age plays an important role in the variability of LOS. Financing long-term care for the elderly is one of the most challenging health care problems facing the health care system today. The dramatic increase in health expenditures for long-term care is straining health care budget and the specter of aging population suggests that the problem will become worse. Frequently overlooked, however, is the fact that financing long-term care is also a significant drain on private insurance and that the options for privately insuring against such expenditures are extremely limited. Elderly persons with resources who need long-term care must pay for such services out-of-pocket. Since such care can be quite expensive, particularly if it is at a level that requires long stay, people who need it become candidates for early discharge and not receiving the needed care.

We have divided the methods of statistical analyses into descriptive methods and predictive modeling approach. The motivation is, if a patient is diagnosed with a certain disease, then we should be able to predict the length of his/her stay in order to better manage the hospital resources. Descriptive and exploratory analyses seek to understand from multiple angles, the current circumstances surrounding the hospital LOS. We included three detailed analysis items in this analysis category: LOS analysis according to DRG, and analysis for patients Age-At-Admission (AAA). **Table 1** shows the summary measures of LOS including number of patients, mean and median LOS the inter-quartile range (IQR) and the minimum and maximum LOS for each disease. Cancer has the highest mean LOS (mean LOS for Leukemia is about 17 days, while the mean LOS for Lymphoma is 11 days). An important feature in this table is that the variance of LOS is much higher than the mean LOS, a phenomenon known as “Over Dispersion” causing the distribution of LOS to be rightly skewed. As is also depicted in **Figure 1**, the histogram of LOS.

Table 2 has important information as well. It shows the percentage of patients whose LOS is above the mean LOS for the corresponding DRG. As can be seen higher percentage of patients in the cancer groups are long stayer.

Table 1. Summary measures for LOS for each DRG.

DRG	N	Mean	SD	Median	IQR	Minimum	Maximum
Acute Leukemia	1308	16.7	18.8	9.5	21	1	218
Endocrine Metabolic Diseases	1127	4.17	3.03	4	1	1	314
Kidney Diseases	1173	5.71	4.2	5	2	1	48
Lymphoma	1273	11	12	7	9	1	174
Respiratory System Diseases	1013	9	10.14	6	7	1	107

In **Table 3** the same summary measures are presented for AAA according to each disease group. As can be seen, younger patients fall in the Acute Leukemia group, with mean age 17 and median age 12. The other four groups have much older patients.

Although our modeling strategy is to use age (transformed on the logarithmic scale) as a covariate in the Cox model, we shall investigate its potential effect when it is appropriately categorized into groups.

K-Means Clustering of AAA

Research clinicians prefer having many demographic data of their patients to be reported on categorical scale. In order to produce unbiased categorization of age, we used the k-means clustering algorithm, a form of unsupervised learning, to obtain meaning categories for AAA.

This k-means algorithm tries to cluster (group) data based on their similarity.

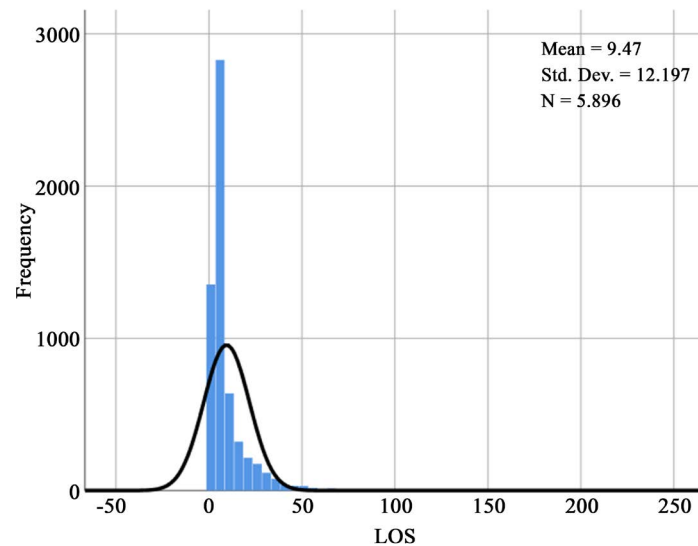


Figure 1. Histogram of LOS.

Table 2. Percentage of patients whose LOS is above the mean LOS of the corresponding DRG.

% above respective mean LOS	57.5%	6.2%	9%	28.2%	20%
DRG	Acute leukemia	Endocrine diseases	Kidney diseases	Lymphoma	Respiratory diseases

Table 3. Summary measures of Patients Age at Admission for each DRG.

DRG	N	Mean	SD	Median	IQR	Minimum	Maximum
Acute Leukemia	1308	17	15	12	16	1	81
Endocrine Metabolic Diseases	1127	38	18.2	39	23	1	98
Kidney Diseases	1173	32	15	30	13	1	100
Lymphoma	1273	38	22	35	39	1	96
Respiratory System Diseases	1013	39	30	37	61	1	100

This form of unsupervised machine learning means that there is no outcome to be predicted, and the algorithm just tries to find patterns in the data. In k-means clustering, we have to specify the number of clusters we want the data to be grouped into. The algorithm randomly assigns each observation to a cluster, and finds the centroid of each cluster. Then, the algorithm iterates through two steps:

- Reassign data points to the cluster whose centroid is closest.
- Calculate new centroid of each cluster.

These two steps are repeated till the within cluster variation cannot be reduced any further. The within cluster variation is calculated as the sum of the Euclidean distance between the data points and their respective cluster centroids. On using this algorithm, we specified $k = 4$, to be the number of groups of for AAA. The choice of 4 age groups is not arbitrary, but is in accordance with clinical classifications introduced in [16]. The results are shown in **Table 4**.

Figure 2 gives the bar chart of the data presented in **Table 4**. We note that the young patients form the highest number in the leukemia group. As well, middle age (mean age 29.4 years) forms the majority of the kidney diseases group.

In **Figure 3** we plot the 95% confidence levels for LOS by each DRG and the age grouping.

We can now combine the information in **Table 1** and **Table 3** and **Figure 3**, so

Table 4. K-means clustering algorithm for AAA.

Age group	Number of patients	Mean Age	Standard Deviation	Range
1	2060	9.29	5.56	1 - 20
2	1835	29.40	5.16	21 - 39
3	1208	49.53	5.97	40 - 60
4	791	71.98	8.12	61 - 100
Total	5894	32.21	22.27	1 - 100

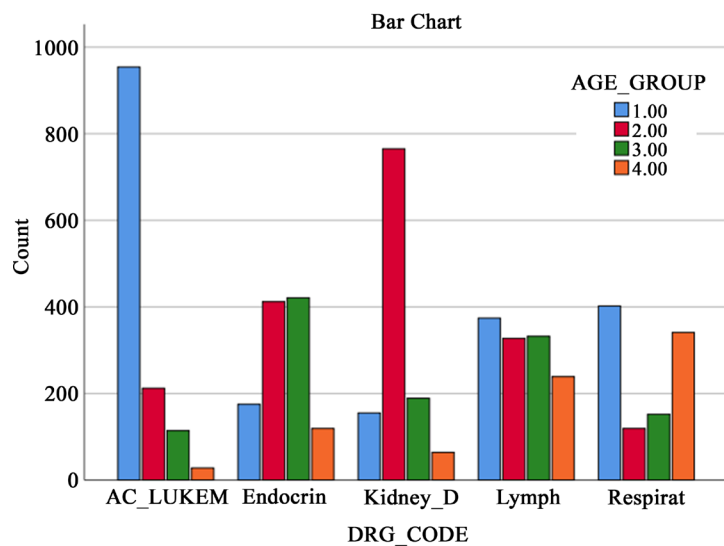


Figure 2. Cross classification of DRG by age groups.

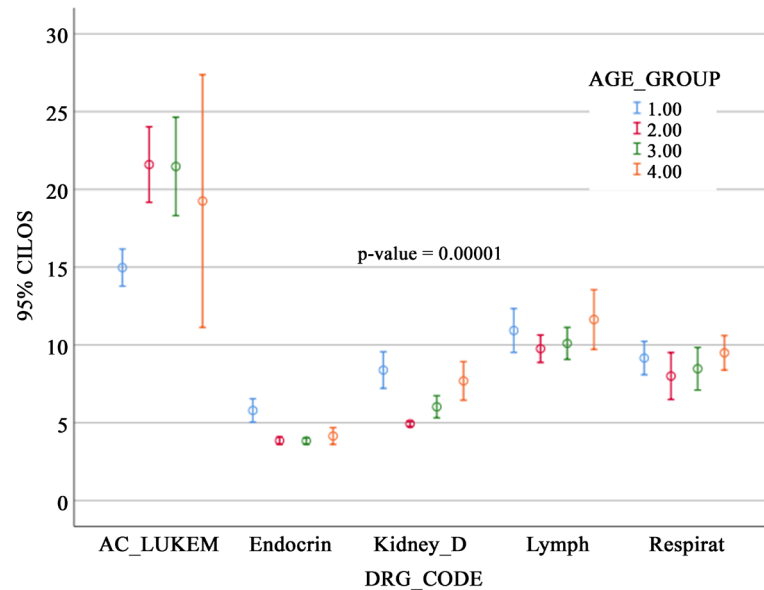


Figure 3. Error bars plot of LOS for each DRG stratified by age groups.

Table 5. Estimated coefficients of Cox regression analysis.

Variable	β _coefficient	SE	P-Value
CODE (1) Acute Leukemia	-0.398	0.098	0.00001
CODE (2) Endocrine diseases	0.640	0.274	0.019
CODE (3) Kidney diseases	0.536	0.176	0.002
CODE (4) Lymphoma	-0.10	0.108	0.358
Logarithm_AAA	0.105	0.031	0.001

that we can detect the variability in the LOS and its association with both age and DRG. It is clear that, after adjusting for age at admission, there are fundamentally two sub-clusters of the disease groups with respect to the LOS. Moreover, from **Figure 3**, we see that the mean LOS for the acute leukemia differs significantly from that of all other DRG's, but there seem to be no age differences within each DRG, as the 95% confidence intervals are overlapping.

In the second stage of the analysis, we develop a predictive model for the risk of over staying using the Cox proportional hazard regression model. The Cox model is one of the most accurate method belonging to the class of semiparametric statistical models. The Cox model has the advantage is that it can use different types of independent variables (continuous, ordered categorical, and nominal variables). The regression coefficients, their standard errors, and the corresponding p-values are presented in **Table 5**. The 21 days (three weeks) baseline survival function when all covariates are set equal to zero is $S(21) = 0.802$. In **Figure 4** we present the hazard function at each time point, with separate curve for each DRG after correcting for the effect of age at admission. The endocrine diseases, and acute leukemia have the highest hazard of over staying, and the other three disease groups have significantly lower hazards (p -value = 0.00001).

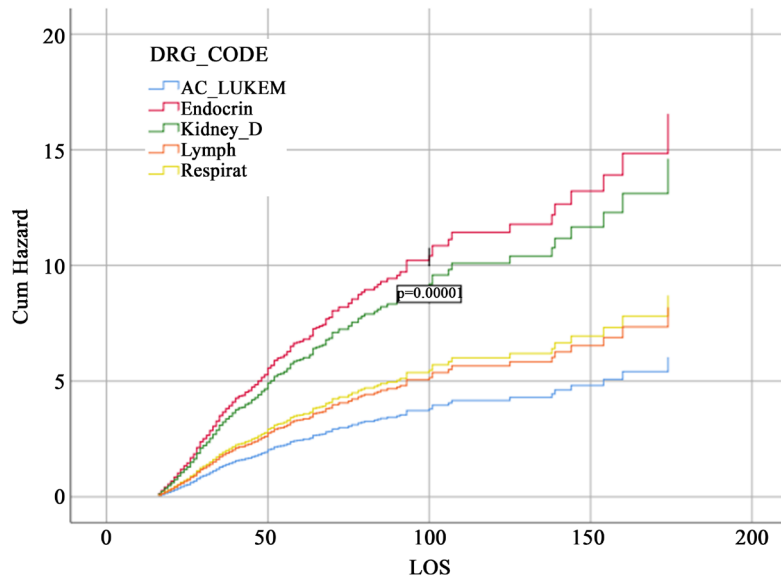


Figure 4. Hazard plots based on the Cox regression analysis.

The risk prediction equation is given by:

$$\text{Risk}(t) = 1 - S(t)^{h(x)} \tag{1}$$

The components of the prediction analyses are:

$$h(x) = \exp\left(\sum_{j=1}^5 \beta_j x_j\right) \tag{2}$$

- $x_1 = 1$ if patient belongs to Acute Leukemia group, and 0 otherwise.
- $x_2 = 1$ if patient belongs to Endocrine Diseases group, and 0 otherwise.
- $x_3 = 1$ if patient belongs to Kidney disease group, and 0 otherwise.
- $x_4 = 1$ if patient belongs to the Lymphoma disease group, and 0 otherwise.
- $x_5 = \text{Log_Age}$.

Hence,

$$h(x) = \exp(-0.398 * (x_1) + 0.640 * (x_2) + 0.536 * (x_3) - 0.10 * (x_4) + 0.105 * (x_5)) \tag{3}$$

The Equations (1)-(3) complete the specifications of the Cox regression prediction model. To clarify the utility of the above approach we consider two examples below.

The above coding of the DRG’s means that the “Respiratory system disease group” is taken to be the reference group. Moreover, from the Cox regression model we estimate the 21 days survival probability to be:

$$S(21) = 0.802$$

It is known that the validity of the prediction depends heavily on the proportionality assumption needed for the applicability of the Cox regression model. There several approaches to verify the proportionality assumption. One approach is to plot the complementary-log of the survival function against the log-event time [23]. A straight line passing through the scatter plot would

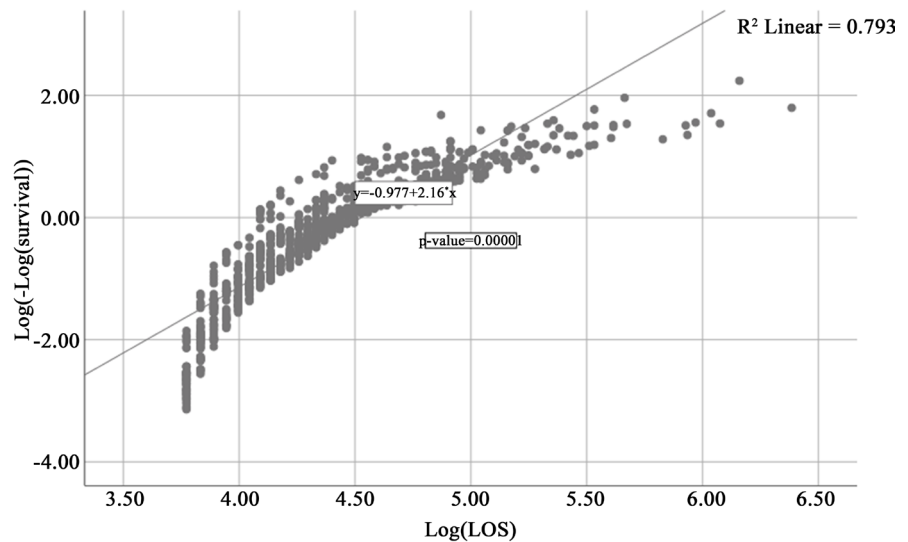


Figure 5. Plotting the complementary log of the survival function estimate against the log of LOS to verify that the proportionality assumption is satisfied [24] [25].

indicate that the proportionality assumption is satisfied. **Figure 5** shows the scatter plot and the approximate straight line passing through the scatters. The R-square is 80% indicating an excellent fit. We therefore conclude that the proportionality assumption is approximately satisfied.

To illustrate the utility of the prediction we consider two examples.

Example 1: Suppose that we have two Leukemia patients ($x_1 = 1$) the age of one patient is 15 years and the age of the other patient is 60 years. That is $h(x) = \exp(-0.398 + 0.105 * \log(15)) = 0.892$, for the first patient, and the risk of staying over 21 days is

$$R_1 = 1 - (0.802)^{0.892} = 17.9\% .$$

For the other patient we have $h(x) = \exp(-0.398 + 0.105 * \log(60)) = 1.032$, and the risk of staying over 21 days is:

$$R_2 = 1 - (0.802)^{1.032} = 20.3\%$$

Therefore, the relative risk is $R_2/R_1 = 1.13$.

This means that a 60 years old leukemia patient has a 13% increase in the risk of overstay relative to a 15 years old leukemia patient.

The second example compares the risk for two patients of the same age but they belong to two different DRG. We shall assume that the first patient has the same risk profile as in the previous example. We assume that the second patient has the same age (15 years) but is from the respiratory diseases group. In this case:

$$h(x) = \exp(0 + 0.105 * \log(15)) = 1.33$$

$$R_2 = 1 - (0.802)^{1.33} = 25.41\%$$

Therefore, the relative risk is $R_2/R_1 = 25.41/17.9 = 1.42$. This means that a pa-

tient with respiratory illness has a 42% increase in the risk of overstaying beyond 21 days as compared to a leukemia patient of the same age.

4. Discussion and Study Limitations

Due to the COVID-19 pandemic, the demand for hospital resources is at all time high. Therefore, controlling hospital stay should be one of the priorities of in hospital admissions. In this study, we investigated two important variables correlated with LOS using the KFSHRC electronic medical records data. Research on the duration of hospital stay is important because it helps hospitals to more effectively manage their resources for efficient delivery of health care. Specifically, identifying factors which are associated with the LOS in order to accurately predict and manage the number of inpatient days, and enabling the development of a Clinical Pathway useful for inpatient treatment. Needless to say, reducing unnecessary hospital stays is a strategy to reduce overall national medical expenses.

There were some limitations to this study which should be addressed. First, the analysis of patient process correlating with the LOS was based on data from a single hospital. As there are differences between hospitals in the admission process and treatment plans, generalizability was limited and it is important to collect and analyze data from multiple hospitals. Furthermore, data analysis was largely confined to the main hospitalization events of the EHR system; the general characteristics of the individual patients and the hospital's environmental factors were not considered in the analysis. The LOS may also be related with month of the year or day of the week of admission/discharge date, for example, admissions on Friday not being discharged until Monday due to lack of senior staff on weekends. Despite these limitations, this study analyzed the LOS based on objective EHR data that included all medical events for each inpatient rather than some specific patients. Importantly, this study is of value as it analyzed the factors correlating with LOS and identified solutions to reduce this time.

In future studies related to hospital stay, it may be necessary to collect multi-institutional data, as well as the general characteristics of individual subjects, their environmental factors, medical insurance status, and seasonal and date/time factors, which were not considered in this study.

Acknowledgements

The authors are grateful to Dr. Maha Al-Eid of the Research Center of KFSHRC for reviewing the final draft of the manuscripts.

This work was completed while the corresponding author was Principal Scientist in the department of cell biology at KFSHRC.

“The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.”

Both authors contributed equally to the research constituting this paper.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Bueno, H., Ross, J.S., Wang, Y., Chen, J., Vidan, M.T., Normand, S.L., *et al.* (2010) Trends in Length of Stay and Short-Term Outcomes among Medicare Patients Hospitalized for Heart Failure, *JAMA*, **303**, 2141-2147. <https://doi.org/10.1001/jama.2010.748>
- [2] Rotter, T., Kinsman, L., James, E., Machotta, A., Gothe, H., Willis, J., *et al.* (2010) Clinical Pathways: Effects on Professional Practice, Patient Outcomes, Length of Stay and Hospital Costs. *Cochrane Database of Systematic Reviews*, **3**, Article ID: Cd006632. <https://doi.org/10.1002/14651858.CD006632.pub2>
- [3] Gemmel, P., Vandaele, D. and Tambeur, W. (2008) Hospital Process Orientation (HPO): The Development of a Measurement Tool. *Total Quality Management*, **19**, 1207-1217. <https://doi.org/10.1080/14783360802351488>
- [4] Krumholz, H.M. (2008) Outcomes Research: Generating Evidence for Best Practice. *Circulation*, **118**, 309-318. <https://doi.org/10.1161/CIRCULATIONAHA.107.690917>
- [5] Verhaak, R.G., Wouters, B.J., Erpelinck, C.A., Abbas, S., Beverloo, H.B., Lugthart, S., Löwenberg, B., Delwel, R. and Valk, P.J. (2009) Prediction of Molecular Subtypes in Acute Myeloid Leukemia Based on Gene Expression Profiling. *Haematologica*, **94**, 131-134. <https://doi.org/10.3324/haematol.13299>
- [6] Fitzmaurice, C., Dicker, D., Pain, A., Hamavid, H., Moradi-Lakeh, M., MacIntyre, M.F., Allen, C., Hansen, G., Woodbrook, R. and Wolfe, C. (2015) The Global Burden of Cancer 2013. *JAMA Oncology*, **1**, 505-527. <https://doi.org/10.1001/jamaoncol.2015.0735>
- [7] Gulf Centre for Cancer Control and Prevention (2011) Ten-Year Cancer Incidence among Nationals of the GCC States 1998-2007. Gulf Centre for Cancer Control and Prevention, Riyadh.
- [8] Saudi Cancer Registry and Saudi Health Council (2019) Cancer Incidence Report: Saudi Arabia. 2016. Saudi Cancer Registry and Saudi Health Council, National Health Information Center, Riyadh.
- [9] Bawazir, A., Al-Zamel, N., Amen, A., Akiel, M.A., Alhawiti, N.M. and Alshehri, A. (2019) The Burden of Leukemia in the Kingdom of Saudi Arabia: 15 Years Period (1999-2013). *BMC Cancer*, **19**, Article No. 703. <https://doi.org/10.1186/s12885-019-5897-5>
- [10] Alghamdi, I.G, Hussain, I.I., Alghamdi, M., Dohal, A., Alghamdi, M.M. and El-Sheemy, M. (2014) Incidence Rate of Non-Hodgkin's Lymphomas among Males in Saudi Arabia: An Observational Descriptive Epidemiological Analysis of Data from the Saudi Cancer Registry, 2001-2008. *International Journal of General Medicine*, **2014**, 311-317. <https://doi.org/10.2147/IJGM.S63645>
- [11] Alsuwaida, A.O., Farag, Y.M.K., Al Sayyari, A.A., Mousa, D., Alhejaili, F., Al-Harbi, A., Housawi, A., Mittal, B.V. and Singh, A.K. (2010) Epidemiology of Chronic Kidney Disease in the Kingdom of Saudi Arabia (SEEK-Saudi Investigators)—A Pilot Study. *Saudi Journal of Kidney Disease and Transplantation*, **21**, 1066-1072.

- [12] Al-Rubeaan, K., Bawazeer, N., Al Farsi, Y., Youssef, A.M., Al-Yahya, A.A., Al-Qumaidi, H., Al-Malkil, B.M., Naji, K.A., Al-Shehri, K. and Al Rumaih, F. (2018) Prevalence of Metabolic Syndrome in Saudi Arabia—A Cross Sectional Study. *BMC Endocrine Disorders*, **18**, Article No. 16. <https://doi.org/10.1186/s12902-018-0244-4>
- [13] Alsubaiei, M.E., Cafarella, P.A., Frith, P.A., Doug McEvoy, R. and Effing, T.W. (2018) Factors Influencing Management of Chronic Respiratory Diseases in General and Chronic Obstructive Pulmonary Disease in Particular in Saudi Arabia: An Overview. *Annals of Thoracic Medicine*, **13**, 144-149. https://doi.org/10.4103/atm.ATM_293_17
- [14] World Health Organization (2015) Noncommunicable Diseases (Fact Sheet). <http://www.who.int/mediacentre/factsheets/fs355/en/>
- [15] Khan, J.H., Lababidi, H.M., Al-Moamary, M.S., Zeitouni, M.O., Al-Jahdali, H.H., Al-Amoudi, O.S., *et al.* (2014) The Saudi Guidelines for the Diagnosis and Management of COPD. *Annals of Thoracic Medicine*, **9**, 55-76. <https://doi.org/10.4103/1817-1737.128843>
- [16] Ministry of Health (2014) Statistics Year Book. <http://www.moh.gov.sa/Ministry/Statistics/Book/Pages/default.aspx>
- [17] World Health Organization (2015) Chronic Obstructive Pulmonary Disease (COPD). <http://www.who.int/mediacentre/factsheets/fs315/en/>
- [18] Brasel, K.J., Lim, H.J., Nirula, R. and Weigelt, J.A. (2007) Length of Stay: An Appropriate Quality Measure? *The Archives of Surgery*, **142**, 461-465. <https://doi.org/10.1001/archsurg.142.5.461>
- [19] Epstein, A.M., Bogen, J., Dreyer, P. and Thorpe, K.E. (1991) Trends in Length of Stay and Rates of Readmission in Massachusetts: Implications for Monitoring Quality of Care. *Inquiry*, **28**, 19-28.
- [20] McAleese, P. and Odling-Smee, W. (1994) The Effect of Complications on Length of Stay. *Annals of Surgery*, **220**, 740-744. <https://doi.org/10.1097/00000658-199412000-00006>
- [21] Rosen, H., Saleh, F., Lipsitz, S., Rogers Jr., S. O. and Gawande, A.A. (2009) Downwardly Mobile: The Accidental Cost of Being Uninsured. *The Archives of Surgery*, **144**, 1006-1011. <https://doi.org/10.1001/archsurg.2009.195>
- [22] Doyle, J.J. (2005) Health Insurance, Treatment and Outcomes: Using Auto Accidents as Health Shocks. *Review of Economics and Statistics*, **87**, 256-270. <https://doi.org/10.1162/0034653053970348>
- [23] Khaliq, A.A., Broyles, R.W. and Robertson, M. (2003) The Use of Hospital Care: Do Insurance Status, Prospective Payment, and the Unit of Payments Make a Difference? *Journal of Health and Human Services Administration*, **25**, 471-496.
- [24] Bradburn, M.J., Clark, T.G., Love, S.B. and Altman, D.G. (2003) Survival Analysis Part II: Multivariate Data Analysis—An Introduction to Concepts and Methods. *British Journal of Cancer*, **89**, 431-436. <https://doi.org/10.1038/sj.bjc.6601119>
- [25] Steyeberg, E.W. (2010) Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer, New York.