Digital Repository Universitas Jember

ISSN : 2349-6495(P) | 2456-1908 (O)

# International Journal of Advanced Engineering Research and Science

(IJAERS)

An Open Access Peer Reviewed International Journal

Journal DOI: 10.22161/ijaers

Issue DOI: 10.22161/ijaers.4.7

AI PUBLICATIONS

Vol.- 4| Issue - 7 | July, 2017

editor@ijaers.com | http://www.ijaers.com/

Digital Repository Universitas Jember

# FOREWORD

I am pleased to put into the hands of readers Volume-4; Issue-7: 2017 (July, 2017) of **"International Journal of Advanced Engineering Research and Science (IJAERS) (ISSN: 2349-6495(P)| 2456-1908(O)" ,** an international journal which publishes peer reviewed quality research papers on a wide variety of topics related to Science, Technology, Management and Humanities. Looking to the keen interest shown by the authors and readers, the editorial board has decided to release print issue also, but this decision the journal issue will be available in various library also in print and online version. This will motivate authors for quick publication of their research papers. Even with these changes our objective remains the same, that is, to encourage young researchers and academicians to think innovatively and share their research findings with others for the betterment of mankind. This journal has DOI (Digital Object Identifier) also, this will improve citation of research papers.

I thank all the authors of the research papers for contributing their scholarly articles. Despite many challenges, the entire editorial board has worked tirelessly and helped me to bring out this issue of the journal well in time. They all deserve my heartfelt thanks.

Finally, I hope the readers will make good use of this valuable research material and continue to contribute their research finding for publication in this journal. Constructive comments and suggestions from our readers are welcome for further improvement of the quality and usefulness of the journal.

With warm regards.

**Dr. Swapnesh Taterh**
Editor-in-Chief
Date:5th , Aug, 2017

Digital Repository Universitas Jember

## Editorial Board

**Dr. C.M. Singh**
*BE., MS(USA), PhD(USA),Post-Doctoral fellow at NASA (USA)*
*Professor, Department of Electrical & Electronics Engineering, INDIA*

**Dr. Ram Karan Singh**
*BE.(Civil Engineering), M.Tech.(Hydraulics Engineering), PhD(Hydraulics & Water Resources Engineering),BITS- Pilani*
*Professor, Department of Civil Engineering,King Khalid University, Saudi Arabia.*

**Dr. Asheesh Kumar Shah**
*IIM Calcutta, Wharton School of Business, DAVV INDORE, SGSITS, Indore*
*Country Head at CrafSOL Technology Pvt.Ltd, Country Coordinator at French Embassy, Project Coordinator at IIT Delhi, INDIA*

**Dr. Swapnesh Taterh**
*Ph.d with Specialization in Information System Security*
*Associate Professor, Department of Computer Science Engineering*
*Amity University, INDIA*

**Dr. Ebrahim Nohani**
*Ph.D.(hydraulic Structures), Department of hydraulic Structures,Islamic Azad University, Dezful, IRAN.*

**Dr. Dinh Tran Ngoc Huy**
*Specialization Banking and Finance, Professor,Department Banking and Finance*
*Viet Nam*

**Dr.Sameh El-Sayed Mohamed Yehia**
*Assistant Professor, Civil Engineering(Structural), Higher Institute of Engineering -El-Shorouk Academy, Cairo, Egypt*

**Dr.AhmadadNabihZaki Rashed**
*Specialization Optical Communication System,Professor,Department of Electronic Engineering, Menoufia University*

**Dr. Alok Kumar Bharadwaj**
*BE(AMU), ME(IIT, Roorkee), Ph.D (AMU),Professor, Department of Electrical Engineering, INDIA*

**Dr. M. Kannan**
*Specialization in Software Engineering and Data mining*
*Ph.D, Professor, Computer Science,SCSVMV University, Kanchipuram, India*

**Dr.Sambit Kumar Mishra**
*Specialization Database Management Systems*
*BE, ME, Ph.D,Professor, Computer Science Engineering*
*Gandhi Institute for Education and Technology, Baniatangi, Khordha, India*

Digital Repository Universitas Jember

**Dr. M. VenkataRamana**
*Specialization in Nano Crystal Technology*
*Ph.D,Professor, Physics,Andhara Pradesh, INDIA*

**DR. C. M. Velu**
*Prof.& HOD, CSE, Datta Kala Group of Institutions, Pune, India*

**Dr.RabindraKayastha**
*Associate Professor, Department of Natural Sciences*
*School of Science, Kathmandu University, Nepal*

**Dr. P. Suresh**
*Specialization in Grid Computing and Networking, Associate Professor, Department of Information Technology, Engineering College, Erode,Tamil Nadu ,INDIA*

**Dr. Uma Choudhary**
*Specialization in Software Engineering Associate Professor, Department of Computer Science Mody University, Lakshmangarh, India*

**Dr.Varun Gupta**
*Network Engineer,National Informatics Center , Delhi ,India*

**Dr. Hanuman Prasad Agrawal**
*Specialization in Power Systems Engineering Department of Electrical Engineering, JK Lakshmipat University, Jaipur, India*

**Dr.Hou, Cheng-I**
*Specialization in Software Engineering, Artificial Intelligence, Wisdom Tourism, Leisure Agriculture and Farm Planning, Associate Professor, Department of Tourism and MICE, Chung Hua University, Hsinchu Taiwan*

**Dr. Anil TrimbakraoGaikwad**
*Associate Professor at BharatiVidyapeeth University, Institute of Management , Kolhapur, India*

**Dr. Ahmed Kadhim Hussein**
*Department of Mechanical Engineering, College of Engineering, University of Babylon, Republic of Iraq*

**Dr.GamalAbd El-Nasser Ahmed Mohamed Said**
*Computer Lecturer, Department of Computer and Information Technology, Port Training Institute (PTI), Arab Academy For Science, Technology and Maritime Transport, Egypt*

**Mr. T. Rajkiran Reddy**
*Specialization in Networing and Telecom*
*Research Database Specialist, Quantile Analytics, India*

**M. HadiAmini**
*Carnegie Mellon University, USA*

Digital Repository Universitas Jember




| Sr No. | Detail |
|---|---|
| 1 | **Assessment of contamination of soil by pesticides in Djidja's cotton area in Benin**<br>*Firmin H. Aikpo, Miriac Dimitri S. Ahouanse, Lucien Agbandji, Patrick A. Edorh, Christophe S. Houssou*<br>**DOI:** 10.22161/ijaers.4.7.1<br>**Page No:** 001-005 |
| 2 | **Iron and Nickel -ligand bonding in metallocene: Differentiation between bond Stability and reactivity**<br>*Tesfalem Belay Woldeamanuale*<br>**DOI:** 10.22161/ijaers.4.7.2<br>**Page No:** 006-019 |
| 3 | **Design and Analysis of 8x8 Wallace Tree Multiplier using GDI and CMOS Technology**<br>*Pradeep Kumar Kumawat, Gajendra Sujediya*<br>**DOI:** 10.22161/ijaers.4.7.3<br>**Page No:** 020-024 |
| 4 | **Prevalence of Musculoskeletal Problems among Sugarcane Workers in Uttar Pradesh**<br>*Sushma Gangwar, Seema Kwatra*<br>**DOI:** 10.22161/ijaers.4.7.4<br>**Page No:** 025-028 |
| 5 | **Microcontroller Based Wireless Controlled Pick & Place Robot**<br>*Md. Faiz Akram*<br>**DOI:** 10.22161/ijaers.4.7.5<br>**Page No:** 029-033 |
| 6 | **Information and Communication Technologies for Veterinary Sciences and Animal Husbandry in Jammu and Kashmir**<br>*T. A. Raja, Azmat Alam Khan, N A Ganai andImtiyaz Ahmad Najar*<br>**DOI:** 10.22161/ijaers.4.7.6<br>**Page No:** 034-038 |
| 7 | **Design, Buckling and Fatigue Failure Analysis of Connecting Rod: A Review**<br>*Manish Kumar, Shiv N Prajapati*<br>**DOI:** 10.22161/ijaers.4.7.7<br>**Page No:** 039-044 |
| 8 | **The Effect of Ketapang Leaf Extracts (Terminalia catappa L.) on the Cholesterol Levels of Male Mice (Mus musculus L.) Hypercholesterolmia**<br>*Joko Waluyo, Dwi Wahyuni*<br>**DOI:** 10.22161/ijaers.4.7.8<br>**Page No:** 045-049 |

Digital Repository Universitas Jember


| | |
|---|---|
| 9 | **Construction Productivity Estimation Model Using Artificial Neural Network for Foundations Works in Gaza Strip Construction Sites**<br>*Dr. Eyad Haddad*<br>**DOI: 10.22161/ijaers.4.7.9**<br>**Page No:** 050-062 |
| 10 | **Implementation of Invisible Digital Watermarking Technique for Copyright Protection using DWT-SVD and DCT**<br>*Devasis Pradhan*<br>**DOI: 10.22161/ijaers.4.7.10**<br>**Page No:** 063-069 |
| 11 | **Climate Variability and Its implication on Wheat Productivity; Farmers' Adaptation strategies: in Robe Town and Surrounding Area, Oromia Regional State, Ethiopia**<br>*Aliyi Mama Gasu*<br>**DOI: 10.22161/ijaers.4.7.11**<br>**Page No:** 070-081 |
| 12 | **Structural and Electronic Properties of Cis-platin Metal Complex: B3LYP-SDD/DFT Calculations**<br>*Faeq A. Mohammed, Hamid I. Abbood*<br>**DOI: 10.22161/ijaers.4.7.12**<br>**Page No:** 082-086 |
| 13 | **Energy and Exergy Analysis of a CI engine fuelled with biodiesel fuel from palm kernel oil and its blends with petroleum diesel**<br>*Iortyer H. A, Likita Bwonsi*<br>**DOI: 10.22161/ijaers.4.7.13**<br>**Page No:** 087-097 |
| 14 | **The Effect of various Sterilization Procedure on Number of Colony-Forming units of isolated endophytic bacteria from Cosmos caudatus Kunth. leaf**<br>*Fiqih Ramadhan, Fikri Ainur Risma Hardianti Oktavia, Luluk Mukarramah, Ria Yulian, Nurul Hilyatun, Iis Nur Asyiah*<br>**DOI: 10.22161/ijaers.4.7.14**<br>**Page No:** 098-100 |
| 15 | **Studies on Preparation of Custard Apple Vinegar**<br>*S. J. Raichurkar, R. A. Dadagkhair*<br>**DOI: 10.22161/ijaers.4.7.15**<br>**Page No:** 101-106 |
| 16 | **Health benefits of Green and Black Tea: A Review**<br>*K.H.G.K. Kodagoda, I. Wickramasinghe*<br>**DOI: 10.22161/ijaers.4.7.16**<br>**Page No:** 107-112 |
| 17 | **Removal of Zinc Metal Ions from Electroplating Industrial Waste Water by Using Bio-Sorbent** |

Digital Repository Universitas Jember


| | |
|---|---|
| | *Mr. Vivek S. Damal, Mrs. V. U. Khanapure*<br>DOI: 10.22161/ijaers.4.7.17<br>**Page No:** 113-117 |
| 18 | **Antioxidants and antioxidant activity common eight banana varieties in Kerala**<br>*Siji .S, Nandini. P.V*<br>DOI: 10.22161/ijaers.4.7.18<br>**Page No:** 118-123 |
| 19 | **Theoretical Study of Electronic and Electrical Properties of Pure and Doped Graphene Sheets**<br>*Hamid I. Abbood, Hakiema S. Jabour*<br>DOI: 10.22161/ijaers.4.7.19<br>**Page No:** 124-127 |
| 20 | **Nutritional Evaluation of Different Mango Varieties available in Sri Lanka**<br>*Kothalawala S.G., Jayasinghe J.M.J.K.*<br>DOI: 10.22161/ijaers.4.7.20<br>**Page No:** 128-131 |
| 21 | **Five New Ways to Prove a Pythagorean Theorem**<br>*Nurul Laily, Hobri, Dafik*<br>DOI: 10.22161/ijaers.4.7.21<br>**Page No:** 132-137 |
| 22 | **Survival Analysis in Patients with Dengue Hemorrhagic Fever (DHF) Using Cox Proportional Hazard Regression**<br>*Luluk Handayani, Mohamat Fatekurohman, Dian Anggraeni*<br>DOI: 10.22161/ijaers.4.7.22<br>**Page No:** 138-145 |
| 23 | **An Efficient Energy Savings Schemes using Adjacent Lossless Entropy Compression for WSN**<br>*Cisil Baby, Dr .H. N. Suresh*<br>DOI: 10.22161/ijaers.4.7.23<br>**Page No:** 146-150 |
| 24 | **Reactive Power Control in Power System using Modified UPFC**<br>*Engr. M. Bilal, Engr. Dr. Ejaz Hassan, Engr. Hamza Zafar*<br>DOI: 10.22161/ijaers.4.7.24<br>**Page No:** 151-155 |
| 25 | **Revolutionary Automatic Traffic Controller**<br>*Satyendra Pratap Singh, Rajan Tiwari*<br>DOI: 10.22161/ijaers.4.7.25<br>**Page No:** 156-163 |
| 26 | **Short-term rhGH increases PIIINP, a biomarker of endothelial dysfunction**<br>*Graham M. R., Gu Y., Evans P. J., Cooper S. M., Davies B., Baker J. S.* |

Digital Repository Universitas Jember


| | |
|---|---|
| | DOI: 10.22161/ijaers.4.7.26<br>**Page No:** 164-173 |
| 27 | **Data modeling diagnostics for share price performance of Islamic Bank in Malaysia using Computational Islamic Finance approach**<br>*Nashirah Abu Bakar, Sofian Rosbi*<br>DOI: 10.22161/ijaers.4.7.27<br>**Page No:** 174-179 |
| 28 | **Characterization of alpha-Amylase Produced by the Endophytic Strain of Penicillium digitatum in Solid State Fermentation (SSF) and Submerged Fermentation (SmF)**<br>*Sideney Becker Onofre, Ivan Carlos Bertoldo, Dirceu Abatti, Amarildo Antonio Tessaro, Douglas Refosco*<br>DOI: 10.22161/ijaers.4.7.28<br>**Page No:** 180-187 |

Digital Repository Universitas Jember

https://dx.doi.org/10.22161/ijaers.4.7.22 ISSN: 2349-6495(P) | 2456-1908(O)

# Survival Analysis in Patients with Dengue Hemorrhagic Fever (DHF) Using Cox Proportional Hazard Regression

Luluk Handayani, Mohamat Fatekurohman, Dian Anggraeni

Mathematics Depart. University of Jember, Indonesia

***Abstract*—***Indonesia is a tropical country that has two seasons: the rainy season and dry season. In the rainy season frequent flooding or puddles of water that could become mosquito breeding and the spread of various diseases, one of which is the dengue fever. Dengue Hemorrhagic Fever (DHF) is the cause of public health problems with a very rapid deployment and can lead to death within a short time. This causes dengue become one of the attractions to be investigated further. This study discusses the survival analysis and the factors that affect the healing rate of dengue patients using Cox proportional hazard regression based on data from the medical records of hospitalized dengue patients at the Jember Klinik Hospital. The results showed that the factors of age, gender, hemoglobin, trombonist, and hematocrit affect the healing rate of DHF patients.*

***Keywords*—survival analysis; DHF; cox proportional hazard.**

## I. INTRODUCTION

Indonesia is a tropical country that has two seasons, namely the rainy season and the dry season. In the rainy season, for various reasons, there are floods or puddles that can become a mosquito breeding and spread various diseases, one of which is dengue fever disease. Dengue Haemorrhagic Fever (DHF) is an acute febrile illness caused by dengue virus that enters the bloodstream through the bite of Aedes Aegypti mosquito (Wikipedia, 2012), which spreads very quickly and can result in death in a short time. Dengue Hemorrhagic Fever (DHF) is one of the diseases that almost always cause public health problems and the number is always there, even tended to increase. It is known that since the first DBD appeared in Indonesia, precisely in Surabaya in 1968, it quickly spread to other areas so that in 1980 all provinces in Indonesia have been infected with DHF (Darmowandowo, 2006). Indonesia is the country with the highest incidence of DHF in Southeast Asia since 1968-2009 (WHO, 2009). This causes DBD to be one of the interesting objects to be studied and studied further, for example, to know the length of time to survive the DHF patients to recover. The application of statistical methods that can be used to analyze the case is survival analysis.

Survival analysis is a statistical analysis that is specifically used to analyze data or cases related to the time or length of time until a particular event occurs. This survival analysis is usually used in the health field (Kleinbaum and Klein, 2012). Observational data for survival analysis are survival data, ie observation data about the time period from the beginning of observation until the occurrence of an event, the event could be death, healing or other symptoms (Lee, 1992). According to Collett (1994), survival data do not meet the statistical standard procedure requirements used in data analysis, since survival data is usually not symmetrically distributed. The histogram model of survival time in a group of individuals will tend to be skewed to the right, so it is possible that survival data is not only normally distributed. There are several other distributions that are commonly used in survival analysis, ie, exponential distributions, Weibull distributions and lognormal distributions. The distribution used in a survival analysis can be determined from the estimation of the survival data distribution with some statistical test methods, such as Anderson Darling test and Chi-Squared test. While some of the models known in this analysis include binary logistic regression, Bayesian Mixture Survival model, Random Survival Forests (RSF), Multivariate Adaptive Regression Splines and Cox Proportional Hazard regression.

Several previous studies have been conducted on DHF cases using survival analysis, such as Nisa '(2012) about survival analysis with Multivariate Adaptive Regression Splines approach in DHF case in the district.

## II. LITERATURE REVIEW

### *A. Survival Analysis*

The survival analysis or survival analysis is a time-related data analysis, from the beginning to the occurrence of a specific event (Collett, 2003). Duration

Digital Repository Universitas Jember

https://dx.doi.org/10.22161/ijaers.4.7.22 ISSN: 2349-6495(P) | 2456-1908(O)

from the beginning of the observation (time origin) until the occurrence of a special event (end point or failure event) is called the time of survival. The particular event (failure event) may be a failure, death, relapse or recovery from an illness, a response from an experiment, or another event chosen according to the researcher's interest. The survival analysis has a special characteristic, namely the distribution of data in the form of long life time (skewed) right because the value will always be positive and the data is censored (Lee, 1992).

1. SurvivalTime

Survival time can be defined as the time from the beginning of observation to the occurrence of events, can be in days, months, and years. Such events may be the development of a disease, the response to treatment, the recurrence of an illness, death or other event chosen in accordance with the interests of the researcher. Therefore, the time of survival can be the time of recovery from the disease, the time from start of treatment to the occurrence of response and time to death (Lee and Wang, 2003). In determining survival time, there are three factors that are needed (Le, 1997):

a.Time of origin (time origin or startingpoint)
b. End events / end time (ending event ofinterest)
c.Time scale as the measurement scale for the passage oftime.

2. CencoredData

The difference between survival analysis and other statistical analysis is the presence of censored data. Censored data is recorded data when there is information about individual survival times, but does not know the exact time of survival (Kleinbaum & Klein, 2012). Censorship is one of the steps that must be taken to overcome the incompleteness of an observation data. The data is said to be censored if the data can not be observed completely because the research subject is lost or resigned or until the end of the research the subject has not experienced a certain incident, while the data can be observed completely until the end of research called unencensored data (Lee & Wang, 2003). The causes of censored data are:

a. Loss to follow up, occurs when objects move, dies or
b. refuses toparticipate.Drop out, occurs when the treatment is stopped
c. for somereason.Termination of study, occurs when the study period ends intermediate object observed has not reached the failureevent.

The 3 (three) types of sensors used in survival analysis (Collet, 1994) are as follows:

a. RightSensorSensors that occur when a failure event has not occurred until the end of the study.
b. LeftSensorSensors that occur when a failure event occurs before the research begins.
c. IntervalSensorSensors that if termination in data collection and event failure occur between these time intervals.

3. SurvivalFunction and HazardFunction

The function of survival and hazard function is a fundamental function in survival analysis. In theory, the survival function can be described with a smooth curve and has the following characteristics (Kleinbaum & Klein, 2012):

1. Not increase, curve tends to decrease when tincreases
2. For t = 0, S (t) = S (0) = 1 is the beginning of the study, since no object has an event, the probability of survival time 0 is 1
3. For t = ∞, S (t) = S (∞) = 0 theoretically, if the period of study increases without limit then none survives so that the survival curve approaches zero.

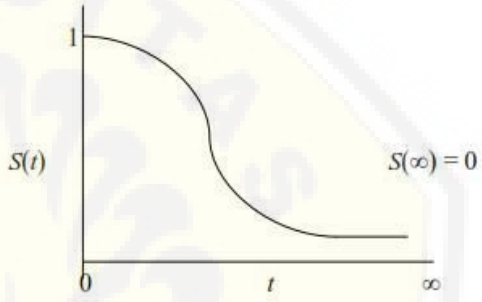


*Fig.2.2: Survival Function Curve (Kleinbaum & Klein, 2005)*

The survival function is essential in survival analysis, since there is a survival probability for various t values that is important information from survival data. The survival function is used to represent the individual probability of surviving from the initial time to some time. The survival function, *S(t)*, is defined as the probability that an individual survives greater than t time (Le, 1997), thus:

$$S(t)=P(T\square t)=1-P(T\square t)=1-F(t) \qquad (2.1)$$

with *F(t)* is the Cumulative Distribution Function (CDF) of the data distribution.

In contrast to survival functions that focus on non-occurrence of events, hazard functions focus on the occurrence of events. Therefore, hazard function can be viewed as the information giver that is opposite to the survival function. Similar to the survival survival curve, the hazard function curve also has characteristics, namely (Kleinbaum and Klein, 2012):

1. Always nonnegative, ie equal to or greater thanzero
2. Has no upper limit
3. In addition the hazard function is also used for reasons:
4. Giving an overview of the failurerate
5. Identify the specific model shape

Digital Repository Universitas Jember

*https://dx.doi.org/10.22161/ijaers.4.7.22* *ISSN: 2349-6495(P) | 2456-1908(O)*

6. Creating a mathematical model for survival analysis is usually written in the form of a hazardfunction

The hazard function *h(t)* is the probability of a person failing after a given time unit, which is the opposite of the survival function *S(t)*. The hazard formula can be interpreted as the probability of occurrence at a time interval between t and t + Δt where the survival time T is greater than or equal to t.

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t} \quad (2.2)$$

$$h(t) = \lim_{\Delta t \to 0} \frac{P(\text{an individual } t \text{ age dies at an interval up } t + \Delta t)}{\Delta t} \quad (2.3)$$

$$h(t) = dt = P(t \leq T \leq t + \Delta t | T \geq t \quad (2.4)$$

In other words, the hazard function *h(t)* estimates the proportion of deaths of individuals or individuals experiencing an event in time t (Kleinbaum and Klein, 2012). When the hazard function is always constant, itwill get a constant-risk model (exponential). The following is the functional relationship between cumulative hazard function, *H(t)*, and survival function, *S(t)* (Le, 1997) are:

$$H(t) = -lnS(t) \quad (2.5)$$

***B. Kaplan Meier***

The purpose of survival analysis is to estimate and interpret survival function. In this research, the method used is Kaplan-Meier method. Kaplan-Meier method is a technical type of survival analysis that is often used. This product is often referred to as product limit method, the method is not made a certain interval and the effect is calculated exactly when it happens. The length of each subject's observations is composed of the shortest to the longest, with the censored records included or calculated, this is considered to be proportional to numerical measurements. This research is nonparametric statistical research with censored data, so use Kaplan-Meier's method is the best.

Actually, a life-table method is the same as Kaplan-Meier, but in the life-table object is classified based on certain characteristics which each character is arranged with interval by considering the chance of effect during the interval period is constant so that the data obtained will be more general. While the Kaplan-Meier method is analyzed according to their original time. This results in a definite proportion of survival because it uses precise time survival in order to obtain more accurate data. In addition, Kaplan-Meier is a method used when no model is feasible for survival data (Sari, 2011).

**Log Rank Test**

The log rank test is a significance test for comparing survival functions between the two groups. This test is a nonparametric statistic test and is suitable to be used when data is not symmetrical ie data tilted to the right. In addition, the Log Rank test is widely used in clinical trials to look at the efficiency of a new treatment compared to the old treatment when measured is the time until an event occurs. To calculate the log rank there are several stages,namely:

a. Count the number of risky subjects in each group at the time of failure$(n_{ij})$.
b. Calculates the number of subjects who experienced events in each group at the time of failure$(m_{ij})$.
c. Calculates the number of subjects who experienced the expected event for each group at the time of failure $(e_{ij})$.

$$e_{ij} = \frac{n_{ij}}{n_{1j} + n_{2j}} \times (m_{1j} + m_{2j}) \quad (2.6)$$

d. CalculThe Log Rank

$$x^2 = \frac{(O_i - E_j)^2}{var(O_i - E_j)} \quad (2.7)$$

With :

$$(O_i - E_j) = \Sigma(m_{ij} - e_{ij}), i = 1,2, \ldots \quad (2.8)$$

*C. Cox Regression (Cox Proportional Hazard Regression)*

The function of survival and hazard function is an analysis used to see the difference of 2 or more groups. However, if there are covariate variables that want to be controlled or if using some explanatory variables in explaining the relationship between survival time then cox regression is used. Thus cox regression can be used to create a model that describes the relationship between time of survival as a dependent variable with a set of independent variables. This independent variable can be either continuous or categorical.

Cox proportional hazard is the model used in survival analysis which is a semiparametric model. Cox proportional hazard regression is used when the observed outcame is the length of time of an event. Initially this modeling was used in the branch of statistics, especially biostatistika, which is used to analyze the death or life expectancy of a person. But over the development of the modeling era is widely used in various fields, including academic, medical, social, science, engineering, agriculture and so on (Sari, 2011).

When investigating a case in medicine for example the case of a patient with a particular disease, a relationship between the patient's survival time and the clinical characteristics of the patient's medical data is required. By denoting the average hazard function h0 (t) can be determined hazard *h(t)* of a given patient, by:

Digital Repository Universitas Jember

*https://dx.doi.org/10.22161/ijaers.4.7.22* *ISSN: 2349-6495(P) | 2456-1908(O)*

$h(t) = (t)h_0(t)$ (2.9)

The cox model formula is the multiplication of two magnitudes, the baseline hazard function and the exponential form for the linear summation of $\beta_i X_i$, ie the sum of the independent variables *X* (Kleinbaum and Klein, 2012). The Bazeline hazard function is the hazard rate when it is an unknown function because thedistribution of survival time (*T*) is unknown. This function is time-dependent only and does not contain *X*. This exponential quantity depends only on *X* called time independent covariate. This is because X does not depend on time, so *X* is called an independent time covariate. However, when X is time dependent, a different method is needed to model the hazard.

In the general regression model hazard function h depends on t and dependent covariates $X_1$, $X_2$,..., $X_m(t)$. And on a simple proportional hazard cox model, with the covariates $X_1$, $X_2$, ..., $X_m$is independent of t the hazard function is as follows:

$$h(t, v_1, \dots, x_m, \beta_1, \beta_2, \dots, \beta_m) = h_0(t)\exp\{\beta_1 x_1 + \beta_1 x_1 + \beta_m x_m\} \quad (2.10)$$

$$h(t) = h_0(t)exp\{\beta_1 x_1 + \beta_1 x_1 + \beta_m x_m\} = h_0(t)\exp\{\beta' x\} \quad (2.11)$$

The most specific of these formulas is the proportional hazard assumption of baseline hazard is a function of t but does not involve the variable *X*. Unlike the exponential form involving variable *X* but does not involve t. *X* is said to be time independent (independent of time). The assumption on the proportional hazard cox model is the hazard ratio that compares the two categories of independent variables is constant at all times or is independent of time. If this assumption is not met or time dependent (*X* depends on time) then the model used is extended cox model. Another important characteristic of the cox model is that the baseline hazard, $h_0(t)$, is an unspecified function. This is what makes cox proportional hazard a parametric model.

The cox proportional hazard model is a well-known model for survival analysis. According to Kleinbaum and Klein (2012) the causes of this model are well known and widely used include:

1. The cox model is a non-parametric model.
2. Can estimate hazard ratio without the need to know $h_0(t)$ or baseline hazardfunction.
3. Can estimate $h_0(t)$, $h(t,X)$ and survival function. Although, $h_0(t)$ is notspecific.
4. Is a robust model so that the results of the model cox almost the same with the parametric modelresults.
5. The safe model is chosen when in doubt to determine its parametric model, so there is no fear about the choice of the wrong parametric model.
6. Better than the logistics model when it comes to timing survival andcensorship.

The purpose of coxregression:

1. Estimate the hazardratio
2. Testing thehypothesis
3. Looking at the confidence interval of the hazardratio

The assumption in the Cox Proportional Hazard model is that the hazard ratio comparing the two categories of predictors is constant at all times or is independent of time. In general, there are three approaches to assessing proportional hazard assumptions, namely:

1. With the graphapproach
   You do this by plotting the Log Minus Log (LML) plot of the survival function. In this plot for each strata must be parallel / parallel. This method can only be used for categorical variables. For continuous variables should be converted into categorical (2 or 3 groups). If each strata of the tested variable is parallel, then the assumption is satisfied. If it is not parallel then the proportional assumption is not met.
2. Using time dependent variable in extended cox model
   The trick is to make the interaction between independent variables with survival time then see the significance value. The proportional assumption is met when the value
3. Using goodness of fittest.

The trick is to look at the value of p (Chi-square). If the value then the proportional assumption is met.

These three ways have advantages and disadvantages, for that a researcher should use at least two ways to test the proportional assumptions.

Candidate variable included in the interaction test is the independent variable that influences the survival time ( $p$ 0,25). Next test the interaction between independent variables by using likelihood ratio test. If thevalue included in the model. Cox (Cox Proportional Hazard) regression model is:

$$h(t) = h_0(t)exp\{\beta_1 x_1 + \beta_1 x_1 + \beta_i x_i\} \quad (2.12)$$

If the assumption is not met then the model used cox regression is recommended with time dependent covariate or extended cox model and can also use cox stratification model.

*D. Dengue Hemorrhagic Fever (DHF)*

Dengue Hemorrhagic Fever (DHF), a medical language called Dengue Hemorrhagic Fever (DHF), is a disease caused by dengue virus that is transmitted through the bite of Aedes Aegypti and Aedes Albopictus mosquitoes, which causes disruption of the capillary blood vessels and the blood clotting system, resulting in bleeding. The disease is found in many tropical regions

Digital Repository Universitas Jember

https://dx.doi.org/10.22161/ijaers.4.7.22 ISSN: 2349-6495(P) | 2456-1908(O)

such as Southeast Asia, India, Brazil, America, including all over Indonesia, except in places more than 1.000 meters above sea level.

Incubation period for 3 - 15 days since someone attacked by dengue virus, furthermore patient will show various signs and symptoms of dengue as follows:

1. High fever suddenly 2-7 days (38 - 40 degreesCelcius).
2. On examination torniquet test, it appears there larva (puspura)bleeding.
3. The presence of bleeding in the inner eye (conjunctiva), nosebleeds (epitaxis), defecation with the feces (mucous) mucus mixed with blood (melena), andothers.
4. Enlarged liver (hepatomegaly).
5. Blood pressure decreases causingshock.
6. On the laboratory examination (blood) day 3 - 7 there was a decrease in platelet count below 150.000/$mm^3$ (thrombocytopenia), an increase in hematocrit value above 20% of the normal value(hemoconcentration).
7. Occurrence of some accompanying clinical symptoms such as nausea, vomiting, decreased appetite (anorexia), abdominal pain, diarrhea, chills, seizures andheadaches.
8. Experiencing bleeding on the nose (nosebleeds) andgums.
9. Fever felt by the sufferer causing complaints of sore / pain in thejoints.
10. The appearance of red spots on the skin due to rupture of bloodvessels.

## III. METHODOLOGY

### A. ResearchData

Data used in this research is secondary data in the form of patient data of Dengue Hemorrhagic Fever (DHF) at Jember Clinic Hospital in September until December 2016 obtained from medical record of Jember Clinic Hospital as many as 100 patients

### B. ResearchVariabel

In this study the variables used are response variables and predictor variables, with the description as follows:

1. VariableResponse

Response variable in this research is survival time, that is time needed by patient to survive from time of early patient of DHF patient in hospital (start point) until patient stated recover (failure/end point). The response variable in this study is denoted by the letter T with time units in days.

2. Predictor Variables

Predictive variables in this study are data predicted to affect the survival time of DHF patients obtained from medical record data at Jember Clinic Hospital, namely:

a. Age ($X_1$)

The age variable is the age of DHF patients at the beginning of admission to hospitalization at Jember Clinic Hospital.

b. Gender ($X_2$)

The gender variable is the sex of DHF patients who are hospitalized in Jember Clinic Hospital, which is classified into two categories:

1 = Male

2 = Female

c. Hemoglobin($X_3$)

Hemoglobin variable that is hemoglobin level of DHF patients during undergoing hospitalization at Jember Clinic Hospital.

d. Trombosit/Platelets($X_4$)

Platelet variable is the number of thrombocyte of DHF patients during undergoing inpatient at Jember Clinic Hospital, which is divided into two categories, namely:

1 = below normal (<150.000 / mm3)

2 = normal (150,000 - 400,000 / mm3)

(Gandasoebrata, 2009)

e. Hematocrit($X_5$)

Hematocrit variable that is the percentage of hematocrit of DHF patients during undergoing inpatient at Jember Clinic Hospital.

### C. ResearchSteps

The steps taken in this research are as follows:

1. Conducting a descriptive statistical analysis to determine the characteristics of DHFpatients.
2. Make a Kaplan Meier Curve and perform a Log-Rank test on a categorical independent variable(predictor).
3. Conduct analysis of factors that affect the healing rate of DHFpatients.

## IV. RESULTS AND DISCUSSION

### Descriptive StatisticsAnalysis

The following is the result of descriptive statistical analysis on the characteristics of DHF patients during undergoing inpatient at Jember Clinic Hospital as many as 100 patients.

a. Survival Time, Age, Hemoglobin, Platelets andHematocrit

Descriptive statistical analysis of DHF patients in RSP Jember Clinic for survival time, age, hemoglobin, platelets and hematocrit are presented in Table 4.1. Especially for hemoglobin, thrombocyte and hematocrit are divided into two, namely at the time of admission and during the last medical record of each DHF patient.

Digital Repository Universitas Jember

*https://dx.doi.org/10.22161/ijaers.4.7.22* *ISSN: 2349-6495(P) | 2456-1908(O)*

*Table 4.1 Descriptive Statistics Analysis Time Survival, Age, Hemoglobin, Platelets and Hematocrit*

| Variabel | Min | | Max | | Mean | | Normal |
|---|---|---|---|---|---|---|---|
| Survival Time | 1 | | 9 | | 3,093 | | |
| Age | 1 | | 59 | | 19,13 | | |
| | Saat Awal Masuk Rawat Inap | | | Saat Rekam Medis Terakhir | | | |
| | Min | Max | Mean | Min | Max | Mean | |
| Hemoglobin | 10,40 | 18,10 | 13,55 | 9,80 | 17,30 | 12,94 | Male (14 – 18 gr/dL) Female (12 – 16 gr/dL) Children (10 – 16 gr/dL) |
| Trombosit | 8.000 | 363.000 | 103.750 | 19.000 | 325.000 | 131.520 | 150.000 – 400.000 sel/mm$^3$ |
| Hematokrit | 31,00 | 53,30 | 41,02 | 31,00 | 52,40 | 39,28 | 3 times Hemoglobin |

Table 4.1 provides information that the average length of hospital stay (survival time) is 3 days with the shortest hospitalization is 1 day and the longest is 9 days. While the patient age is 19 years on average, the youngest patient is one year old and the oldest is 59 years (patients aged 0 - 12 months are assumed to be 1 year old). DHF patients declared cured is indicated by the results of a medical examination under normal circumstances. Medical examination in question includes the examination of hemoglobin levels, platelets and hematocrit.

At baseline admission to hospitalization, the average hemoglobin of patients tended to be normal at 13.55 g/dL. But not all patients have normal hemoglobin levels. The lowest hemoglobin level was 10.40 g/dL and the highest hemoglobin level was 18.10 g/dL. While for platelets, the average platelet count of patients below normal is 103.750/mm$^3$ (normal platelets 150,000 - 400,000/mm$^3$). At baseline admission the lowest platelet count is 8,000 / mm3 and the highest platelet level is 363.000/mm$^3$. As for hematocrit, hematocrit level is said to be normal if the percentage is 3 times the hemoglobin level or according to Kusriastuti (2011), the normal percentage of hematocrit for children 33-38%, adult male 40-48% and adult female 37-43%. FromTable 4.1 the mean percentage of hematocrit level of DHF patients is 41.02%, it shows that the mean hematocrit level of DHF patients is normal, although not all patients have normal hematocrit level. At the beginning of admission hospitalization hematocrit lowest level 31.00% and highest 53.30%.

Table 4.1 also provides information on hemoglobin levels, platelets and hematocrit of DHF patients when the last medical record. At the time of the last medical record, the average patient's hemoglobin was lower when compared to the initial admission to hospitalization, which was 12.94 gr/dL, but still relatively normal. As for platelets, the average patient had elevated platelet levels when compared to early entry. According to Kusriastuti (2011), new patients can be discharged one of the conditions is if the platelet count is greater than50.000/mm3. From Table 4.1, the average platelet count of DHF patients was 131.520/mm$^3$, so the average patient's platelet level was normal. As for hematocrit, the average hematocrit level of patients is normal that is 39.28%. Thus, from Table 4.1 it can be seen that during the last medical record, hemoglobin, tombocyte and hematocrit levels of DHF patients were normal.

Gender

b. Data of in-patient DHF in Jember Clinic Hospital from September to December of 2016 were 100 patients, obtained from medical record of Jember Clinic Hospital, consist of 58 male patients and 42 female patients.

**Kaplan Meier Analysis and Log-RankTest**

The characteristics of survival time of DHF patients at Jember Clinic Hospital during inpatient care can be shown with Kaplan Meier's survival curve. Kaplan Meier's survival curve is a curve describing the probability of surviving DHF patients. In this study, the event studied was the recovery of dengue patients (patients had clinical improvement). Thus the Kaplan Meier curve explains the probability of the patient not recovering (not undergoing clinical improvement). Based on the duration of hospitalization of DHF patients obtained from medical record data of Jember Clinic Hospital, Kaplan Meier survival curve was obtained as in Figure4.1.

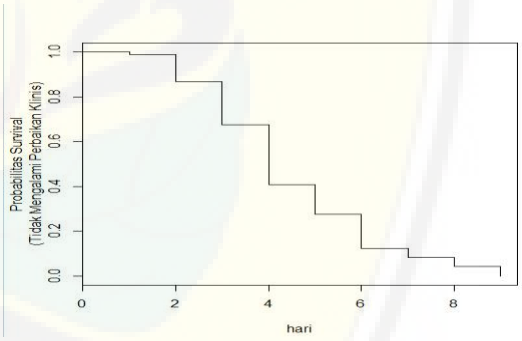


*Fig.4.1: Kaplan Meier Survival Curve Based Survival Time*

Figure 4.1 informs visually that the longer DHF patients undergo inpatient (t), the probability of a DHF patient not recovering (not undergoing clinical improvement) until the time t is less close to zero, meaning the longer the patient is hospitalized (the longer Get medical treatment) then the greater the probability of patients to recover (clinical improvement). Description of probability of healing DHF patients can be see in Table 4.2.

*Table.4.2: Probability of DHF Patients*

| *time* | *n.risk* | *n.event* | *survival* | *std.err* | *lower* | *upper* |
|---|---|---|---|---|---|---|
| 1 | 100 | 1 | 0,9900 | 0,00995 | 0,9707 | 1,0000 |
| 2 | 98 | 12 | 0,8688 | 0,03392 | 0,8048 | 0,938 |
| 3 | 85 | 19 | 0,6746 | 0,04728 | 0,5880 | 0,774 |

Digital Repository Universitas Jember

*https://dx.doi.org/10.22161/ijaers.4.7.22* *ISSN: 2349-6495(P) | 2456-1908(O)*

| | | | | | | |
|---|---|---|---|---|---|---|
| 4 | 66 | 26 | 0,4088 | 0,04967 | 0,3222 | 0,519 |
| 5 | 40 | 13 | 0,2760 | 0,04517 | 0,2002 | 0,380 |
| 6 | 27 | 15 | 0,1227 | 0,03316 | 0,0722 | 0,208 |
| 7 | 12 | 4 | 0,0818 | 0,02770 | 0,0421 | 0,159 |
| 8 | 8 | 4 | 0,0409 | 0,02002 | 0,0157 | 0,107 |
| 9 | 4 | 4 | 0,0000 | NaN | NA | NA |

In Table 4.2, time indicates days, n.risk indicates the number of patients diagnosed with DHF, n.event shows the number of dengue fever patients and survival indicates the probability of patients not recovering. Further explained about the characteristics of DHF patients in Jember Klinik Hospital based on sex which is suspected to affect Kaplan Meier survival curve, as in Figure 4.2.

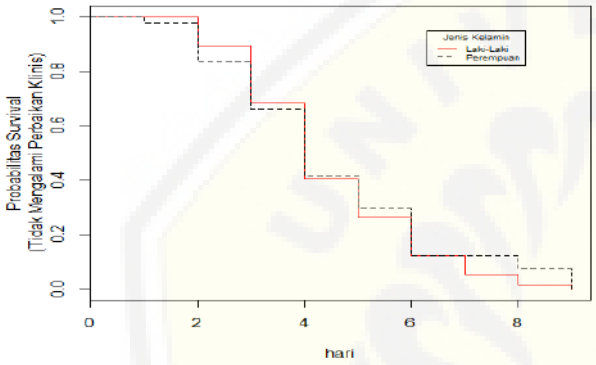


*Fig.4.2: Kaplan Meier Survival Curve By Gender*

Figure 4.2 shows a dotted line (black) more dominating above a straight line (red) indicating that the probability of not recovering female patients is greater than that of male patients, meaning that the survival time of the male gender is better than patients with female gender.

The next step is to do a log-rank test to find out the differences between the survival probability curves in Figure 4.2. From the log-rank test results obtained p-value value of 0,604. When compared with chi-square of 0.3 with degrees of freedom 1, it shows that the probability of survival of DBD patients of either male or female sex does not differ significant.

**Factors Affecting the Healing Rate of DHFPatients**

To know what factors influence the rate of healing of DHF patients, then modeling between response variables (survival time) and predictor variables used. Modeling with all predictor variables is shown in Table 4.3.

*Table.4.3: Parameter Significance Test*

| Variabel | Parameter Estimates | z | *P-value* |
|---|---|---|---|
| age | −0,0256 | −3,088 | 0,0020 |
| gender | −0,0548 | −0,232 | 0,8164 |
| hemoglobin | 0,1328 | 0,874 | 0,3819 |
| trombosit | 0,0000 | 3,524 | 0,0004 |
| hematokrit | −0,0444 | −0,935 | 0,3496 |

Table 4.4 The value of Akaike's Information Criterion (AIC)

| Variabel | AIC |
|---|---|
| All Independent Variable | 882,1166 |
| Age, Gender and Hematokrit | 891,0849 |
| Gender, Trombosit and Hematokrit | 888,4697 |
| Age, Hemoglobin and Trombosit | 878,9539 |
| **Age andTrombosit** | **877,0109** |

Table 4.4 shows that the smallest AIC value is 877,0109 when doing cox proportional hazard regression without gender, hemoglobin and hematocrit variables. So the best model is obtained by doing regression cox proportional hazard using age and trombosit variables.

Estimation of the parameters for each variable based on the best model can be seen in Table 4.5.

Table 4.5 Estimation of Cox Regression Parameters in the Best Model

| Variabel | Parameter Estimates | z | *P-value* |
|---|---|---|---|
| age | −0,0256500 | −3,5460000 | 0,0003910 |
| trombosit | 0,0000054 | 3,6480000 | 0,0002640 |

So, the best model can be written as follows:

$$\hat{h}(t) = exp\,(\hat{\beta}'x)$$
$$= exp(-0{,}02565\,age + 0{,}0000054\,trombosit)$$
$$= exp(-0{,}02565X_1 + 0{,}0000054\,X_4)$$

## V. CONCLUSION

Descriptive statistical analysis showed that from 100 DHF patients admitted to RSP Jember Klinik, the average clinical condition improved within 3 days, with 53 patients belonging to productive age group (15- 59 years) and 58 male sexually transmitted Men. At the time of admission was 83 patients whose platelet count was below normal (less than 150,000 / mm3). While during the last medical record there was a decrease in the number of patients with under-normal platelets to 64 patients. The mean hemoglobin and thrombocyte levels of DHF patients at baseline were 13.55 g / dL and 41.02% respectively, while during the last medical record the mean hemoglobin and platelet counts were 12.94 grams / DL and 39.28%. While the factors that significantly influence the rate of improvement of clinical condition of DHF patients are age andplatelets.

Digital Repository Universitas Jember

*https://dx.doi.org/10.22161/ijaers.4.7.22* *ISSN: 2349-6495(P) | 2456-1908(O)*

Suggestions that can be given from the results of this study, for the medical team is expected to control the factors that affect the rate of improvement of the clinical condition of DHF patients. While for further research, it is suggested to add factors that theoretically can influence the rate of improvement of clinical condition of DHF patients and can develop survival analysis in other fields, such as in industry or education.

## REFERENCES

[1] Collett, D. (2003). *Modelling Survival Data in Medical Research.* 2nd edision . Chapman and Hall/CRC, London.

[2] Darmowandowo, W. (2006). *Continuing Education*. Divisi Tropik & Infeksi Bagian Ilmu Kesehatan Anak FK Unair RSU Dr. Soetomo,Surabaya.

[3] Gandasoebrata, R. (2009). *Penuntun Laboratorium Klinik*. Dian Rakyat, Jakarta.

[4] Kleinbaum, D.G. & Klein, M. (2012). *Survival Analysis: A Self-Learning Text*. Springer, London. Kusriastuti, R. (2011). *Modul Pengendalian Demam Berdarah Dengue*. Kementrian Kesehatan RI, Jakarta. Le, C.T. (1997). *Applied Survival Analysis.* John Wiley and Sons Inc, New York.

[5] Lee, E.T. & Wang, J.W. (2003). *Statistical Methods for Survival Data Analysis.* 3rd edision. John Wiley and Sons Inc, Canada.

[6] Mufidah, A.S. (2016). Analisis Survival Pada Pasien Demam Berdarah *Dengue* (DBD) di RSU Haji Surabaya Menggunakan Model Regresi Weibull. *Jurnal Sains dan Seni ITS*, 5(2), p. 2337-3520.

[7] Nisa', S.F. (2012). Analisis Survival Dengan Pendekatan Multivariate Adaptive Regression Splines Pada Kasus Demam Berdarah Dengue (DBD). *Jurnal Sains dan Seni ITS*, 1(1), p.318-323.

[8] Sari, N. (2011). Aplikasi Regresi Cox Propotional Hazard Pada Analisis Kesintasan Dan Identifikasi Faktor Resiko. *Skripsi*. Departemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sumatra Utara, Medan.

[9] Omurlu, I.K., Ture, M. & Tokatli, F. (2009). The Comparisons of Random Survival Forest and Cox Regression analysis with Simulation and an Application Related to Breast Cancer. *Journal International of Expert Systems with Applications*, 36, p. 8582-8588.

[10] WHO. (2009). *Dengue, Guidelines For Diagnosis, Treatment, Prevention, and Control.* Special Programme for Research and Training in Tropical Disease (TDR), France.