International Journal of Contemporary Issues in Social Sciences
Volume 3, Issue 2, 2024 ISSN(P):2959-3808 | 2959-2461

PIECE WISE KAPLAN-MEIER SURVIVAL FUNCTION FOR DEPENDENT DISEASES

Qamruz Zaman[*1], Danish Wasim[2], Syed Habib Shah[3], Neelam[4], Atta Ur Rehman[5], Mohib Ullah Khan[6], Asma Qureshi[7]

[*1,4,5,6,7]Department of Statistics University of Peshawar, Pakistan;
[2]Department of Statistics, Government Superior Science College Peshawar, Pakistan;
[3]Institute of Numerical Sciences Kohat University of Science and Technology

[*1]qamruzzaman@uop.edu.pk; [2]danishwasim.std@icp.edu.pk; [3]habibshah@kust.edu.pk; [4]neelam_arbab@uop.edu.pk; [5]attabaqi11@gmail.com ; [6]mohibullahkhan578@gmail.com ; [7]kamikhan87993@gmail.com

Corresponding Author: *

Received: 11 May, 2024 **Revised:** 13, June, 2024 **Accepted:** 24, June, 2024 **Published:** 29 June, 2024

ABSTRACT
One of the assumptions of censoring in Kaplan-Meier Survival Function is independency i.e. disease under study must be independent but in real world, the situation is entirely different. In this case relative risk factors i.e. dependency plays an important role. To overcome this problem a new estimator called the Piece wise Kaplan-Meier Survival Function was developed and discussed in detail. Similarly, for measuring the variation, new variance estimator based on the Piece wise Kaplan-Meier Survival Function was introduced. Similarly, with the help of Greenwood Variance estimator and new variance estimator 95 Confidence Intervals are constructed. Results of the analysis showed that the Kaplan-Meier gives the overestimate results as compared to the new procedure. New procedures consider the competing risk factor and give more satisfactory results.

1. INTRODUCTION
Survival analysis is the backbone of biostatistics. In survival study, we generally refer to the time variable as survival time, because it gives the time that an individual has" survived" over some follow up period. We also normally refer to the occasion as a failure, because the event of interest usually is death However, survival time might be "an ideal opportunity to come back to work after an elective surgery," in which case failure is a positive event. In survival study, there is an inadequate data on survival time of a patient. The response variable in survival analysis is the time till the event happens [1-2].
This incomplete data about the patient survival time known as censoring. Censoring is accessible when we have some data about a subject's occasion time, yet we don't identify about the specific occasion time. For the analysis strategies we will talk about to be effective, censoring system must be autonomous of the survival component. There are three types of censoring. Censor observation can be right, left or interval censor observations [3]. The most famous approach of survival analysis is the nonparametric modelling.

Nonparametric Modeling
The standard non-parametric Model approaches survey ordinal scale of measurement i.e. ordered/rank data are not directly appropriate in occurrence of censored observations. Particularly, a histogram or an empirical cumulative distribution function cannot be used with censored information. As a method of estimation survivor function used Kaplan-Meier (KM) survival function is a nonparametric modelling technique. These do exclude the impacts of covariates and requires just a requesting of the chance to failure (or censoring).

1. Kaplan-Meier (KM) Estimator
Kaplan-Meier is the product limit estimator introduced by (Kaplan and Meier) in 1958.This is the

most popular technique in nonparametric procedures [4].KM is considered as the backbone of survival analysis. The function is defined as Let $X1, X2, \ldots, Xn$ be the event times having distribution F(x) and $C_1, C_2, C_3, \ldots, C_n$ denote the censoring time having distribution G(x).Then $Ti = min\{Xi, Ci\}$ be the observed survival time. Let n_i and d_i denote the number of persons at risk at time i and the number of deaths respectively. The KM estimator is defined as

$$S^{\wedge}(t) = \prod_{i=1}^{n}\left(1 - \frac{d_i}{n_i}\right)$$

The Kaplan-Meier estimator is the most simple method of recording the survival after some time regardless of every one of these troubles related with subjects or conditions. The survival curve can be made accepting different conditions. It includes recording of probabilities of occurrence of event at one point of time and repeating these succeeding probabilities by any previous calculated probabilities to get the last estimator. This can be determined for two sets of subjects and their statistical difference in the survivals.

The KM technique is used in univariate case and is based on the independent censoring assumption. The Greenwood [5] variance estimator formula is used to calculate the variance of K-M function is given by

$$\hat{V}ar_{Gr}\left[\hat{S}(t)\right] = \hat{S}(t)^2\left(\sum_{i=1}^{n}\frac{d_i}{n_i(n_i - d_i)}\right)$$

1.1 Greenwood's Variance Estimator

Greenwood's formula sets a s error on the K.M estimator utilizing the delta technique at a specific time t through a failure, Suppose N_t be the quantity of subjects on test "at time t−," that is, not long before time t. The likelihood of making due from t− to t+ is evaluated as $/Nt$, where Xt is the number who make due from t− to t+. The Kaplan-Meir estimator is

$$T \to \prod_{t<T}\frac{X_t}{N_t}X_t$$

X_t displayed as independent binomial $B(N_t, P_t)$ variables. Independence is clearly off-base, chance of failure times is overlooked, and concealed irregularity—nonattendance of failure between viewed failure times—is disregarded. Limited example, this is nothing but bad. Asymptotically, under conditions, may be fine. Anyway, we make the displaying assumptions.

Suppose $\hat{K}$ represents the product with K expected value.

$$\frac{\hat{K}}{K} = \prod\frac{X_t}{N_tP_t} = \prod\left(1 + \frac{X_t - N_tP_t}{N_tP_t}\right) \approx 1 + \sum\frac{X_t - N_tP_t}{N_tP_t}$$

If N_tP_t are for the maximum part. So

$$var\left(\frac{\hat{K}}{K}\right) \approx \sum\frac{1-P_t}{N_tP_t} \approx \sum\frac{1-\hat{P}_t}{N_t\hat{P}_t}$$

P_t represents the survival probability, where as $\hat{P}_t$ shows the estimated survival probability.

Therefore,

$$var\,(\hat{K}) \approx K^2\sum\frac{1-\hat{P}_t}{N_t\hat{P}_t} \approx \hat{K}^2\sum\frac{1-\hat{P}_t}{N_t\hat{P}_t}$$

Greenwood's Formula.

$$N(1, \sum\frac{1-\hat{P}_t}{N_t\hat{P}_t})$$

Confidence Intervals.

$$\log\hat{K} = \log K + \sum\log\left(1 + \frac{X_t - N_tp_t}{N_tp_t}\right) \approx \log K + \sum\frac{X_t - N_tp_t}{N_tp_t}(\quad)(\quad)$$

So

$$var(\log\widehat{K}) \approx \sum\frac{1-\hat{P}_t}{N_t\hat{p}_t}$$

1.2 Need of the Study

The techniques of survival analysis are based on the assumptions that diseases are independent i.e. if a person has both weight and cholesterol problems. According to the non-parametric assumption, these are independent and there is no relation between them. In real world, there is a strong relation between them. So far no method is available in literature to deal with the problem of dependent diseases. The Kaplan-Meier survival function and the other extensions of the function are based on the independent assumption. The purpose of this research is to deal with problem.

1.3 Aims and Objectives

Following are the main aims and objectives of the study.

- To develop a Kaplan-Meier survival function for dependent cases
- To develop the variance of the new estimator
- To assess the performance of the estimator using real data set

2. METHODOLOGY

PIECE WISE KAPLAN-MEIER SURVIVAL FUNCTION AND ITS VARIANCE

A Piece wise Kaplan-Meier Estimator and its variance estimator will be discussed in this chapter to overcome the problem of dependent disease. The concept of survival analysis is useless without constructing the survival probabilities using Kaplan-Meier (KM) technique. For the calculation of confidence interval of KM estimator, different variance estimators are used proposed by different researchers. Some of them are Zaman et.al. [5], Zaman et.al. [6], Zaman et.al. [7], Zaman et.al. [8], Borkowf [8]. Zhao [9], Peto [10] and Greenwood [11]. But all these methods based on the assumption of independent censoring.

To overcome the problem , the following procedure is adopted.

Let $d_{11}, d_{12}, \ldots, d_{1n}$ denote the event time of disease 1 and let $C_{11}, C_{,12}, \ldots, C_{1n}$ denote the censoring time due to disease1. $d_{21}, d_{22}, \ldots, d_{2n(n-1)}$ and $C_{21}, C_{22}, \ldots, C_{2(n-2)}$ denote the event and censoring times due to disease 2. Similarly, $d_{31}, d_{32}, \ldots, d_{3(n-2)}$ and $C_{31}, C_{32}, \ldots, C_{3(n-3)}$ are used for event and censoring time for disease 3 and so on. The information is summarized in the following table 1.

Table 1 : Layout for the Piecewise Survival Function

	Disease 1		Disease 2		Disease 3		And so on.
Time	Event	Censoring	Event	Censoring	Event	Censoring	
T_1	d_{11}	C_{11}	d_{21}	C_{21}	d_{31}	C_{31}	
T_2	d_{12}	C_{12}	d_{22}	C_{22}	d_{32}	C_{32}	
T_3	d_{13}	C_{13}	d_{23}	C_{23}	d_{33}	C_{33}	
:	:	:	:	:	:	:	
T_n	d_{1n}	C_{1n}	$d_{2(n-1)}$	$C_{2(n-2)}$	$d_{3(n-2)}$	$C_{3(n-3)}$	

On the basis of the above information an extension of KM is developed. The new method which is called the Piece wise Kaplan-Meier Survival Function gives us the survival probabilities each disease. Similarly, for the new developed survival function, the variance estimator is developed using the commutative property of addition. For detail illustration,

Let t_1, t_2, … t_{11} denote the observed times corresponding to $d_1, d_2, \ldots, d_{11}$ deaths and $c_1, c_2, \ldots, c_{11}$are the observed censoring times respectively. Suppose the censoring is not independent and the data consists of two diseases namely disease I and disease II. Out of these 11 observed times, only 4 deaths namely at d_{11}, d_{14}, d_{16} and d_{110} and c_{12}, c_{15} and c_{19} are independently censored i.e leave the study without any reason along with the number of persons at risks are n_{11}, n_{14}, n_{16} and n_{110}. After the follow-up for period of time, it is observed that 3 deaths/events occurred due to disease II at d_{23}, d_{28} and d_{211} along with one censored observation c_{27} and the number of persons at risks are n_{23}, n_{28} and n_{211} respectively. Survival probability of disease I is calculated, which changes for each event and remains constant for others. The products of theses give the survival probability $S_1(t)$for disease I, similar procedure is adopted for disease II. The products of these probabilities namely; $S_1(t)$ and $S_2(t)$ give the Piecewise Kaplan-.Meier Survival Function. This function gives the probabilities, which must be smaller than the conventional probabilities. The reason is that the KM method considered only one disease ignoring severity of the other disease, which obviously decrease the life expectancy and as a result the survival probabilities decreases. The procedure can also be explained through the following flowchart/ Figure 1.

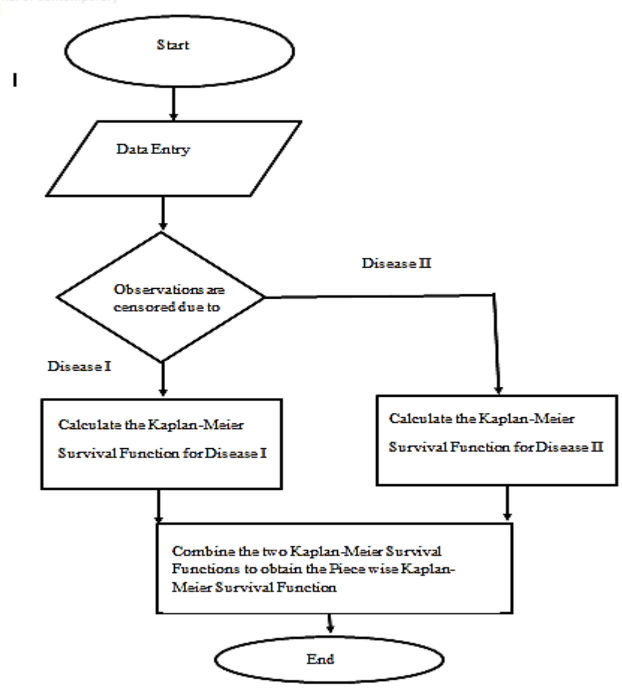
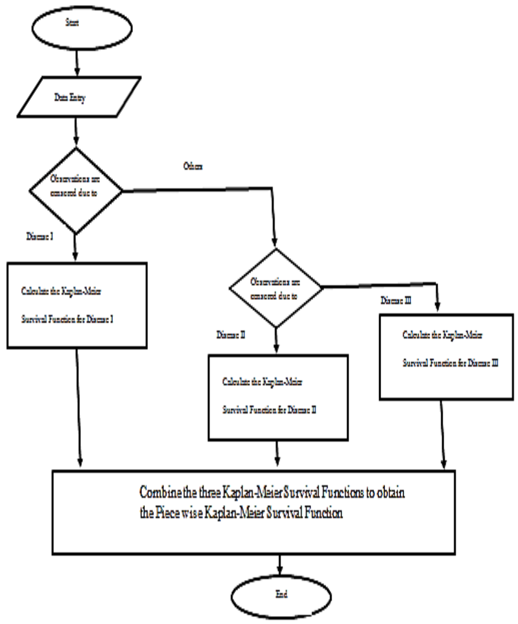


Figure1: Illustration of the Piecewise Kaplan-Meier Survival Function through flowchart.



The procedure can also be extended two three or more diseases too. This can be done as in Figure: 2.

Figure 2: Piece wise Kaplan-Meier Survival Function for three diseases

2.1 Piece wise Variance estimator

For the purpose of illustration the table 2 is shrinked into the following form

Table 2 : Illustration of the new variance estimator through eleven observations.

			Disease I					Disease I						
Time	d_{1i}	c_{1i}	n_{1i}	$S_1(T)$	$\frac{d_{1i}}{n_{1i}\,(n_{1i}-d_{1i})}$	Sum1	GW_v	d_{2i}	c_{2i}	n_{2i}	$\frac{d_{2i}}{n_{2i}\,(n_{2i}-d_{2i})}$	Sum2	$S_c(t)^2$	Piece wise Variance Estimator
t_1	d_{11}		n_{11}	$S_{11}(T)$	$\frac{d_{11}}{n_{11}\,(n_{11}-d_{11})}$	Sum11= $\frac{d_{11}}{n_{11}\,(n_{11}-d_{11})}$	V_{G1}					Sum21	$S_{c1}(T)^2$	V_1=Sc11$(T)^2$*(Sum11+ Sum21)
t_2		c_{12}		$S_{12}(t)$		Sum12=sum11+0	V_{G2}					Sum22= Sum21+..	$S_{c2}(T)^2$	$V_2=S_{c12}(T)^2$*(Sum11+ Sum21)
t_3				$S_{13}(t)$		Sum13	V_{G3}	d_{23}		n_{23}	$\frac{d_{23}}{n_{23}\,(n_{23}-d_{23})}$	Sum23= Sum22+ $\frac{d_{23}}{n_{23}\,(n_{23}-d_{23})}$	$S_{c3}(T)2$	$V_3=S_{c13}(T)2$*(Sum11+ Sum21)
t_4	d_{14}		n_{14}	$S_{14}(t)$	$\frac{d_{14}}{n_{14}(n_{14}-d_{14})}$	Sum14=sum13+ $\frac{d_{14}}{n_{14}(n_{14}-d_{14})}$	V_{G4}					Sum24	$S_{c4}(T)^2$	$V_4=S_{c14}(T)^2$*(Sum11+ Sum21)
t_5		c_{15}		$S_{15}(t)$		Sum15	V_{G5}					Sum25	$S_{c5}(T)^2$	$V_5=S_{c15}(T)^2$*(Sum11+ Sum21)
t_6	d_{16}		n_{16}	$S_{16}(t)$	$\frac{d_{16}}{n_{16}(n_{16}-d_{16})}$	Sum16	V_{G6}					Sum26	$S_{c6}(T)^2$	$V_6=S_{c16}(T)^2$*(Sum11+ Sum21)
t_7				$S_{17}(t)$		Sum17	V_{G7}		c_{27}			Sum27	$S_{c7}(T)^2$	$V_7=S_{c17}(T)^2$*(Sum11+ Sum21)
t_8				$S_{18}(t)$		Sum18	V_{G8}	d_{28}			$\frac{d_{28}}{n_{28}\,(n_{28}-d_{28})}$	Sum28	$S_{c8}(T)^2$	$V_8=S_{c18}(T)^2$*(Sum11+ Sum21)
t_9		c_{19}		$S_{19}(t)$		Sum19	V_{G9}					Sum29	$S_{c9}(T)^2$	$V_9=S_{c19}(T)^2$*(Sum11+ Sum21)
t_{10}	d_{110}		n_{110}	$S_{110}(t)$	$\frac{d_{110}}{n_{110}(n_{110}-d_{110})}$	Sum10	V_{G10}					Sum210	$S_{c10}(T)^2$	$V_{10}=S_{c110}(T)^2$*(Sum11+ Sum21)
t_{11}				$S_{111}(t)$		Sum11	V_{G11}	d_{211}		n_{211}	$\frac{d_{211}}{n_{211}\,(n_{211}-d_{211})}$	Sum211	$S_{c111}(T)^2$	$V_{11}=S_{c111}(T)^2$*(Sum11+ Sum21)

For the calculation of the Piece wise variance estimator in case of two diseases, the calculations are divided into fifteen columns. For the calculations of variance, the ratio of deaths to the product of number of persons at risk into the difference of the number of Persons at risks to the number of deaths at that time was summarized into two columns. Column VI is used for the calculations of Greenwood Variance estimator and column XII for the Piecewise variance estimator. Column VII gives the cumulative sum of column VI. Similarly, the cumulative sum of disease II is written in column XIII. The square of Piecewise Survival Function is given in column XIV. The last column gives the variance of Piecewise Estimator, which is equal to the product of the square of piecewise Survival Function into the sum of columns VII and XIII. This can be easily explained through the following flowchart/Figure 3.

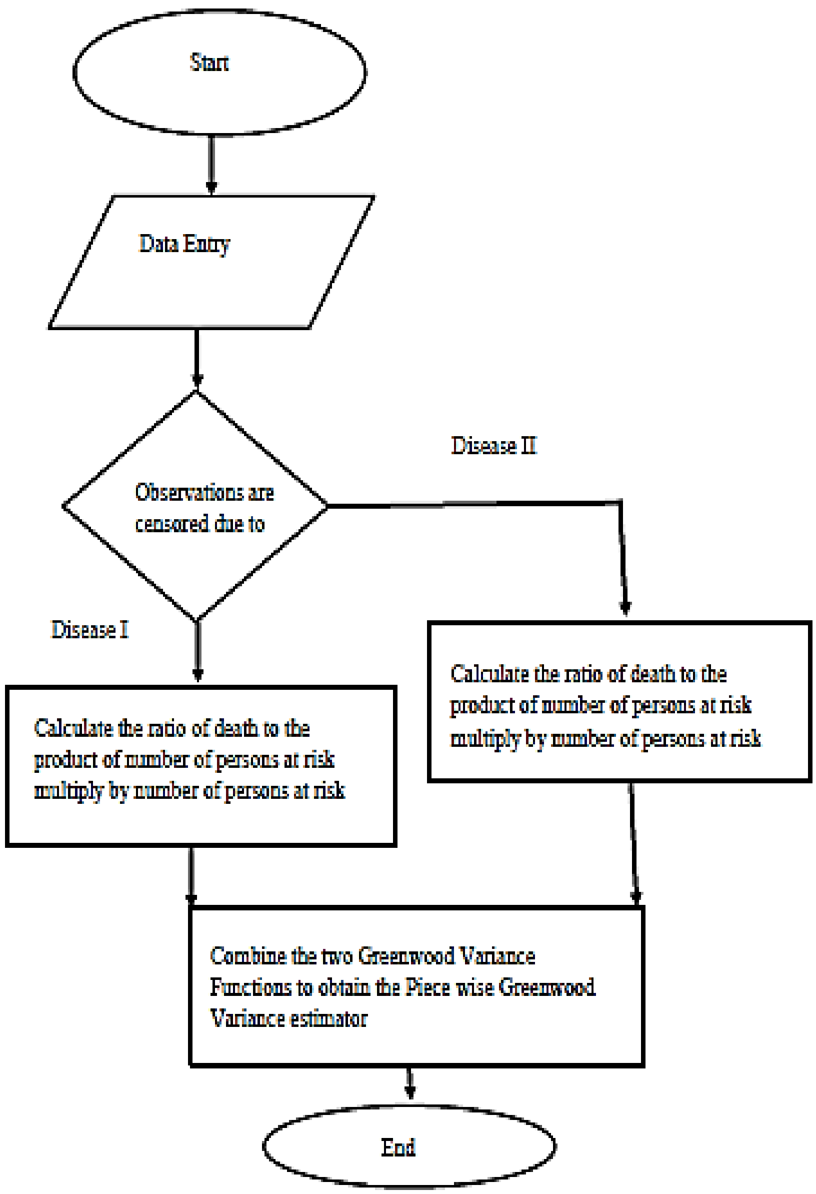


Figure 3: Illustration of the Piecewise Variance estimator through flowchart

The procedure can also be extended two three or more diseases too. This can be done as in Figure: 4.

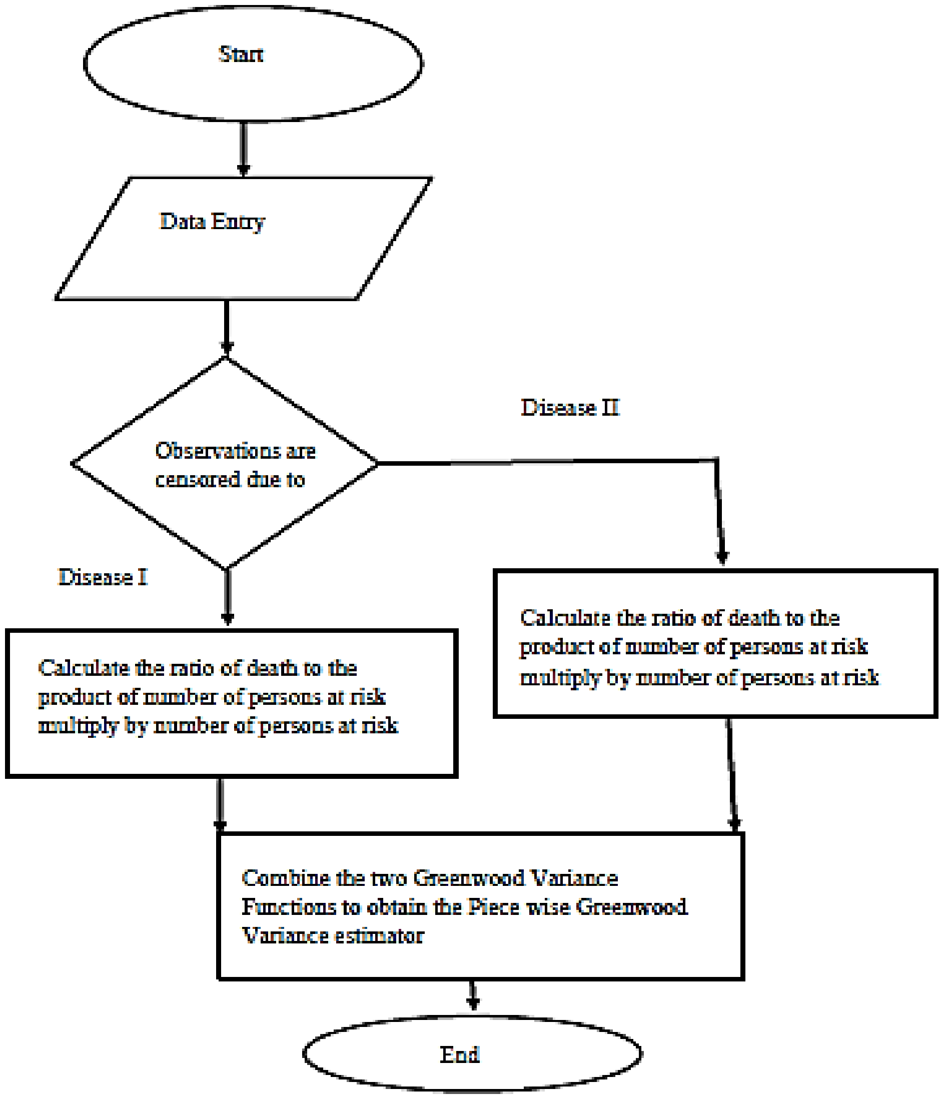


Figure 4: Illustration of the Piecewise Variance estimator for three diseases

REAL DATA ANALYSIS

A data of 119 thalassemia patients was collected during April 2019 to September 2019 from Fatimid Foundation [12] Peshawar branch, with the help of staff members, doctors, patients, medical records and parents. The major disease of interest was the thalassemia but after detailed study of the disease, it came out that the disease creates the heart problem i.e. the disease is not independent but related with heart disease too. So, the competing risk plays an important role in this case.

5 DESCRIPTION OF 119 PATIENTS

Fatimid Foundation, A symbol of hopes Pakistan. A Charitable Organization. Established **1981**.

For the analysis purpose, the data from the Fatimid Foundation was collected based on 119 patients.

The complete data is summarized below (Table 3):

Table 3: Thalassemia Data Set

t	d_1	c_1	d_2	c_2	t	d_1	c_1	d_2	c_2	t	d_1	c_1	d_2	c_2
12	1		0	0	15+	0	1	0	0	17	1		0	0
19	0		1		22	1		0	0	24	1		0	0
25	0		0	1	27	1		0	0	29+	0	1	0	0
30	0		1		33	1		0	0	36	1		0	0
37	1		0	0	38	0		0	1	39+	0	1	0	0
42	0		1		44	1		0	0	47	1		0	0
49+	0	1	0	0	53	0		1	0	57	1		0	0
59	1		0	0	67	0		0	1	69	1		0	0
77+	0	1	0	0	78+	1	1	0	0	79	0		1	0
80	1		0	0	82	1		0	0	83+	0	1	0	0
85	0		0	1	86	1		0	0	87	1		1	0
88+	0	1	0	0	90	0		1	0	91	0			1
92	1		1	0	94	1		0	1	95	1		0	0
97	1		0	0	98+	0	1	0	0	99	0		1	0
100	1	1	0	0	102	0		0	1	103	1		0	0
105	0		1	0	107	1		1	0	109	1		0	0
110	1		0	0	111	0		0	1	112	1	1	1	1
116	1		0	0	117	1		0	0	119	1		0	0
121		1	0	0	123	1		1	0	125		1	0	0
127			0	1	128	1		0	0	129		1	0	0
130			1	0	132	1		0	0	134			0	1
136	1		0	0	137	1		0	0	139	1		0	0
141	1		0	0	143		1	0	0	145			1	1
146	1		0	0	149		1	0	1	150	1		1	0
153		1	0	0	157	1		0	0	159			1	0
160	1		0	0	162		1	0	0	163		1	0	0
165	1		1	0	168			0	1	169		1	0	0
170	1		1	0	172		1	0	0	175	1		0	0
176			1	0	178		1	0	1	179	1		0	0
180		1	0	0	182	1		0	0	184	1		0	0
185	1		0	0	187		1	0	0	188			1	1
189	1		0	0	193		1	0	0	195	1		0	0
197		1	0	0	199			1	0	204		1	0	0
223	1		0	0	237			1	1					

In table 3, t denotes the time of study period in months, d1,c1 and n1 denote the number of deaths, number of censored and number of persons at risks due to thalassemia. Similarly, d2, c2 and n2 symbols are used for the number of deaths, number of censored and number of persons at risks due to heart problem. The minimum observed time is 12 months and the maximum time is 223 months. Out of 119 patients, 52 deaths were observed due to thalassemia and 27 patients were censored not due to heart problem but may be they moved to other places or due to some other reasons. Out of 119 observed patients, 40 were observed to be heart patients. Further these 40 divided into 23 deaths due to heat problem and 17 were censored. Table. 3 give the survival probabilities of K-M function and Piecewise Survival Function.

Table. 4 Survival probabilities of K-M function and Piecewise Survival Function.

Time	Kaplan-Meier Survival Function	Piece wise Kaplan-Meier Survival Function
12	0.9916	0.9916
15	0.9916	0.9916
17	0.9831	0.9831
19	0.9831	0.9585
22	0.9746	0.9502
24	0.9660	0.9419
25	0.9660	0.9419
27	0.9574	0.9335
29	0.9574	0.9335
30	0.9574	0.9089
33	0.9486	0.9006
36	0.9398	0.8922
37	0.9310	0.8839
38	0.9310	0.8839
39	0.9310	0.8839
42	0.9310	0.8593
44	0.9220	0.8510
47	0.9130	0.8426
49	0.9130	0.8426
53	0.9130	0.8186
57	0.9037	0.8103
59	0.8945	0.8020
67	0.8945	0.8020
69	0.8852	0.7937
77	0.8852	0.7937
78	0.8758	0.7852
79	0.8758	0.7614
80	0.8662	0.7531
82	0.8565	0.7447
83	0.8565	0.7447
85	0.8565	0.7447
86	0.8467	0.7361
87	0.8369	0.7041
88	0.8369	0.7041
90	0.8369	0.6806
91	0.8369	0.6806
92	0.8265	0.6482
94	0.8161	0.6400
95	0.8055	0.6317
97	0.7949	0.6234
98	0.7949	0.6234
99	0.7949	0.5994
100	0.7840	0.5912
102	0.7840	0.5912
103	0.7728	0.5828
105	0.7728	0.5585
107	0.7614	0.5263
109	0.7499	0.5184
110	0.7383	0.5104
111	0.7383	0.5104
112	0.7266	0.4784
116	0.7143	0.4703
117	0.7020	0.4622
119	0.6897	0.4541
121	0.6897	0.4541
123	0.6771	0.4223
125	0.6771	0.4223
127	0.6771	0.4223
128	0.6639	0.4141
129	0.6639	0.4141
130	0.6639	0.3897
132	0.6500	0.3816
134	0.6500	0.3816
136	0.6359	0.3733
137	0.6218	0.3650
139	0.6076	0.3567
141	0.5935	0.3484
143	0.5935	0.3484
145	0.5935	0.3252

146	0.5783	0.3168
149	0.5783	0.3168
150	0.5622	0.2824
153	0.5622	0.2824
157	0.5452	0.2738
159	0.5452	0.2489
160	0.5276	0.2409
162	0.5276	0.2409
163	0.5276	0.2409
165	0.5088	0.2091
168	0.5088	0.2091
169	0.5088	0.2091
170	0.4876	0.1753
172	0.4876	0.1753
175	0.4643	0.1670
176	0.4643	0.1431
178	0.4643	0.1431
179	0.4370	0.1347
180	0.4370	0.1347
182	0.4079	0.1257
184	0.3788	0.1167
185	0.3496	0.1077
187	0.3496	0.1077
188	0.3496	0.0862
189	0.3108	0.0766
193	0.3108	0.0766
195	0.2664	0.0657
197	0.2664	0.0657
199	0.2664	0.0438
204	0.2664	0.0438
223	0.1776	0.0292
237	0.1776	0.0146

Table 4. can be summarized into the following figure 5.

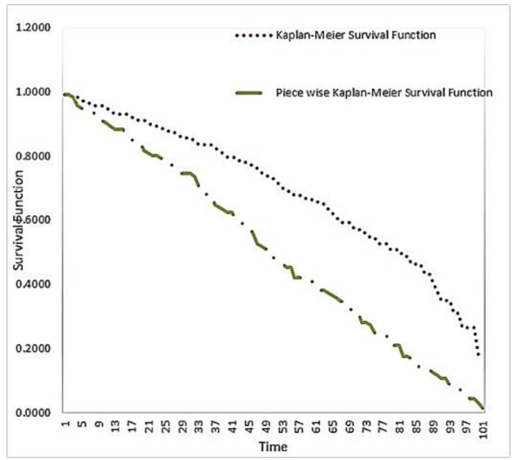


Figure. 5 Survival probabilities of K-M function and Piecewise Survival Function.

From the above table and figure, it revealed that without considering the competing risk factor, the conventional survival function gives the overestimate the results. Piecewise Kaplan-Meier Survival Function overcome this problem and considers heart disease too. The performance of new method is more satisfactory than the K-M procedure. The complete description of the data is analyzed in appendix- A. Table. 4 give the comparison of Greenwood Variance estimates of K-M function and Piecewise Survival Function Variance estimates. Graphical comparison of the two methods is summarized in Figure 5. Greenwood variance estimates the increasing pattern, giving zero weight to disease II, while the new method importance to each and every noted observation, it gives the inverse tub i.e. it increases in the start reaches to the peak and then decreases. Disease II changes the pattern of the variance and gives more satisfactory results.

Table. 5 Comparison of Greenwood Variance estimates of K-M function and Piecewise Survival Function Variance estimates.

Time	Greenwood Variance	Piecewise Variance
12	0.00007	7.00231E-05
15	0.00007	7.00231E-05
17	0.00014	0.00014
19	0.00014	0.00072
22	0.00021	0.00078
24	0.00028	0.00083
25	0.00028	0.00083

27	0.00035	0.00089
29	0.00035	0.00089
30	0.00035	0.00143
33	0.00042	0.00147
36	0.00049	0.00151
37	0.00055	0.00156
38	0.00055	0.00156
39	0.00055	0.00156
42	0.00055	0.00206
44	0.00062	0.00209
47	0.00069	0.00211
49	0.00069	0.00211
53	0.00069	0.00256
57	0.00076	0.00257
59	0.00083	0.00259
67	0.00083	0.00259
69	0.00090	0.00260
77	0.00090	0.00260
78	0.00097	0.00262
79	0.00097	0.00301
80	0.00104	0.00302
82	0.00111	0.00302
83	0.00111	0.00302
85	0.00111	0.00302
86	0.00118	0.00302
87	0.00125	0.00337
88	0.00125	0.00337
90	0.00125	0.00368
91	0.00125	0.00368
92	0.00132	0.00396
94	0.00140	0.00392
95	0.00147	0.00389
97	0.00154	0.00386
98	0.00154	0.00386
99	0.00154	0.00412
100	0.00162	0.00407
102	0.00162	0.00407
103	0.00170	0.00403
105	0.00170	0.00426
107	0.00177	0.00440
109	0.00185	0.00433

110	0.00193	0.00426
111	0.00193	0.00426
112	0.00200	0.00434
116	0.00208	0.00426
117	0.00216	0.00418
119	0.00223	0.00410
121	0.00223	0.00410
123	0.00231	0.00413
125	0.00231	0.00413
127	0.00231	0.00413
128	0.00239	0.00404
129	0.00239	0.00404
130	0.00239	0.00413
132	0.00248	0.00403
134	0.00248	0.00403
136	0.00257	0.00392
137	0.00265	0.00382
139	0.00273	0.00371
141	0.00280	0.00361
143	0.00280	0.00361
145	0.00280	0.00365
146	0.00288	0.00353
149	0.00288	0.00353
150	0.00297	0.00347
153	0.00297	0.00347
157	0.00308	0.00333
159	0.00308	0.00332
160	0.00318	0.00317
162	0.00318	0.00317
163	0.00318	0.00317
165	0.00330	0.00293
168	0.00330	0.00293
169	0.00330	0.00293
170	0.00346	0.00267
172	0.00346	0.00267
175	0.00365	0.00248
176	0.00365	0.00231
178	0.00365	0.00231
179	0.00394	0.00212
180	0.00394	0.00212
182	0.00422	0.00192

184	0.00443	0.00173
185	0.00456	0.00155
187	0.00456	0.00155
188	0.00456	0.00136
189	0.00494	0.00116
193	0.00494	0.00116
195	0.00532	0.00095
197	0.00532	0.00095
199	0.00532	0.00074
204	0.00532	0.00074
223	0.00762	0.00047
237	0.00762	0.00022

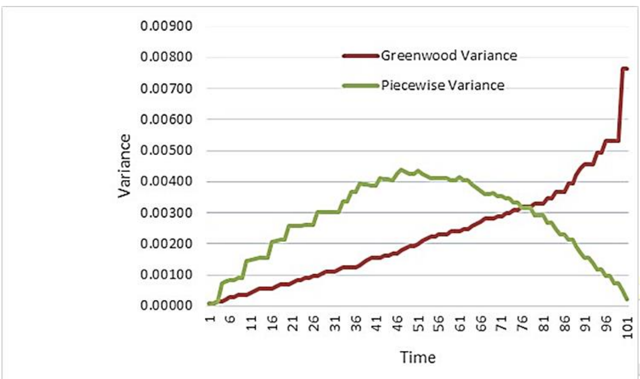


Figure 6: Comparison of the Greenwood Variance estimates and the Piecewise Kaplan-Meier Variance Estimates

CONCLUSIONS

The present study was based on exploring the phenomena of dependent diseases. As one of the assumptions of censoring in Kaplan-Meier Survival Function is independency i.e. disease under study must be independent but in real world, it is observed that most of the diseases are dependent. e.g. Overweight causes cholesterol, which is the main reason of high blood pressure and heart attack. In this research the attempt was made to develop the simple survival estimator and its variance estimator to overcome the problem of competing risk factor. The Piecewise Kaplan-Meier Estimator and Piecewise Estimator was developed. New methods are also described through the flowcharts.

For the comparison purpose detailed information based on 119 thalassemia patients was collected from the Fatimid Foundation Peshawar. Along with the major disease of interest, a competing risk i.e. heart problem was also considered. The complete data set was analyzed into three steps to check the overall performance of the new methods. Similarly, with the help of Greenwood Variance estimator and new variance estimator 95 Confidence Intervals were constructed. Results of the analysis showed that the Kaplan-Meier Survival Function give over estimates, while the Piecewise Kaplan-Meier Survival Function considered the dependency of censoring and provided the satisfactory results.

Limitation: Due to limitation of the data, methods were applied to only two diseases but it can be extended to more than two diseases.

REFERENCES

Ajeba O. B., Kabir. Z. O'Connor. T. Survival Analysis of Adult Tuberculosis Disease.PLOS ONE. (2014). 9(11).

Kleinbaum, D. G., & Klein, M. Survival Analysis. New York: SpringerScience+Business Media, Inc.(1996)

Leemis, L. M. *Reliability: Probabilistic Models and Statistical Methods-II*. Prentice Hall, New Jersey, 1997

Zaman Q, Strasak A, Pfeiffer KP. (2011). *Exact Waiting Time Survival Function. Journal of Biometrics & Biostatistics:* 2(3):100017.1-9.

Kaplan, E.L., and Meier, P. Non Parametric Estimation from incomplete observations. Journal of the American Statistical Association 1958, Vol.53(282);457-481.

Greenwood, M. A report on the natural duration of cancer. *In Reports on Public Health and Medical Subjects* 1926, 33; 1-26.

Peto, R.; Peto, J. Asymptotically efficient Rank invariant test procedures. Journal of the royal statistical society, 1972,185-207.

Danken, E., killer, D.,Fischer,S., Marti, H.R(1975).Incidence of thalassemia in Switzerland. *Schweiz-Med-Wochenschr*. 105(4): 102-5

Peto, R., Pike, M. C., Armitage, P. et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. Part II: Analysis and Examples. *British Journal of Cancer* 1977, 35(1); 1-39

Politis, C., Richardson, C., Yfantopoulos, J.G., (1991).Public knowledge of thalassemia in Greece and current concepts of the social status of thalassemic patients.*Soc-Sci-Med.32(1):* 59-64

Zaman Q, Raziq F, Iqbal M, Salahuddin, Hussain M. (2013). *Ideal Bootstrap Survival Function and Its Variance Estimator.* Journal of Science & Technology: 37(1):55-68.

Zaman Q, Atif M, Iqbal M, Bibi A, Shah SH, Farooq M, Rafiq M, Pfeiffer KP. (2014). Estimation of Survival Probabilities in the Presence of Ties. Life Science Journal. 11(10s): 155-164.

Fatimid Foundation, A symbol of hopes Pakistan. A Charitable Organization. Established 1981.

Appendix-A : Based on 119 Observations

time	d1	c1	n1	s1	GW	d2	c2	n2	s2	sc	PWV	se	95lowerlimit	95upperlimit
12	1		119	0.9916	7E-05	0	0	40	1	0.9916	7E-05	0.0084	0.9752	1.0080
15	0	1	118	0.9916	7E-05	0	0	40	1	0.9916	7E-05	0.0084	0.9752	1.0080
17	1		117	0.9831	0.00014	0	0	40	1	0.9831	0.00014	0.0118	0.9599	1.0063
19	0		116	0.9831	0.00014	1		40	0.975	0.9585	0.000722	0.0269	0.9059	1.0112
22	1		115	0.9746	0.00021	0	0	39	0.975	0.9502	0.000778	0.0279	0.8955	1.0049
24	1		114	0.9660	0.000279	0	0	39	0.975	0.9419	0.000834	0.0289	0.8853	0.9985
25	0		113	0.9660	0.000279	0	1	39	0.975	0.9419	0.000834	0.0289	0.8853	0.9985
27	1		112	0.9574	0.000348	0	0	38	0.975	0.9335	0.000889	0.0298	0.8750	0.9919
29	0	1	111	0.9574	0.000348	0	0	38	0.975	0.9335	0.000889	0.0298	0.8750	0.9919
30	0		110	0.9574	0.000348	1		38	0.9493	0.9089	0.00143	0.0378	0.8348	0.9830
33	1		109	0.9486	0.000418	0	0	37	0.9493	0.9006	0.001473	0.0384	0.8253	0.9758
36	1		108	0.9398	0.000486	0	0	37	0.9493	0.8922	0.001515	0.0389	0.8159	0.9685
37	1		107	0.9310	0.000554	0	0	37	0.9493	0.8839	0.001556	0.0394	0.8066	0.9612
38	0		106	0.9310	0.000554	0	1	37	0.9493	0.8839	0.001556	0.0394	0.8066	0.9612
39	0	1	105	0.9310	0.000554	0	0	36	0.9493	0.8839	0.001556	0.0394	0.8066	0.9612
42	0		104	0.9310	0.000554	1		36	0.9230	0.8593	0.002056	0.0453	0.7704	0.9482
44	1		103	0.9220	0.000624	0	0	35	0.9230	0.8510	0.002086	0.0457	0.7615	0.9405
47	1		102	0.9130	0.000693	0	0	35	0.9230	0.8426	0.002114	0.0460	0.7525	0.9328
49	0	1	101	0.9130	0.000693	0	0	35	0.9230	0.8426	0.002114	0.0460	0.7525	0.9328
53	0		100	0.9130	0.000693	1	0	35	0.8966	0.8186	0.002558	0.0506	0.7194	0.9177
57	1		99	0.9037	0.000763	0	0	34	0.8966	0.8103	0.002574	0.0507	0.7109	0.9097
59	1		98	0.8945	0.000832	0	0	34	0.8966	0.8020	0.00259	0.0509	0.7023	0.9018
67	0		97	0.8945	0.000832	0	1	34	0.8966	0.8020	0.00259	0.0509	0.7023	0.9018
69	1		96	0.8852	0.0009	0	0	33	0.8966	0.7937	0.002605	0.0510	0.6936	0.8937
77	0	1	95	0.8852	0.0009	0	0	33	0.8966	0.7937	0.002605	0.0510	0.6936	0.8937
78	1	1	94	0.8758	0.000969	0	0	33	0.8966	0.7852	0.00262	0.0512	0.6849	0.8856
79	0		92	0.8758	0.000969	1	0	33	0.8694	0.7614	0.003013	0.0549	0.6539	0.8690
80	1		91	0.8662	0.00104	0	0	32	0.8694	0.7531	0.003016	0.0549	0.6454	0.8607
82	1		90	0.8565	0.001108	0	0	32	0.8694	0.7447	0.003019	0.0549	0.6370	0.8524

83	0	1	89	0.8565	0.001108	0	0	32	0.8694	0.7447	0.003019	0.0549	0.6370	0.8524
85	0		88	0.8565	0.001108	0	1	32	0.8694	0.7447	0.003019	0.0549	0.6370	0.8524
86	1		87	0.8467	0.001179	0	0	31	0.8694	0.7361	0.003022	0.0550	0.6284	0.8439
87	1		86	0.8369	0.001247	1	0	31	0.8414	0.7041	0.003366	0.0580	0.5904	0.8178
88	0	1	84	0.8369	0.001247	0	0	30	0.8414	0.7041	0.003366	0.0580	0.5904	0.8178
90	0		83	0.8369	0.001247	1	0	30	0.8133	0.6806	0.003678	0.0606	0.5618	0.7995
91	0		82	0.8369	0.001247		1	29	0.8133	0.6806	0.003678	0.0606	0.5618	0.7995
92	1		81	0.8265	0.001322	1	0	28	0.7843	0.6482	0.003957	0.0629	0.5249	0.7715
94	1		79	0.8161	0.001397	0	1	27	0.7843	0.6400	0.003924	0.0626	0.5173	0.7628
95	1		77	0.8055	0.001472	0	0	26	0.7843	0.6317	0.00389	0.0624	0.5095	0.7540
97	1		76	0.7949	0.001544	0	0	26	0.7843	0.6234	0.003857	0.0621	0.5017	0.7451
98	0	1	75	0.7949	0.001544	0	0	26	0.7843	0.6234	0.003857	0.0621	0.5017	0.7451
99	0		74	0.7949	0.001544	1	0	26	0.7541	0.5994	0.004119	0.0642	0.4736	0.7252
100	1	1	73	0.7840	0.001619	0	0	25	0.7541	0.5912	0.004073	0.0638	0.4661	0.7163
102	0		71	0.7840	0.001619	0	1	25	0.7541	0.5912	0.004073	0.0638	0.4661	0.7163
103	1		70	0.7728	0.001697	0	0	24	0.7541	0.5828	0.004028	0.0635	0.4584	0.7072
105	0		69	0.7728	0.001697	1	0	24	0.7227	0.5585	0.004264	0.0653	0.4305	0.6865
107	1		68	0.7614	0.001774	1	0	23	0.6913	0.5263	0.004396	0.0663	0.3964	0.6563
109	1		66	0.7499	0.001852	0	0	22	0.6913	0.5184	0.004326	0.0658	0.3895	0.6473
110	1		65	0.7383	0.001927	0	0	22	0.6913	0.5104	0.004257	0.0652	0.3825	0.6383
111	0		64	0.7383	0.001927	0	1	22	0.6913	0.5104	0.004257	0.0652	0.3825	0.6383
112	1	1	63	0.7266	0.002001	1	1	21	0.6584	0.4784	0.004343	0.0659	0.3492	0.6075
116	1		59	0.7143	0.002083	0	0	19	0.6584	0.4703	0.004262	0.0653	0.3423	0.5982
117	1		58	0.7020	0.002161	0	0	19	0.6584	0.4622	0.004181	0.0647	0.3354	0.5889
119	1		57	0.6897	0.002235	0	0	19	0.6584	0.4541	0.0041	0.0640	0.3286	0.5796
121	0	1	56	0.6897	0.002235	0	0	19	0.6584	0.4541	0.0041	0.0640	0.3286	0.5796
123	1		55	0.6771	0.002308	1	0	19	0.6237	0.4223	0.004129	0.0643	0.2964	0.5483
125	0	1	53	0.6771	0.002308	0	0	18	0.6237	0.4223	0.004129	0.0643	0.2964	0.5483
127	0		52	0.6771	0.002308	0	1	18	0.6237	0.4223	0.004129	0.0643	0.2964	0.5483
128	1		51	0.6639	0.002392	0	0	17	0.6237	0.4141	0.004036	0.0635	0.2895	0.5386
129	0	1	50	0.6639	0.002392	0	0	17	0.6237	0.4141	0.004036	0.0635	0.2895	0.5386
130	0		49	0.6639	0.002392	1	0	17	0.5870	0.3897	0.004133	0.0643	0.2637	0.5157

132	1		48	0.6500	0.00248	0	0	16	0.5870	0.3816	0.004027	0.0635	0.2572	0.5060
134	0		47	0.6500	0.00248	0	1	16	0.5870	0.3816	0.004027	0.0635	0.2572	0.5060
136	1		46	0.6359	0.002569	0	0	15	0.5870	0.3733	0.003921	0.0626	0.2505	0.4960
137	1		45	0.6218	0.002651	0	0	15	0.5870	0.3650	0.003816	0.0618	0.2439	0.4861
139	1		44	0.6076	0.002727	0	0	15	0.5870	0.3567	0.003712	0.0609	0.2373	0.4761
141	1		43	0.5935	0.002797	0	0	15	0.5870	0.3484	0.003609	0.0601	0.2307	0.4661
143	0	1	42	0.5935	0.002797	0	0	15	0.5870	0.3484	0.003609	0.0601	0.2307	0.4661
145	0		41	0.5935	0.002797	1	1	15	0.5479	0.3252	0.003647	0.0604	0.2068	0.4435
146	1		39	0.5783	0.002881	0	0	13	0.5479	0.3168	0.00353	0.0594	0.2004	0.4333
149	0	1	38	0.5783	0.002881	0	1	13	0.5479	0.3168	0.00353	0.0594	0.2004	0.4333
150	1		36	0.5622	0.002974	1	0	12	0.5022	0.2824	0.003471	0.0589	0.1669	0.3978
153	0	1	34	0.5622	0.002974	0	0	11	0.5022	0.2824	0.003471	0.0589	0.1669	0.3978
157	1		33	0.5452	0.003078	0	0	11	0.5022	0.2738	0.003335	0.0577	0.1606	0.3870
159	0		32	0.5452	0.003078	1	0	11	0.4566	0.2489	0.003319	0.0576	0.1360	0.3618
160	1		31	0.5276	0.003182	0	0	10	0.4566	0.2409	0.003171	0.0563	0.1305	0.3513
162	0	1	30	0.5276	0.003182	0	0	10	0.4566	0.2409	0.003171	0.0563	0.1305	0.3513
163	0	1	29	0.5276	0.003182	0	0	10	0.4566	0.2409	0.003171	0.0563	0.1305	0.3513
165	1		28	0.5088	0.003301	1	0	10	0.4109	0.2091	0.002932	0.0541	0.1029	0.3152
168	0		26	0.5088	0.003301	0	1	9	0.4109	0.2091	0.002932	0.0541	0.1029	0.3152
169	0	1	25	0.5088	0.003301	0	0	8	0.4109	0.2091	0.002932	0.0541	0.1029	0.3152
170	1		24	0.4876	0.003462	1	0	8	0.3596	0.1753	0.002666	0.0516	0.0741	0.2765
172	0	1	22	0.4876	0.003462	0	0	7	0.3596	0.1753	0.002666	0.0516	0.0741	0.2765
175	1		21	0.4643	0.003654	0	0	7	0.3596	0.1670	0.002484	0.0498	0.0693	0.2646
176	0		20	0.4643	0.003654	1	0	7	0.3082	0.1431	0.002313	0.0481	0.0488	0.2374
178	0	1	19	0.4643	0.003654	0	1	6	0.3082	0.1431	0.002313	0.0481	0.0488	0.2374
179	1		17	0.4370	0.003939	0	0	5	0.3082	0.1347	0.002115	0.0460	0.0445	0.2248
180	0	1	16	0.4370	0.003939	0	0	5	0.3082	0.1347	0.002115	0.0460	0.0445	0.2248
182	1		15	0.4079	0.004223	0	0	5	0.3082	0.1257	0.001918	0.0438	0.0399	0.2115
184	1		14	0.3788	0.00443	0	0	5	0.3082	0.1167	0.001729	0.0416	0.0352	0.1982
185	1		13	0.3496	0.004558	0	0	5	0.3082	0.1077	0.001547	0.0393	0.0306	0.1848
187	0	1	12	0.3496	0.004558	0	0	5	0.3082	0.1077	0.001547	0.0393	0.0306	0.1848
188	0		11	0.3496	0.004558	1	1	5	0.2466	0.0862	0.001362	0.0369	0.0139	0.1585

189	1		9	0.3108	0.004943	0	0	3	0.2466	0.0766	0.001158	0.0340	0.0099	0.1433
193	0	1	8	0.3108	0.004943	0	0	3	0.2466	0.0766	0.001158	0.0340	0.0099	0.1433
195	1		7	0.2664	0.005321	0	0	3	0.2466	0.0657	0.000953	0.0309	0.0052	0.1262
197	0	1	6	0.2664	0.005321	0	0	3	0.2466	0.0657	0.000953	0.0309	0.0052	0.1262
199	0		5	0.2664	0.005321	1	0	3	0.1644	0.0438	0.000743	0.0273	-0.0096	0.0972
204	0	1	4	0.2664	0.005321	0	0	2	0.1644	0.0438	0.000743	0.0273	-0.0096	0.0972
223	1		3	0.1776	0.007621	0	0	2	0.1644	0.0292	0.000472	0.0217	-0.0134	0.0718
237	0		2	0.1776	0.007621	1	1	2	0.0822	0.0146	0.000225	0.0150	-0.0148	0.0440