GLOBAL JOURNAL OF MATHEMATICAL SCIENCES VOL 9, NO. 1, 2010: 37-43
 ISSN 1596-6208
www.globaljournalseries.com; Email: info@globaljournalseries.com

ON THE DETERMINATION OF RISK FACTORS ASSOCIATED WITH THE RECOVERY OF CATARACT PATIENTS.

F. I. UGWUOWO AND E. F. UDOUMOH

(Received 14, January 2010; Revision Accepted 4, May 2010)

ABSTRACT

Markov modeling approach is adapted in a multi-stage process involving acuity levels of patients. The multiple binary logistic regression models is used to identify some risk factors associated with the recovery process. Result shows that Blood pressure and Sugar in the blood form the significant risk factors associated with recovery.

KEYWORDS: Recovery process; transition probability matrix, multiple logistic regression; risk factors; multi-stage process.

1.1 INTRODUCTION

Research have shown that about 2 million people become blind annually and cataract account for more than 50% of this blindness. Population projections also suggest that the number of cataract blind persons could reach 40 million by year 2025 as the population of elderly persons increase. But with more successful cataracts surgery in the developing world, the figure could be reduced (see Dandora *et al* (2001). Kupfer (1984), Stark *et al* (1989)).
Many scholars are involved in cataract related research in the areas of predicting best corrected acuities after surgery and identifying risk factors associated with cataract formation. Some of the risk factors identified include: sex, age, active cigarette smoking, exposure to severe sunlight, race, diabetes, family history, nutrition, geographic factors, and cooking smoke (see Nirmalan et al (2004), Kahn & Moorhead (1973), Kahn *et al* (1977), Hiller *et al* (1986), Chatterjee et al (1982)). Post operation visual ability of cataract patients receives general comments of improved vision. Hence, some best corrected acuities of 3/60 or less, and 6/60 or less have been predicted (e.g. Lundstorm et al (2002), Bekibele et al (2004), Kupfer (1984), Kahn & Moorhead (1973)). Recently, Ugwuowo & Udoumoh (2009) adapted a Semi-Markov modeling approach to study post operation acuity levels of cataract patients. They used Parametric and Non-parametric methods to estimate the sojourn times. The inter-event times were assumed to follow Weibull distribution. The results presented clear view of the recovery process showing the proportion of patients having a particular visual acuity given the immediate past state, with the length of time taken for such transitions. In this paper, we will explore further to ascertain the risk factors associated with these transitions. We present the transition probability matrix (TPM) of the recovery process and then use logistic regression model to identify risk factors associated with these transitions. Recovery here is measured in terms of the probability of patients' transition from one level of visual acuity to the other. Hence, we think that the probability of a patient transiting from one vision state to the other depend only on the immediate past state.

2.0 MODELS

2.1 The Markov Model

Consider the stochastic process { X_n, n = 1,2,…} with countable finite values. If X_n = i, the process is in state i at time n. Whenever the process is in state i, there is a fixed probability P_{ij} that the next transition will be into state j, that is

$$P\{X_{n+1} = j / X_n = i\} = P_{ij}$$

For all states (i,j) and n ≥ 1.

2.1.1 The Transition probability Matrix

i) For the transition probabilities compute;

$$P_{ij} = \frac{\text{Number of observed transitions from i to j}}{\text{Number of observed transitions from } i}$$

ii) Stationary distribution

Letting $\pi_j = \lim_{n\to\infty} P_{ij}^n$; j ≥ 0

F. I. Ugwuowo, Department of Statistics, University of Nigeria, Nsukka, Enugu State, Nigeria.
E. F. Udoumoh, Department of Mathematics, Statistics, & Computer Science, University of Agriculture, Makurdi, Benue State, Nigeria.

Then the stationary distribution π_j is

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij} \quad ; j \geq 0$$

π_i is the proportion of patients in state I.

$$\sum_{j=0}^{\infty} \pi_j = 1$$

(Ross, S. M. 1987)

2.2 The Binary Logistic Regression Model

Assume that there are k explanatory variable $X = \{x_1, x_2, \cdots, x_k\}$. The response variable y is a binary variable indicating whether a patient transit (y = 1) or do not transit (y = 0). If $\pi(x)$ is the conditional probability given the explanatory variables, then,

$\pi(x) = P(y_i = 1 \setminus x_{i_1}, \cdots, x_{i_k})$

$1 - \pi(x) = P(y_i = 0 \setminus x_{i_1}, \cdots, x_{i_k})$

A standard regression model is formulated as follows;

and $\pi(x) = \dfrac{\exp\{\sum_{j=0}^{k} \beta_j x_{ij}\}}{1 + \exp\{\sum_{j=0}^{k} \beta_j x_{ij}\}}$, (2.1)

$$1 - \pi(x) = \frac{1}{1 + \{\sum_{j=0}^{k} \beta_j x_{ij}\}} \tag{2.2}$$

Where $\beta = \{\beta_0, \beta_1, \cdots, \beta_k\}$ are unknown parameters to be estimated.

The logit transform of equation (2.1) in terms of $\pi(x)$ is

g(x) = $\beta_0 + \sum_{i=1}^{k} \beta_j x_{ij}$ (2.3)

To estimate coefficient (β_j 's) of the explanatory variables we apply the MLE method. The log-likelihood function for n individuals is given as

$$InL(\beta_0, \beta_1, \cdots, \beta_n) = \sum_{k} \beta_j \sum_{n} x_{ij} y_i - \sum_{n} In\{\exp(\sum_{k} \beta_j x_{ij})\} \tag{2.4}$$

(Hosmer and Lemeshow, 1989)

Differentiating the log-likelihood equation with respect to (k + 1) β_j 's result in k + 1 likelihood equations, which could be solved simultaneously using special purpose software.

2.2.1 Model adequacy

To test the adequacy of the logistic regression model we use Wald test, Hosmer-Lemeshow goodness of fit test, and Cook's influence plots. These can be seen in tables 1 & 2 and Figs 2-40 respectively.

3.0 MODELLING CATARACT DATA

Data was obtained from 150 patients' files systematically sampled. It spanned through January, 2000 to December, 2005. Measurement of best visual acuity of patients taken after surgery form the state space namely: A-light perception (LP), B-hand movement (HM), C-counting of fingers(CF), D-6/60, E-6/36, F-6/24, G-6/18, H-6/12, I-6/9, J-6/6 and are irreducible (see fig 1). The response variable is patients' vision state which is 'transition' or 'no transition' denoted by 1 or 0 respectively, taken through the study period. Transition here means movement of patient(s) from one vision state to the other.

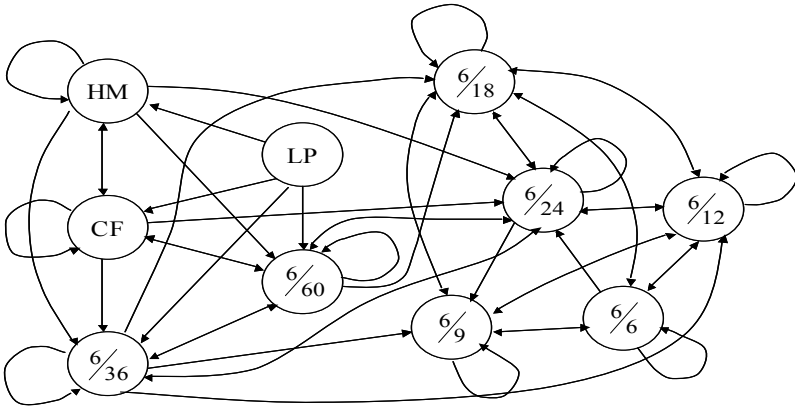

Fig. 1: A schematic representation of the recovery process

The explanatory variables include factors that potentially influenced the likelihood of transition and were available in the data. They are; Sex (male -1 or female-0) blood pressure (BP) (Normal -1or abnormal -0), family history (FM) (trait -1 or no trait -0), any other ocular disease (A.O.O.D) (present -1 or not present -0), and sugar level (SL) (Normal-1 or abnormal-0).Presented below are the number of observed transitions (t_{ij}) in the data set with row sum ($t_{i.}$)

T =

	A	B	C	D	E	F	G	H	I	J	t_i
A	0	11	25	15	11	0	0	0	0	0	62
B	0	8	37	33	6	1	0	0	0	0	85
C	0	13	54	59	13	3	0	0	0	0	142
D	0	0	19	102	63	21	4	0	0	0	209
E	0	0	0	25	76	32	13	10	3	0	159
F	0	0	0	8	9	27	20	10	8	0	82
G	0	0	0	0	0	4	16	18	11	2	51
H	0	0	0	0	0	5	8	15	6	1	35
I	0	0	0	0	0	0	5	8	15	7	35
J	0	0	0	0	0	2	3	4	2	1	12

The transition probabilities are given below;

$\hat{P}_{ij}$ =

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
A	0	0.1774	0.4032	0.2419	0.1774	0	0	0	0	0
B	0	0.0941	0.4353	0.3882	0.0706	0.0118	0	0	0	0
C	0	0.0915	0.3803	0.4155	0.0915	0.0211	0	0	0	0
D	0	0	0.0909	0.4880	0.3014	0.1005	0.0191	0	0	0
E	0	0	0	0.1572	0.4780	0.2013	0.0818	0.0629	0.0189	0
F	0	0	0	0.0976	0.1096	0.3293	0.2439	0.1220	0.0976	0
G	0	0	0	0	0	0.0784	0.3137	0.3529	0.2157	0.0392
H	0	0	0	0	0	0.1429	0.2286	0.4286	0.1714	0.0286
I	0	0	0	0	0	0	0.1429	0.2286	0.4286	0.2000
J	0	0	0	0	0	0.1667	0.2500	0.3333	0.166	0.0833

The stationary distribution of the transition probability matrix is π_A = 0, π_B = 0.001, π_C = 0.008, π_D = 0.055, π_E = 0.070, π_F = 0.136, π_G = 0.211, π_H = 0.277, π_I =0.189, and π_J = 0.054;

On logistic regression, we present in table 1 Walds test results for pairs of transitions that have only significant covariates. From the results, column 1 presents pairs of transitions while column 2 presents the significant covariates. Column 3 presents the coefficient (B) of significant covariates and column 4 presents the significant values with all $\leq$ 0.05. Column 5 has exp (B) which is the odd ratio (OR). Odd ratio is an important tool for interpreting results of logistic

regression analysis. For instance, the odd ratio of 0.089 indicates that the odd of making a transition from A(LP) to D(6/60) is delayed by 8.9% with each (mm/hg) of blood pressure hike. Also the odd ratio of 13.169 indicates that the odd of remaining in state B (HM) is 13.2 times higher with each (mm/Hg) blood pressure hike. Odd ratio of 5.908 means that the odd of moving from H(6/12) to G(6/18) is 5.9 times higher with each (mg/ml) sugar in the blood. Table 2 presents Hosmer-Lemeshow goodness of fit test for all pairs of transitions with all the significant values indicating non-significance (sig. value > 0.05). This is an indication that the predicted values are not significantly different from the observed. Figures 2-40 present Cook's influence plots for all pairs of transitions, plotting analog of Cook's influence statistics against predicted probability. Like the graphical approach of interpreting the value of the diagnostics in linear regression, large values of diagnostics either appear as spikes or at the extreme corners of the plots. The high-leverage points could be influential to the results of the analysis. Hence, those cases which fall far away from the balance should be noted for further investigation. For example, in fig.2, one point lies at the extreme upper end of the plot. In fig 15, there are three high leverage values. In fig 38, there is one high leverage point and three moderate ones. More high-leverage points indicate poor model. In the above plots, we have few high-leverage with more low and moderate leverage points. This is an indication of a fair model, Hosmer & Lemoshow (1989). The analyses were done using SPSS.

Table 1: Wald's Test Results

Transitions (i,j)	Covariates	B	Sig. Values	Exp(B)
AD	BP	-2.422	0.024	0.089
AE	BP	-2.195	0.046	0.111
BB	BP	2.578	0.022	13.169
BD	BP	-1.057	0.037	0.347
CB	BP	1.584	0.016	4.731
CC	SL	0.961	0.050	2.615
CD	BP	-0.792	0.043	0.453
CE	BP	-2.362	0.030	0.094
DC	SL,BP	1.214,1.453	0.045,0.018	3.366,4.275
DD	SL	1.030,	0.025	2.800
DE	SL	-1.107	0.040	0.331
DF	BP	-1.706	0.028	0.182
ED	BP,AOOD	1.003,1.173	0.037,0.028	2.727,3.23
EE	SL	-1.911	0.013	0.148
EF	BP	-1.038	0.043	0.354
EG	BP	-2.101	0.050	0.122
FD	BP	1.988	0.023	7.302
FE	BP	1.927	0.025	6.867
FF	Gen	-1.109	0.029	0.330
FG	BP	-1.446	0.032	0.236
FH	Gen , SL	2.362,2.040	0.039,0.012	10.608,7.693
GG	BP	-2.146	0.046	0.117
GI	BP	-2.229	0.045	0.108
HG	SL	1.776	0.022	5..908
HH	BP	-1.589	0.049	0.204
IG	SL	3.540	0.004	34.480
IH	SL	2.266	0.008	9.641

Table 2: Hosmer-Lemeshow Test Results

Transitions(i,j)	Chi-Square	df	Sig. Values	Transition (i,j)	Chi-Square	df	Sig. values
AB	4.416	7	0.695	EG	1.868	6	0.931
AC	6.368	7	0.497	EH	9.165	8	0.241
AD	2.472	7	0.979	FD	3.553	8	0.895
AE	1.607	8	0.991	FE	2.638	7	0.916
BB	5.729	6	0.451	FF	13.171	7	0.68
BC	4.774	8	0.781	FG	3.297	7	0.856
BD	4.630	7	0.705	FH	8.206	7	0.315
BE	8.636	8	0.374	FI	5.950	7	0.546
CB	2.717	7	0.910	GF	1.567	7	0.980
CC	5.901	8	0.658	GG	8.567	8	0.380
CD	12.383	8	0.135	GH	4.274	7	0.748
CE	1.774	6	0.939	GI	6.991	8	0.538
DC	6.289	8	0.615	HF	8.588	8	0.378
DD	3.989	6	0.678	HG	4.023	7	0.777
DE	4.965	7	0.664	HH	2.852	6	0.827

DF	4.837	7	0.680	HI	3.634	8	0.889
DG	2.108	8	0.978	IG	1.395	8	0.994
ED	5.229	7	0.632	IH	3.168	8	0.923
EE	8.690	7	0.276	II	1.607	8	0.991
EF	3.171	8	0.923				

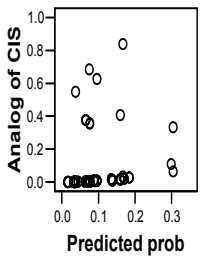

Fig.2: CIP for AB

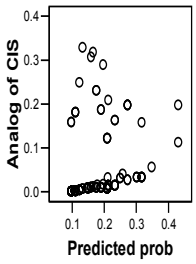

Fig.3: CIP for AC

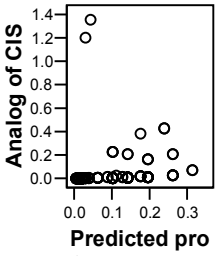

Fig.$:CIP for AD

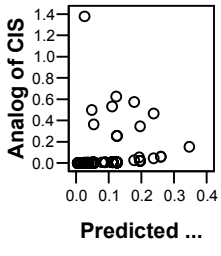

Fig.5: CIP for AE

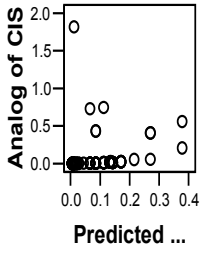

Fig.6: CIP for BB

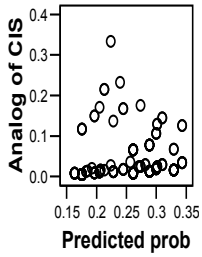

Fig.7: CIP for BC

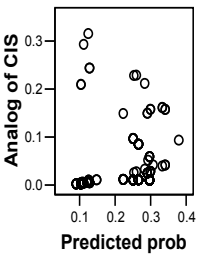

Fig 8:CIP for BD

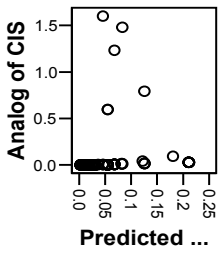

Fig 9:CIP for BE

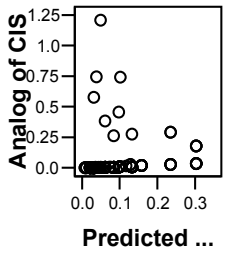

Fig 10: CIP for CB

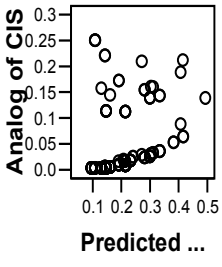

Fig 11: CIP for CC

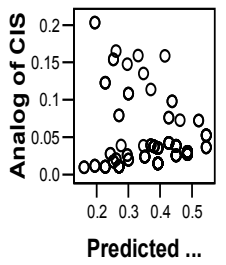

Fig 11: CIP for CD

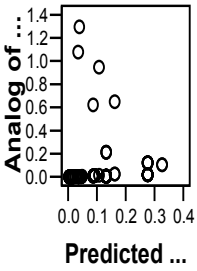

Fig.13:CIP for CE

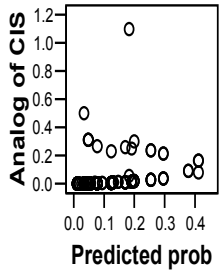

Fig 14: CIP for DC

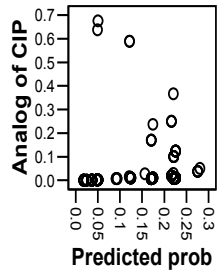

Fig.15:CIP for DF

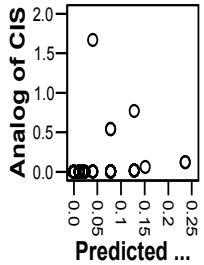

Fig 16: CIp for DG

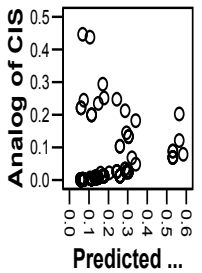

Fig 17: CIP for ED

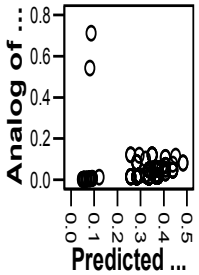

Fig 18: CIP for EE

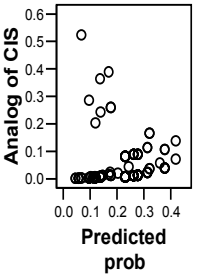

Fig 19 CIP for EF

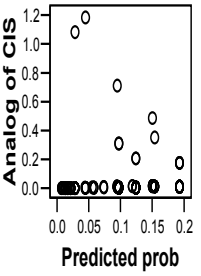

Fig 20:CIP for EG

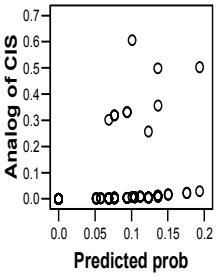

Fig 21: CIP for EH

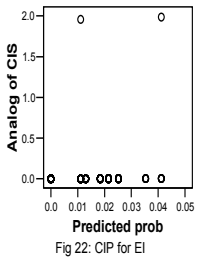

Fig 22: CIP for EI

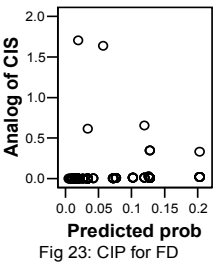

Fig 23: CIP for FD

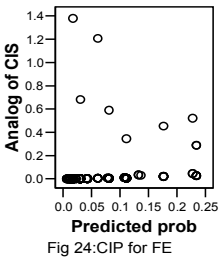

Fig 24:CIP for FE

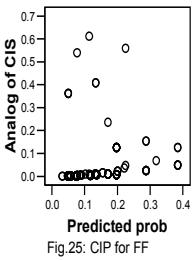

Fig.25: CIP for FF

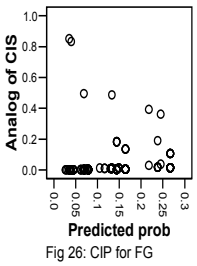

Fig 26: CIP for FG

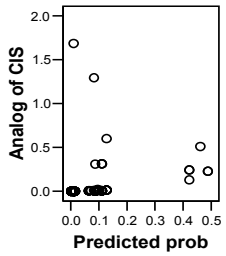

Fig.27: CIP for FH

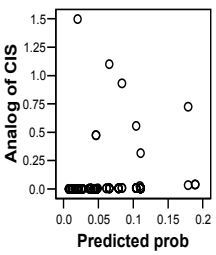

Fig. 28: CIP for FI

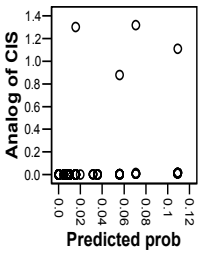

Fig.29: CIP for GF

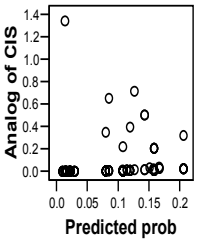

Fig.30: CIP for GG

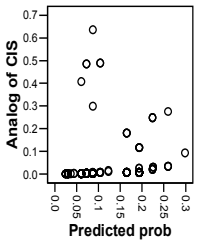

Fig.31: CIP for GH

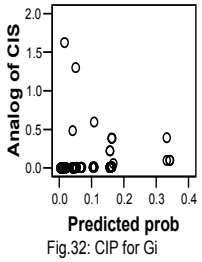

Fig.32: CIP for Gi

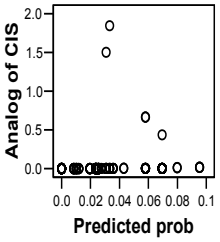

Fig.33: CIP for HF

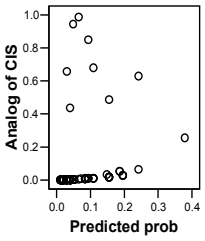

Fig.34: CIP for HG

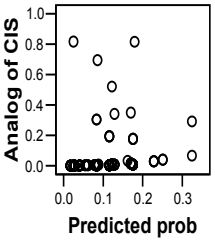

Fig.35: CIP for HH

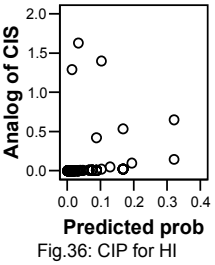

Fig.36: CIP for HI

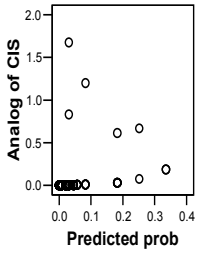

Fig.37: CIP for IG

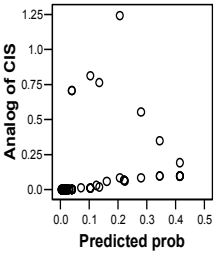

Fig.38: CIP for IH

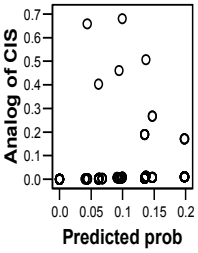

Fig.39:CIP for II

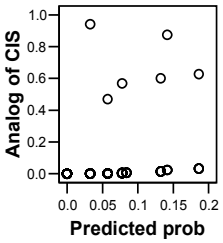

Fig.40: CIP for IJ

4.0 DISCUSSION

Proportions of patients having particular visual acuities given the immediate past state could easily be ascertained with the transition probability matrix (TPM). The logistic regression results reveal that blood pressure and sugar in the blood (diabetes) are the dominant risk factors that influence patients' recovery. Previous findings have shown that one of the first signs of diabetes is the sudden change in eyeglasses prescription due to poor vision. Also, diabetic retinopathy develops in people with diabetes. This causes blood vessel abnormalities in the retina which have the potential to diminish vision. High blood pressure is also a risk factor for the development of retinopathy and age-related macular degeneration, SerVaas (2004).

5.0 CONCLUSION

Despite the interesting results obtained, a more detailed hospital based research, with the introduction of more risk factors could improve the quality of results.

REFERENCES

Bekibele, C.O., Ubah, J.N. and Fasina, O.A., 2004. Comparative evaluation of outcome of cataract surgery at Ago-Iwoye, Ogun State. Nigerian Journal of Surgical Research, Vol. 6 No. 1-2, 25-29.

Chatterjee, A., Milton, R.C. and Thyle, S., 1982. Prevalence and etiology of cataract. Punjab Br. J. Ophthalmol. 66:35.

Dandora, R., Dandona, L. and John, J. K., 2001. Awareness of eye diseases in urban population in Southern India. Bull World Health Organs. 79:No.2.

Hiller, R., Sperduto, R. D. and Ederer, F., 1986. Epidemiologic association with nuclear, Cortical and posterior subcapsular cataracts. *Am. J. Epidemiol.* 124:916.

Hosmer, D. W. and Lemoshow, S., 1989. Applied Logistic Regression. John Wiley&Sons Inc. New York.

Kahn, H, A. and Moorhead, H. B., 1973. Statistics on Blindness in the Model Reporting Area, 1969-1970. Washington,D.C. USAHUW, PHS Publication No. (NIH) 73-427.

Kahn, H. A., Leibowitz, H. M. and Ganley, J. P., 1977. The Framingham Eye study Outline and Major Prevalence Findings. *Am. J. Epidemiol*. 106:17.

Kupfer, C., 1984. Bowman Lecture. The Conquest of Cataract: Global Challenge Trans Ophthal Soc. UK 104:1.

Lundstorm, M. and Brege, K.G., 2002. Cataract surgery and quality of life in patients with age related macular degeneration. *Br. S. Ophthalmol* , 86; 1330-1335.

Nirmalan, P.K., Robin, A.L, Katz, J., 2004. Risk factors for age related cataract in rural population of Southern India; the Aravind comprehensive Eye study. *Br J Ophthalmol*; 88; 989-994

Ross, S., 1983.Stochastic Processes. John Wiley & Sons, New York.

SerVaas, C., 2004. Hypertension and Vision Loss. Neighborhood Heart Watch, Vol. III, No.10.

Stark, W. J., Sommer, A. and Smith, R. E., 1989. Challenging trends in intraocullan lens implantation. Arch. Ophthalmol 107: 1441.

Ugwuowo, F.I. and Udoumoh, E, F., 2009. Multi-Stage Semi-Markov Models for Recovery Process. The 13th International Conference “Applied Stochastic Models and data analysis” (ASMDA-2009) Vilnius, Lithuania.441-444.