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. Lundstrom 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 \( \{ X_n, n = 1,2,\ldots\} \) 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 \( n \geq 1 \).

2.1.1 The Transition probability Matrix

i) For the transition probabilities compute;

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

ii) Stationary distribution

Letting \( \pi_j = \lim_{n \to \infty} P^n_{ij} \); \( j \geq 0 \)

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 \( \pi_j \) is

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

\( \pi_i \) 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 variables \( X = \{x_1, x_2, \ldots, x_k\} \). 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,

\[
\pi(x) = P(y_i = 1 | x_1, \ldots, x_k)
\]

\[
1 - \pi(x) = P(y_i = 0 | x_1, \ldots, x_k)
\]

A standard regression model is formulated as follows;

\[
\pi(x) = \frac{\exp(\sum_{j=0}^{k} \beta_j x_j)}{1 + \exp(\sum_{j=0}^{k} \beta_j x_j)},
\]

\[
1 - \pi(x) = \frac{1}{1 + \exp(\sum_{j=0}^{k} \beta_j x_j)}
\]

Where \( \beta = \{\beta_0, \beta_1, \ldots, \beta_k\} \) are unknown parameters to be estimated.

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

\[
g(x) = \beta_0 + \sum_{j=1}^{k} \beta_j x_j
\]

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

\[
InL(\beta_0, \beta_1, \ldots, \beta_n) = \sum_{k} \beta_j \sum_{n} x_j y_i - \sum_{n} \ln(\exp(\sum_{k} \beta_j x_j))
\]

(Hosmer and Lemeshow, 1989)

Differentiating the log-likelihood equation with respect to \( (k + 1) \) \( \beta_j \)'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.
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 -1 or 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 \((i_j)\) in the data set with row sum \((i_t)\)

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

The transition probabilities are given below;

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

The stationary distribution of the transition probability matrix is \(\pi_A = 0, \pi_B = 0.001, \pi_C = 0.008, \pi_D = 0.055, \pi_E = 0.070, \pi_F = 0.136, \pi_G = 0.211, \pi_H = 0.277, \pi_I = 0.189, \) and \(\pi_J = 0.054;\)

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 & Lemeshow (1989). The analyses were done using SPSS.

Table 1: Wald’s Test Results

<table>
<thead>
<tr>
<th>Transitions (i,j)</th>
<th>Covariates</th>
<th>B</th>
<th>Sig. Values</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>BP</td>
<td>-2.422</td>
<td>0.024</td>
<td>0.089</td>
</tr>
<tr>
<td>AE</td>
<td>BP</td>
<td>-2.195</td>
<td>0.046</td>
<td>0.111</td>
</tr>
<tr>
<td>BB</td>
<td>BP</td>
<td>2.578</td>
<td>0.022</td>
<td>13.169</td>
</tr>
<tr>
<td>BD</td>
<td>BP</td>
<td>-1.057</td>
<td>0.037</td>
<td>0.347</td>
</tr>
<tr>
<td>CB</td>
<td>SL</td>
<td>1.584</td>
<td>0.016</td>
<td>4.731</td>
</tr>
<tr>
<td>CC</td>
<td>SL</td>
<td>0.961</td>
<td>0.050</td>
<td>2.615</td>
</tr>
<tr>
<td>CD</td>
<td>BP</td>
<td>-0.792</td>
<td>0.043</td>
<td>0.453</td>
</tr>
<tr>
<td>CE</td>
<td>BP</td>
<td>-2.362</td>
<td>0.030</td>
<td>0.094</td>
</tr>
<tr>
<td>DC</td>
<td>SL,BP</td>
<td>1.214,1.453</td>
<td>0.045,0.018</td>
<td>3.366,4.275</td>
</tr>
<tr>
<td>DD</td>
<td>SL</td>
<td>1.030</td>
<td>0.025</td>
<td>2.800</td>
</tr>
<tr>
<td>DE</td>
<td>SL</td>
<td>-1.107</td>
<td>0.040</td>
<td>0.331</td>
</tr>
<tr>
<td>DF</td>
<td>BP</td>
<td>-1.706</td>
<td>0.028</td>
<td>0.182</td>
</tr>
<tr>
<td>ED</td>
<td>BP,AOOD</td>
<td>1.003,1.173</td>
<td>0.037,0.028</td>
<td>2.727,3.23</td>
</tr>
<tr>
<td>EE</td>
<td>SL</td>
<td>-1.911</td>
<td>0.013</td>
<td>0.148</td>
</tr>
<tr>
<td>EF</td>
<td>BP</td>
<td>-1.038</td>
<td>0.043</td>
<td>0.354</td>
</tr>
<tr>
<td>EG</td>
<td>BP</td>
<td>-2.101</td>
<td>0.050</td>
<td>0.122</td>
</tr>
<tr>
<td>FD</td>
<td>BP</td>
<td>1.988</td>
<td>0.023</td>
<td>7.302</td>
</tr>
<tr>
<td>FE</td>
<td>BP</td>
<td>1.927</td>
<td>0.025</td>
<td>6.867</td>
</tr>
<tr>
<td>FF</td>
<td>Gen</td>
<td>-1.109</td>
<td>0.029</td>
<td>0.330</td>
</tr>
<tr>
<td>FG</td>
<td>BP</td>
<td>-1.446</td>
<td>0.032</td>
<td>0.236</td>
</tr>
<tr>
<td>FH</td>
<td>Gen, SL</td>
<td>2.362,2.040</td>
<td>0.039,0.012</td>
<td>10.608,7.693</td>
</tr>
<tr>
<td>GG</td>
<td>BP</td>
<td>-2.146</td>
<td>0.046</td>
<td>0.117</td>
</tr>
<tr>
<td>GI</td>
<td>BP</td>
<td>-2.229</td>
<td>0.045</td>
<td>0.108</td>
</tr>
<tr>
<td>HG</td>
<td>SL</td>
<td>1.776</td>
<td>0.022</td>
<td>5.908</td>
</tr>
<tr>
<td>HH</td>
<td>BP</td>
<td>-1.589</td>
<td>0.049</td>
<td>0.204</td>
</tr>
<tr>
<td>IG</td>
<td>SL</td>
<td>3.540</td>
<td>0.004</td>
<td>34.480</td>
</tr>
<tr>
<td>IH</td>
<td>SL</td>
<td>2.266</td>
<td>0.008</td>
<td>9.641</td>
</tr>
</tbody>
</table>

Table 2: Hosmer-Lemeshow Test Results

<table>
<thead>
<tr>
<th>Transitions (i,j)</th>
<th>ChijSquare</th>
<th>df</th>
<th>Sig. Values</th>
<th>Transition (i,j)</th>
<th>ChijSquare</th>
<th>df</th>
<th>Sig. values</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>4.416</td>
<td>7</td>
<td>0.695</td>
<td>EG</td>
<td>1.868</td>
<td>6</td>
<td>0.931</td>
</tr>
<tr>
<td>AC</td>
<td>6.368</td>
<td>7</td>
<td>0.497</td>
<td>EH</td>
<td>9.165</td>
<td>8</td>
<td>0.241</td>
</tr>
<tr>
<td>AD</td>
<td>2.472</td>
<td>7</td>
<td>0.979</td>
<td>FD</td>
<td>3.553</td>
<td>8</td>
<td>0.895</td>
</tr>
<tr>
<td>AE</td>
<td>1.607</td>
<td>8</td>
<td>0.991</td>
<td>FE</td>
<td>2.638</td>
<td>7</td>
<td>0.916</td>
</tr>
<tr>
<td>BB</td>
<td>5.729</td>
<td>6</td>
<td>0.451</td>
<td>FF</td>
<td>13.171</td>
<td>7</td>
<td>0.288</td>
</tr>
<tr>
<td>BC</td>
<td>4.774</td>
<td>8</td>
<td>0.781</td>
<td>FG</td>
<td>3.297</td>
<td>7</td>
<td>0.856</td>
</tr>
<tr>
<td>BD</td>
<td>4.630</td>
<td>7</td>
<td>0.705</td>
<td>FH</td>
<td>8.206</td>
<td>7</td>
<td>0.315</td>
</tr>
<tr>
<td>BE</td>
<td>8.636</td>
<td>8</td>
<td>0.374</td>
<td>FI</td>
<td>5.950</td>
<td>7</td>
<td>0.546</td>
</tr>
<tr>
<td>CB</td>
<td>2.717</td>
<td>7</td>
<td>0.910</td>
<td>GF</td>
<td>1.567</td>
<td>7</td>
<td>0.980</td>
</tr>
<tr>
<td>CC</td>
<td>5.901</td>
<td>8</td>
<td>0.658</td>
<td>GG</td>
<td>8.567</td>
<td>8</td>
<td>0.380</td>
</tr>
<tr>
<td>CD</td>
<td>12.383</td>
<td>8</td>
<td>0.135</td>
<td>GH</td>
<td>4.274</td>
<td>7</td>
<td>0.748</td>
</tr>
<tr>
<td>CE</td>
<td>1.774</td>
<td>6</td>
<td>0.939</td>
<td>GI</td>
<td>6.991</td>
<td>8</td>
<td>0.538</td>
</tr>
<tr>
<td>DC</td>
<td>6.289</td>
<td>8</td>
<td>0.615</td>
<td>HF</td>
<td>8.588</td>
<td>8</td>
<td>0.378</td>
</tr>
<tr>
<td>DD</td>
<td>3.989</td>
<td>6</td>
<td>0.678</td>
<td>HG</td>
<td>4.023</td>
<td>7</td>
<td>0.777</td>
</tr>
<tr>
<td>DE</td>
<td>4.965</td>
<td>7</td>
<td>0.664</td>
<td>HH</td>
<td>2.852</td>
<td>6</td>
<td>0.827</td>
</tr>
</tbody>
</table>
ON THE DETERMINATION OF RISK FACTORS

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
<td>4.837</td>
<td>7</td>
<td>0.680</td>
<td>HI</td>
</tr>
<tr>
<td>DG</td>
<td>2.108</td>
<td>8</td>
<td>0.978</td>
<td>IG</td>
</tr>
<tr>
<td>ED</td>
<td>5.229</td>
<td>7</td>
<td>0.632</td>
<td>IH</td>
</tr>
<tr>
<td>EE</td>
<td>8.690</td>
<td>7</td>
<td>0.276</td>
<td>II</td>
</tr>
<tr>
<td>EF</td>
<td>3.171</td>
<td>8</td>
<td>0.923</td>
<td></td>
</tr>
</tbody>
</table>

Fig.2: CIP for AB

Fig.3: CIP for AC

Fig.4: 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.12: CIP for CD

Fig.13: CIP for CE

Fig.14: CIP for CF

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 FI
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


