International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

Prognostic Factors for Evolution of Non Alcoholic Fatty Liver Disease Patients Utilizing Poisson Regression and Continuous Time Markov Chains

Iman M. Attia *

**Department of Mathematical Statistics, Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt*

Submitted: 15-10-2021 Revised: 26-10-2021 Accepted: 28-10-2021

ABSTRACT: In the present paper, the deleterious effects of obesity, type 2diabetes and insulin resistance, systolic and diastolic hypertension on the rate of progression of fibrosis in non-alcoholic fatty liver disease (NAFLD) patients are illustrated using a new approach utilizing the Poisson regression to model the transition rate matrix. The observed counts in the transition counts matrix are used as response variables and the covariates are the risk factors for fatty liver. Then the estimated counts from running the Poisson regression are used to estimate the transition rates using the continuous time Markov chains (CTMC) followed by exponentiation of the estimated rate matrix to obtain the transition probability matrix at specific time points.Using a hypothetical data of 150 participants followed up every year for a total of 28 years recording their demographic characteristics and their timeline of follow up are demonstrated. The findings revealed that insulin resistance expressed by MOMA-IR 2 has the most deleterious effects among other factors for increasing the rate of forward progression of patients from state 1 to state 2 as well as from state 2 to state 3 and from state 3 to state 4.The higher the level of HOMA-IR is, the more rapid the rate of progression is.

Key words: Continuous time Markov chains, Life expectancy, Maximum Likelihood estimation, Mean Sojourn Time, Non-Alcoholic Fatty Liver Disease, Panel Data.

I. INTRODUCTION

Continuous time Markov chains (CTMC) are valuableand of great potentiality mathematical and statistical toolsto be used for evaluation of disease progression over time. CTMCs are a subtype of multistate models to be utilized to study this progression in NAFLD patients, with its characteristic phenotypes NAFLD and NASH, hand in hand with the presence of associated fibrosis and its stages. The prevalence of NAFLD is quickly growing worldwide, and matches the epidemics of obesity and type2 diabetes. Metabolic syndrome is a well-known risk factor which requires the presence of abdominal obesity distinguished by waist circumference >94 cm for males and >80 cm for females in eastern countries while it is >120 cm for males and >88 cm for females in the western countries, plus 2 or more of the following: blood glucose ≥100 mg/dL or drug treating diabetes, arterial blood pressure ≥ 130/85 mmgh or drug treating hypertension, triglyceride levels ≥150mg/dL or drug treating increased levels in blood or high density lipoprotein (HDL) levels <40 mg/dL for males and <50 mg/dL for females or drug treating this condition.

NAFLD can be modeled using the simplest form for health, disease, and death model, with one state for susceptible individuals with risk factors, such as: type 2 diabetes, dyslipidemia and hypertension, the other state is the NAFLD phenotypes, and two competing states for death: one for liver-related mortality as a complication of NAFLD, and the other death state is death causes unrelated to liver disease[1]. This is shown in figure 1:

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

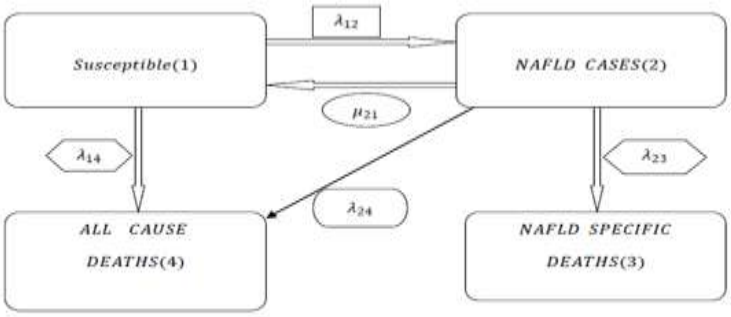

Figure 1: General Model Structure

In addition, NAFLD is modeled in more elaborative expanded form, which includes nine states: the first eight states are the states of disease progression as time elapses, while the ninth state is the death state[1], as illustrated in figure 2:

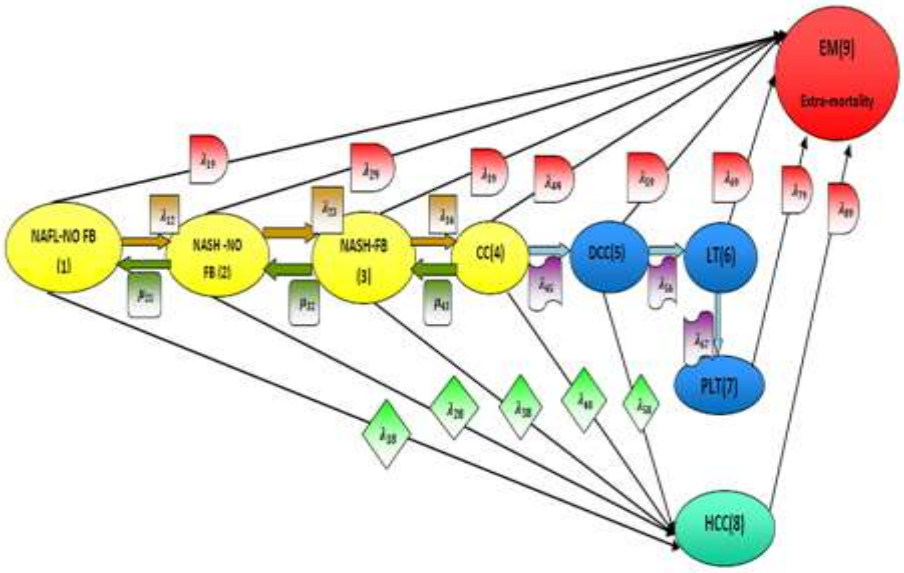

Figure 2: disease model structure:

NAFLD-NO FB: nonalcoholic fatty liver disease with no fibrosis (stage 1). NASH-NO FB: nonalcoholic steatohepatitis with no fibrosis (stage 2). NASH-FB: nonalcoholic steatohepatitis with fibrosis (stage 3). CC: compensated cirrhosis (stage 4). DCC: decompensated cirrhosis (stage 5). LT: liver transplant(stage 6). PLT: post liver transplant (stage 7). HCC:hepato-cellular carcinoma (stage 8). EM:extramortality (stage 9).

Moreover, a subset of the states that explicitly illustrates the phases of fibrosis process, which develops early in disease evolution cycle if the risk factors are not treated or eliminated, is modeled with CTMC to demonstrate: how covariates incorporated in a log-linear model can relate these predictors to transition rates among states, as illustrated in figure 3[2],[3]. The presence of fibrosis is considered an ominous predictor for disease progression. This subset is a subset of states from the expanded model especially early phases or stages where reversibility of conditions in each stage can be achieved if properly treated and controlled so as to prevent reaching the irreversible damaged state which is liver cirrhosis or F4.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

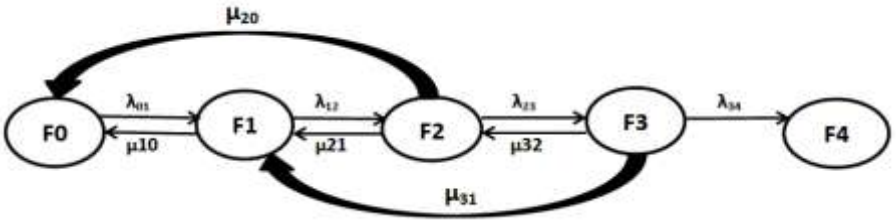

Figure 3: NAFLD with the evolving fibrosis stages.

F0= no fibrosis (stage 0) whether hepatic steatosis is present or not. NASH-FB-1: nonalcoholic steatohepatitis with mild fibrosis (stage 1). NASH -FB-2: NASH with moderate fibrosis (stage 2). NASH -FB-3: NASH with advanced or severe fibrosis (stage 3). CC: compensated cirrhosis (stage 4) which is the more severe or advanced form of fibrosis.

Singh et al. 2015 conducted a meta-analysis to evaluate the rate of fibrosis progression and thus searched multiple databases through a thoroughly systematic manner associated with author contact and found 11 cohort studies on NAFLD adult patients having at least one year apart paired liver biopsy specimens, from which they calculated a pooled-weighted annual fibrosis progression rate (number of stages changed between the 2 biopsy samples) with 95% confidence interval (CIs), and characterized the clinical risk factors accompanying this progression. They identified 411 patients with biopsy-proven NAFLD (150 with NAFL and 261 withNASH) included in those studies. Initially, the distribution of fibrosis for stages 0,1,2,3 and 4 was 35.8%, 32.5%, 16.7 %, 9.3% and 5.7% respectively, and over 2145.5 person-years of follow-up evaluation, 33.6% had fibrosis progression, 43.1% had stable fibrosis, and 22.3% had an improvement in fibrosis stage. The annual fibrosis progression rate in patients with NAFL who had stage 0 fibrosis at baseline was .07 stages (95% CI, 0.02-0.11 stages), compared with 0.14 stages in patients with NASH(95% CI, 0.07-0.21 stages). These findings correspond to 1 stage of progression over 14.3 years for patients with NAFL (95% CI, 9.1-50.0 y) and 7.1 years for patients with NASH (95% CI, 4.8-14.3 y).

Kalbfleisch and Lawless [4] related the instantaneous rate of transitions from state i to state j to covariates, by regression modeling of the Q transition ratematrix using log-linear model for the Markov rates.

In the present study, Poisson regression is used to model the rates among states. The counts of each transition can be modeled as a function of some explanatory variables reflecting the characteristics of the patients. This can be accomplished by using Poisson regression model or log-linear model. The Poisson regression model specifies that each response y_i is drawn from a Poisson population with parameter λ_i, which is related to the regressors or the covariates. The primary equation of the model is

$$P(Y = y_i|x_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}$$

The most common formulation for the λ_i is the log-linear model:

$$\ln \lambda_i = x_i'B$$

And the expected number of events per period is given by:

$$E[y_i|x_i] = var[y_i|x_i] = \lambda_i = e^{x_i'B}$$

The observed counts in the transition counts matrix is used as response variables and the covariates are the risk factors for fatty liver. Then the estimated counts obtained from the Poisson regression model are used to estimate the rates using the CTMC, as the initially observed transition rates approximately equal the estimated transition rates among states, as illustrated by the author in previous 2 papers, followed by exponentiation of the estimated rate matrix.To expound this procedure a hypothetical example is used, and it is in the form of a study conducted on 150 participants over 28 years to follow the progression of the NAFLD from F0 to F4.

The paper is divided into 3 sections. In section 1, illustration of the study design is clarified. In section 2, the results and discussion of running the Poisson regression model is elucidated. In section 3, conclusion of the running this model is expounded. Supplementary materials are complementary to this paper as some information are strictly presented in these materials and not in this main paper, such materials are table1,6,8,23, and figures from figure 13to figure 21.

1. Study Design

One hundred fifty participants were followed up every year for 28 years, and at each visit the characteristics of the participants were recorded like sex(0=female,1=male),age, BMI, LDL-chol, HOMA2_IR, systolic blood pressure as well as the diastolic pressure as shown in the table(1) (see

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

supplementary materails). For each participant the recorded value in the table is the mean of the follow up measurements. Fitting the Poisson regression and the estimated counts for each transition were calculated using Stata 14. A summary statistics for the patients' characteristics is shown in table(2). The participants were categorized according to these demographic characteristics as shown in table(3), while in table (4) summary of the categorical groups according to the participants' characteristics like: age category BMI category, LDL-chol category, systolic and diastolic blood pressure category. There are high correlations between the continuous predictor variablesas shown in table (5). In table (6) (see supplementary materials) the transition counts accomplished by each participant in these 28 years are illustrated. Summary of transition counts among the states in these 28 years is clarified in table (7).The timeline for each participant is shown in table(8)(see supplementary materials) with first column is t=0 and the last column is t=28 and in each of these column(year) the state of the patient was recorded.The observed transition counts are illustrated in table (9).

The distribution of the transition counts among the states is Poisson as illustrated in the following successive figures using Statgraphics-19 software. For transition from 0 to 1, see figure (4). For transition from 1 to 2, see figure (5). For transition from 2 to 3, see figure (6). For transition from 3 to 4, see figure (7). For transition from 1 to 0, see figure (8). For transition from 2 to 1, see figure (9). For transition from 3 to 2, see figure (10). For transition from 2 to 0, see figure (11). For transition from 3 to 1, see figure (12).

Table (2): statistical summary of the patients' characteristics

Variable	Observations	mean	Std. Dev.	Min	Max
Gender:	150				
Female=0	69(0.46)				
Male=1	81(0.54)				
Age	150	40.2	4.93	27	53
LDL-chol	150	94.81	15.41	59.89	133.1
HOMA2-IR	150	2.28	.71	.49	4.36
BMI	150	28.28	2.991	20.3	35.16
Sys.Bl.Pr.	150	149.73	10.434	123.4	175.75
Dias.Bl.Pr.	150	94.25	11.39	70	124

Table (3): table summarizing the categorical groups of patients according to the previous characteristics

Variable	Group1 (desirable)	Group2 (borderline)	Group3 (high)
Age	Age ≤ 35	35 < age ≤ 45	Age > 45
LDL-chol	LDL ≤ 70	70 < LDL < 100	LDL ≥ 100
HOMA2-IR	HOMA < 1.22	1.22 ≤ HOMA < 2.7	HOMA ≥ 2.7
BMI	BMI ≤ 25	25 < BMI < 30	BMI ≥ 30
Systolic blood pressure	Sys.Pr. ≤ 130	130 <Sys.Pr. < 160	Sys.Pr. ≥ 160
Diastolic blood pressure	Dias. Pr. ≤ 85	85 < Dias. Pr. < 100	Dias. Pr. ≥ 100

Table (4): summary of categorical groups of the patients' characteristics regarding age category, BMI category, LDL-chol category, systolic and diastolic blood pressure category:

category	Age			BMI			LDL-chol		
	Frequency	**Percent**	**Cum.**	**Frequency**	**Percent**	**Cum.**	**Frequency**	**Percent**	**Cum.**
1	22	14.67	14.67	22	14.67	14.67	5	3.33	3.33
2	109	72.67	87.33	83	55.33	70	93	62.00	65.33
3	19	12.67	100	45	30	100	52	34.67	100
total	150	100		150	100		150	100	
category	**HOMA2-IR**			**Systolic Blood Pressure**			**Systolic Blood Pressure**		
	Frequency	**Percent**	**Cum.**	**Frequency**	**Percent**	**Cum.**	**Frequency**	**Percent**	**Cum.**

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

1	10	6.67	6.67	4	2.67	2.67	33	22	22
2	93	62.00	68.67	123	82.00	84.67	69	46	68
3	47	31.33	100	23	15.33	100	48	32	100
total	100	100		150	100		150	100	

Table (5): correlation between continuous predictor variables

	age	**LDL-chol**	**HOMA2-IR**	**BMI**	**Sys. Bl.Pr.**	**Dias. Bl.Pr.**
Age	1					
LDL-chol	.9919	1				
HOMA2-IR	.9941	.9947	1			
BMI	.9938	.9948	.996	1		
Sys. Bl.Pr.	.9958	.9953	.9958	.9962	1	
Dias. Bl.Pr.	.9915	.9951	.9962	.9945	.9949	1

Table (7): summary transition counts between the states

Counts	**Transition 0→1**	**Transition 1→2**	**Transition 2→3**	**Transition 3→4**	**Transition 1→0**	**Transition 2→1**	**Transition 3→2**	**Transition 2→0**	**Transition 3→1**
0	63	96	121	128	121	127	130	138	139
1	58	43	23	22	24	17	17	11	9
2	25	9	4		3	5	3	1	2
3	4	2	2		2	1			

Table (9): Observed transitions counts of the patients over the 28 years

	State 0	**State1**	**State2**	**State3**	**State4**	**total**
State0	1909	120	15	6	0	2050
State1	36	1116	67	28	0	1247
State2	13	30	703	37	0	783
State3	11	14	23	50	22	120
State4	0	0	0	0	0	0
						4200

Initial observed rates are:

$$\lambda_{01} = \frac{120}{2050} = .059 \ , \lambda_{12} = \frac{67}{1247} = .0537 \ , \ \lambda_{23} = \frac{37}{783} = .047 \ , \lambda_{34} = \frac{22}{120} = .183$$

$$\mu_{10} = \frac{36}{1247} = .0288 \ , \mu_{21} = \frac{30}{783} = .0383 \ , \mu_{32} = \frac{23}{120} = .191 \ , \mu_{20} = \frac{13}{783} = .016 \ , \mu_{31} = \frac{14}{120} = .116$$

Using CTMC, the estimated rates approximately equal the initially observed rates, as illustrated by the author ImanAttiain previous 2 papers utilizing the simplest small model and the expanded model, where no covariates were included in the analysis.[5]

The distribution of the transition counts is Poisson as illustrated in the following figures using the Statgraphics-19 software.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

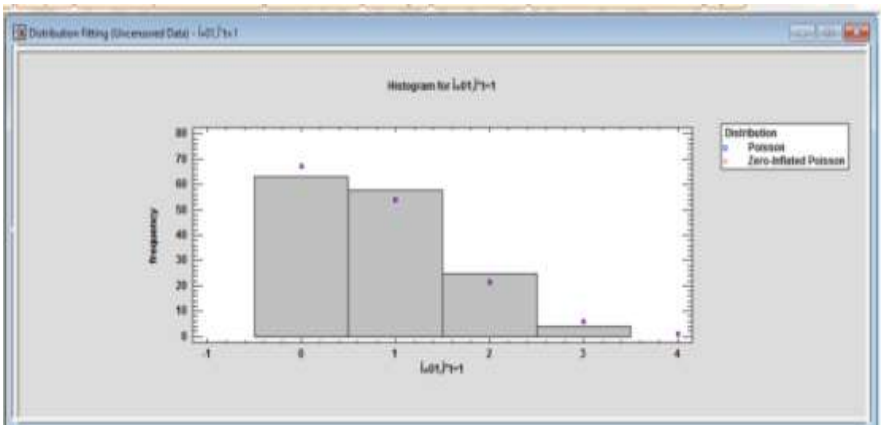

Figure 4: transition from 0 to 1

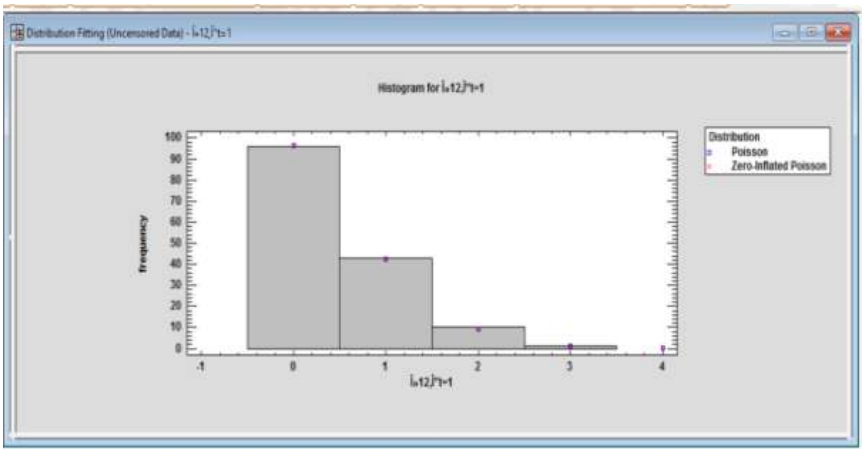

Figure 5: transition from 1 to 2

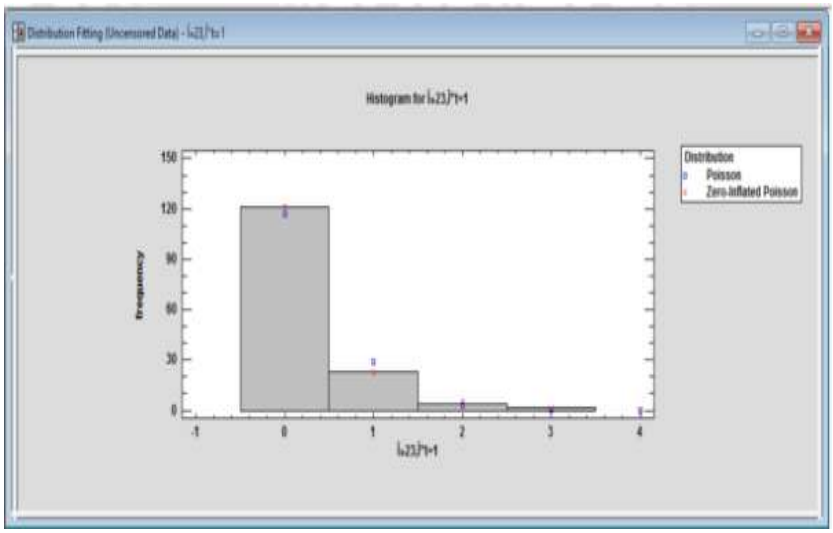

Figure 6: Transition from 2 to 3

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

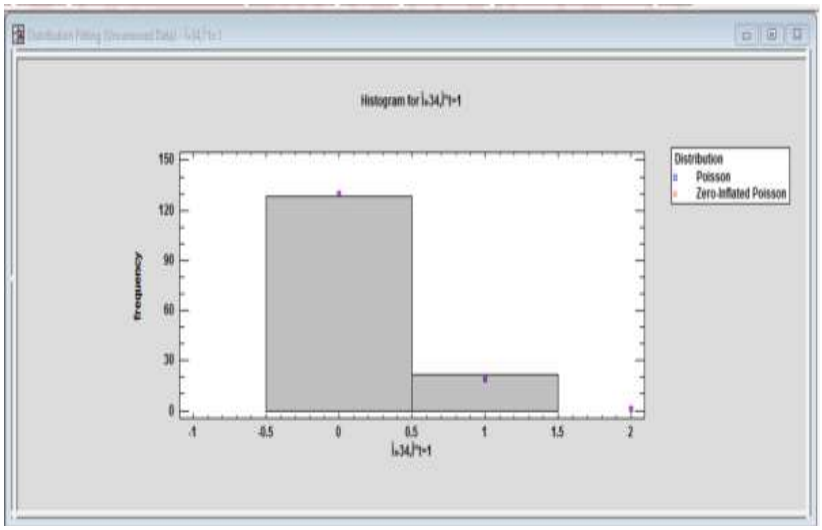

Figure7: transition from 3 to 4

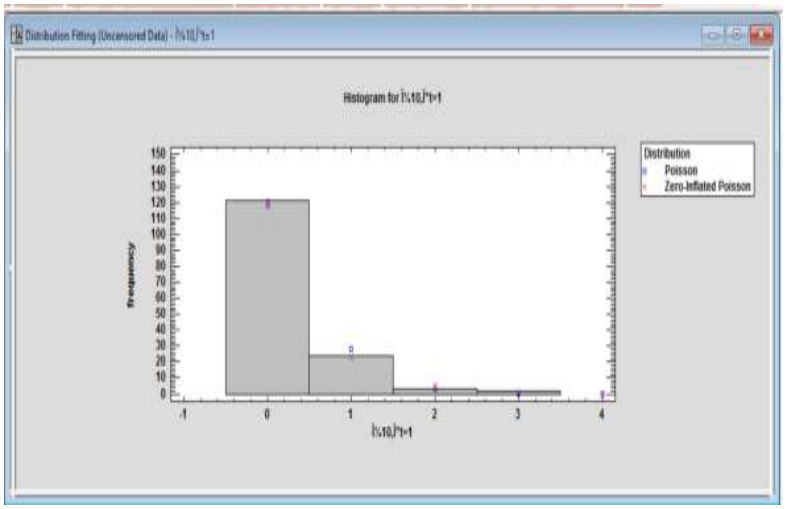

Figure 8: transition from 1 to 0

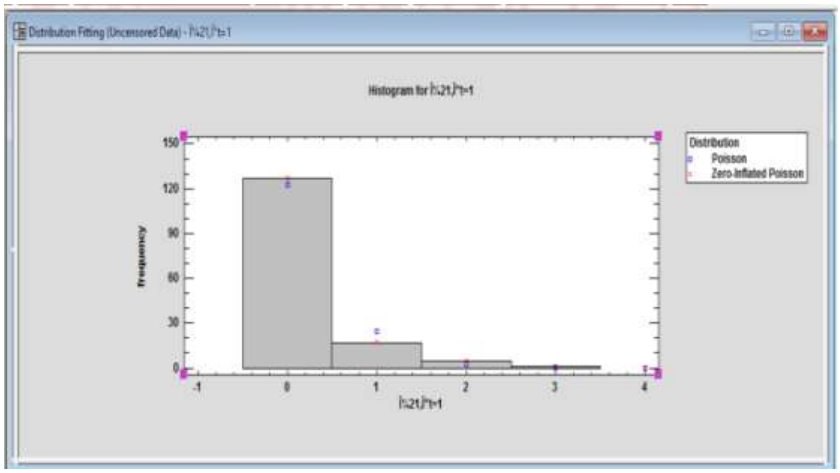

Figure 9: transition from 2 to 1

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

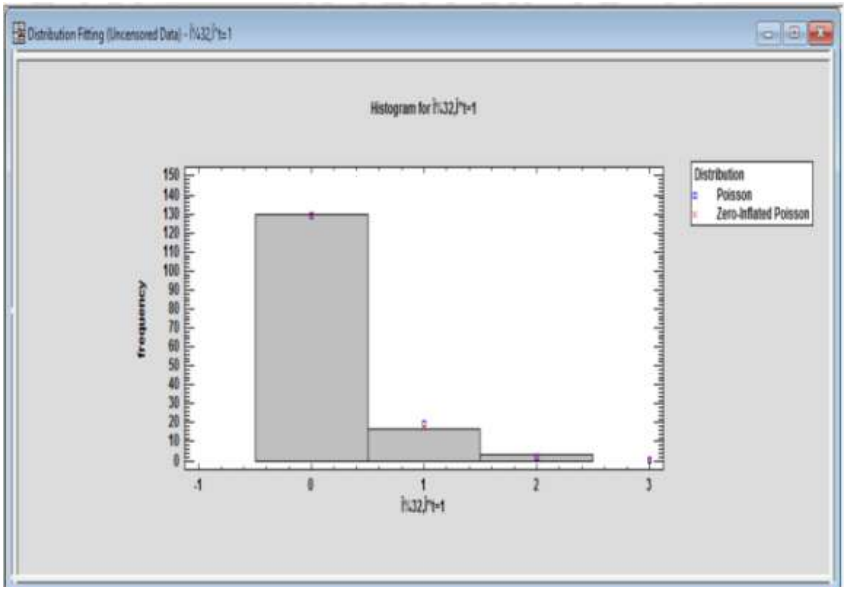

Figure 10: transition from 3 to 2

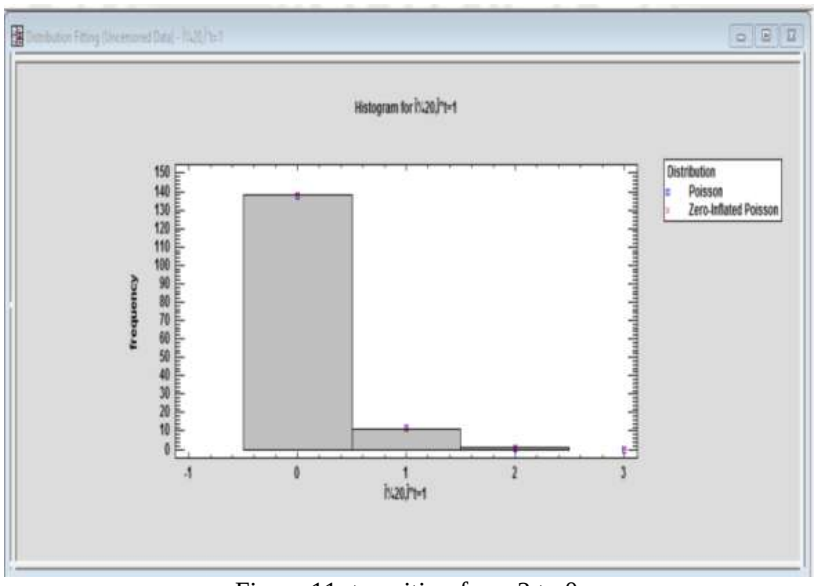

Figure 11: transition from 2 to 0

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

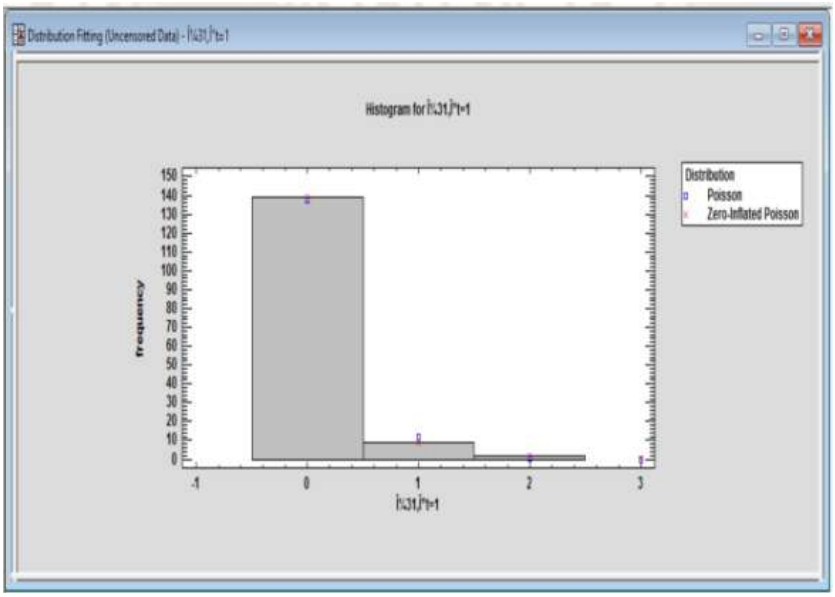

Figure 3:transition from 3 to 1

Lowess smoother illustrates that the relationships between each of the response rate and each variable is not strictly linear, but it is curvilinear relationship, with initial part of this relation being nearly horizontal and it starts to curve upwards at some predictor point located inside the second category of each predictor. The figures illustrating these relations are in supplementary materials from figure (13) to figure (21) for each response rate to the 7 variables. For example, relationship between number of transitions from state 0 to state 1 starts to bends up where each of the six predictors are located inside the second category; where age is approximately ≥37, BMI is approximately ≥ 26, LDL-chol is approximately ≥ 85 mg/dL, HOMA-IR is approximately ≥1.7, systolic blood pressure is approximately 142 mmHg, and diastolic blood pressure is approximately ≥ 85 mmHg. All these values are located in the second category. This can give good orientation to the functional form of the variables to be used in the regression model and avoid the misspecification resulting from mal-functional form of the predictors. In this work the restricted cubic splines are used for the predictors with 5 knots using Harrell approach which is the default procedure utilized by Stata14 software. The locations of knots are illustrated in table (10)and correlations between the transformed variables are presented in table (11).

Table (10): location of knots for specified variables using Harrell approach (the default used in Stata 14)

	Knot 1	Knot 2	Knot 3	Knot 4	Knot 5
LDL-chol	71.22	83.7	94.62	104.48	124.14
HOMA2-IR	1.09	1.8	2.26	2.75	3.48
sysBloodPr.	133.09	143.88	149.41	255.58	168.04
Dias b lood Pr.	74.45	87.44	94.07	101.11	114.49

Table (11): correlation between thetransformed variables used in the Poisson regression models

	LDLsp2	HOMAsp1	SYSsp2	HOMAsp2	DiasSP2
LDLsp2	1				
HOMAsp1	.8572	1			
sysSP2	.9959	.8674	1	.9908	
HOMAsp2	.9893			1	
DiasSP2	.9944		.9929	.995	1

The Poisson regression wasapplied using the observed counts of the transition counts matrix as response variable, and the following results are obtained as discussed below in the next section.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

2. Results and Discussion:

In the next discussion, the results of running Poisson regression to obtain the following estimated counts are demonstrated. Running Poisson regression on these transformed variables gives the estimated counts shown in table (13):

Table 13: the estimated counts for each transition

Counts	Transition 0→1	Transition 1→2	Transition 2→3	Transition 3→4	Transition 1→0	Transition 2→1	Transition 3→2	Transition 2→0	Transition 3→1
0	75	102	125	133	126	132	135	140	140
1	34	35	18	14	15	12	12	8	7
2	37	11	4	3	7	4	2	2	3
3	4	1	3	0	1	2	1	0	0
4	0	1	0	0	1	0	0	0	0
Total	150	150	150	150	150	150	150	150	150

The results for each transition are demonstrated in supplementarytables from table 14 up to table 22, and accompanied by discussion of these results to illustrate the importance of conducting such regression.[See supplementary materials for tables from table (14) to table (22) and associated discussion].

The comparison between the distribution of the response rates and the estimated rates is illustrated in table(23) (See supplementary materials).

As the estimated rates approximately equal the observed rates obtained by CTMC especially when using the initial rates calculated as $\theta_0 = \frac{n_{ijr}}{n_{i+}}$ where the n_{ijr} is the transition counts from state i to state j and the n_{i+} is the total marginal transition counts out of this state i , as verified by the author; ImanAttia in previous 2 papers, and assuming that the marginal counts are the same, so the estimated Q transition rate matrix according to the estimated counts obtained by fitting Poisson regression is:

$$Q = \begin{bmatrix} -.059 & .059 & 0 & 0 & 0 \\ .029 & -.080 & .051 & 0 & 0 \\ .015 & .033 & -.093 & .045 & 0 \\ 0 & .108 & .158 & -.409 & .167 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \; where$$

$$\lambda_{01} = \frac{120}{2050} = .059, \lambda_{12} = \frac{64}{1247} = .051, \; \lambda_{23} = \frac{35}{783} = .045,$$

$$\lambda_{34} = \frac{20}{120} = .167 \, , \mu_{10} = \frac{36}{1247} = .029 \; , \mu_{21} = \frac{26}{783} = .033$$

$$\mu_{32} = \frac{19}{120} = .158, \mu_{20} = \frac{12}{783} = .015 \, , \mu_{31} = \frac{13}{120} = .108$$

Probability transition matrix is obtained from exponentiating this Q matrix after 1 year:

$$\mathrm{P(t = 1)} = \begin{bmatrix} .9435 & .0551 & .0014 & 0 & 0 \\ .0274 & .9247 & .0469 & .0009 & .0001 \\ .0144 & .0327 & .9149 & .0348 & .0032 \\ .0023 & .0863 & .1245 & .6512 & .1357 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

To calculate goodness of fit for multistate model used in this model, it is like the procedure used in contingency table, and it is calculated in each interval and then summed:

Step1:$\mathrm{H_0} =$ future state does not depend on the current state$\mathrm{H_1} =$ future state does depend on the current state.

Step2:Calculatethe $\mathrm{p_{ij}(\Delta t = 1)} = \begin{bmatrix} .9435 & .0551 & .0014 & 0 & 0 \\ .0274 & .9247 & .0469 & .0009 & .0001 \\ .0144 & .0327 & .9149 & .0348 & .0032 \\ .0023 & .0863 & .1245 & .6512 & .1357 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

Using exponentiation of the estimated Q matrix

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

Step3: Calculate the expected counts in this interval by multiplying each row in the probability matrix with the corresponding total marginal counts in the observed transition counts matrix in the same interval to get the expected counts as in the following table (24)

Table(24): expected counts of transition

	State 0	State 1	State 2	State 3	State 4	Total
State 0	1934.175	112.955	2.87	0	0	2050
State 1	34.1678	1153.101	58.4843	1.1223	0.1247	1247
State 2	11.2752	25.6041	716.3667	27.2484	2.5056	783
State 3	.276	10.356	14.94	78.144	16.284	120
State 4	0	0	0	0	0	0

Step4: apply $\sum_{i=1}^{5}\sum_{j=1}^{5}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} = 1140.097 \sim \chi^2_{(5-1)(5-1)(.05)}$

Therefore,from the above results the null hypothesis is rejected while the alternative hypothesis is accepted and the model fits the data that is to mean the future state depends on the current state with the estimated transition rates and probability matrices as obtained.

Of those patients starting at F0 ,only 5.51% will move to F1 in one year, this declines to 4.69% of patients starting at F1 moving to F2 ,while 3.48% of patients starting at F2 will move to F3 ; however, 13.57% of patients starting in F3 will move to F4, and this high percentage of patients moving towards advanced fibrosis may be due to the fact that advanced fibrosis is considered to be F3 and F4 and once the patient reaches F3 , his chance to progress to F4 is higher than being in any starting stage considered less advanced fibrosis including F0 to F2 (by definition), and this is obvious as shown by incidence rate ratio of this transition being the highest (5.237e+6) . It is shown that progression from F0 to F1 and from F1 to F2 is approximately equal, while transition from F2 to F3 is less and this may be to more aggressive intervention taken by the patients to hinder the progression of fibrosis by applying more intensive lifestyle modifications, but once the patient reaches stage F3 the progression to F4 is by far the most among the forward transitions. There are 2.74% of patients starting at F1 will move to F0 while this percentage decreases to 1.44% if starting at F2, and it is even less if starting at F3 (only .23 % of patients can achieve this task); hence it is more feasible to move from F1 to F0 than to move from F2 to F0 than to move from F3 to F0; that is to mean, the more advanced the stage of fibrosis the patient experiences, the less likely movement to F0 he affords to do. There is a paradox if the starting stage is F2 or F3 to F1. The movement to F1 is more obvious if the patient is in F3(8.63% of patients move to F1) than if he is in F2 (3.27 % of patients move to F1); therefore, the more advanced fibrosis stage the patient recognizes , the more likely movement to F1 he can do, and may be this is due to the extensive lifestyle modification he performs to achieve less degree of fibrosis, but it remains a little bit difficult to reach F0 (only .23 % of patient can move from F3 to F0). It is also noted that 2.74% of patients move from F1 to F0 , 3.27% of patients move from F2 to F1 while 12.45% of patients move from F3 to F2 ; in other words the more advanced the fibrosis stage is, the more likely the movement to the immediately previous stage is. Moreover if the starting stage is F3, then 13.57% of these patients move to F4, a little bit higher than moving to F2 (12.45% of the patients); whereas, movement to F1 and F0 declines (8.63% of the patients and .23% of the patients respectively, approximately movement to F0 is 2.66% that to F1). Of those patients starting in F2, 3.48% move to F3, a little bit more than moving to F1 (3.27 % of patients); nevertheless, movement to F0 is almost 44% that to F1 (1.44% of the patients move to F0).

Mean time spent by the patient in state 0 is approximately 17 years that declines to 12 years and 6 months spent in state 1, which further declines to approximately 10 years and 9 months spent in state 2, and ultimately reaching 2 years and 3.7 months spent in state 3. It is shown that, there is decrease in time spent in each stage as the disease process evolves over time. This huge rapid decline in time spent in state 3 is due to advanced fibrosis induced by dead hepatocytes, especially if no treatment is introduced like:lifestyle modification ,risk factors treatment, as well as anti-inflammatory and anti-fibrotic drugs, and if so, it is a matter of time to reach state 4, which is irreversible stage of damaged liver cells that will soon manifest with reduction in liver cell functions, and may be to hepatocellular carcinoma, and eventually death, if not managed with liver transplantation.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 10 Oct 2021, pp: 1242-1254 www.ijaem.net ISSN: 2395-5252

CONCLUSIONS:

Insulin resistance is a key stone for triggering all these abnormalities, the more sensitive the body cells is to insulin, the less likely the complications of NALFD will develop. The effect of risk factors or covariates as a mainstay players, like: increased insulin resistance, hyperlipidemia with increased LDL-cholesterol, high systolic and diastolic blood pressure are thoroughly explained using the Poisson regression model combined with CTMC. As concluded from the hypothetical model that for every unit increase in the transformed HOMA, the incidence rate ratio for transition from state 0 to state 1 is increased by 5909.7% and this elevation is kept rising while moving forward from subsequent state to the immediately next state, that is to mean, for every unit increase in the transformed HOMA, the incidence rate ratio (IRR) for transition from state 1 to state 2 is increased by 24017.9%, while for the transition from state 2 to state 3, it is increased by 47931.8% , and for transition from state 3 to state 4 it is increased by 5237498.4%. This increment is almost always highly statistically significant. This is in comparison with transformed LDL, as for every unit increase in the transformed LDL, the IRR for transition from state 0 to state 1 is increased by 68.7%, while for the transition from state 1 to state 2, it is increased by 36.4% , and for transition from state 3 to state 4 it is increased by 57.1%. And it is only highly statistically significant for transition from state 3 to state 4. However the systolic blood pressure is almost highly statistically significant for the transition from state 2 to state 3 as obvious by for every unit increase in the transformed systolic pressure, the IRR for this transition to occur is increased by 1114.3%. Moreover, for every unit decrease in the transformed HOMA, the IRR for transition from state 1 to state 0 is increased by 1.1%,for transition from state 2 to state 1 it is increased by 3.7%, for transition from state 3 to state 2 it is increased by 0.5%, for transition from state 2 to state 0 it is increased by 6.6%, and for transition from state 3 to state 1 it is increased by 8.4%. This emphasizes that better control of insulin resistance helps the patient to reverse his condition. To sum up, the precipitating factors should be rigorously and extensively treated and controlled by life style modifications represented by dietary restriction of high calorie diet and sedentary life,thus the predisposed persons should consume healthy diets and regularly practicing physical exercises suitable for their medical conditions. The newly discovered drugs like anti-fibrotic drugs that treat the fibrotic changes in the liver are promising drugs and await further longitudinal studies, to reveal the most effective protocol, by which they are administered to the patients, for better control of the rate of progression of liver fibrosis. This control keeps the patient out of loss of liver functions, and subsequently away from end stage liver disease, which necessitates liver transplantation with all its accompanying post transplantation complications.

Hint (programs and supplementary materials):
The above example is published with Stata data, accompanied do file, as well as the supplementary materials file on the code ocean sit with the following URL :
Codeocean.com/capsule/4752445/tree/v1

Abbreviations:
CC:compensated cirrhosis (stage 4),CTMC: continuous time Markov chains, DCC:decompensated cirrhosis (stage 5),EM: extramortality (stage9),HCC:hepato-cellular carcinoma(stage 8),LT: liver transplant(stage 6),NAFLD: nonalcoholic fatty liver disease, NAFLD-NO FB: nonalcoholic fatty liver disease with no fibrosis (stage1),NASH: nonalcoholic steatohepatitis, NASH-NO FB: nonalcoholic steatohepatitis with no fibrosis (stage2), NASH-FB: nonalcoholic steatohepatitis with fibrosis (stage3),PLT : post liver transplant (stage7),T2DM: type 2 diabetes mellitus.

Declarations:
Ethics approval and consent to participate
Not applicable.
Consent for publication
Not applicable
Availability of data and material
Not applicable. Data sharing were not applicable to this article as no datasets were generated or analyzed during the current study.
Competing interests
The author declares that I have no competing interests.
Funding
No funding resource. No funding roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript are declared
Authors' contribution
I am the author who has carried the mathematical analysis as well as applying these mathematical statistical concepts on the hypothetical example.
Acknowledgement
Not applicable

REFERENCES

[1]. Younossi, Zobair M., Deirdre Blissett, Robert

Blissett, Linda Henry, Maria Stepanova, Youssef Younossi, Andrei Racila, Sharon Hunt, and Rachel Beckerman. 2016. "The Economic and Clinical Burden of Nonalcoholic Fatty Liver Disease in the United States and Europe." Hepatology 64(5):1577–86.

[2]. Younossi, Zobair M., Radhika P. Tampi, Andrei Racila, Ying Qiu, Leah Burns, Issah Younossi, and Fatema Nader. 2020. "Economic and Clinical Burden of Nonalcoholic Steatohepatitis in Patients with Type 2 Diabetes in the US." Diabetes Care 43(2):283–89.

[3]. Singh, Siddharth, Alina M. Allen, Zhen Wang, Larry J. Prokop, Mohammad H. Murad, and Rohit Loomba. 2015. "Fibrosis Progression in Nonalcoholic Fatty Liver vs Nonalcoholic Steatohepatitis: A Systematic Review and Meta-Analysis of Paired-Biopsy Studies." Clinical Gastroenterology and Hepatology 13(4):643–54.

[4]. Kalbfleisch, J. D., and Jerald Franklin Lawless. 1985. "The Analysis of Panel Data under a Markov Assumption." Journal of the American Statistical Association 80(392):863–71.