An Overview of Statistical Models for Recurrent Events Analysis: A Review

Chander Prakash Yadav[1]**, Sreenivas V** [2]**, Khan MA**[2] **and Pandey RM** [2*]

1

Appropriate non-survival approaches for recurrent event analysis

T ese methods can be used in such situations where information on time is not available or the time of the event does not play any role in addressing of research question. Amongst several approaches, two commonly used method for recurrent event analysis are Poisson regression and Negative Binomial regression. Although, recurrent event rate (number of events divided by follow up time for each individual) can be compared using Mann-Whitney U test but adjustment for several confounding variable is not feasible. T erefore, there was a need of a regression model where outcome of interest would number of event or event rate. Poisson regression [9] has come up to overcome this issue which models number of occurrences of an event or event rate as a function of some explanatory variables. Model parameters are estimated based on the principal of maximum likelihood method that provides reasonable good estimate for a parameter- as long as assumption of homogeneous event rate across the subject is valid. Validity of estimates derived from Poisson regression highly depends upon assumption of homogeneous event rate across individuals which is di f cult to achieve in practice. In general, we observe that there are some individuals who are more prone to develop recurrent events than others and assumption of homogeneity events rate gets violated and estimates from Poisson regression are no longer valid. For such situation another model have been used we call it as Negative binomial regression [9] which assume that each patient has recurrent events according to individual Poisson event rate and Poisson rates varies according to Gamma distribution across patients, because of it sometime we call it as Poisson gamma regression. T e phenomena of how negative binomial regressing gives better prediction than Poisson regression when assumption of uniform risk across subject is not valid was discussed by RJ Glynn et al. [5], basically they opted one example from several example discussed by Greenwood and Yule to illustrate the limitation of Poisson regression where propensity rate varies across individual [10]. T ey have shown distribution of number of accidents among 414 machinists (Table 1).

No. of Accidents	No. of Machinists	Expected Event(s)	
		Poisson Distribution	Negative binomial Distribution
0	296 (71.5)	256	299
1	74 (17.)	122	69
2	26 (6.2)	30	26
3	8 (1.9)	5	11
4	4 (1.0)	1	5
5	4 (1.0)	0	2
6	1 (0.002)	0	1
7	0	0	1
8	1 (0.002)	0	0

Table 1: Distribution of number of accidents among 414 machinists.

As, It is seen clearly Negative binomial regression gave better f t as compared to Poisson regression when homogeneous event is violated. Since, variance of negative binomial distribution is always greater than the variance of Poisson distribution, resulting to that negative binomial regression allow for more variability than Poisson regression. Despite many advantages it has few limitations, like it is di f cult to decide the distribution for di f erent propensities rate among individual, Gamma distribution generally used because it is easy to understand and easily approachable by the so f ware. But one should keep one thing in mind that Gamma distribution is not always an appropriate distribution for explaining di f erent propensities rates. Hence, it is advisable that one should try more than one distribution for estimating the propensities rate.

Appropriate survival approaches for recurrent event analysis

Whenever information on time is collected throughout the study and information on event time play an important role in addressing true research question, survival techniques are always better choice than non-survival techniques. For example one may be interested in knowing that whether the intervention is responsible for increasing time between successive events or what is protective e f ect of intervention on the rate of higher order events compared to control [11]. Over the last few decades many powerful survival methods have been invented for recurrent event data by extending Cox's proportional hazard regression, which can be categorized as: variance corrected models and Frailty models. Only di f erence between these two types of model is the way, they deal with within subject correlation.

Variance-corrected Models

In variance correction models [2], within-subject correlation due to heterogeneity is accounted by adjusting variance-covariance matrix using grouped jackknife estimator and correlation due to event dependency is accounted by constructing di f erent risk set [12] which are based on di f erent consensus which one is better) are used in Poisson regression [17,18] while robust group jackknife correction is used in AG model [19]. In general, the Poisson regression with correction for over dispersion had similar coverage probabilities of confidence interval, but slightly higher parametric equation is used for estimating frailty term for the estimation of within subject correlation while in case of AG model within subject correlation is accommodated by adjusting variance covariance matrix. Standard frailty model is computationally very intense required much larger time than AG model and interpretation of frailty model is also not so straightforward. Generally, frailty model is interpreted as keeping frailty term constant across individuals, which is intuitively not acceptable for many researchers.

Conditional frailty model

Many times, it is di f cult to distinguish among sources of within subject correlation i.e. whether it is because of event dependency or heterogeneity or both. In view of this, frailty term was added into PWP-GT model so that within subject correlation due to either of sources could be accommodated in the model and new model is known as Conditional frailty model. Basically, idea was, within subject correlation due to event dependency will be accommodated by conditional nature of model (i.e. a subject is not at the risk for m^{th} event until he/she experience their $(m-1)^{th}$