

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

Chander Prakash Yadav¹, Sreenivas V², Khan MA² and Pandey RM^{2*}

¹ICMR-National Institute of Malaria Research (NIMR), New Delhi, India

²Department of Biostatistics, AIIMS, New Delhi, India

*Corresponding author: Pandey RM, Professor & Head, Department of Biostatistics, AIIMS, New Delhi, India, Tel: + 09811912117; E-mail: rmpandey@yahoo.com

Received date: September 17, 2018; Accepted date: October 10, 2018; Published date: October 19, 2018

Copyright: © 2018 Yadav CP, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Recurrent event may be explained as similar to acute event that can occur to an individual more than once, are common in medical research. Recurrent event data possess two specific characteristics: within subject correlation and time-varying covariate. Because of these underlying characteristics, traditional statistical techniques do not work well on recurrent event data. Over the last few decades, many models proposed for the analysis of recurrent event data that can be grouped into two categories: Survival methods for recurrent event analysis and Non-survival methods for recurrent event analysis. This review intended to discuss primarily survival techniques for recurrent event analysis while non-survival techniques, discussed here, in order to complete the picture.

Despite the fact that many appropriate alternatives are available for recurrent event data, most of researchers are still using traditional statistical techniques for the analysis of recurrent data and loss in terms of, appropriately addressing the true research question and inefficient estimation of standard error associated with parameter estimate. There may be two possible reasons behind this either they are not aware of these newly developed techniques or they are not able to select one among several available techniques.

Hence, in this review all important aspects related to the analysis of recurrent events data (such as what is recurrent event data, characteristic of recurrent event data, why traditional statistical techniques are naïve techniques for the analysis of recurrent event data, appropriate available alternatives for the analysis of recurrent event data and how to select one over others based on research question and available data) are discussed in detail.

Keywords: Recurrent event; Extended Cox; AG; PWP; Frailty; Conditional frailty; Survival analysis; Repeated event

Introduction

Generally, two types of events are encountered in health research: non-reversible events and reversible events. Non-reversible events are those events which are chronic in nature and occur to an individual only once (e.g. hypertension, AIDS, diabetes and cystic fibrosis). While other types of events are reversible event which are acute in nature and can occur to an individual more than once. Reversible event can further be bifurcated as multiple events and recurrent events. Multiple events are those repeated events that are not of exactly same type but somewhat related, such as repeated hospitalization due to different reasons (hospitalization due to road accident, hospitalization due to fall, hospitalization due to fever, etc.). Unlike multiple events, recurrent events are those repeated event, which are of same type such as acute exacerbations in asthmatic children, seizures in epileptics, low back pain in women, skin cancer, myocardial infarctions, migraine pain, and sports injuries.

Recurrent events data have two main characteristics viz. within subject correlation and time varying covariates. Recurrent event within subject are very unlikely independent, they are related and this phenomenon is known as within-subject correlation and there are two possible sources this within subject correlation viz. within-subject

correlation due event dependency and with subject correlation due to heterogeneity. Within-subject correlation due to event dependency refers to a situation where an event itself accelerates or decelerates the rate of subsequent event [1]. For example, as the first heart attack occurs to a subject, chances of happening second heart attack become increase because during the first heart attack some part of the heart get damaged. Within subject correlation due to heterogeneity refers to the situation wherein some subjects are more prone to experiencing a larger number of events than other subjects because of some unknown, unmeasured or immeasurable reasons [1]. This phenomenon also causes within-subject correlation. Proper adjustment of within-subject correlation (either of source) is essential in order to correct estimation of standard error; if we treat correlated observation as uncorrelated we would overstate the amount of information each observation provided leading to incorrect estimates of standard errors [2].

Another important concern related to recurrent event analysis is how to deal with time-varying covariates. In many studies there are some covariates which are subject to change over time for example in case of asthma management, dose and type of drugs are subject to change during course of time which have direct effect on outcome (asthma exacerbation in this case).

Over the last few decades, a lot of statistical advancement took place in the field of recurrent event data analysis and several approaches have proposed for the analysis of recurrent event data. These newly developed techniques are far better than traditional statistical

techniques (such as t-test, logistic regression, multiple linear regression, Cox's Proportional Hazard regression) for addressing recurrent event process more appropriately. Despite several powerful techniques available for analysis of the recurrent event data, most researchers are still using traditional statistical techniques for analyzing their research questions where outcome of interest is recurrent in nature. In a systematic review, Meghan et al. [3] have revealed that less than one-third of 83 research articles having outcome of interest as a recurrent event- namely fall in elderly, used an appropriate statistical method [3]. Application of sub-optimal suitability could lead to loss in terms of internal validity and precision of the results. Possible explanations of why most researchers continue to use naive statistical techniques despite availability of appropriate alternatives are as, either they are not aware of these techniques because most of these techniques are discussed in specific literature which are hard to understand for those who are not from statistical or mathematical background [4] or possible no clear cut guidelines are available regarding selection of an appropriate alternative, based on research question and type of data.

Therefore, aims of this review are to describe the several aspects of recurrent events analysis as to what are recurrent events; characteristics of recurrent events; why traditional statistical techniques are not appropriate choice for the analysis of recurrent event data; what all appropriate alternatives are available for recurrent event analysis; and how to select one over the other approach based on research question and nature of data.

Why conventional statistical methods are not appropriate contender for the analysis of recurrent events data

We start our discussion with t-test, which is used for comparing number of recurrent events between two populations in cohort studies or treatment arm and control arm in RCT but often it is observed that there are some subjects who are more prone to experience larger number of events than others which may distort the assumption of normality and result in inappropriate estimation of standard error. In this situation, one can go for the upgraded non-parametric counterpart i.e. Mann Whitney (Wilcoxon's Ranksum) test which does not require normality assumption. As we are aware, none of the observational study are free from confounding effect, the only way to assess confounding in t-test and rank-sum test is to do subgroup analysis, which is feasible up to two-three confounding factors only. If we have more number of confounding factors, we have to go for multiple regression but in case of recurrent event analysis, assumption that residuals follow normal distribution get violated. Also, multiple regression assumes uniform risk across the events but in case of recurrent event analysis, risk of subsequent events may be different from risk of previous events. Additionally, there is no way to deal with time varying covariate in this procedure.

Another traditional contender which is used most frequently for the analysis of recurrent event data is logistic regression. Logistic regression divides all subjects into two groups as those who experienced any event and those who did not experience any event and then the proportion of subjects with and without any event are compared between treatment/exposure groups, adjusting for confounding variables. Such an analytical treatment to the data makes a subject that experienced only one event during the follow up equal to another subject that experienced more than one event. In other words, logistic regression does not distinguish the subjects with different number of events and puts them all in one basket, ignoring the number

of events in the analysis. This is a grossly inadequate description of the recurrent event process. Secondly the logistic regression is unable to accommodate time-dependent covariate naturally in the analysis (either baseline information or end line information are used in analysis) which is again lead to incomplete and/or inappropriate description of the process.

Cox's Proportional Hazard (PH) regression is another traditional technique, used most frequently for the analysis of recurrent event data whenever information on event time is available. It is a type of survival technique which is somewhat better than logistic regression if information of time is collected and time plays an important role in addressing true research question. Though it has the ability to handle time varying covariate, it is still not appropriate for recurrent event analysis. This is because it uses information up to first event only, and all information after first event is not used in analysis. It also avoids the methodological complication that can occurs if first event is not representative of subsequent events or risk of first event affects risk of a sub-sequent event [5,6]. Thus, consideration of only first events may lead to an inaccurate evaluation of the efficacy of a treatment. In particular, it can substantially underestimate potential benefits in terms of event prevented by a treatment.

Till now we have discussed inappropriate contender for recurrent event analysis, now we would discuss how each piece of information is critical for addressing true research question. This was well explained by RJ Glynn et al. [5]. They have analyzed a data of clinical trial in three ways, where objective was to assess effect of cranberry juice on bacteriuria and pyuria [7]. They had randomized 153 patients into two groups (treatment arm: consume 300 ml/day cranberry juice and control arm: consume 300 ml/day placebo which was indistinguishable in taste, appearance). After randomization six clean voided urine samples were collected at roughly monthly interval. And primary outcome variable of interest was bacteriuria (organism numbering $\geq 10^5$ /ml, regardless of organism) with pyuria in a given study month [5]. In first analysis, proportion of bacteriuria and pyuria between two groups was compared and no statistically significance difference was observed. As a second analysis, these two groups were compared based on first event only and still no significance difference was observed. Finally, all urine samples collected throughout the study were considered and compared between study groups using the method given by Zeger and Liang [8] and substantial difference between cranberry group and placebo group was observed. Further investigation revealed that discrepancy between rates of first event and overall rates between groups because women in the cranberry group were far more likely to recover from their bacteriuria-pyuria than women in the placebo group. The average one month probability of change from a bacteriuric-pyuric sample to a non-infected sample was 0.54 in the cranberry group and 0.28 in the placebo group ($P=0.006$). Thus, restriction of analyses to only first events would have obscured important clinical differences in this trial [5].

Some appropriate methods for recurrent event analysis

In this section, we are discussing methods that were developed for recurrent event analysis. These methods can be categorized into two categories: non-survival methods for recurrent event analysis and survival methods for recurrent events analysis.

Appropriate non-survival approaches for recurrent event analysis

These 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. Therefore, 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 difficult 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. The 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]. They 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 fit 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 difficult to decide the distribution for different propensities rate among individual, Gamma distribution generally used because it is easy to understand and easily approachable by the software. But one should keep one thing in mind that Gamma distribution is not always an appropriate distribution for explaining different 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 effect 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 difference 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 different risk set [12] which are based on different risk interval [12]. Under variance corrected models, varieties of model have been discussed in literature such as Andersen and Gill (AG) model [13], Wei, Lin and Weissfeld (WLW) model [14], Prentice, Williams and Peterson-Counting Process (PWP-CP) model [15], Gap time-unrestricted (GT-UR) model [12]; Total time-restricted (TT-R) model [12] and Multi-state models [16]. Amongst these most widely used methods in field of recurrent event analysis are as AG, WLW, PWP-CP, PWP-GT, standard frailty and conditional frailty model each of which is briefly described below with their pros and cons, so that an user can select the appropriate model when dealing with an outcome variable which is recurrent in nature.

Anderson-Gill (AG) Independent Increment Model

Anderson and Gill [13] model is the simplest extension of Cox's proportional hazard regression using counting process time interval [12,13]. It assumes that recurrent events within subjects are independent and share common baseline hazard. It gives more efficient estimates of regression coefficient than traditional Cox's proportional hazard regression as long as assumption of independent event within subject is valid. AG model shows similarity with Poisson regression because both of these are based on independent increment assumption. AG model has the advantage over Poisson regression whereas Poisson regression can only be used for uniform risk over time while AG model can also be used for non-constant but proportional hazard risk [9] as well. In practice independent increment assumption (i.e. recurrent event within subject are independent) is hardly fulfill. To overcome of this, some corrective measures are used for both the methods. Deviance correction and Pearson correction (but no

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 type I error rates compared to the robust Andersen-Gill model [9]. Another advantage with AG model is, being a survival approach it uses more information (event as well as time of the event) resulted to that AG model address research question more appropriately than Poisson regression.

It could be better understood with the help of an example. Suppose the primary aim in a placebo controlled trial is to reduce the number of asthma exacerbations or increase time gap between consecutive exacerbations. Let a subject 'A' from treatment group has experienced two exacerbations at month 6 and month 15 and was followed till month 18 without further exacerbations. Let another subject 'B' from placebo group also experienced two exacerbations on month 6 and month 34 and was followed up month 50 without further exacerbations. In this case, both Poisson regression and Negative Binomial regression do not make any distinction between these two subjects and underestimate the effect of drug. In such situations AG model gives better result than these two. Another attractive feature of AG model is that it can incorporate time-varying covariates in the model, especially when time-varying covariates can distort the true relationship between outcome and explanatory variables, such as smoking status which can change over time. This information is naturally accommodated in AG model to draw valid estimates.

Prentice, Williams and Peterson (PWP) Conditional Model

In 1981, Prentice, Williams and Peterson [15] had proposed two models for recurrent event analysis which are considered as first extension of Cox's proportional hazard regression. Unlike to AG model which assumed that recurrent events within a subject are independent and baseline risk are same for all events, both PWP models assumes that recurrent event within subject are related and baseline hazard is vary from event to event. For example, baseline risk for second heart attack is always higher than baseline risk of first heart attack because during first heart attack some part of heart got damage. This feature of analysis where baseline risk is vary from event to event and occurrence of subsequent event is affected by previous event is very well incorporated in both PWP models. A subject is not at risk for m^{th} event until he/she experienced his/her $(m-1)^{\text{th}}$ event at time 't' (risk set of second event consider only those subjects who have already experienced their first event at time 't'). PWP models have one additional advantage over other models, since they have event-specific baseline hazard, so one can estimate an overall effect or event specific effect for each covariate. These two PWP models are exactly same, the only difference between them is, first one is based on counting process time interval and known as Prentice, Williams and Peterson-Counting Process (PWP-CP) model while second one is based on gap time interval and known as Prentice, Williams and Peterson-Gap time (PWP-GT) model. PWP-CP model can be used if one is interested in knowing the effect of intervention on the outcome variable from the beginning of the study while PWP-GT should be used if one is interested in knowing effect from previous event(s). Though both models are very appropriate for recurrent event analysis, they suffer from some limitations. One limitation is that they can give unreliable estimates for higher order events because as the event order increases, number of subjects in the risk set is decreased.

Wei, Lin and Weissfeld (WLW) Marginal Model

WLW [14] is the only variance-corrected model which can be applied to multiple failures of same type of events or multiple failures of different types of events as well. It considers each recurrence as a separate process and there is no ordering among events within subject. For example, during neonatal intensive care unit (NICU) stay, a neonate is at the risk of several events simultaneously such as infection due to gram positive organism, infection due to gram negative organism, necrotizing enterocolitis, meningitis, jaundice, and diarrhea etc. Each of these can occur more than once in any order. WLW simultaneously analyzes time to first/second/third/more incident detection of several types of events either in same or different clinical visit separately. Risk set for m^{th} event in WLW model contains all individuals who have not yet experienced their m^{th} event and remain in follow up at time t. For example, risk set for second event would contain all the individual who have not experience their 2nd events and remained in the follow up at time t, in other words risk set includes those who have not experienced any event and are in follow up at time t; and those who have experienced only one event and are in follow up at time t. This model provides reliable estimates of regression coefficient when data do not follow specific ordering, especially more useful in competing event analysis rather than recurrent event analysis because in general recurrent event follow an order. Whereas order matter it exaggerated the true effect because it allow a subject to be at risk for several time for same event. Because of this characteristic, many researchers [2,12] have criticize the use of this model in the field of recurrent event analysis.

Frailty Models

Frailty models [20,21] are another class of models, extended on traditional Cox's proportional hazard model. Contrary to variance corrected models, frailty models assumes that correlation among recurrent events is due to tendency that some individuals are more prone to develop recurrent event as compared to others because of some unobserved/unknown factors. They may be their socio-demographic factors, environment factors, behavioral factors or genetic factors. Many times these factors are unknown to researcher and hard to accommodate into analysis.

In order to correct estimation of regression parameters, a frailty term, which follow a specific frailty distribution, is directly incorporate into the model estimation while in case of variance corrected model, it is done by adjusting variance covariance matrix. Estimates from frailty distribution may be more efficient than variance corrected models if we could correctly identify the frailty distribution. Presently there is no guidelines, how to select appropriate frailty distribution for a given scenario. In general, Gamma distribution is most commonly used in the estimation of frailty term because of its easy to use feature. Other distributions for frailty estimation are normal, log normal and uniform distribution.

Standard frailty model

Standard frailty model is the simplest frailty extension of Cox's proportional hazard model and it is very similar to AG model. Like AG model it also assumes that there is no within subject correction due to event dependency and whatever correlation is present among recurrent events is only because of heterogeneity. Similar to AG model, common baseline hazard for all events is assumed in standard frailty model but frailty term is directly incorporated into model and structure

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 difficult 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 $(m-1)^{\text{th}}$ event) and within subject correlation due to heterogeneity will be accommodated by incorporating frailty term in model estimation process itself. This model is relatively newer and till now Box-Steffensmeier et al. [1,22] have only worked on it. Hence some more work is needed to know the actual efficacy of this model on real data.

Conclusion and Recommendation

Choice of appropriate alternative always depends upon, types of data we have and research question of interest. If event time information has not collected or does not add anything to address research question, one should go for non-survival approach. Between Poisson and Negative regression, the latter is preferable because it allows for more variability. If event time information plays a role in addressing a research question, the obvious choice is survival models. Under these, we can choose simply AG model if we are confident that there is no correlation among recurrent event within subject. Otherwise one should prefer PWP-GT and conditional frailty model over AG, PWP-CT and standard frailty model because these two models address recurrent event process naturally by assuming that recurrent events are not independent and risk of subsequent event is not same as previous event.

References

1. Box-Steffensmeier JM, De Boef S (2006) Repeated events survival models: the conditional frailty model. *Stat Med* 25: 3518-3533.
2. Box-Steffensmeier JM, Zorn C (2002) Duration Models for Repeated Events. *J Polit* 64: 1069-1094.
3. Donaldson MG, Sobolev B, Cook WL, Janssen PA, Khan KM (2009) Analysis of recurrent events: a systematic review of randomised controlled trials of interventions to prevent falls. *Age Ageing* 2009; 38: 151-155.
4. Twisk DJWR (2003) *Applied longitudinal data analysis for epidemiology*. Cambridge University Press.
5. Glynn RJ, Buring JE (1996) Ways of measuring rates of recurrent events. *BMJ* 312: 364-367.
6. Glynn RJ, Buring JE (2001) Counting recurrent events in cancer research. *J Natl Cancer Inst* 93: 488-489.
7. Avorn J, Monane M, Gurwitz JH, Glynn RJ, Choodnovskiy I, et al. (1994) Reduction of bacteriuria and pyuria after ingestion of cranberry juice. *JAMA* 271: 751-754.
8. Liang K-Y, Zeger SL (1986) Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika* 73: 13-22.
9. Jahn-Eimermacher A (2017) Comparison of the Andersen-Gill model with poisson and negative binomial regression on recurrent event data. *Computational Statistics & Data Analysis* 52: 4989-4997.
10. Greenwood M, Yule GU (1920) An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents. *J R Stat Soc* 83: 255-279.
11. Kuramoto L, Sobolev BG, Donaldson MG (2008) On reporting results from randomized controlled trials with recurrent events. *BMC Med Res Methodol* 8: 35.
12. Kelly PJ, Lim LL (2000) Survival analysis for recurrent event data: an application to childhood infectious diseases. *Stat Med* 19: 13-33.
13. Andersen PK, Gill RD (1982) Cox's Regression Model for Counting Processes: A Large Sample Study. *Ann Stat* 10: 1100-1120.
14. Wei LJ, Lin DY, Weissfeld L (1989) Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. *J Am Stat Assoc* 84: 1065-1073.
15. Prentice RL, Williams BJ, Peterson AV (1981) On the regression analysis of multivariate failure time data. *Biometrika* 68: 373-379.
16. Andersen PK, Keiding N (2002) Multi-state models for event history analysis. *Stat Methods Med Res* 11: 91-115.
17. McCullagh P, Nelder JA (1989) *Generalized Linear Models*. London: Chapman and Hall.
18. Keene ON, Calverley PMA, Jones PW, Vestbo J, Anderson JA (2008) Statistical analysis of exacerbation rates in COPD: TRISTAN and ISOLDE revisited. *Eur Respir J* 32: 17-24.
19. David G. Kleinbaum DG, Klein M (2012) *Survival Analysis: A self Learning Text., Second.* Springer.
20. Therneau TM, Grambsch PM, Pankratz VS (2003) Penalized Survival Models and Frailty. *J Comput Graph Stat* 12: 156-175.
21. Oakes D (2017) *Frailty Models For Multiple Event Times - Springer.* *Survival Analysis: State of the Art* 211: 371-379.
22. Box-Steffensmeier JM, De Boef S, Joyce KA. Event Dependence and Heterogeneity in Duration Models: The Conditional Frailty Model. *Polit Anal* 2007; 15: 237-256.