Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Survival analysis in longitudinal studies for recurrent events: Applications and challenges

Teerapong Aramruang

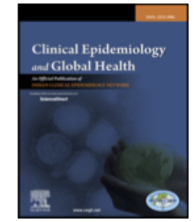
Wisdom of the Land

Literature

Contents lists available at ScienceDirect

Clinical Epidemiology and Global Health

journal homepage: www.elsevier.com/locate/cegh



Survival analysis in longitudinal studies for recurrent events: Applications and challenges

Mani Thenmozhi[a], Visalakshi Jeyaseelan[a,*], Lakshmanan Jeyaseelan[a], Rita Isaac[b], Rupa Vedantam[c]

[a] Department of Biostatistics, Christian Medical College, Vellore, India
[b] Department of RUSHA, Christian Medical College and Hospital, Vellore, India
[c] Department of ENT Unit 3, Christian Medical College & Hospital, Vellore, 632004. India

Literature
(cont.)

IJEA

International Journal of Epidemiology, 2015, 324–333
doi: 10.1093/ije/dyu222
Advance Access Publication Date: 9 December 2014
Original article

Education Corner

Modelling recurrent events: a tutorial for analysis in epidemiology

Leila DAF Amorim[1]* and Jianwen Cai[2]

[1]Department of Statistics, Institute of Mathematics, Federal University of Bahia, Brazil and [2]Department of Biostatistics, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

*Corresponding author. Departmento de Estatística, Instituto de Matemática, Universidade Federal da Bahia, Av.Adhemar de Barros s/n, Campus de Ondina, Salvador-Bahia, Brazil CEP 40170-115. E-mail: leiladen@ufba.br

Accepted 27 October 2014

Introduction

Types of events in health research

Non-reversible events

- Chronic in nature
- Occur to an individual only once
- E.g., Hypertension, AIDS, Diabetes, and Cystic fibrosis

Reversible events

- Acute in nature
- Occur to an individual more than once

<u>Multiple events</u>

- Repeated events, which are **not exactly same type but somewhat related**
- E.g., repeated hospitalization due to different reasons (hospitalization due to road accident, hospitalization due to fall, hospitalization due to fever)

<u>Recurrent events</u>

- Repeated event, which are **the same type**
- E.g., acute exacerbations in asthmatic children, seizures in epileptics, low back pain in women, skin cancer, myocardial infarctions, migraine pain, and sports injuries

Int J Epidemiol. 2015;44(1):324-33; Epidemiology: Open Access. 2018;08

Introduction
(cont.)

Characteristics of recurrent events

Within subject correlation

Event dependency

- An event itself accelerates or decelerates the rate of subsequent event.
- E.g., 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.

Heterogeneity

- Some subjects are more prone to experiencing a larger number of events than other subjects because of some unknown, unmeasured or immeasurable reasons.

Time varying covariates

- Covariates whose value can change during follow-up.

Int J Epidemiol. 2015;44(1):324-33; Epidemiology: Open Access. 2018;08

Introduction
(cont.)

Why conventional statistical methods are not appropriate?

T-test or Mann Whitney (Wilcoxon's Rank-sum) test

- Limited up to 2-3 confounding factors
- Violated normal distribution and uniform risk across the events
- Unable to accommodate time-dependent covariate

Logistic regression

- Not distinguishing the subjects with different number of events and puts them all in one basket, ignoring the number of events
- Unable to accommodate time-dependent covariate

Cox proportional hazard regression

- Usually uses information up to first or last event only, and all information after first event is not used in analysis.
- First event is not representative of subsequent events or risk of first event affects risk of a sub-sequent event.

Int J Epidemiol. 2015;44(1):324-33; Epidemiology: Open Access. 2018;08

Introduction
(cont.)

Available approaches for modeling recurrent events

Non-survival approaches

- Poisson regression
- Negative Binomial regression

Survival approaches

Means/Rates models

- Lin, Wei, Yang, and Ying (JRSS-B, 2000)

Hazards/Intensity models

❖ **Variance correction methods**

- Andersen-Gill (AG) model (Annals of Statistics, 1982)
- Prentice, Williams and Peterson (PWP) model (Biometrika, 1981)
- Wei, Lin and Weissfeld (WLW) Marginal model (JASA, 1989)

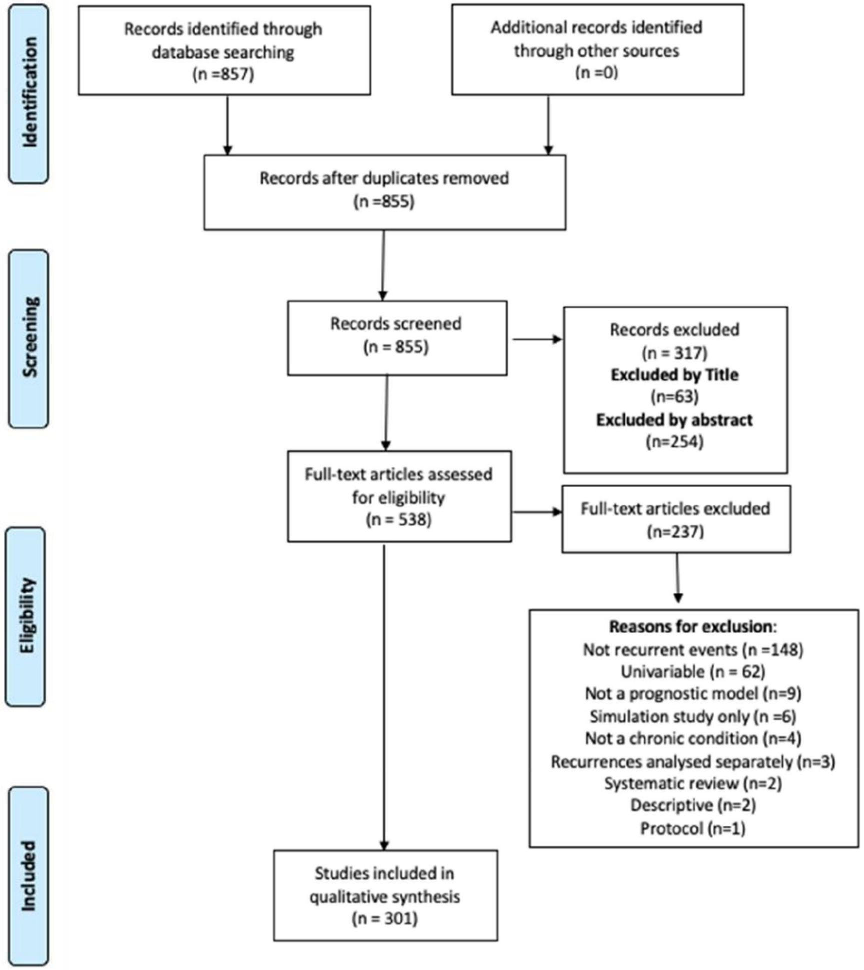
❖ **Frailty methods**

- Standard frailty model (Lifetime Data Anal, 1995)
- Conditional frailty model (Stat Med, 2006)

Commonly used models

Systematic review

- **Aims:** To identify and describe existing methodology being applied for the development and validation of prediction models for recurrent event outcome data.
- MEDLINE, inception to 24 October 2019
- **Inclusion:** (1) developed or validated a multivariable prediction model for recurrent event data predicting the risk of future recurrences, (2) included both the number of recurrent events and the timing between them as part of the model.
- **Exclusion:** (1) analyzed the time to the first event, (2) analyzed only the number of events using a Poisson or Negative Binomial model, (3) considered only one prognostic factor.



Diagn Progn Res. 2024;8(1):13

Commonly used models (cont.)

Systematic review

Table 1 Summary of methods identified from the data extraction

Method	Number (%) of included studies
Recurrent event methods	
Andersen-Gill (AG) [26]	152 (50.5%)
Frailty Model [27]:[a]	116 (38.5%)
Gamma	63 (20.9%)
Unspecified	35 (11.6%)
Gaussian	18 (6.0%)
Log-Normal	15 (5.0%)
Weibull	10 (3.3%)
Exponential	8 (2.7%)
Log-Logistic	3 (1.0%)
Poisson	3 (1.0%)
Compound Poisson	1 (0.3%)
Gompertz	1 (0.3%)
Logistic	1 (0.3%)
Prentice, Williams and Peterson Models [29]:[b]	41 (13.6%)
Prentice, Williams and Peterson-Total Time (PWP-TT)	27 (9.0%)
Prentice, Williams and Peterson-Gap Time (PWP-GT)	22 (7.3%)
Wei, Lei and Weissfeld (WLW) [30]	33 (11.0%)
Bayesian Methods	11 (3.7%)
Multi-State Model (MSM)	9 (3.0%)
Lin, Wei, Ying and Yang (LWYY) [31]	2 (0.7%)
Lee, Wei and Amato (LWA) [32]	1 (0.3%)
Lawless and Nadeau marginal model (LN) [33]	1 (0.3%)
Liang, Self and Chang (LSC) [34]	1 (0.3%)
Multilevel Survival Model [35]	1 (0.3%)
Papers which used multiple recurrent event methods	48 (15.9%)

[a] Some papers applied more than one type of frailty model
[b] Some papers applied both the PWP-TT and PWP-GT variation

Table 13 Frequency of clinical area recurrent event methods applied in

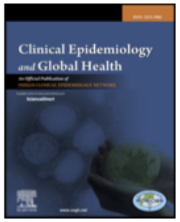
Clinical area	Number (%) of included studies[1,2]
Cardiology	62 (20.6%)
Oncology	45 (15.0%)
Human immunodeficiency virus (HIV)/Acquired immunodeficiency syndrome (AIDS)	30 (10.0%)
Mental health	24 (8.0%)
Neurology	21 (7.0%)
Kidney disease	13 (4.3%)
Respiratory	11 (3.7%)
Drugs & alcohol abuse	9 (3.0%)
Elderly people & accidents	9 (3.0%)
Infectious diseases	9 (3.0%)
Other illness	9 (3.0%)
Sexually transmitted infection (STI)/Sexually transmitted disease (STD)	9 (3.0%)
Paediatrics	8 (2.7%)
Haematology	5 (1.7%)
Hospital admissions	5 (1.7%)
Arthritis	4 (1.3%)
Diabetes	4 (1.3%)
Maternal health	4 (1.3%)
Chronic injuries	3 (1.0%)
Surgeries	3 (1.0%)
Bacterial infections	2 (0.7%)
Gastroenterology	2 (0.7%)
Optometry	2 (0.7%)
Osteoarthropathy	2 (0.7%)
Autoimmune Disease	1 (0.3%)
Gynecology	1 (0.3%)
Inflammatory bowel disease	1 (0.3%)
Ophthalmology	1 (0.3%)
Podiatry	1 (0.3%)

[1] Some studies applied recurrent event analysis models to more than one clinical area
[2] Results are sorted in descending frequency

Diagn Progn Res. 2024;8(1):13

Literature

Contents lists available at ScienceDirect
Clinical Epidemiology and Global Health
journal homepage: www.elsevier.com/locate/cegh



Survival analysis in longitudinal studies for recurrent events: Applications and challenges

Mani Thenmozhi[a], Visalakshi Jeyaseelan[a,*], Lakshmanan Jeyaseelan[a], Rita Isaac[b], Rupa Vedantam[c]

[a] Department of Biostatistics, Christian Medical College, Vellore, India
[b] Department of RUSHA, Christian Medical College and Hospital, Vellore, India
[c] Department of ENT Unit 3, Christian Medical College & Hospital, Vellore, 632004. India

❖ **Objective:** To summarize various methods for modelling recurrent event data and would show the differences in estimation and interpretation of recurrent event approaches, as well as to sensitize appropriate models, based on research objectives for the longitudinal study.

Materials and methods

Dataset*

Objective

- To study the incidence and age-related prevalence and **risk factors associated with upper respiratory infection (URI)** in a birth cohort of Indian **infants** in the first year of life

Study design

- Prospective cohort study
- Conducted in Christian Medical College, Vellore, India
- Between February 2009 to August 2010

Participants

- **210 babies with newborn** were recruited between February 2009 and August 2009.

Data collection

- Patient information was obtained from their parents.
- Socio-demographic history and related risk factors (e.g., smoke exposure, daily care)
- At birth and at **monthly scheduled visits**

Outcome

- **URI** using nasopharyngeal swabbing with a calcium alginate swab stick
- Criteria: mucoid or mucopurulent secretion in the nasal cavity + mother's history of the child having a runny nose with or without cough or fever + not feeding well

Clinical Epidemiology and Global Health. 2019;7(2):253-60
* International Journal of Pediatric Otorhinolaryngology. 2012;76(12):1835-9

Materials and methods
(cont.)

Models

Standard methods

- Cox proportional hazard model

Variance correction methods

- Andersen-Gill (AG) model
- Prentice, Williams and Peterson-Counting Process (PWP-CP) model
- Prentice, Williams and Peterson-Gap time (PWP-GT) model
- Wei, Lin and Weissfeld (WLW) Marginal model

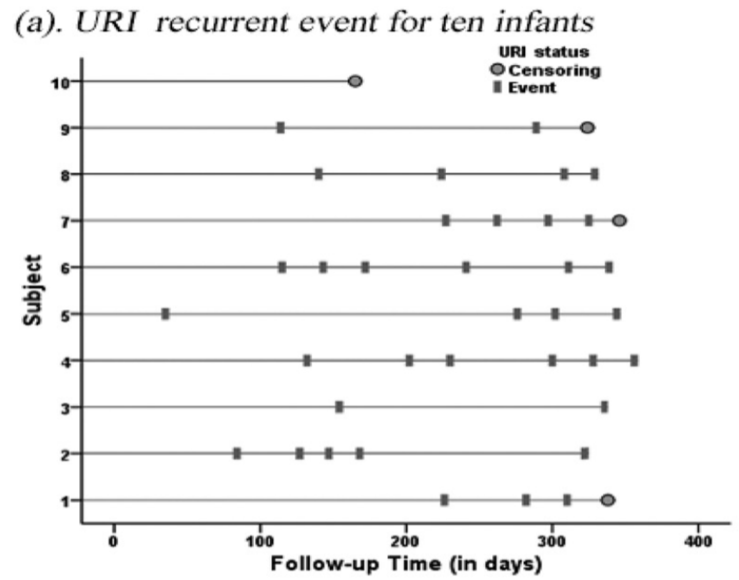
Frailty methods

- Standard frailty model

Results

Rate of URI and visits

- **17 infants (8.1%)** from 210 infants did not return to the study after birth.
- Number of recurrence experienced by infants **ranged between 0 to 10** during the follow-up.
- The URI recurred at least **once in 193 subjects** and highest recurrence events **(9 and 10 times) were observed in 7 patients.**
- A total of 163 infants (77.6%) had 6-13 visits whereas 30 infants (14.3%) made < 5 visits.
- The **median number of visits** for these 193 infants was **9 visits.**
- In thousand days of life, 845 records from 747 upper respiratory patients were followed-up during the study period and **three infants died during the period of the study.**



Results
(cont.)

Socio-demographic and baseline characteristics

Variables	Baseline (n = 210)
	n%
Sex	
Male	121 (57.6)
Female	89 (42.4)
Type of House	
Thatched	70 (33.3)
Tiled/Terraced/Group House	140 (66.7)
Parental Occupation	
Nil/Laborer	130 (61.9)
Petty Business/Professional/Others	80 (38.1)
Father's Education	
Illiterate/Primary	32 (15.2)
High/Higher Secondary and above	178 (84.8)
Mother's Education	
Illiterate/Primary	28 (13.3)
High/Higher Secondary and above	182 (86.7)
Birth weight (Grams)	
≤ 2500	76 (36.2)
> 2500	134 (63.8)
Smoke	
Yes	15 (7.1)
No	195 (92.9)
No. of members in the house	
≤4	48 (23.1)
> 4	160 (76.9)
Firev	
Yes	85 (40.5)
No	125 (59.5)
Water	
Bore well	124 (59.0)
River/Open Well	86 (41.0)
Nasophryngeal Swab Report	
Positive	8 (3.8)
Negative	201 (96.2)

Clinical Epidemiology and Global Health. 2019;7(2):253-60

Results
(cont.)

Socio-demographic and baseline characteristics by URI recurrent events

Variable	Event 1	Event 2	Event 3	Event 4	Event 5
	n (%)	n (%)	n (%)	n (%)	n (%)
Sex					
Female	85 (41.9)	73 (41.7)	62 (42.2)	50 (41.3)	93 (46.7)
Male	118 (58.1)	102 (58.3)	85 (57.8)	71 (58.7)	106 (53.3)
Type of house					
Tiled/Terraced/Grouped Houses	136 (67.0)	120 (68.6)	104 (70.7)	82 (67.8)	134 (67.3)
Thatched	67 (33.0)	55 (31.4)	43 (29.3)	39 (32.2)	65 (32.7)
Occupation					
Farmer/Bigbusiness/Petty business	75 (36.9)	63 (36.0)	52 (35.4)	39 (32.2)	54 (27.1)
Nil/Labourer	128 (63.1)	112 (64.0)	95 (64.6)	82 (67.8)	145 (72.9)
Father Education					
High school/Secondary and above	173 (85.2)	150 (85.7)	126 (85.7)	102 (84.3)	164 (82.4)
Illiterate/Primary	30 (14.8)	25 (14.3)	21 (14.3)	19 (15.7)	35 (17.6)
Mother Education					
High school/Secondary and above	186 (91.6)	161 (92.0)	135 (91.8)	111 (91.7)	180 (90.5)
Illiterate/Primary	17 (8.4)	14 (8.0)	12 (8.2)	10 (8.3)	19 (9.5)
Birth weight					
> 2.5 kg	130 (64.0)	109 (62.3)	90 (61.2)	71 (58.7)	116 (58.3)
≤ 2.5 kg	73 (36.0)	66 (37.7)	57 (38.8)	50 (41.3)	83 (41.7)
Smoking					
No	189 (93.1)	163 (93.1)	138 (93.9)	112 (92.6)	180 (90.5)
Yes	14 (6.9)	12 (6.9)	9 (6.1)	9 (7.4)	19 (9.5)
Mem5					
≤ 4	188 (93.5)	173 (98.9)	145 (98.6)	120 (99.2)	196 (98.5)
> 4	13 (6.5)	2 (1.1)	2 (1.4)	1 (0.8)	3 (1.5)
Fire					
No	123 (60.6)	107 (61.1)	91 (61.9)	75 (62.0)	101 (50.8)
Yes	80 (39.4)	68 (38.9)	56 (38.1)	46 (38.0)	98 (49.2)
Water					
Bore well	121 (59.6)	106 (60.6)	95 (64.6)	78 (64.5)	123 (61.8)
Open well/River	82 (40.4)	69 (39.4)	52 (35.4)	43 (35.5)	76 (38.2)
Swab					
Negative	139 (68.8)	102 (58.3)	80 (54.4)	83 (68.6)	166 (83.4)
Positive	63 (31.2)	73 (41.7)	67 (45.6)	38 (31.4)	33 (16.6)
Season(Months)					
March to June	24 (11.8)	28 (16.0)	42 (28.6)	56 (46.3)	128 (64.3)
July to October	118 (58.1)	56 (32.0)	23 (15.6)	8 (6.6)	21 (10.6)
November to February	61 (30.0)	91 (52.0)	82 (55.8)	57 (47.1)	50 (25.1)

Clinical Epidemiology and Global Health. 2019;7(2):253-60

Results
(cont.)

Risk factors for URI recurrent events in the first year of life

Variable	Model 1 (AG Model)			Model 2 (PWP Total Time Model)			Model 3 (PWP Gap time Model)			Model 4 (Marginal Model)			Model 5 (Cox Frailty Model)		
	HR	95% CI	P Value	HR	95% CI	P Value	HR	95% CI	P Value	HR	95% CI	P Value	HR	95% CI	P Value
Season															
March–June	1.00			1.00			1.00			1.00			1.00		
July to October	2.27	1.70–3.03	< 0.001	2.60	1.14–5.94	0.023	2.22	1.66–3.00	< 0.001	1.58	1.05–2.37	0.027	2.30	1.84–2.86	< 0.001
November to February	1.43	1.19–1.71	< 0.001	1.50	1.12–2.02	0.007	1.37	1.11–1.69	0.003	2.50	1.11–1.69	< 0.001	1.44	1.17–1.76	< 0.001
Sex															
Male	0.95	0.83–1.09	0.492	0.93	0.86–1.01	0.098	0.93	0.81–1.07	0.308	1.09	0.81–1.07	0.504	0.95	0.82–1.11	0.530
Female	1.00			1.00			1.00			1.00			1.00		
Swap															
Positive	1.23	1.07–1.42	0.003	1.18	1.08–1.39	0.039	1.13	0.97–1.31	0.127	1.49	1.14–1.95	0.003	1.22	1.03–1.44	0.014
Negative	1.00			1.00			1.00			1.00			1.00		
Mem5															
≤4	1.01	0.48–2.16	0.966	1.02	0.49–2.11	0.954	0.91	0.42–1.97	0.811	0.97	0.42–1.97	0.849	1.66	0.84–3.31	0.960
> 4	1.00			1.00			1.00			1.00			1.00		
Smoking															
Yes	1.04	0.86–1.25	0.693	1.04	0.83–1.30	0.737	1.05	0.84–1.31	0.658	0.99	0.84–1.31	0.963	1.04	0.79–1.37	0.790
No	1.00			1.00			1.00			1.00			1.00		
Water															
Open well/River	0.97	0.84–1.14	0.748	0.96	0.80–1.16	0.691	1.01	0.87–1.16	0.947	1.08	0.87–1.16	0.562	0.97	0.83–1.12	0.750
Borewell	1.00			1.00			1.00			1.00			1.00		
Fire															
Yes	1.14	0.98–1.33	0.087	1.15	1.04–1.27	0.004	1.18	1.02–1.35	0.022	1.22	1.02–1.35	0.131	1.14	0.98–1.34	0.920
No	1.00			1.00			1.00			1.00			1.00		
Father Education															
Illterate/Primary	1.07	0.89–1.28	0.483	1.07	0.92–1.25	0.372	1.02	0.86–1.23	0.790	1.01	0.86–1.23	0.960	1.07	0.87–1.32	0.540
High school & above	1.00			1.00			1.00			1.00			1.00		
Mother Education															
Illterate/Primary	0.97	0.74–1.28	0.850	0.89	0.71–1.12	0.330	0.99	0.76–1.31	0.977	0.78	0.96–1.27	0.223	0.98	0.75–1.27	0.850
High school & above	1.00			1.00			1.00			1.00			1.00		
Birth_weight															
≤2.5 kg	1.11	0.96–1.28	0.151	1.14	1.04–1.27	0.009	1.11	0.96–1.27	0.145	1.30	1.08–1.27	0.035	1.11	0.95–1.30	0.190
> 2.5 kg	1.00			1.00			1.00			1.00			1.00		
Parent's Occupation															
Nil/Labourer	1.11	0.95–1.30	0.177	1.11	0.97–1.27	0.122	1.08	0.91–1.27	0.367	0.92	0.86–1.16	0.576	1.10	0.94–1.29	0.190
Professional & Others	1.00			1.00			1.00			1.00			1.00		
Type of House															
Pucca/Kacha	0.99	0.85–1.16	0.963	1.01	0.82–1.23	0.943	0.99	0.86–1.16	0.975	0.98	0.86–1.16	0.908	1.02	0.87–1.20	0.960
Tiled/Terraced/Grouped house	1.00			1.00			1.00			1.00			1.00		
Frailty Variance													0.00		0.870
Log likelihood	−3712.08			−2578.19			−2918.27			−5627.17			−3706.00		
R Square	0.083			0.092			0.072			0.129			0.083		
		4			1			5			2			3	

Clinical Epidemiology and Global Health. 2019;7(2):253-60

Results
(cont.)

AG model and frailty model comparison

- **AG model**

$$\lambda_i(t) = \lambda_0(t)\exp\{\beta_k\, x_i(t)\}$$

- **Frailty model**

$$\lambda_{ik}(t) = \lambda_{0k}(t) Z_i \{\exp\{x_i(t)\beta_k\},\ \ t > 0$$

Frailty **Zi** is the unobserved (random) common risk factors shared by all subjects in cluster 'i' and is assumed to be i.i.d random variable with unit mean and unknown variance θ.

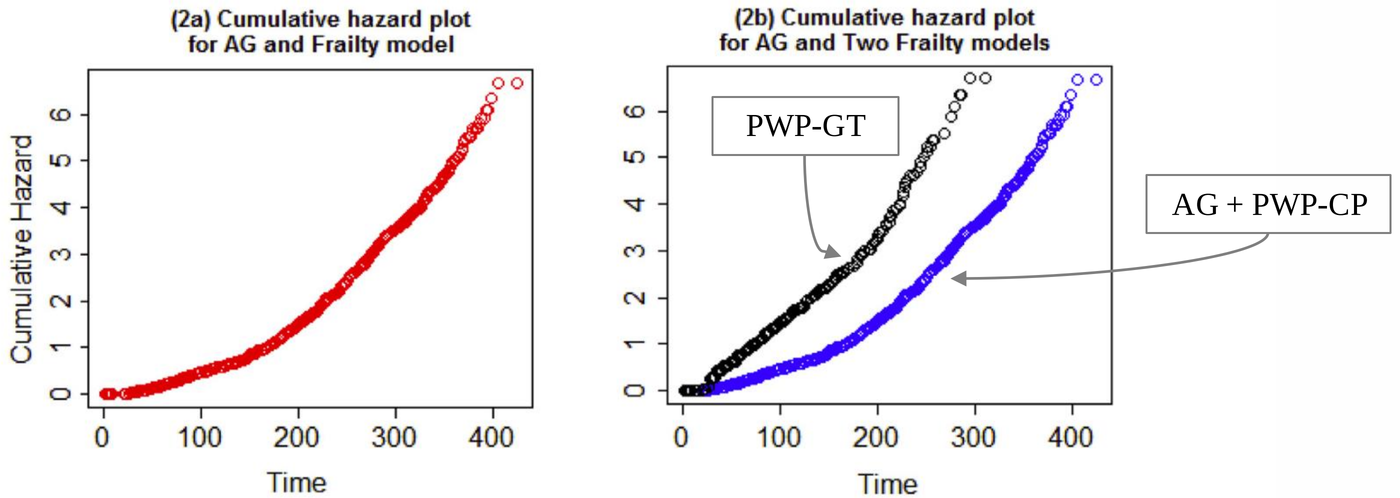


Fig. 2. Cumulative hazard plot for upper respiratory infection recurrence over a time of follow-up for AG model and frailty models.

Results
(cont.)

Risk factors for URI recurrent events in the first year of life

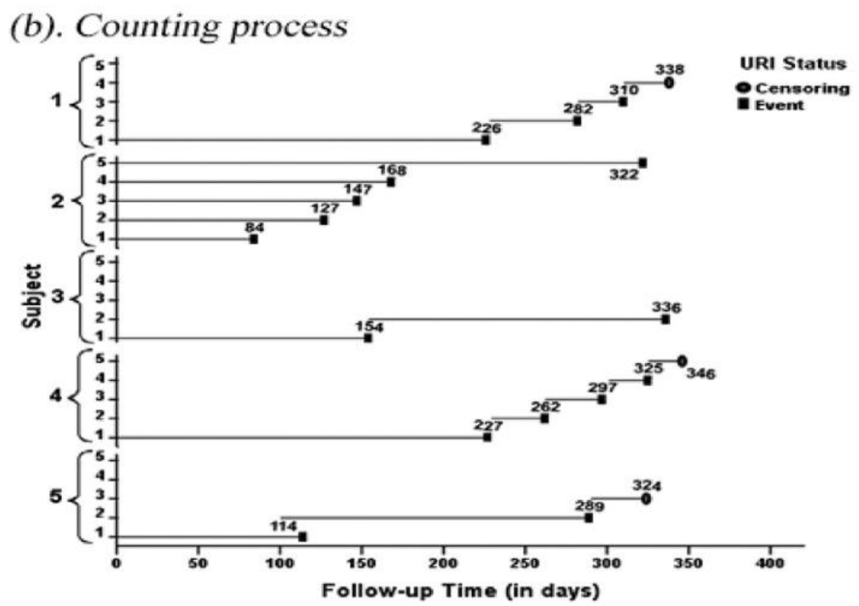
Variable	Model 1 (AG Model)			Model 2 (PWP Total Time Model)			Model 3 (PWP Gap time Model)			Model 4 (Marginal Model)			Model 5 (Cox Frailty Model)		
	HR	95% CI	P Value	HR	95% CI	P Value	HR	95% CI	P Value	HR	95% CI	P Value	HR	95% CI	P Value
Season															
March–June	1.00			1.00			1.00			1.00			1.00		
July to October	2.27	1.70–3.03	< 0.001	2.60	1.14–5.94	0.023	2.22	1.66–3.00	< 0.001	1.58	1.05–2.37	0.027	2.30	1.84–2.86	< 0.001
November to February	1.43	1.19–1.71	< 0.001	1.50	1.12–2.02	0.007	1.37	1.11–1.69	0.003	2.50	1.11–1.69	< 0.001	1.44	1.17–1.76	< 0.001
Sex															
Male	0.95	0.83–1.09	0.492	0.93	0.86–1.01	0.098	0.93	0.81–1.07	0.308	1.09	0.81–1.07	0.504	0.95	0.82–1.11	0.530
Female	1.00			1.00			1.00			1.00			1.00		
Swap															
Positive	1.23	1.07–1.42	0.003	1.18	1.08–1.39	0.039	1.13	0.97–1.31	0.127	1.49	1.14–1.95	0.003	1.22	1.03–1.44	0.014
Negative	1.00			1.00			1.00			1.00			1.00		
Mem5															
≤4	1.01	0.48–2.16	0.966	1.02	0.49–2.11	0.954	0.91	0.42–1.97	0.811	0.97	0.42–1.97	0.849	1.66	0.84–3.31	0.960
> 4	1.00			1.00			1.00			1.00			1.00		
Smoking															
Yes	1.04	0.86–1.25	0.693	1.04	0.83–1.30	0.737	1.05	0.84–1.31	0.658	0.99	0.84–1.31	0.963	1.04	0.79–1.37	0.790
No	1.00			1.00			1.00			1.00			1.00		
Water															
Open well/River	0.97	0.84–1.14	0.748	0.96	0.80–1.16	0.691	1.01	0.87–1.16	0.947	1.08	0.87–1.16	0.562	0.97	0.83–1.12	0.750
Borewell	1.00			1.00			1.00			1.00			1.00		
Fire															
Yes	1.14	0.98–1.33	0.087	1.15	1.04–1.27	0.004	1.18	1.02–1.35	0.022	1.22	1.02–1.35	0.131	1.14	0.98–1.34	0.920
No	1.00			1.00			1.00			1.00			1.00		
Father Education															
Illterate/Primary	1.07	0.89–1.28	0.483	1.07	0.92–1.25	0.372	1.02	0.86–1.23	0.790	1.01	0.86–1.23	0.960	1.07	0.87–1.32	0.540
High school & above	1.00			1.00			1.00			1.00			1.00		
Mother Education															
Illterate/Primary	0.97	0.74–1.28	0.850	0.89	0.71–1.12	0.330	0.99	0.76–1.31	0.977	0.78	0.96–1.27	0.223	0.98	0.75–1.27	0.850
High school & above	1.00			1.00			1.00			1.00			1.00		
Birth_weight															
≤2.5 kg	1.11	0.96–1.28	0.151	1.14	1.04–1.27	0.009	1.11	0.96–1.27	0.145	1.30	1.08–1.27	0.035	1.11	0.95–1.30	0.190
> 2.5 kg	1.00			1.00			1.00			1.00			1.00		
Parent's Occupation															
Nil/Labourer	1.11	0.95–1.30	0.177	1.11	0.97–1.27	0.122	1.08	0.91–1.27	0.367	0.92	0.86–1.16	0.576	1.10	0.94–1.29	0.190
Professional & Others	1.00			1.00			1.00			1.00			1.00		
Type of House															
Pucca/Kacha	0.99	0.85–1.16	0.963	1.01	0.82–1.23	0.943	0.99	0.86–1.16	0.975	0.98	0.86–1.16	0.908	1.02	0.87–1.20	0.960
Tiled/Terraced/Grouped house	1.00			1.00			1.00			1.00			1.00		
Frailty Variance													0.00		0.870
Log likelihood	−3712.08			−2578.19			−2918.27			−5627.17			−3706.00		
R Square	0.083			0.092			0.072			0.129			0.083		

Clinical Epidemiology and Global Health. 2019;7(2):253-60

Discussion

Andersen-Gill (AG) model

- Assumes: Recurrent events within subject are **independent and share common baseline hazard.**
- The risk of the repeated infections remains constant, irrespective of the number of previous infections.
- This assumption is usually untenable.



Clinical Epidemiology and Global Health. 2019;7(2):253-60; Int J Epidemiol. 2015;44(1):324-33; Epidemiology: Open Access. 2018;08

Discussion
(cont.)

Andersen-Gill (AG) model

❖ **Data structure and analysis (R program)**

Study ID	Start	Stop	URI Status	Gap	Sex	Swab	Months	URI count
1	0	226	1	226	1	0	3	1
1	226	282	1	56	1	0	3	2
1	282	310	1	28	1	1	3	3
1	310	338	0	28	1	0	1	4
2	0	84	1	84	1	0	1	1
2	84	127	1	43	1	0	2	2
2	127	147	1	20	1	0	2	3
2	147	168	1	21	1	0	2	4
2	168	322	1	154	1	0	3	5
3	0	132	1	132	2	1	1	1
3	132	202	1	70	2	1	2	2
3	202	230	1	28	2	1	2	3
3	230	300	1	70	2	0	3	4
3	300	328	1	28	2	0	3	5
3	328	356	1	28	2	1	3	5
4	0	154	1	154	1	0	1	1
4	154	336	1	182	1	0	3	2
5	0	35	1	35	1	0	1	1
5	35	276	1	241	1	0	3	2
5	276	302	1	26	1	0	3	3
5	302	344	1	42	1	1	3	4

```
AG_Model<-coxph (Surv (Start, Stop,URI_status)~Mon_R + Sex_r + Swap_r.
+smk r + water r + fire r + bwt r + Pocc r2+Toh r + cluster (StudyID), data=uri)
```


Clinical Epidemiology and Global Health. 2019;7(2):253-60

Discussion
(cont.)

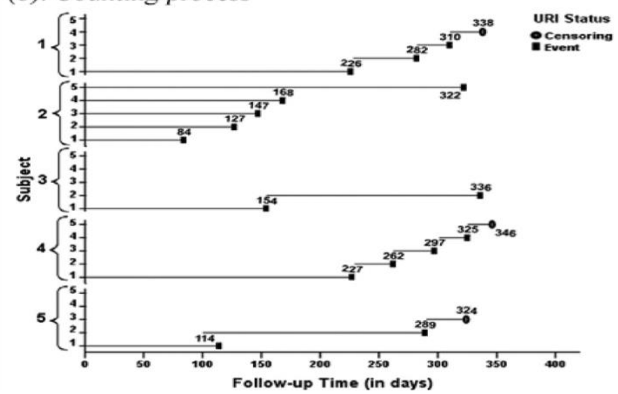
Prentice, Williams and Peterson (PWP) model

- Assumes: Recurrent events within subject are **related and baseline hazard is varied from event to event** or subject is not at risk of a second event until the first event has occurred.
- Limitation: might give **unreliable estimates for higher order events** because as the event order increases, number of subjects in the risk set is decreased.

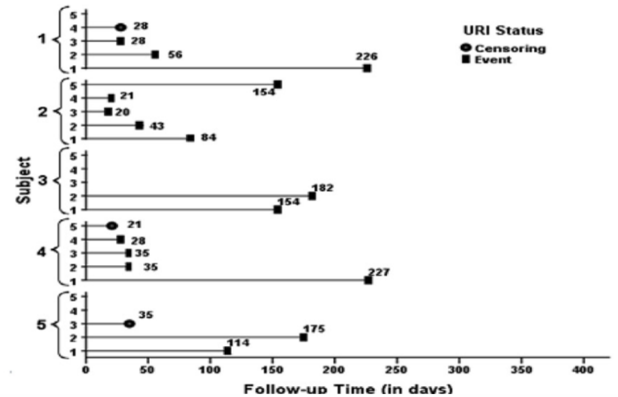
PWP-CP model
- Time from entry (time to each event is measured from entry time)
- Interested in knowing the effect of intervention on the outcome from the beginning

PWP-GT model
- Time from previous event (time to each event is measured from the previous event)
- Interested in knowing effect from previous event(s)

(b). Counting process



(d). Gap time



Clinical Epidemiology and Global Health. 2019;7(2):253-60; Int J Epidemiol. 2015;44(1):324-33; Epidemiology: Open Access. 2018;08

Discussion
(cont.)

Prentice, Williams and Peterson (PWP) model

❖ **Data structure and analysis (R program) of PWP-CP model**

Study ID	Start	Stop	URI Status	Gap	Sex	Swab	Months	URI count
1	0	226	1	226	1	0	3	1
1	226	282	1	56	1	0	3	2
1	282	310	1	28	1	1	3	3
1	310	338	0	28	1	0	1	4
2	0	84	1	84	1	0	1	1
2	84	127	1	43	1	0	2	2
2	127	147	1	20	1	0	2	3
2	147	168	1	21	1	0	2	4
2	168	322	1	154	1	0	3	5
3	0	132	1	132	2	1	1	1
3	132	202	1	70	2	1	2	2
3	202	230	1	28	2	1	2	3
3	230	300	1	70	2	0	3	4
3	300	328	1	28	2	0	3	5
3	328	356	1	28	2	1	3	5
4	0	154	1	154	1	0	1	1
4	154	336	1	182	1	0	3	2
5	0	35	1	35	1	0	1	1
5	35	276	1	241	1	0	3	2
5	276	302	1	26	1	0	3	3
5	302	344	1	42	1	1	3	4

```
PWP_TT< -coxph (Surv (Start, Stop,URI_status) ~ Mon_R + Sex_r + Swap_r.
+smk_r + water_r + fire_r + bwt_r + Pocc_r2+Toh_r + cluster (StudyID)+Strata (URI_Count),data = uri)
```

Discussion
(cont.)

Prentice, Williams and Peterson (PWP) model

❖ **Data structure and analysis (R program) of PWP-GT model**

Study ID	Start	Stop	URI Status	Gap	Sex	Swab	Months	URI count
1	0	226	1	226	1	0	3	1
1	226	282	1	56	1	0	3	2
1	282	310	1	28	1	1	3	3
1	310	338	0	28	1	0	1	4
2	0	84	1	84	1	0	1	1
2	84	127	1	43	1	0	2	2
2	127	147	1	20	1	0	2	3
2	147	168	1	21	1	0	2	4
2	168	322	1	154	1	0	3	5
3	0	132	1	132	2	1	1	1
3	132	202	1	70	2	1	2	2
3	202	230	1	28	2	1	2	3
3	230	300	1	70	2	0	3	4
3	300	328	1	28	2	0	3	5
3	328	356	1	28	2	1	3	5
4	0	154	1	154	1	0	1	1
4	154	336	1	182	1	0	3	2
5	0	35	1	35	1	0	1	1
5	35	276	1	241	1	0	3	2
5	276	302	1	26	1	0	3	3
5	302	344	1	42	1	1	3	4

Gap time = Stop - Start time

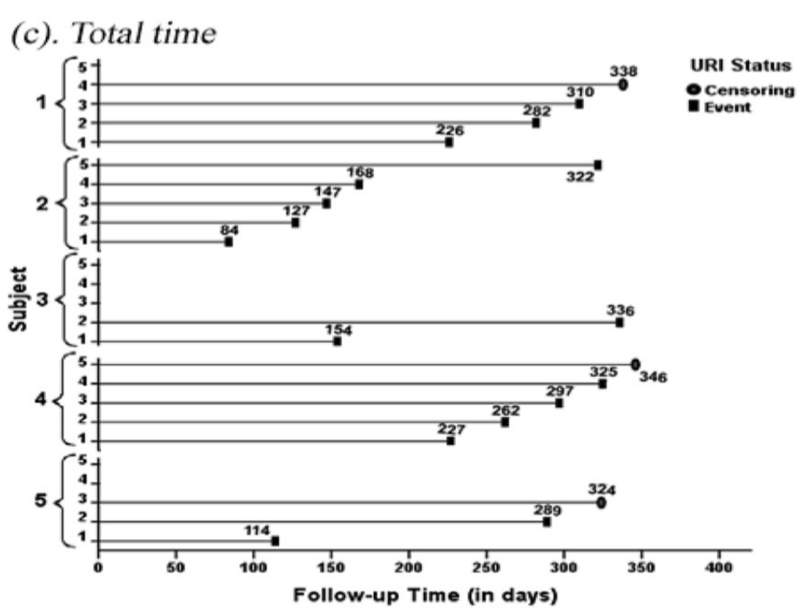
```
PWP_GT<-coxph (Surv (Stop-Start,URI_status)~Mon_R + Sex_r+
Swap_r + smk_r + water_r + fire_r + bwt_r + Pocc_r2+Toh_r + cluster.
(StudyID)+Strata (URI_Count),data = uri)
```


Clinical Epidemiology and Global Health. 2019;7(2):253-60

Discussion
(cont.)

Wei, Lin and Weissfeld (WLW) Marginal model

- ❖ Assumes: Each recurrence as **separated process** and **no ordering** among events within subject.
- ❖ The marginal risk set at time t for event k is made up of all subjects under observation at time t regardless of whether they had experienced or not events 1, ⋯ , k−1.
- ❖ Only variance-corrected model which can be applied to **multiple events of same type** of events or **multiple events of different types** of events.
- ❖ E.g., 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.



Clinical Epidemiology and Global Health. 2019;7(2):253-60; Int J Epidemiol. 2015;44(1):324-33; Epidemiology: Open Access. 2018;08

Discussion
(cont.)

Wei, Lin and Weissfeld (WLW) Marginal model

- **Data structure and analysis (R program)**

Study ID	Start	Stop	URI Status	Sex	Swab	Months	URI count
1	0	226	1	1	0	3	1
1	0	282	1	1	0	3	2
1	0	310	1	1	1	3	3
1	0	338	0	1	0	1	4
1	0	338	1	1	0	1	5
1	0	338	1	1	0	1	6
1	0	338	1	1	0	1	7
1	0	338	1	1	0	1	8
1	0	338	1	1	0	1	9
1	0	338	1	1	0	1	10

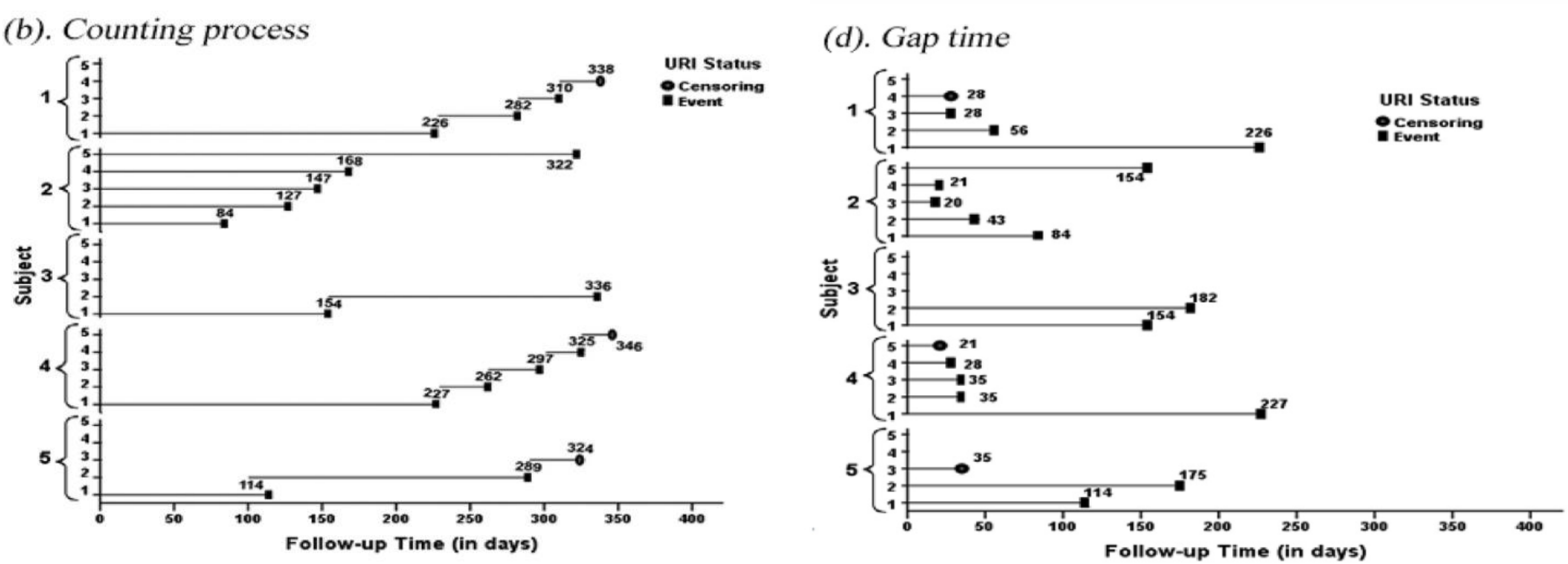
```
Marginal<-coxph (Surv (Start, Stop,URI_status)~Mon_R + Sex_r+
Swap_r + smk_r + water_r + fire_r + bwt_r + Pocc_r2+Toh_r + cluster (StudyID)+Strata (URI_Count),data = uri)
```


Clinical Epidemiology and Global Health. 2019;7(2):253-60

Discussion
(cont.)

Frailty model

- Assumes: 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.
- The association between recurrent events is explicitly modeled as a **random-effect term**, called the frailty shared by all members of the cluster.
- Factors may be socio-demographic, environment, behavioral or genetic.



Clinical Epidemiology and Global Health. 2019;7(2):253-60; Int J Epidemiol. 2015;44(1):324-33; Epidemiology: Open Access. 2018;08

Discussion
(cont.)

Frailty model

❖ **Data structure and analysis (R program)**

Study ID	Start	Stop	URI Status	Gap	Sex	Swab	Months	URI count
1	0	226	1	226	1	0	3	1
1	226	282	1	56	1	0	3	2
1	282	310	1	28	1	1	3	3
1	310	338	0	28	1	0	1	4
2	0	84	1	84	1	0	1	1
2	84	127	1	43	1	0	2	2
2	127	147	1	20	1	0	2	3
2	147	168	1	21	1	0	2	4
2	168	322	1	154	1	0	3	5
3	0	132	1	132	2	1	1	1
3	132	202	1	70	2	1	2	2
3	202	230	1	28	2	1	2	3
3	230	300	1	70	2	0	3	4
3	300	328	1	28	2	0	3	5
3	328	356	1	28	2	1	3	5
4	0	154	1	154	1	0	1	1
4	154	336	1	182	1	0	3	2
5	0	35	1	35	1	0	1	1
5	35	276	1	241	1	0	3	2
5	276	302	1	26	1	0	3	3
5	302	344	1	42	1	1	3	4

```
Frailty<-coxph (Surv (Start, Stop,URI_status)~Mon_R + Sex_r + Swap_r.
+mem5_r + smk_r + water_r + fire_r + Fathedu_r + MothEdu_r + bwt_r + Pocc_r2+Toh_r + frailty (StudyID, dist="gamma"),
data=uri)
```


Clinical Epidemiology and Global Health. 2019;7(2):253-60

Discussion
(cont.)

Table 2. Standard statistical models for recurrent events analyses.

Model	Components and specificities
AG	Conditional model, accounts for the counting process as a time scale and unrestricted set for subjects at risk Recurrent events within individuals are independent and share a common baseline hazard function Intensity of the model: $\lambda_i(t) = Y_i(t) \times \lambda_0(t) \times exp(\beta^t X_i)$
PWP	Conditional model, counting process as time scale and restricted set for subjects at risk Stratified AG, stratum k collects all the kth events of the individuals Hazard function for each event Hazard function: $\lambda_{ik}(t) = Y_i(t) \times \lambda_{0k}(t) \times exp(\beta_k^t X_i)$
WLW	Marginal model, also stratified, calendar time scale and semi-restricted set for subjects at risk Intra-subject dependence Hazard function: $\lambda_{ik}(t) = Y_i(t) \times \lambda_{0k}(t) \times exp(\beta_k^t X_i)$
Frailty	Extension of AG model Random term z_i for each individual to account for unobservable or unmeasured characteristics Hazard function: $\lambda_i(t) = Y_i(t) \times \lambda_0(t) \times z_i \times exp(\beta^t X_i)$

Notes: AG, Andersen–Gill; PWP, Prentice, William and Peterson; WLW, Wei-Lin-Weissfeld.

Biostatistics & Epidemiology. 2023;7(1):e2283650.

Discussion
(cont.)

Continuous vs Discontinuous risk intervals

- Objective: Comparing different approaches analyzing recurrent malaria episodes, with **continuous** and **discontinuous risk interval** models.
- The discontinuous risk interval analysis was found to be the more appropriate approach.

Table 1 Data structures for modelling recurrent time-to-event outcomes

ID	Start	End	Episode	Order	Time	Treatment	Age (Years)	Quarter
1	0	28	1	1	28	AS + SP	3.93	1
1	42	52	1	2	10	AS-SP	3.93	1
1	476	700	0	10	224	AS + SP	3.93	1
2	0	77	1	1	77	AS + AQ	1.15	1
2	91	375	1	2	284	AS + AQ	1.15	1
2	417	700	0	4	283	AS + AQ	1.15	1
3	0	28	1	1	28	AL	1.48	1
3	42	78	1	2	36	AL	1.48	1
3	150	700	0	5	550	AL	1.48	1

Conclusion

Key factors for selecting models to analyze recurrent events

- ❖ Number of events
- ❖ Relationships among subsequent events
- ❖ **Biological understanding for a particular disease** (e.g., infections).
- ❖ Within-subject correlation
- ❖ Varying covariates
- ❖ Sample size

Study findings

- ❖ Selection of an appropriate method should **not rely solely on statistical criteria** (i.e., high log likelihood value).
- ❖ It should also be guided by the **research question** and a **clinical knowledge** on the events of interest.
- ❖ The PWP-CP model fit the data appropriately while the biological process also suggested the same model.

Importance of data structure

- ❖ The structure of the data significantly influences the analysis of recurrent events.

ML/DL models

ML/DL models for modeling recurrent events

- **DL for survival analysis: a review**

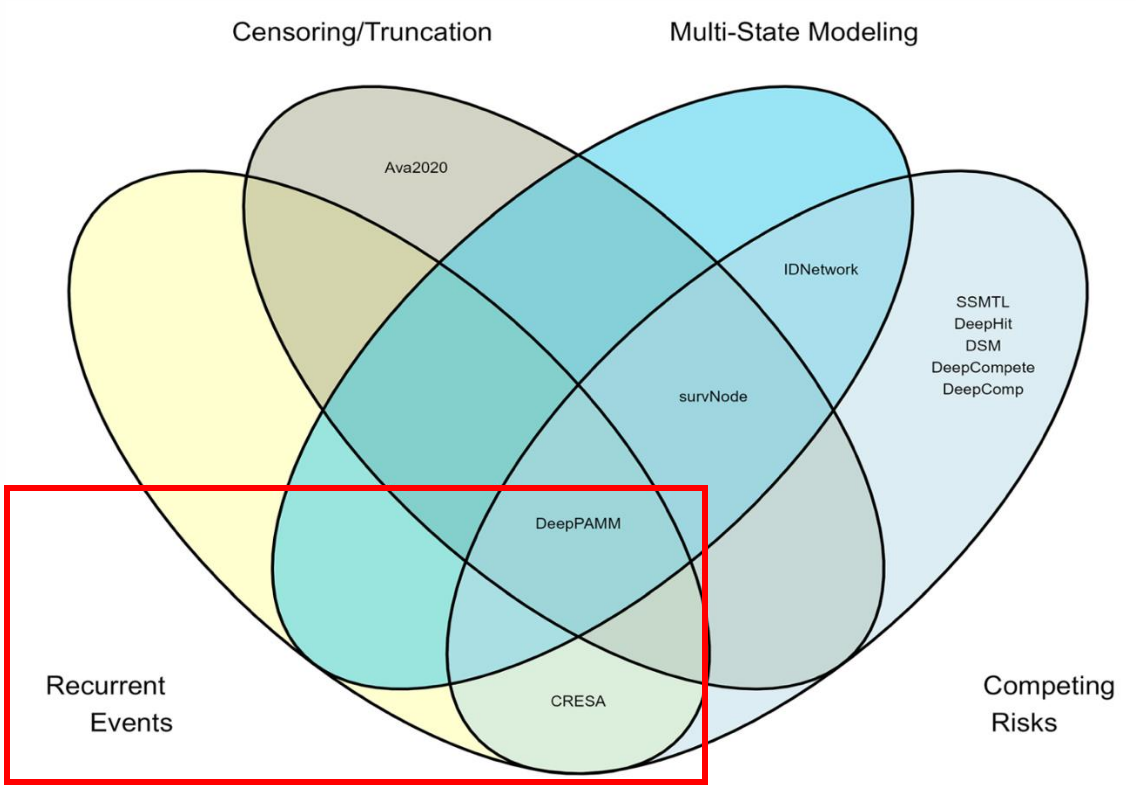


Fig. 6 Venn diagram illustrating which methods can handle the distinct survival outcome types

Artificial Intelligence Review. 2024;57(3):65.

ML/DL models
(cont.)

ML/DL models for modeling recurrent events

(2019) CRESA:
A Deep Learning Approach to Competing Risks, Recurrent Event[1]

(2022) DeepPAMM:
Deep Piecewise Exponential Additive Mixed Models for Complex Hazard Structures in Survival Analysis[2]

(2024) Random survival forests
for the analysis of recurrent events for right-censored data, with or without a terminal event[3]

(2024) LSTM-COX Model:
A Concise and Efficient Deep Learning Approach for Handling Recurrent Events[4]

1. Gupta G, et al. 2019; 2. Kopper P, et al. 2022; 3. Murris et al. 2024; 4. Runquan, et al. 2024

Thank You