An estimator for correlation between two survival score statistics

Zakiyah Zain
Lancaster University, U.K.

KEY SUMMARY

- Successful work for binary and ordinal, but a challenge for survival.
- Application to logrank is difficult, hence interval censored approach.
- Impact: -minimize sample size,
  -facilitate multiple testing procedures
- Objectives: -derive an estimator for the correlation
  -investigate properties
  -evaluate accuracy

Interval-censored survival

- Survive past time *t*? Yes or No.

Responses during interval (0, t)	Experimental	Control	Total
No. of events (Failure)	o_E	o_C	o
No. of patients survived (Success)	$r_E - o_E$	$r_C - o_C$	r - o
Total	r_E	r_C	r

- The log hazard ratio,

$$\theta = -\log\{-\log S_E(t)\} + \log\{-\log S_C(t)\}$$

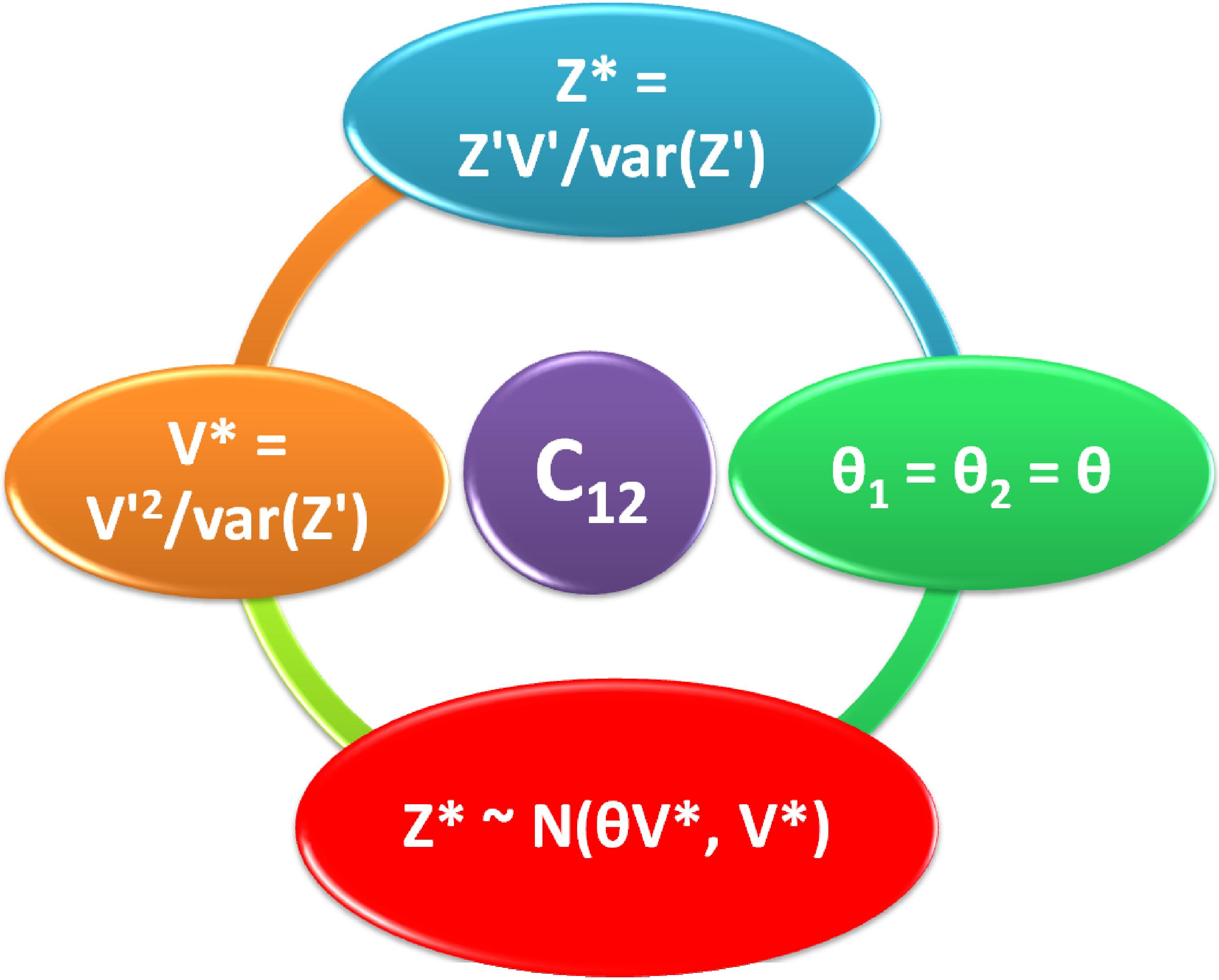
where S_E is the survivor function of patients on experimental.

$$Z = \frac{q}{o}\{r_E o_C - r_C o_E\} \qquad V = \frac{q^2(r-o)r_C r_E}{o(r-1)}$$

where q = -log(1-o/r).

- **For large samples and small θ, Z ~ N(θV, V)**

Bivariate score tests



Estimator for Correlation

- Correlation, $\rho_{12} = C_{12} / \sqrt{V_1 V_2}$
- Covariance derived from cloglog link for uncensored data,

$$C_{12} = \frac{q_1 q_2 r_{1E} r_{1C}}{o_1 o_2 r_1}(r_1 o_{12} - o_1 o_2)$$

where $r_1 = r_2 = r_{12} = n$ *(sample size).*

Interval-censored bivariate survival

- Without censoring, straightforward..... 

For control	**T_1**			
T_2		Failure	Success	Total
	Failure	FF	SF	o_{2C}
	Success	FS	SS	r_{2C}-o_{2C}
	Total	o_{1C}	r_{1C}-o_{1C}	$r_{1C} = r_{2C} = r_{12C}$

- With censoring, complication arises..... 

For control		**T_1**			
		Failure	Success	Missing	Total
T_2	Failure	FF	SF	MF	o_{2C}
	Success	FS	SS	MS	r_{2C}-o_{2C}
	Missing	FM	SM	MM	m_2
	Total	o_{1C}	r_{1C}-o_{1C}	m_1	?

Interval-censored bivariate survival

double failure, o_{12}		T_1									
		(0, 1)		(1, 3)		(3, 6)		(6, 12)		>12	
treatment		E	C	E	C	E	C	E	C	E	C
T_2	(0, 1)	15	10	5	10	1	2	2	1	0	0
	(1, 3)	18	26	3	10	2	1	6	0	0	0
	(3, 6)	6	2	5	5	1	1	1	2	0	0
	(6, 12)	3	9	2	1	3	2	0	0	0	0
	>12	0	0	0	0	0	0	0	0	0	0

$cov(Z_{1i}, Z_{2j})$		T_1			Z_{2j}	V_{2j}		
		(0, 1)	(1, 3)	(3, 6)				
T_2	(0, 1)	1.274	1.477	0.137	-1.219	11.482		
	(1, 3)	5.366	0.664	-0.303	1.464	16.275		
	(3, 6)	-0.775	1.469	-0.237	-2.969	5.717		
	(6, 12)	1.103	-0.196	.	0.687	4.917		
	Z_{1i}	9.383	2.654	2.640			$\sum Z_1$	14.676
	V_{1i}	30.515	18.199	7.088			$\sum V_1$	55.802
					$\sum Z_2$	$\sum V_2$	C_{12}	
					-2.037	38.391		9.980

Simulation Study

Distribution

- $T_1 \sim Exp(\lambda e^s)$, $T_2 \sim Exp(\lambda e^s)$,
- $s \sim N(0, \sigma^2)$; $\sigma = c(\log \lambda_C - \log \lambda_E)$, c = 1, 2 and 4.

Theory

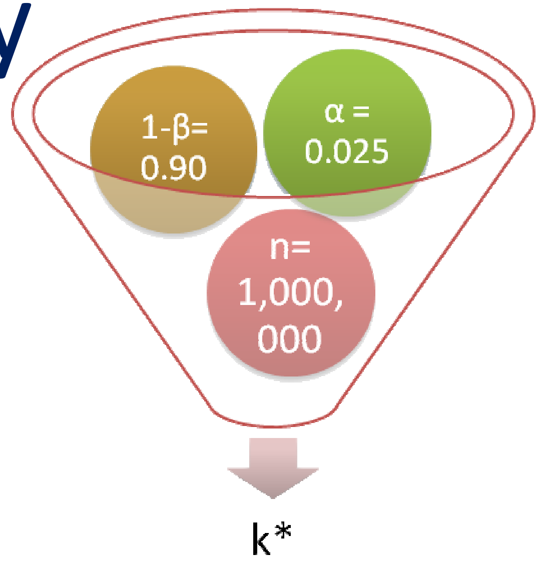
- $V^* = k^* n$
- $V^* = \{(u_\alpha + u_\beta)/\theta_R\}^2$
- $\theta_R^2 = 10.51/k^* n$

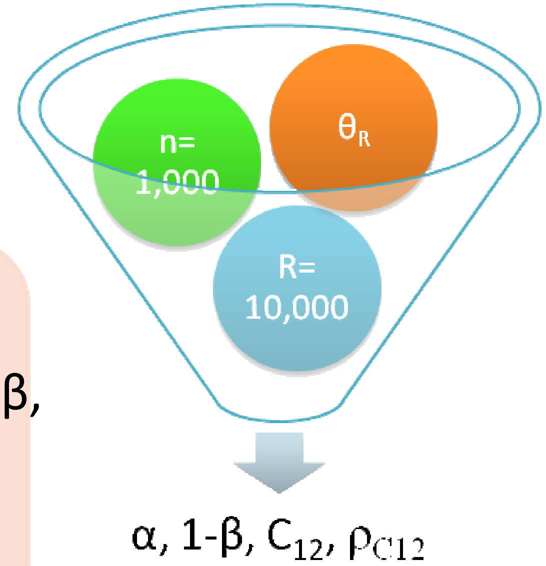
Part I

- Generate n = 1 million to get k^*
- T_1 and T_2 split into 5 intervals of (0, 1), (1, 3), (3, 6) , (6, 12) and (>12) months.

Part II

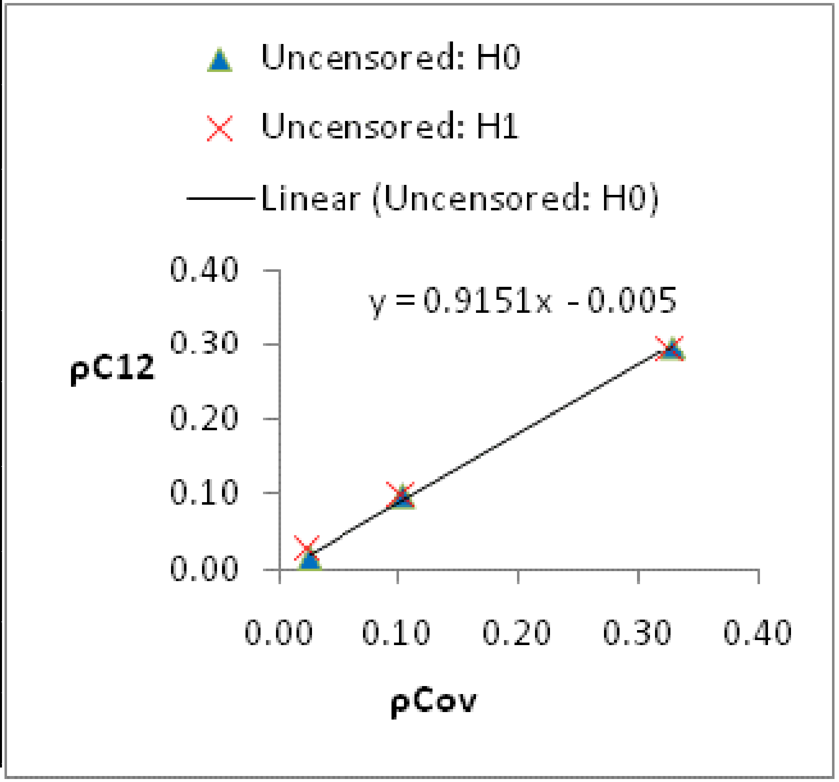
- Fix n=1,000 to get θ_R
- Run 10,000 replicates under H_0, H_1 to get α, $1-\beta$, C_{12}, ρ_{C12}
- Compute sample covariance, $Cov(Z_1, Z_2)$, ρ_{cov}





Case 1: Uncensored

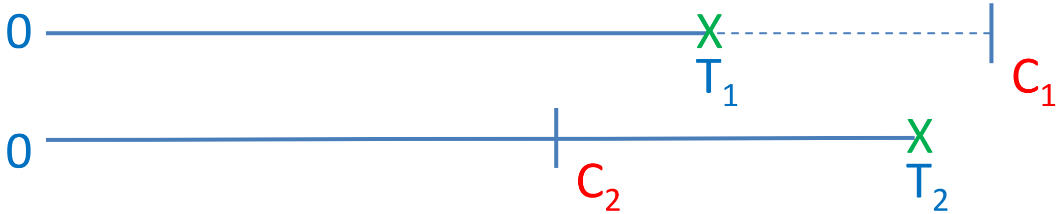
c	H_0				
	α_1	α_2	α_*	ρ_{C12}	ρ_{Cov}
1	0.026	0.023	0.024	0.014	0.026
2	0.024	0.021	0.024	0.097	0.105
4	0.024	0.024	0.027	0.295	0.329
c	H_1				
	$1-\beta_1$	$1-\beta_2$	$1-\beta_*$	ρ_{C12}	ρ_{Cov}
1	0.68	0.69	0.93	0.029	0.024
2	0.70	0.70	0.92	0.100	0.103
4	0.59	0.59	0.78	0.295	0.326



- Type I error within 95% CI (0.022, 0.028).
- Accurate power, hence worth including both times; 95% CI (0.089, 0.91).

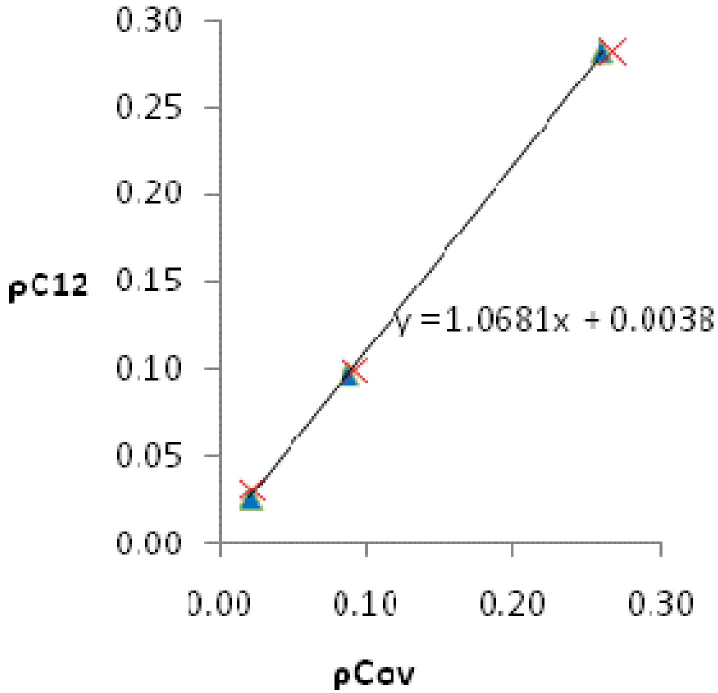
Case 2: General

- For subjects who are related, e.g. twins, married couples,



where $r_1 \neq r_2 \neq r_{12} \neq n$

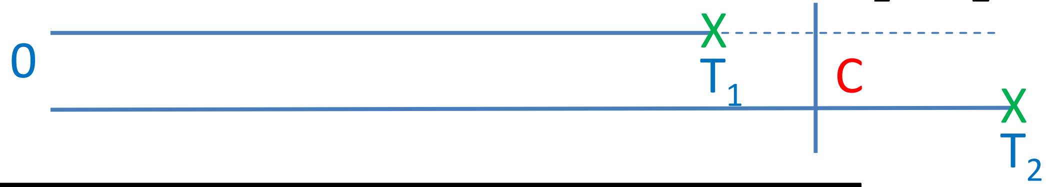
c	H_0				
	α_1	α_2	α_*	ρ_{C12}	ρ_{Cov}
1	0.025	0.024	0.024	0.026	0.021
2	0.024	0.025	0.025	0.097	0.087
4	0.025	0.025	0.025	0.282	0.260
c	H_1				
	$1-\beta_1$	$1-\beta_2$	$1-\beta_*$	ρ_{C12}	ρ_{Cov}
1	0.60	0.66	0.90	0.031	0.022
2	0.61	0.65	0.87	0.100	0.092
4	0.51	0.53	0.71	0.282	0.267



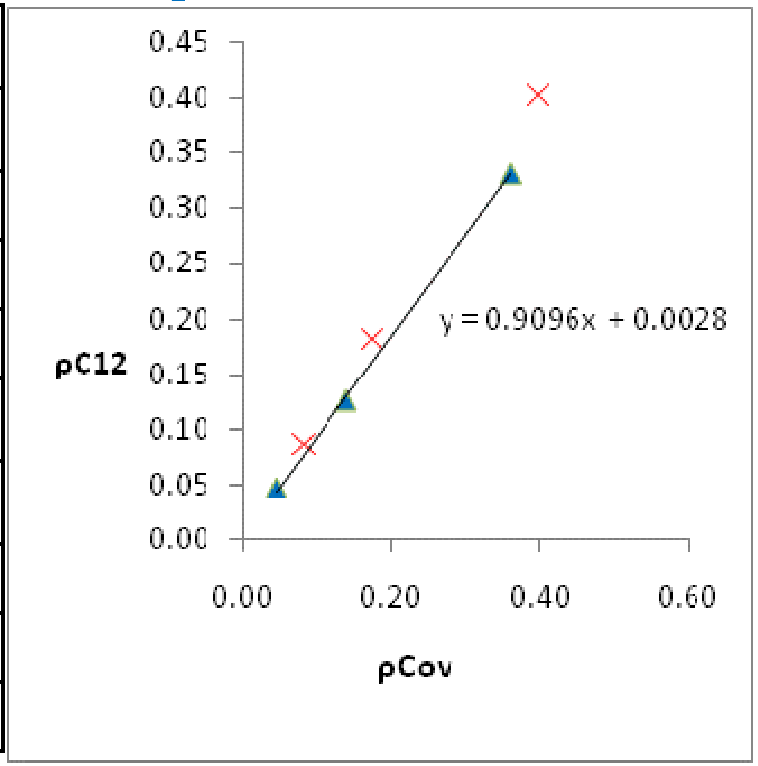
z.zain@lancaster.ac.uk

Case 3: Paired Organs

- For similar physically related parts of a person, e.g. right/left eye or knee or kidney; *where* $r_1 = r_2 = r_{12}$



c	H_0				
	α_1	α_2	α_*	ρ_{C12}	ρ_{Cov}
1	0.024	0.025	0.026	0.047	0.046
2	0.026	0.025	0.028	0.126	0.140
4	0.026	0.024	0.027	0.332	0.361
c	H_1				
	$1-\beta_1$	$1-\beta_2$	$1-\beta_*$	ρ_{C12}	ρ_{Cov}
1	0.63	0.63	0.89	0.052	0.047
2	0.63	0.63	0.86	0.130	0.128
4	0.52	0.53	0.69	0.330	0.348

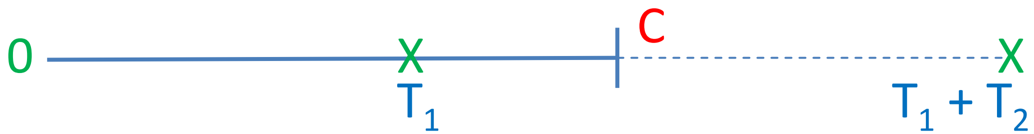


- Accurate power, but slightly inflated Type I error.

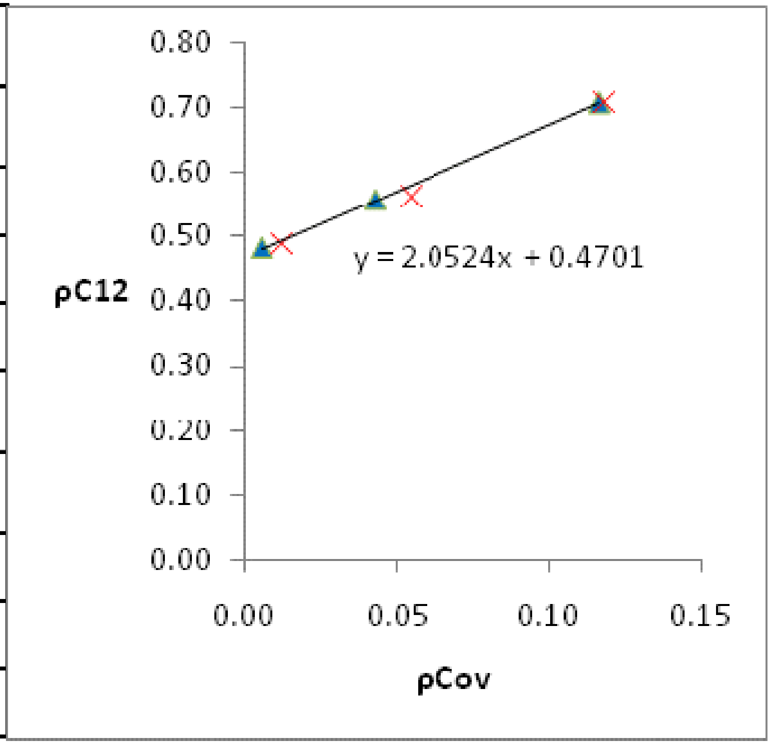
z.zain@lancaster.ac.uk

Case 4: Successive Events

- For recurrent events; e.g. 1st, 2nd seizures ; *where* $r_1 > r_2 = r_{12}$



c	H_0				
	α_1	α_2	α_*	ρ_{C12}	ρ_{Cov}
1	0.025	0.023	0.009	0.482	0.006
2	0.024	0.026	0.008	0.558	0.043
4	0.024	0.024	0.007	0.708	0.116
c	H_1				
	$1-\beta_1$	$1-\beta_2$	$1-\beta_*$	ρ_{C12}	ρ_{Cov}
1	0.85	0.60	0.91	0.490	0.012
2	0.79	0.50	0.82	0.563	0.055
4	0.60	0.28	0.51	0.708	0.118



- Power is accurate, but type I error rate is too conservative.
- Derived correlation is doubled!

Conclusions/Further Work

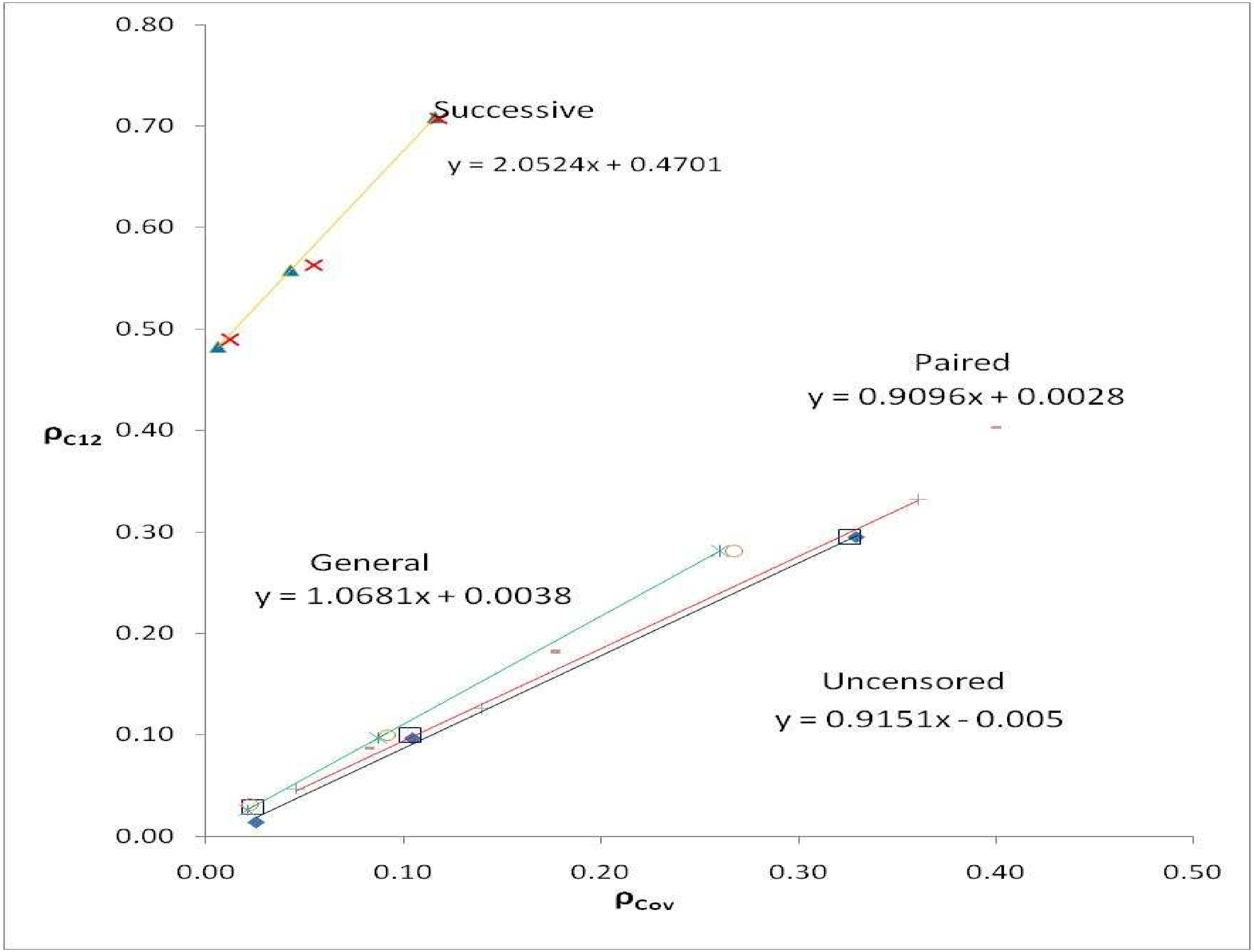
- Guaranteed Type I error rates.
- Achieved accurate power of test.
- Good estimation of ρ_{C12} for bivariate correlated interval censored survival data, *except for successive events.*
- Further work:
  - Overcome inflation of Type I error
  - Resolve issues of ρ_{C12} for successive events
  - Compare with Wei, Lin & Weissfeld method
  - Adjust for covariates

References

- Bolland, K., Whitehead, J., Cobo, E., and Secades, J. J. (2009). Evaluation of a sequential global test of improved recovery following stroke as applied to the ICTUS trial of citicoline. *Pharmaceutical Statistics* **8**, 136-149.
- Pocock SJ, Geller NL, and Tsiatis AA. The Analysis of Multiple Endpoints in Clinical Trials. *Biometrics* 1987; **43**: 487-498.
- Wei, L. J., Lin, D.Y. and Weissfeld, L. (1989). Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. *Journal of the American Statistical Association* **84**, 1065-1073.
- Whitehead J. *The Design and Analysis of Sequential Clinical Trials (Revised 2nd Edition).* Wiley: Chichester, 1997.
- Whitehead J, Bolland K, Branson M, Todd S. Combining binary and ordinal endpoints from clinical trials: a non-GEE approach. *In preparation.*

Thank You

All Results



0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
ρ_{C12}
0.00
0.10
0.20
0.30
0.40
0.50
ρ_{Cov}
Successive
y = 2.0524x + 0.4701
Paired
y = 0.9096x + 0.0028
General
y = 1.0681x + 0.0038
Uncensored
y = 0.9151x - 0.005