CLASSIFICATION PROBABILITIES OF A MARKER FOR CENSORED FAILURE TIME OUTCOME: NONPARAMETRIC METHODS

Laura Antolini

Dipartimento di Medicina Clinica e Prevenzione
Università Milano-Bicocca

V Congresso Nazionale SISMEC 16-19 Settembre 2009, Università di Pavia

TPF and FPF for censored survival time outcome

Context

Y binary marker measured at *t = 0* (e.g. predictor from a regression model)

T survival time τ reference time

Data : cohort with variable follow-up $i = 1,\ldots, n$

observed time $X_i = \min(T_i, C_i)$ C_i censoring time

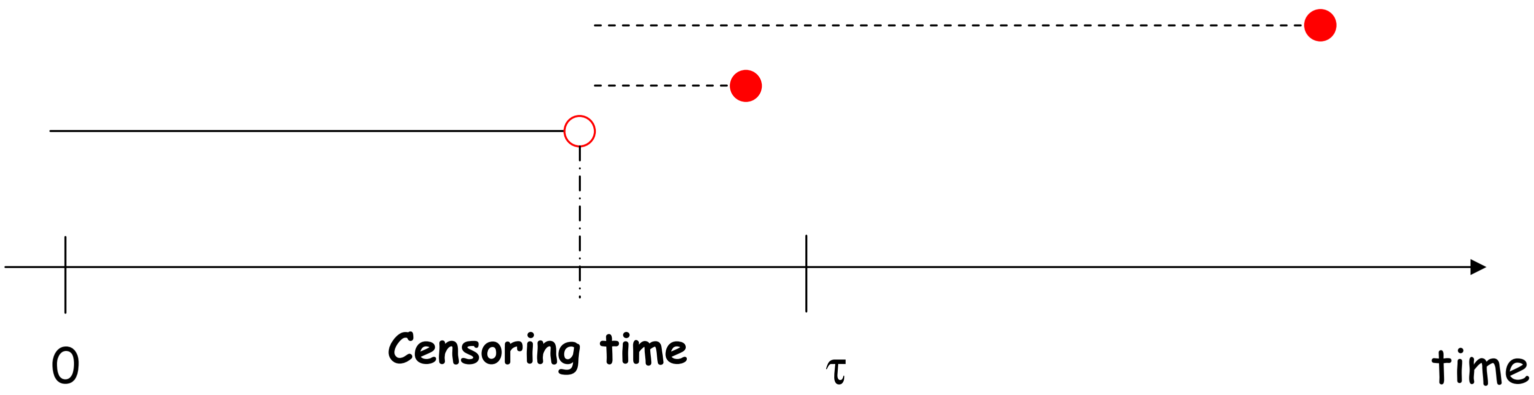
event indicator $\Delta_i = I(T_i \le C_i)$

Goal : estimate

$$TPF = P(Y = 1 \mid T \le \tau)$$

$$FPF = P(Y = 1 \mid T > \tau)$$

TPF and FPF for censored survival time outcome



$$X_i \; ; \quad \Delta_i = 0$$

Issue : accommodating censoring

A simulated sample (n = 20)

$$Y \approx Ber(0.4) \qquad T|_{Y=0} \approx Exp(1/15) \qquad T|_{Y=1} \approx Exp(1/5) \qquad C \approx Unif[0,20] \qquad \tau = 8$$

Cases (10)

X_i	Y_i
1.3	0
3.5	0
7.4	0
0.1	1
0.8	1
2.1	1
2.3	1
2.9	1
6.6	1
7.7	1

Controls (6)

X_i	Y_i
8.2	0
10.7	1
13.8	1
14.7	0
16.7	0
17.4	0

Censored (4)

X_i	Y_i
2	0
2.2	0
4.1	0
6.8	1

<u>case</u>

$$X_i \leq \tau; \Delta_i = 1$$

<u>control</u>

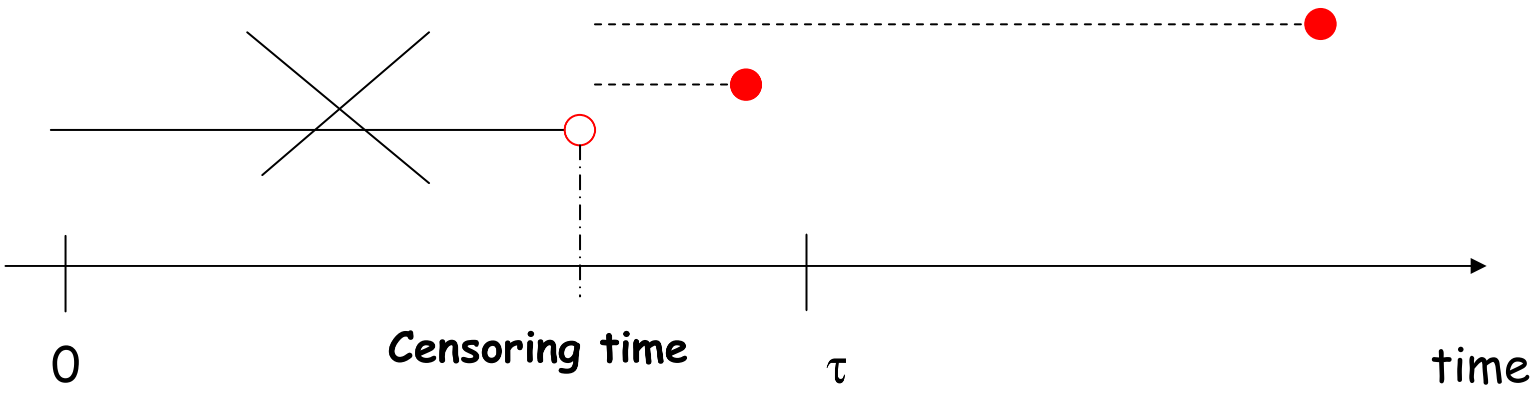
$$X_i > \tau$$

<u>censored</u>

$$X_i \leq \tau; \Delta_i = 0$$

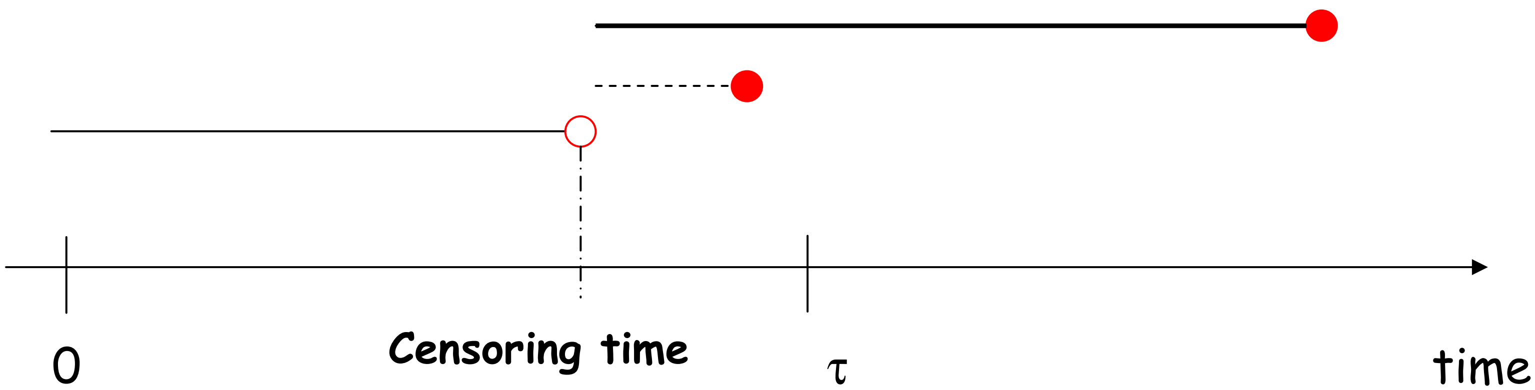
Standard of practice

- ignore censored observations "complete case" analysis



$$X_i \; ; \quad \Delta_i = 0$$

- treat censored observations as controls "misclassified"



$$X_i \; ; \quad \Delta_i = 0$$

Standard of practice

- ignore censored observations "complete case" analysis

original

		marker				
		0	1	total		
status	T ≤ τ	4	8	12	TPF=	0.67
	T > τ	6	2	8	FPF=	0.25
	total	10	10	20		

selected

		marker				
		0	1	total	TPF=	0.70
status	T ≤ τ	3	7	10	FPF=	0.33
	T > τ	4	2	6		
	total	7	9	16		

Proposed solutions to accommodate censoring

TPF and FPF as function of joint distribution of (T,Y)

'PARTIAL' BAYES THEOREM

PLUG IN OF KM ESTIMATES

$$T\hat{P}F = \hat{P}(Y=1 \mid T \leq \tau) = \frac{\hat{P}(T \leq \tau \mid Y=1) \cdot \hat{P}(Y=1)}{\hat{P}(T \leq \tau)}$$

$$F\hat{P}F = \hat{P}(Y=1 \mid T > \tau) = \frac{\hat{P}(T > \tau \mid Y=1) \cdot \hat{P}(Y=1)}{\hat{P}(T > \tau)}$$

Heagerty, Lumley, Pepe , Biometrics 2000

Proposed solutions to accommodate censoring

TPF and FPF as function of joint distribution of (T,Y)

'FULL' BAYES THEOREM

PLUG IN OF KM ESTIMATES

$$T\hat{P}F = \ldots = \frac{\hat{P}(T \leq \tau \mid Y = 1) \cdot \hat{P}(Y = 1)}{\hat{P}(T \leq \tau \mid Y = 1) \cdot \hat{P}(Y = 1) + \hat{P}(T \leq \tau \mid Y = 0) \cdot \hat{P}(Y = 0)}$$

$$F\hat{P}F = \ldots = \frac{\hat{P}(T > \tau \mid Y = 1) \cdot \hat{P}(Y = 1)}{\hat{P}(T > \tau \mid Y = 1) \cdot \cdot \hat{P}(Y = 1) + \hat{P}(T > \tau \mid Y = 0) \cdot \hat{P}(Y = 0)}$$

Heagerty, Lumley, Pepe , Biometrics 2000

Proposed solutions to accommodate censoring

original

		marker 0	marker 1	total		
status	$T \leq \tau$	4	8	12	TPF=	0.67
	$T > \tau$	6	2	8	FPF=	0.25
	total	10	10	20		

selected

		marker 0	marker 1	total		
status	$T \leq \tau$	3	7	10	TPF=	0.70
	$T > \tau$	4	2	6	FPF=	0.33
	total	7	9	16		

'PARTIAL' BAYES

TPF = 0.64 FPF = 0.32

'FULL' BAYES

TPF = 0.66 FPF = 0.30

Verification bias

Y : marker observed on *n*

D : disease status (diseased / non diseased) verfied on *m* < *n*

if ascertainment of *D* is costly

- one may select subjects for ascertaiment of *D* depending on *Y*

  (e.g. PSA value and Biopsy for prostate cancer)

- a negative result may affect willingness to ascertain *D*

  (e.g. neonatal hearing test and 12 month gold standard assessment)

- IMPUTATION of *D* for non verfied subjects

- WEIGHTING cases/controls according to 1/verification probability to represent the status of non verified subjects

Censoring as a kind of verification bias

Y : marker observed on *n*

$D = I(T \leq \tau)$: disease status verfied on *m* < *n*

the ascertainment of *D* depends on the action of censoring that could act before *T*

- IMPUTATION of *D* for censored subjects
- WEIGHTING cases/controls according to the verification probability to represent the status of censored subjects

Censoring as a kind of verification bias

Goal

to estimate classification matrix of the *n* data from the *m* selected data

selected

		marker 0	marker 1	total
status	$T \le \tau$	3	7	10
	$T > \tau$	4	2	6
	total	7	9	16

- IMPUTATION
- WEIGHTING

original

		marker 0	marker 1	total
status	$T \le \tau$	4	8	12
	$T > \tau$	6	2	8
	total	10	10	20

WEIGHTING the observed cases/controls

selected

		marker 0	marker 1	total
status	$T \leq \tau$	3	7	10
	$T > \tau$	4	2	6
	total	7	9	16

X_i	Y_i
2	0
2.2	0
4.1	0
6.8	1

estimate of the original

		marker 0	marker 1
status	$T \leq \tau$	3 x ?	7 x ?
	$T > \tau$	4 x ?	2 x ?
	total	7	9
		10	**10**

WEIGHTING

Cases/controls are inversely weigthed by the verification probability

case $X_i \le \tau; \Delta_i = 1$	verified if the censoring time is greater than the survival time	$\pi_i(X_i) = P(C_i > X_i)$
control $X_i > \tau$	verified if the censoring time is greater than τ	$\pi_i(\tau) = P(C_i > \tau)$

$$\hat{\#}(T \le \tau; Y = 0) = \sum_{i=1}^{n} I(X_i \le \tau)\Delta_i I(Y_i = 0) \cdot \boxed{\frac{1}{\hat{\pi}_i(X_i)}}$$

$$\hat{\#}(T > \tau; Y = 0) = \sum_{i=1}^{n} I(X_i > \tau) I(Y_i = 0) \cdot \boxed{\frac{1}{\hat{\pi}_i(\tau)}}$$

WEIGHTING

How to estimates the weights?

case $\pi_i(X_i) = P(C_i > X_i)$

control $\pi_i(\tau) = P(C_i > \tau)$

Overall estimate ?

$$\hat{\pi}_i(t) = \hat{P}_{KM}\left(C > t\right)$$

Conditional estimate ?

$$\hat{\pi}_i(t) = \hat{P}_{KM}\left(C > t \mid Y = Y_i\right)$$

OVERALL WEIGHTS

TPF= 0.67

FPF= 0.25

selected

		marker 0	marker 1	total
status	T ≤ τ	3	7	10
	T > τ	4	2	6
	total	7	9	16

estimate of the original

		marker 0	marker 1	total
status	T ≤ τ	4.214	7.333	11.547
	T > τ	5.786	2.667	8.453
	total	10	10	20.000

TPF= 0.64

FPF= 0.32

Further solutions

CONDITIONAL WEIGHTS

TPF= 0.67

FPF= 0.25

selected

		marker 0	marker 1	total
status	$T \leq \tau$	3	7	10
status	$T > \tau$	4	2	6
	total	7	9	16

estimate of the original

		marker 0	marker 1	total
status	$T \leq \tau$	3.829	7.333	11.162
status	$T > \tau$	6.171	2.667	8.838
	total	10	10	20.000

TPF= 0.66

FPF= 0.30

We obtain the same estimates given by the 'full' Bayes method

Result 1

The estimates of FPF and TPF obtained by

- imputation
- conditional weigthing
- ‘full’ Bayes

are all equivalent

- conditional or unconditional weigthing?
- ‘partial’ Bayes?

Simulation exercise

Random censoring

	FPF			TPF		
	Bias	st dev	mse	Bias	st dev	mse
TRUE	0.00076	0.02670	0.00071	-0.00091	0.02873	0.00083
CC	-0.00085	0.03429	0.00118	0.00791	0.03170	0.00107
Missclass	0.05884	0.02650	0.00416	0.00791	0.03170	0.00107
Unc. Weigth	-0.00085	0.03429	0.00118	-0.00187	0.03209	0.00103
Partial Bayes	0.00005	0.03028	0.00092	-0.00044	0.03053	0.00093
Cond. Weight	0.00019	0.03024	0.00091	-0.00082	0.03039	0.00092

Marker dependent censoring

	FPF			TPF		
	Bias	st dev	mse	Bias	st dev	mse
TRUE	-0.00017	0.02554	0.00065	0.00013	0.02940	0.00086
CC	0.05342	0.03543	0.00411	0.03376	0.03091	0.00210
Missclass	0.01410	0.02451	0.00080	0.03376	0.03091	0.00210
Unc. Weigth	0.05342	0.03543	0.00411	0.02821	0.03129	0.00177
Partial Bayes	0.00551	0.02710	0.00076	-0.01261	0.02920	0.00101
Cond. Weight	-0.00017	0.02681	0.00072	-0.00002	0.03107	0.00097

Result 2

The estimates of FPF and TPF one could obtain by

- 'partial' Bayes
- overall weigthing

- are inefficient under random censoring
- biased under marker dependent censoring

Result 3

Variability of FPF and TPF estimates can be obtained by

- 'Naive' application of the variance of the proportion obtained on the reweigthed/imputed matrix

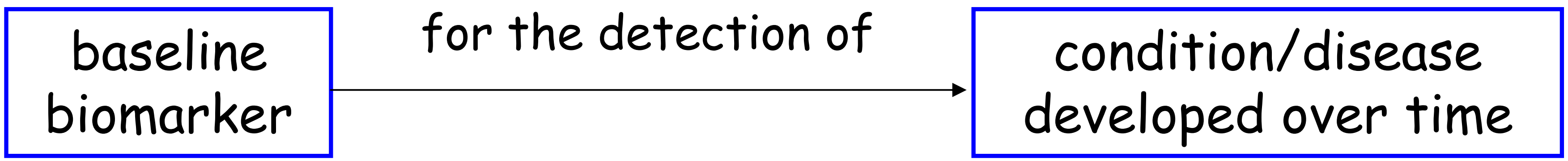
- Delta method since

$$T\hat{P}F(\tau) = f\left(\hat{S}_{KM}(\tau \mid Y=1);\ \hat{S}_{KM}(\tau \mid Y=0);\ \hat{P}(Y=1)\right)$$

$$F\hat{P}F(\tau) = \ldots$$

Summary

* The commonly used performance indicators of a diagnostic/ prognostic test (sensitivity and specificity) can apply to the case of



* Nonparametric estimation and testing is still not established.

* It can be addressed by approaches based on inference on simple proportions corrected for verification bias by conditional weigthing

* This enables to address confidence interval estimation and testing by asymptotics based on delta method

References

Heagerty PJ, Lumley T, Pepe MS
Time-dependent ROC curves for censored survival data and a diagnostic marker
Biometrics 2000; 56(2):337-344

Pepe MS
The Statistical Evaluation of Medical Tests for Classification and Prediction
New York: Oxford University Press, 2003

Pepe MS, Zheng Y, Jin Y, Huang Y, Parikh CR, Levy WC
Evaluating the ROC performance of markers for future events
Lifetime data analysis 2008; 14(1):86-113

ACKNOWLEDGMENTS

Maria Grazia Valsecchi, Università Milano-Bicocca

Margaret Pepe, Fred Hutchinson CCR - Seattle