

Absolute Risk Estimation in a Case Cohort Study of Prostate Cancer

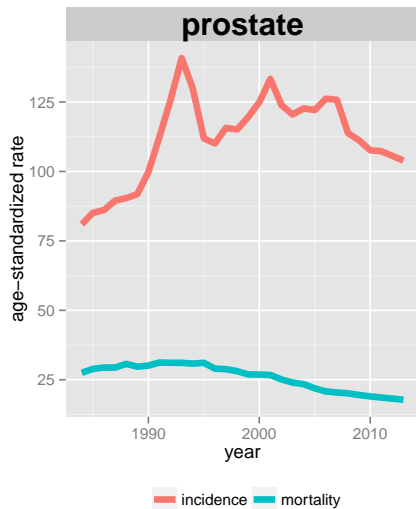
Sahir Rai Bhatnagar

Queen's University

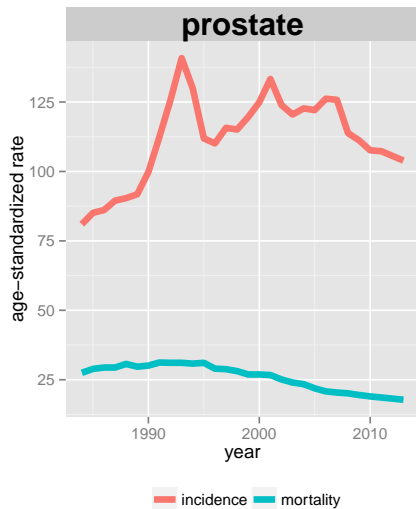
sahir.bhatnagar@queensu.ca

August 30, 2013

Incidence vs. Mortality of Prostate Cancer in Canada



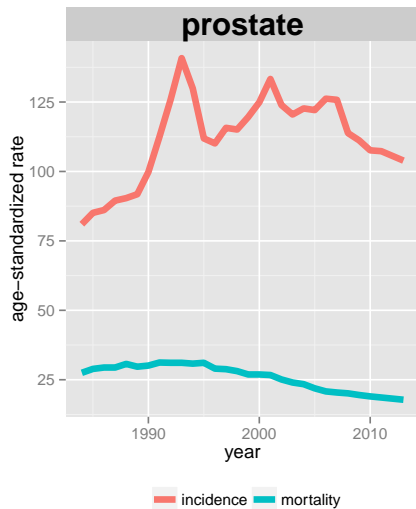
Incidence vs. Mortality of Prostate Cancer in Canada



Possible reasons for gap:

- Better treatment outcomes

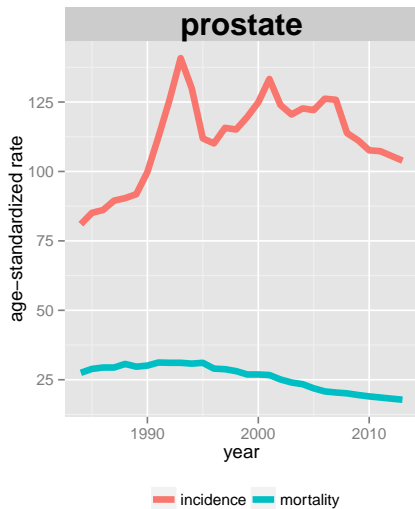
Incidence vs. Mortality of Prostate Cancer in Canada



Possible reasons for gap:

- Better treatment outcomes
- Death from other causes (comorbidities)

Incidence vs. Mortality of Prostate Cancer in Canada



Possible reasons for gap:

- Better treatment outcomes
- Death from other causes (comorbidities)
- Overdiagnosis

Why is this Gap a Problem?

Why is this Gap a Problem?

Overdiagnosis

- 1 Screen detected cancer that would have otherwise not been diagnosed before death from other causes

Why is this Gap a Problem?

Overdiagnosis

- 1 Screen detected cancer that would have otherwise not been diagnosed before death from other causes
- 2 Can lead to overtreatment

Why is this Gap a Problem?

Overdiagnosis

- 1 Screen detected cancer that would have otherwise not been diagnosed before death from other causes
- 2 Can lead to overtreatment

Overtreatment

Treatment for a cancer that is clinically not significant → low tumour grade

Why is this Gap a Problem?

Overdiagnosis

- 1 Screen detected cancer that would have otherwise not been diagnosed before death from other causes
- 2 Can lead to overtreatment

Overtreatment

Treatment for a cancer that is clinically not significant → low tumour grade

- Morbidity associated with being diagnosed with cancer

Why is this Gap a Problem?

Overdiagnosis

- 1 Screen detected cancer that would have otherwise not been diagnosed before death from other causes
- 2 Can lead to overtreatment

Overtreatment

Treatment for a cancer that is clinically not significant → low tumour grade

- Morbidity associated with being diagnosed with cancer
- Quality of life effects, incontinence, erectile dysfunction

Why is this Gap a Problem?

Overdiagnosis

- 1 Screen detected cancer that would have otherwise not been diagnosed before death from other causes
- 2 Can lead to overtreatment

Overtreatment

Treatment for a cancer that is clinically not significant → low tumour grade

- Morbidity associated with being diagnosed with cancer
- Quality of life effects, incontinence, erectile dysfunction
- Lower life expectancy

Solution to Overdiagnosis and Overtreatment

Solution to Overdiagnosis and Overtreatment

Individualised Decision Making

Solution to Overdiagnosis and Overtreatment

Individualised Decision Making

- Active surveillance for patients with favorable risk disease (Gleason score ≤ 6)

Solution to Overdiagnosis and Overtreatment

Individualised Decision Making

- Active surveillance for patients with favorable risk disease (Gleason score ≤ 6)
- Clinical tools that calculate life expectancy based on predisposing factors such as age and comorbidities

Objectives

Research Statement

- 1 Using a semiparametric model based on both age and comorbidity, obtain absolute risk estimates of death from prostate cancer in a population based case cohort study

Objectives

Research Statement

- 1 Using a semiparametric model based on both age and comorbidity, obtain absolute risk estimates of death from prostate cancer in a population based case cohort study
- 2 Create a web tool for life expectancy based on this model, for use in a clinical setting

Wykes (2011)

Wykes (2011) proposed an *ad hoc* approach to estimating the relative risk and baseline hazard function for the case cohort design:

Wykes (2011)

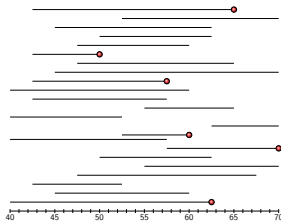
Wykes (2011) proposed an *ad hoc* approach to estimating the relative risk and baseline hazard function for the case cohort design:

Modified PH Model

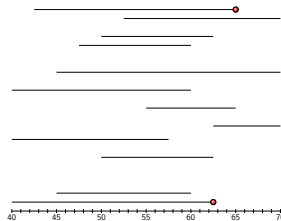
$$S(t) = (S_0(t))^{\alpha} \exp \{ \alpha \beta_1 (\text{age} - \mu_{\text{age}}) + \beta_2 (\text{CIRS-G}_{\text{pros}} - \mu_{\text{CIRS-G}_{\text{pros}}}) \}$$

- $S_0(t), \beta_1, \mu_{\text{age}}$ are derived from the full cohort
- $\beta_2, \mu_{\text{CIRS-G}_{\text{pros}}}$ are from case cohort
- $\alpha = \frac{\beta_1 \text{ from case cohort}}{\beta_1 \text{ from full cohort}}$

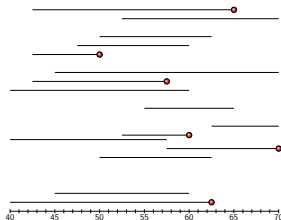
Case Cohort Study Design



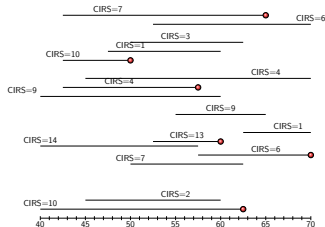
(a) Cohort, no exposure info



(b) Case cohort sampling



(c) Add back failures



(d) Get exposure information

Figure: Each line represents how long a subject is at risk. A failure is represented by a red dot. Subjects without a red dot are censored individuals. The axis represents age

Study Population

Study Population

- Population-based case cohort study of 2,740 patients diagnosed with prostate cancer between 1990 and 1998 in Ontario

Study Population

- Population-based case cohort study of 2,740 patients diagnosed with prostate cancer between 1990 and 1998 in Ontario
- Stratified by CCOR random sample

Study Population

- Population-based case cohort study of 2,740 patients diagnosed with prostate cancer between 1990 and 1998 in Ontario
- Stratified by CCOR random sample
- Cases were sampled independently of subcohort

Study Population

- Population-based case cohort study of 2,740 patients diagnosed with prostate cancer between 1990 and 1998 in Ontario
- Stratified by CCOR random sample
- Cases were sampled independently of subcohort
- Death from prostate cancer and death from other causes assessed → death clearance date December 31, 1999

Study Population

- Population-based case cohort study of 2,740 patients diagnosed with prostate cancer between 1990 and 1998 in Ontario
- Stratified by CCOR random sample
- Cases were sampled independently of subcohort
- Death from prostate cancer and death from other causes assessed → death clearance date December 31, 1999
- Death from any cause updated for subcohort only → death clearance date December 31, 2008

Improvements

Stratified by CCOR Cox PH

$$S_j(t) = S_{0j}(t)^{\exp(\beta_1 \cdot \text{Age} + \beta_2 \cdot \text{CIRS-G}_{\text{pros}})}, \quad j = 1, \dots, 8$$

Improvements

Stratified by CCOR Cox PH

$$S_j(t) = S_{0j}(t)^{\exp(\beta_1 \cdot \text{Age} + \beta_2 \cdot \text{CIRS-G}_{\text{pros}})}, \quad j = 1, \dots, 8$$

We improve on Wykes (2011) work in two ways:

Improvements

Stratified by CCOR Cox PH

$$S_j(t) = S_{0j}(t)^{\exp(\beta_1 \cdot \text{Age} + \beta_2 \cdot \text{CIRS-G}_{\text{pros}})}, \quad j = 1, \dots, 8$$

We improve on Wykes (2011) work in two ways:

- we estimate both the baseline survival function and relative risk parameters using the case cohort population

Improvements

Stratified by CCOR Cox PH

$$S_j(t) = S_{0j}(t)^{\exp(\beta_1 \cdot \text{Age} + \beta_2 \cdot \text{CIRS-G}_{\text{pros}})}, \quad j = 1, \dots, 8$$

We improve on Wykes (2011) work in two ways:

- we estimate both the baseline survival function and relative risk parameters using the case cohort population
- we use a stratified by CCOR analysis, which assumes a different baseline hazard for each region

Relative Risk Estimation

Prentice (1986) proposed a pseudolikelihood given by

$$\tilde{\mathcal{L}}(\beta) = \prod_{j=1}^D \frac{\exp(\mathbf{z}_{(j)}^T \beta)}{\sum_{k \in \tilde{R}_{t_j}} \exp(\mathbf{z}_k^T \beta) w_k}$$

Relative Risk Estimation

Prentice (1986) proposed a pseudolikelihood given by

$$\tilde{\mathcal{L}}(\beta) = \prod_{j=1}^D \frac{\exp(\mathbf{z}_{(j)}^T \beta)}{\sum_{k \in \tilde{R}_{t_j}} \exp(\mathbf{z}_k^T \beta) w_k}$$

- t_1, \dots, t_D : D distinct failure times

Relative Risk Estimation

Prentice (1986) proposed a pseudolikelihood given by

$$\tilde{\mathcal{L}}(\beta) = \prod_{j=1}^D \frac{\exp(\mathbf{z}_{(j)}^T \beta)}{\sum_{k \in \tilde{R}_{t_j}} \exp(\mathbf{z}_k^T \beta) w_k}$$

- t_1, \dots, t_D : D distinct failure times
- \tilde{R}_{t_j} : the risk set at failure time t_j

Relative Risk Estimation

Prentice (1986) proposed a pseudolikelihood given by

$$\tilde{\mathcal{L}}(\beta) = \prod_{j=1}^D \frac{\exp(\mathbf{z}_{(j)}^T \beta)}{\sum_{k \in \tilde{R}_{t_j}} \exp(\mathbf{z}_k^T \beta) w_k}$$

- t_1, \dots, t_D : D distinct failure times
- \tilde{R}_{t_j} : the risk set at failure time t_j
- w_k : the weight applied to the risk set

Weighting Schemes

$$\frac{\exp(\mathbf{z}_i^T \boldsymbol{\beta})}{Y_i(t_j)w_i(t_j)\exp(\mathbf{z}_i^T \boldsymbol{\beta}) + \sum_{k \in SC, k \neq i} Y_k(t_j)w_k(t_j)\exp(\mathbf{z}_k^T \boldsymbol{\beta})}$$

Table: Comparing different values of w_k

Outcome type and timing	Prentice (1986)	Self & Prentice (1988)	Barlow (1994)
Case outside SC before failure	0	0	0
Case outside SC at failure	1	0	1
Case in SC before failure	1	1	$1/\alpha$
Case in SC at failure	1	1	1
SC control	1	1	$1/\alpha$

Relative Risk Estimation

Therneau & Li (1999) showed that $\text{Var}(\tilde{\beta})$ can be computed by adjusting the outputted variance from standard Cox model programs:

$$\tilde{\mathcal{I}}^{-1} + (1 - \alpha) D_{SC}^T D_{SC}$$

Relative Risk Estimation

Therneau & Li (1999) showed that $\text{Var}(\tilde{\beta})$ can be computed by adjusting the outputted variance from standard Cox model programs:

$$\tilde{\mathcal{I}}^{-1} + (1 - \alpha) D_{SC}^T D_{SC}$$

- $\tilde{\mathcal{I}}^{-1}$: variance returned by the Cox model program

Relative Risk Estimation

Therneau & Li (1999) showed that $\text{Var}(\tilde{\beta})$ can be computed by adjusting the outputted variance from standard Cox model programs:

$$\tilde{\mathcal{I}}^{-1} + (1 - \alpha) D_{SC}^T D_{SC}$$

- $\tilde{\mathcal{I}}^{-1}$: variance returned by the Cox model program
- D_{SC} : subset of matrix of *dfbeta* residuals that contains subcohort members only

Absolute Risk Estimation

A weighted version of the Breslow estimator for the full cohort with covariate history $\mathbf{z}_0(\mathbf{u})$ is given by

$$\begin{aligned} d\hat{\Lambda}(u, \mathbf{z}_0(\mathbf{u})) &= d\hat{\Lambda}_0(u) \exp \left(\mathbf{z}_0^T(\mathbf{u}) \hat{\beta} \right) \\ &= \sum_{i=1}^n \frac{dN_i(u)}{\left(\frac{n}{m}\right) \sum_{j \in \tilde{R}_i(u)} \exp \left(\mathbf{z}_j^T(\mathbf{u}) \hat{\beta} \right)} \exp \left(\mathbf{z}_0^T(\mathbf{u}) \hat{\beta} \right) \\ &= \sum_{i=1}^n \frac{dN_i(u)}{\left(\frac{n}{m}\right) \sum_{j \in \tilde{R}_i(u)} \exp \left((\mathbf{z}_j(\mathbf{u}) - \mathbf{z}_0(\mathbf{u}))^T \hat{\beta} \right)} \\ &= \sum_{i=1}^n \frac{m}{n} \frac{dN_i(u)}{\sum_{j \in \tilde{R}_i(u)} \exp \left((\mathbf{z}_j(\mathbf{u}) - \mathbf{z}_0(\mathbf{u}))^T \hat{\beta} \right)} \end{aligned}$$

Absolute Risk Estimation

The sum of these increments can be used to calculate the risk between two time points, say s and t , for $s < t$:

$$\hat{A}(s, t, z_0) = \hat{A}(0, t, z_0) - \hat{A}(0, s, z_0) = \int_s^t d\hat{A}(u, z_0(u))$$

Absolute Risk Estimation

The sum of these increments can be used to calculate the risk between two time points, say s and t , for $s < t$:

$$\hat{\Lambda}(s, t, z_0) = \hat{\Lambda}(0, t, z_0) - \hat{\Lambda}(0, s, z_0) = \int_s^t d\hat{\Lambda}(u, z_0(u))$$

For this project, we will be only considering the case when $s = 0$. The corresponding survival probability is

$$\hat{S}(0, t, z_0) = \exp\left(-\hat{\Lambda}(0, t, z_0)\right)$$

Software

- PROC PHREG in SAS/STAT software and coxph in R

Software

- PROC PHREG in SAS/STAT software and coxph in R
- SAS macro developed by Langholz & Jiao (2007)

Computational Methods for the Case Cohort Design

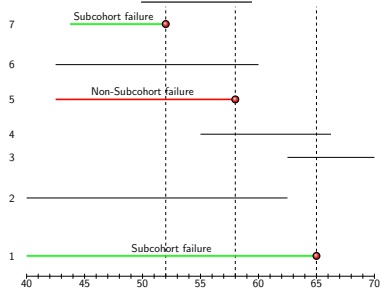
- Standard Cox PH functions in SAS and R cannot directly compute the relative and absolute risk estimates

Computational Methods for the Case Cohort Design

- Standard Cox PH functions in SAS and R cannot directly compute the relative and absolute risk estimates
- Modifications must be made to the dataset to make use of PROC PHREG and coxph functions

Computation of Pseudolikelihood

Manually



$$\tilde{\mathcal{L}}(\beta) =$$

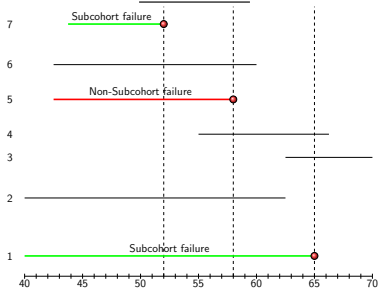
$$\frac{\exp(Z_7\beta)}{\exp(Z_1\beta) + \exp(Z_2\beta) + \exp(Z_6\beta) + \exp(Z_7\beta)}$$

$$\times \frac{\exp(Z_5\beta)}{\exp(Z_1\beta) + \exp(Z_2\beta) + \exp(Z_4\beta) + \exp(Z_5\beta) + \exp(Z_6\beta)}$$

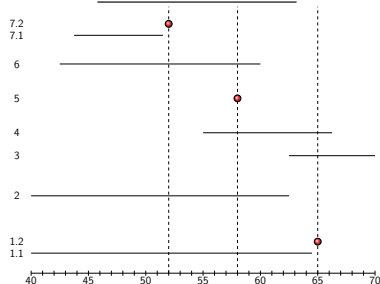
$$\times \frac{\exp(Z_1\beta)}{\exp(Z_1\beta) + \exp(Z_3\beta) + \exp(Z_4\beta)}$$

Computation of Pseudolikelihood

Manually



Computationally



$$\tilde{L}(\beta) =$$

$$\begin{aligned} & \frac{\exp(Z_7\beta)}{\exp(Z_1\beta) + \exp(Z_2\beta) + \exp(Z_6\beta) + \exp(Z_7\beta)} \\ & \times \frac{\exp(Z_5\beta)}{\exp(Z_1\beta) + \exp(Z_2\beta) + \exp(Z_4\beta) + \exp(Z_5\beta) + \exp(Z_6\beta)} \\ & \times \frac{\exp(Z_1\beta)}{\exp(Z_1\beta) + \exp(Z_3\beta) + \exp(Z_4\beta)} \end{aligned}$$

$$\tilde{L}(\beta) =$$

$$\begin{aligned} & \frac{\exp(Z_{7.2}\beta)}{\exp(Z_{1.1}\beta) + \exp(Z_2\beta) + \exp(Z_6\beta) + \exp(Z_{7.2}\beta)} \\ & \times \frac{\exp(Z_5\beta)}{\exp(Z_{1.1}\beta) + \exp(Z_2\beta) + \exp(Z_4\beta) + \exp(Z_5\beta) + \exp(Z_6\beta)} \\ & \times \frac{\exp(Z_{1.2}\beta)}{\exp(Z_{1.2}\beta) + \exp(Z_3\beta) + \exp(Z_4\beta)} \end{aligned}$$

Table: All Cause mortality

Covariate	All Cause		
	1999	2008	Wykes (2011)
Age	1.04 (1.02, 1.05)	1.05 (1.04, 1.06)	1.02 (1.01, 1.03)
CIRS- G_{pros}	1.14 (1.10, 1.18)	1.13 (1.10, 1.16)	1.12 (1.09, 1.15)

Table: All Cause mortality

Covariate	All Cause		
	1999	2008	Wykes (2011)
Age	1.04 (1.02, 1.05)	1.05 (1.04, 1.06)	1.02 (1.01, 1.03)
CIRS-G _{pros}	1.14 (1.10, 1.18)	1.13 (1.10, 1.16)	1.12 (1.09, 1.15)

Table: Cause Specific mortality

Covariate	Cause Specific	
	PCa (1999) ^a	OC (1999)
Age	1.00 (0.99, 1.02)	1.07 (1.05, 1.09)
CIRS-G _{pros}	1.032 (0.99, 1.07)	1.25 (1.20, 1.30)

^a predictors not significant

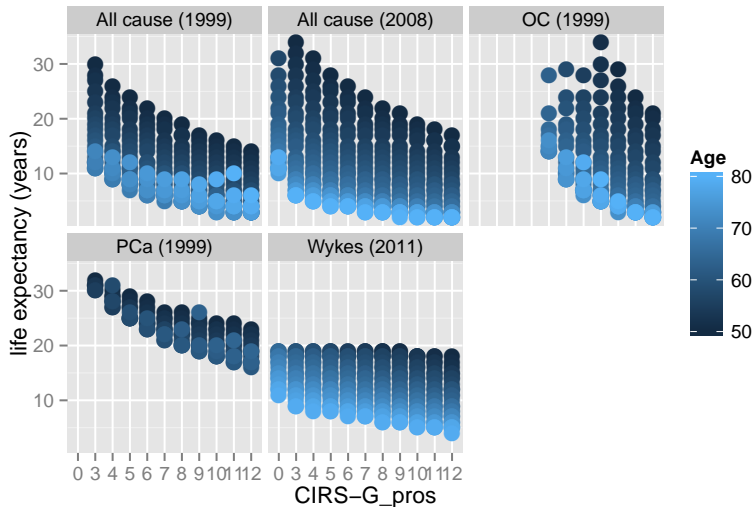


Figure: Estimates from East Region (Wykes estimates based on all cause unstratified subcohort at 2008)

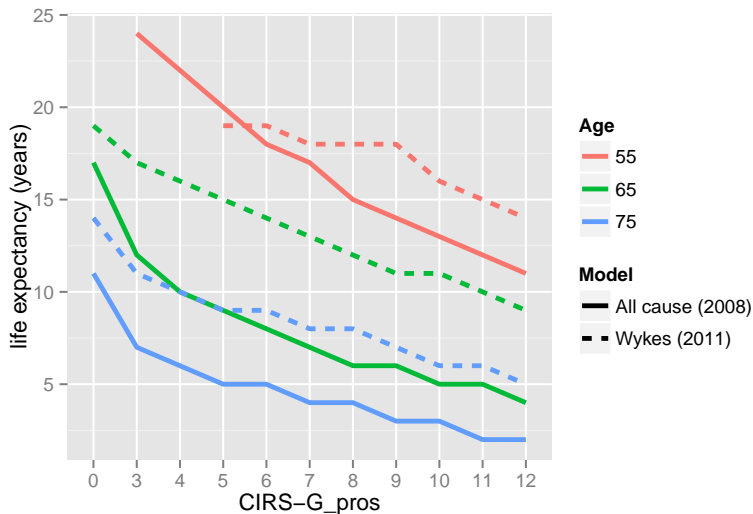
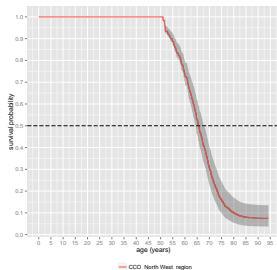
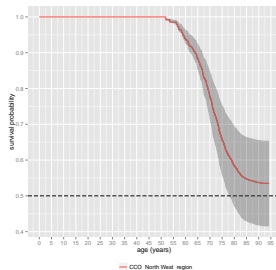


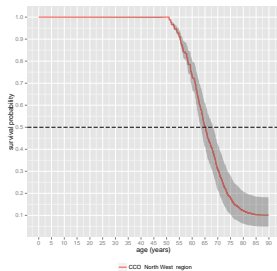
Figure: Estimates from East Region (Wykes estimates based on all cause unstratified subcohort at 2008)



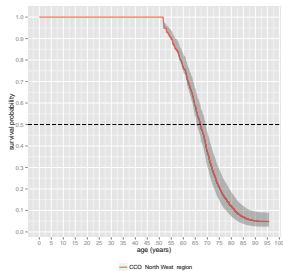
(a) All cause (1999)



(b) Other cause(1999)



(c) Prostate cancer(1999)



(d) All cause (2008)

Tool for Use in a Clinical Setting

- Built using shiny package for R

Tool for Use in a Clinical Setting

- Built using shiny package for R
- Requires R and an internet connection

Tool for Use in a Clinical Setting

- Built using shiny package for R
- Requires R and an internet connection
- Three simple lines of code

Contributions

- 1 Life expectancy estimates based on age and comorbidity score from a population based case cohort study

Contributions

- 1 Life expectancy estimates based on age and comorbidity score from a population based case cohort study
- 2 Improves previous work done with this dataset by using appropriate statistical methodology for case cohort design

Contributions

- 1 Life expectancy estimates based on age and comorbidity score from a population based case cohort study
- 2 Improves previous work done with this dataset by using appropriate statistical methodology for case cohort design
- 3 Easy-to-use interactive web tool for clinical use

Future Work

- Internal validation → risk prediction accuracy, PH assumption, compare expected values from 1999 model to observed 2008 subcohort

Future Work

- Internal validation → risk prediction accuracy, PH assumption, compare expected values from 1999 model to observed 2008 subcohort
- External validation → generalizability

Future Work

- Internal validation → risk prediction accuracy, PH assumption, compare expected values from 1999 model to observed 2008 subcohort
- External validation → generalizability
- Convert Langholz & Jiao (2007) SAS macro to R for absolute risk estimates

Alternative Analysis

- Treat the case cohort design as a cohort study with missing data (NestedCohort and Survey packages for R) (Mark & Katki (2008), Breslow & Lumley (2009))

Alternative Analysis

- Treat the case cohort design as a cohort study with missing data (NestedCohort and Survey packages for R) (Mark & Katki (2008), Breslow & Lumley (2009))
- Competing Risk analysis using the subdistribution hazard of Fine & Gray

References I



M. Center et al.

International Variation in Prostate Cancer Incidence and Mortality Rates

European Association of Urology, 61:2012



Bryan Langholz and Jenny Jiao

Computational methods for case cohort studies

Computational Statistics & Data Analysis, 51:2007



Therneau, T and Li, H

Computing the Cox model for case cohort designs

Lifetime data analysis, 2(5):1999

References II



PRENTICE, R. L.

A case-cohort design for epidemiologic cohort studies and disease prevention trials

Biometrika, 73(1):1986



RStudio and Inc.

shiny: Web Application Framework for R

R package version 0.6.0, 17(25):2012



T. Lumley

survey: analysis of complex survey samples

R package version 3.28-2, 2012

References III



Katki, H.A. and Mark, S.D.

Survival Analysis for Cohorts with Missing Covariate
Information

The R Journal, 2008

Influential Analysis

Dfbeta residuals

Are the approximate changes in the parameter estimates $(\hat{\beta} - \hat{\beta}_{(j)})$ when the j^{th} observation is omitted. These variables are a weighted transform of the score residual variables and are useful in assessing local influence and in computing approximate and robust variance estimates

Standardized *dfbeta* residuals:

$$\frac{(\hat{\beta} - \hat{\beta}_{(j)})}{sd(\hat{\beta} - \hat{\beta}_{(j)})}$$

Why is the PsL not a Partial Likelihood

- Not a PL \rightarrow Asymptotic distribution theory for the PL estimators breaks down because cases outside the subcohort induce non-nesting of the σ -fields, or sets on which the counting process probability measure is defined
- cannot use Likelihood ratio tests since PsL are not real likelihoods

Consistent Estimators

An estimator $\hat{\theta}$ will perform better and better as we obtain more samples. If at the limit $n \rightarrow \infty$ the estimator tends to be always right, it is said to be consistent. This notion is equivalent to convergence in probability

Convergence in Probability

Let X_1, \dots, X_n be a sequence of iid RVs drawn from a distribution with parameter θ and $\hat{\theta}$ an estimator for θ . $\hat{\theta}$ is consistent as an estimator of θ if $\hat{\theta} \xrightarrow{P} \theta$

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| > \epsilon) = 0, \quad \forall \epsilon > 0$$

Competing Risks

- when disease of interest has a delayed occurrence, other events may preclude observation of disease of interest giving rise to a competing risk situation
- to analyze the event of interest taking into account the competing risks, one has to model the hazard of subdistribution. In the presence of competing risks, the probability to observe the event of interest spans the interval $[0, p]$, $p < 1$. Because p does not reach 1, this function is not a proper distribution function and it is called a subdistribution
- Modify the partial likelihood in the Cox model to model hazard of subdistribution, such that the individuals experiencing competing risks are always in the risk set with their contribution regulated by a weight

Competing Risks

Fine and Gray Partial Likelihood:

$$PL(\beta) = \prod_{j=1}^n \left(\frac{\exp \{ \beta x_j \}}{\sum_{r \in \tilde{R}_j} w_{rj} \exp \{ \beta x_r \}} \right)^{\delta_j} \quad (1)$$

$$w_{rj} = \frac{\hat{G}(t_j)}{\hat{G}(\min(t_j, t_r))} \quad (2)$$

$$C_j = \begin{cases} 1 & \text{if the event of interest was observed at time } t_j \\ 2 & \text{if the competing risk event was observed at time } t_j \\ 0 & \text{if no event was observed at time } t_j \end{cases}$$

$$\delta_j = \begin{cases} 1 & C_j = 1 \\ 0 & C_j = 0 \text{ or } C_j = 2 \end{cases}$$

Competing Risks

- $\hat{G}(t_j)$: estimated probability of being censored at time t_j by either the event of interest or a competing risks event
- $\tilde{R}_j = \{r; t_r \geq t_j \text{ or } C_r = 2\}$
- individuals experiencing competing risk event are always considered in the risk set at all time points
- The weight controls how much a specific observation participates in the sum in the denominator
- If the observation at time t_r is censored ($C_r = 0$) or has the event of interest ($C_r = 1$), the weight is either 1 or 0 depending whether the observation is in the risk set ($t_r \geq t_j$) or it is not in the risk set ($t_r < t_j$)
- If the observation has a competing risk event, then it is always in the risk set and the weight is 1 if ($t_r \geq t_j$) or a quantity less than 1 if $t_r < t_j$
- This quantity decreases as the distance between t_r and t_j increases

Case-cohort in the presence of competing risks

$$PsL(\beta) = \prod_{j=1}^n \left(\frac{\exp \{ \beta x_j \}}{\sum_{r \in \tilde{R}_j} I_{rj} w_{rj} \exp \{ \beta x_r \}} \right)^{\delta_j}$$

$$\delta_j = \begin{cases} 1 & C_j = 1 \\ 0 & C_j = 0 \text{ or } C_j = 2 \end{cases}$$

$$I_{rj} = \begin{cases} 1 & r \text{ is in subcohort or } r \text{ is a case outside the subcohort and } t_r = t_j \\ 0 & r \text{ is a case outside the subcohort and } t_r \neq t_j \end{cases}$$

$w_{rj} = \frac{\hat{G}(t_j)}{\hat{G}(\min(t_j, t_r))}$, where $\hat{G}(t_j)$ is the estimation of the probability of censoring regardless of the fact that in a case-cohort study the cases are oversampled