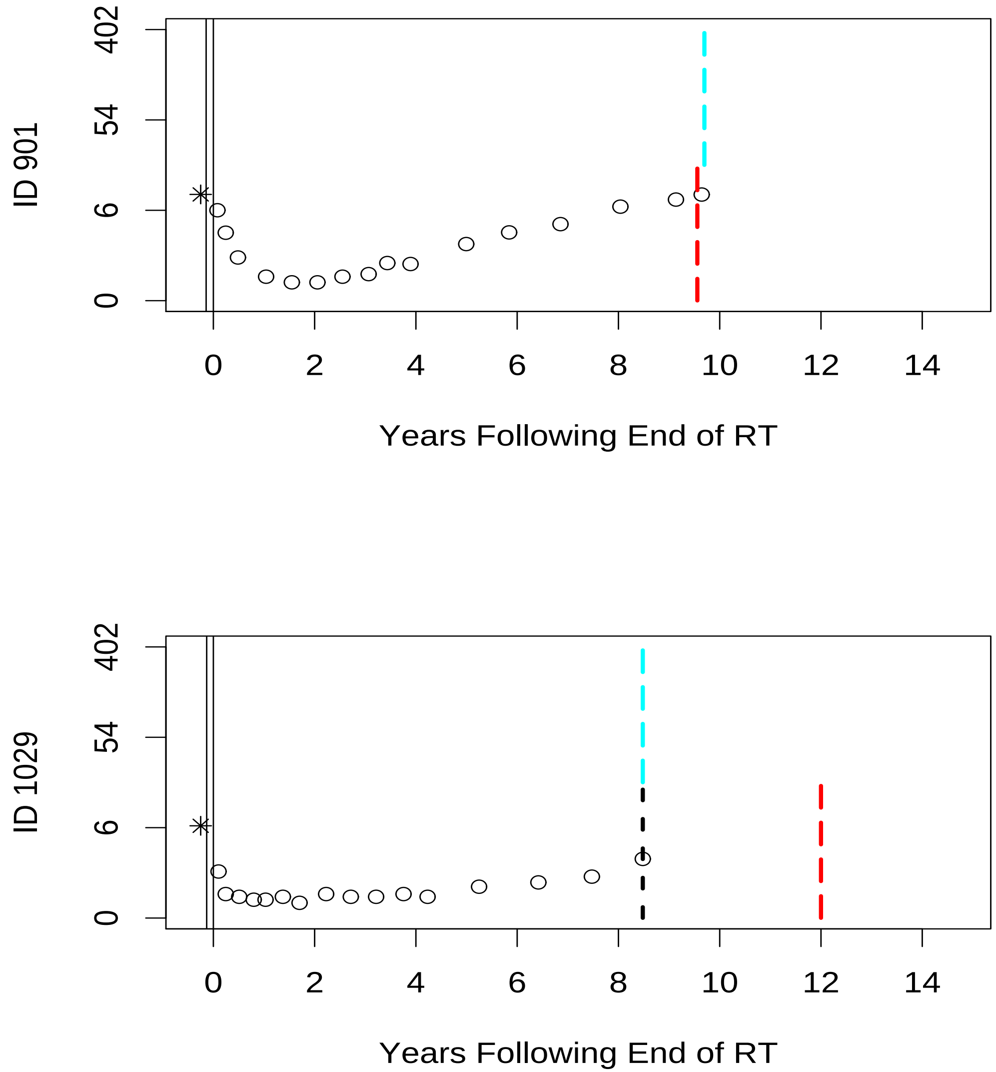
Individual Prediction and Validation Using Joint Longitudinal-Survival Models in Prostate Cancer Studies

Jeremy M G Taylor
University of Michigan

1. INTRODUCTION
   - Prostate Cancer
   - Motivation; Individual Predictions
   - Joint Longitudinal and Survival Models
2. PROSTATE CANCER APPLICATION
   - Datasets, Models and Estimation
   - Individual predictions, longitudinal and survival
3. WEBSITE
4. VALIDATION

PROSTATE CANCER

- Common cancer in older men
- Usually growing slowly, most people diagnosed with prostate cancer die of something else
- Treatment for localised disease
  - Radiation therapy (plus or minus hormones)
  - Surgery
- Following radiation PSA rise suggests cancer is regrowing
- Biochemical recurrence
  - Based on PSA
  - Not the real thing
- Clinical recurrence is the real thing (but PSA is useful)


ID 901
402
54
6
0
0 2 4 6 8 10 12 14
Years Following End of RT
ID 1029
402
54
6
0
0 2 4 6 8 10 12 14
Years Following End of RT

GOAL

- Develop a website for patients and their physicians, psacalc.sph.umich.edu
- The patients were previously treated with radiation therapy for localised prostate cancer
- The patient inputs individual characteristics (stage of disease, treatment dose) and post-treatment measures of health
- The website provides quantitative information about future disease progression

MOCK EXAMPLE

- Patient treated in Oct 2003 for prostate cancer
- Pre-treatment characteristics
  - PSA = 10.3
  - T-stage = 3
  - Gleason grade = 8
  - Treatment dose was 74 Gy
- Patient has not experienced any clinical recurrence of prostate cancer

Table 1: Post treatment PSA measurements

Date	PSA
29 Feb 2004	2.1
1 May 2004	1.5
25 Dec 2004	1.1
4 July 2006	1.3
29 Feb 2008	1.7
1 Jan 2010	2.1

- What is the probability of the prostate cancer coming back within 3 years of today?

- Joint model trained on a large dataset
- Parameter estimates applied to this patient
- P(prostate cancer recurrence within 3 years) = 0.22

Prob(clinical recurrence within 3 years)=0.22

- What should you do?
  - Intervene with salvage hormone therapy?
  - Order another PSA test for X months in the future?
  - Don't change the original plan
- This talk
  - How do we get 0.22
  - Attempts to validate the prediction

JOINT MODELS FOR LONGITUDINAL AND SURVIVAL DATA

- Setting: clinical trial or observational study
- Data
  - (t_i, δ_i), censored event time
  - X_i, time-independent covariates
  - Y_{ij}, time-dependent covariate, biomarker
- Both T and Y are response variables

- Modelling choices for joint distribution of T and Y
  - $[T, Y|X]$
  - Factor as $[T|X]$ and $[Y|X, T]$
  - Factor as $[Y|X]$ and $[T|Y, X]$

- $[Y|X]$ and $[T|Y,X]$
  - Faucett and Thomas (1996), Wulfsohn and Tsiatis (1997), and others
  - Most popular
  - Usually involves latent variables, R_i
  - $[Y_{ij}|X_i, R_i]$
  - $[T_i|Y_i, X_i, R_i] = [T_i|X_i, R_i]$

- Factor $[T, Y \mid X]$ as $[Y \mid X][T \mid Y, X]$
- $[Y \mid X]$, longitudinal model
  - longitudinal random effects model
- $[T \mid Y, X]$
  - time-dependent proportional hazards model
- Use joint model for prediction of future longitudinal and event times for individual patients

PROSTATE CANCER DATASETS

- Prostate cancer patients treated with radiation therapy.
- Training data
  - RTOG, n=615
  - Detroit, n=1268
  - Univ Michigan, n=503
- Testing data
  - Melbourne, n=395
  - Vancouver, n=846

Longitudinal Data (Y). Post-treatment PSA

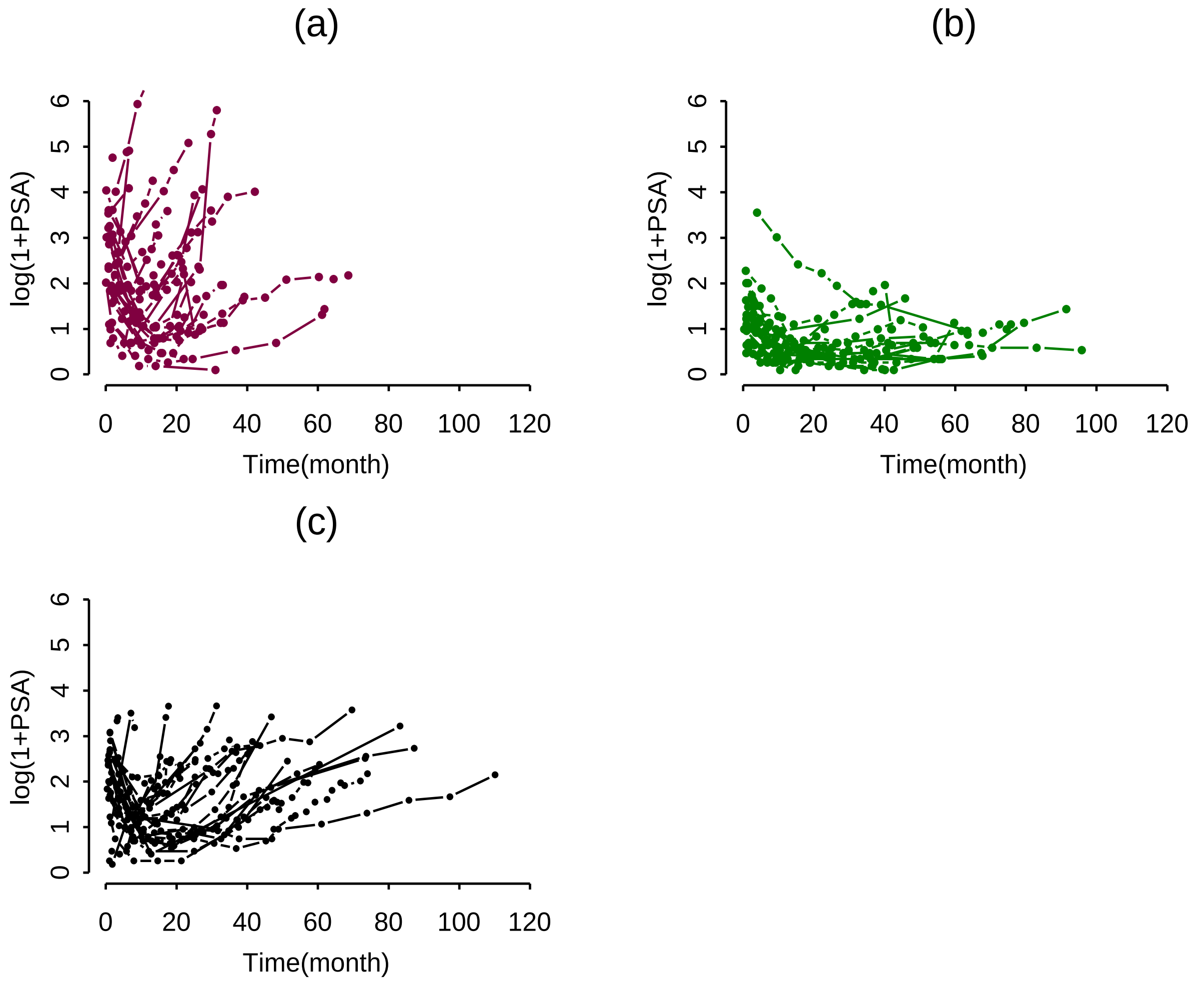
- measured approximately every 6 months
- a total of 46,000 post-treatment PSA values
- median no. of PSA per patient is 8
- 10 year follow-up

Endpoints and Censoring (T)

- 15% events: local/regional recurrence, distant metastasis
- 85% censored patients:
  - 20% are dead not from prostate cancer.
  - 65% are lost to follow-up or censored by the end of the study
- 10% of patients received salvage hormone therapy (HT) prior to recurrence (because of rising PSA).

PSA profiles for 3 groups.

(a) Events, (b) Censored, (c) Hormonal Therapy.

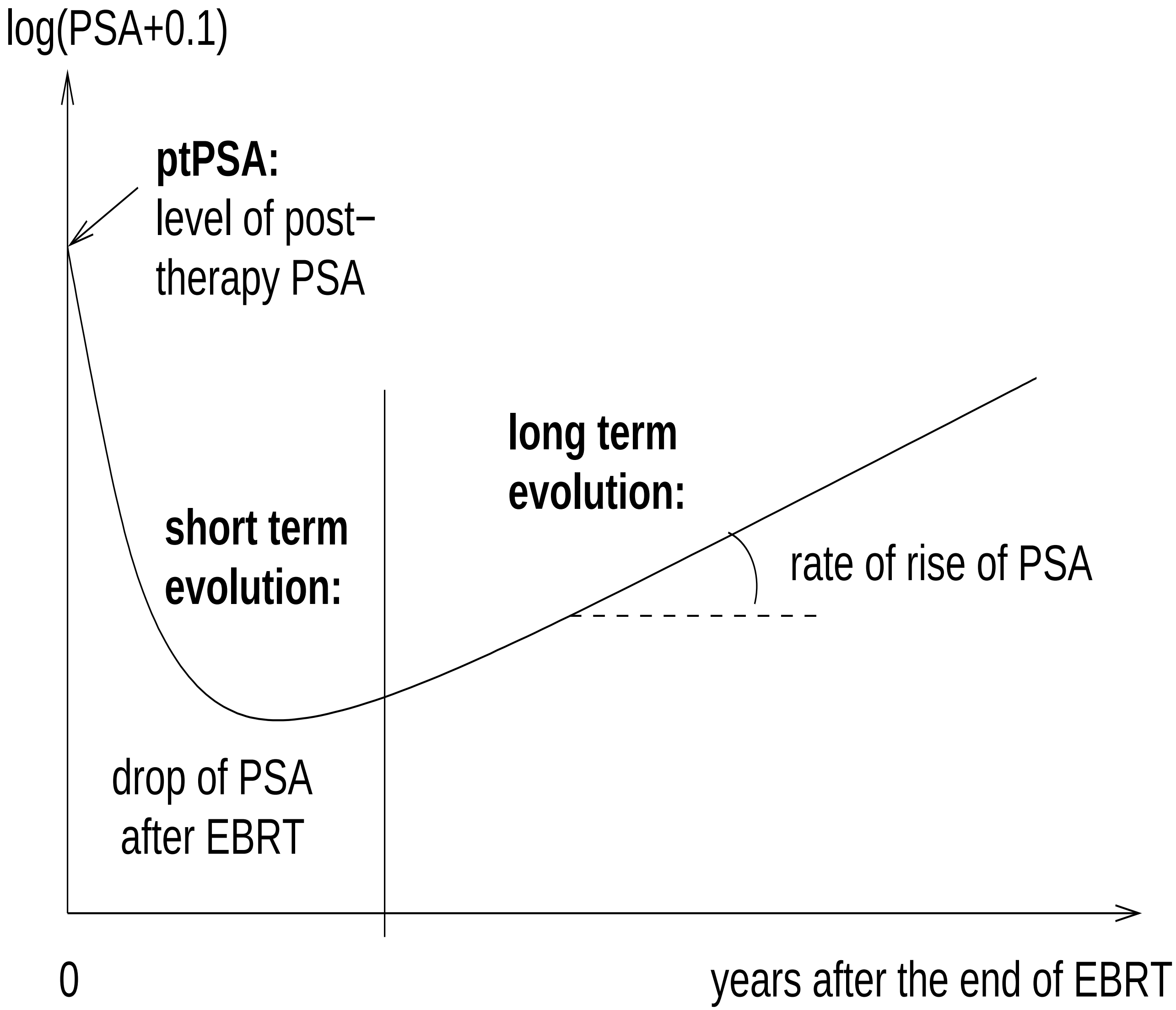


Statistical Model

<u>Notation</u>

- $\mathbf{X_i}$ - baseline covariates.
- $Y_i(t) = PSA_i(t)$ - longitudinal PSA data
- $\mathbf{T_i}$ - time of recurrence.
- $\mathbf{R_i}$ - random effects.
- Assumption - Y_i and T_i conditionally independent given R_i and X_i.

Longitudinal model



Longitudinal. Non-linear random effects models.

$$\log\Big[PSA_i(t) + 0.1\Big] = Z_i(t) + \epsilon_{it}$$

$$Z_i(t) = r_{i0} + r_{i1} f(t) + r_{i2} t$$

where $f(t) = ((1+t)^{-1.5} - 1)$, $\epsilon_{it} \sim$ t-distribution and $R_i = (r_{i0}, r_{i1}, r_{i2})$ are random effects for subject i.

$$[R_i \mid X_i] \sim N(\mu X_i, \Sigma)$$

$$X_i \;\; = \;\; (bPSA_i, Tstage_i, Gleason_i)$$

Hazard model. Time-dependent proportional hazards.

$$\lambda_i(t \mid X_i, Z_i, sl_i, HT_i)$$
$$= \lambda_0(t)$$
$$\exp[\eta g(Z_i(t)) + \omega sl_i(t) + \gamma X_i + \phi HT_i(t)]$$

$$sl_i(t) = \text{ slope of } Z_i(t)$$

$$HT_i(t) = \begin{cases} 0 & \text{if} \quad t < S_i \\ 1 & \text{if} \quad t > S_i \end{cases}$$

$\lambda_0(t)$ is a step function.

- Estimation via MCMC
  - parameters θ
  - latent variables R_i
  - draws of θ, R_i, save for later use
  - over 12 hours of computing to obtain estimates
- Likelihood
  $\prod_i \int [\prod_j P(Y_{ij}|\theta, X_i, R_i)] P(T_i, \delta_i|\theta, X_i, R_i) f(R_i) dR_i$

PREDICT FUTURE PSA VALUES.

From model

$$\log\Big[PSA_i(t) + 0.1\Big] = Z_i(t) + \epsilon_{it}$$

$$Z_i^k(t) = r_{i0}^k + r_{i1}^k f(t) + r_{i2}^k t$$

where k denotes k^{th} draw from posterior distribution (MCMC)

PREDICT RECURRENCE FOR CENSORED PATIENTS IN DATASET

- For patient i, the conditional probability of recurrence within a months
-

$$
\begin{aligned}
& P[\,T < t_i + a \,|\, T > t_i, Y_i, X_i] \\
= (1/K) \sum_k & P[\,T < t_i + a \,|\, T > t_i, X_i, \theta^k, R_i^k]
\end{aligned}
$$

where θ^k, R_i^k are draws from the posterior distribution

Residual time distribution

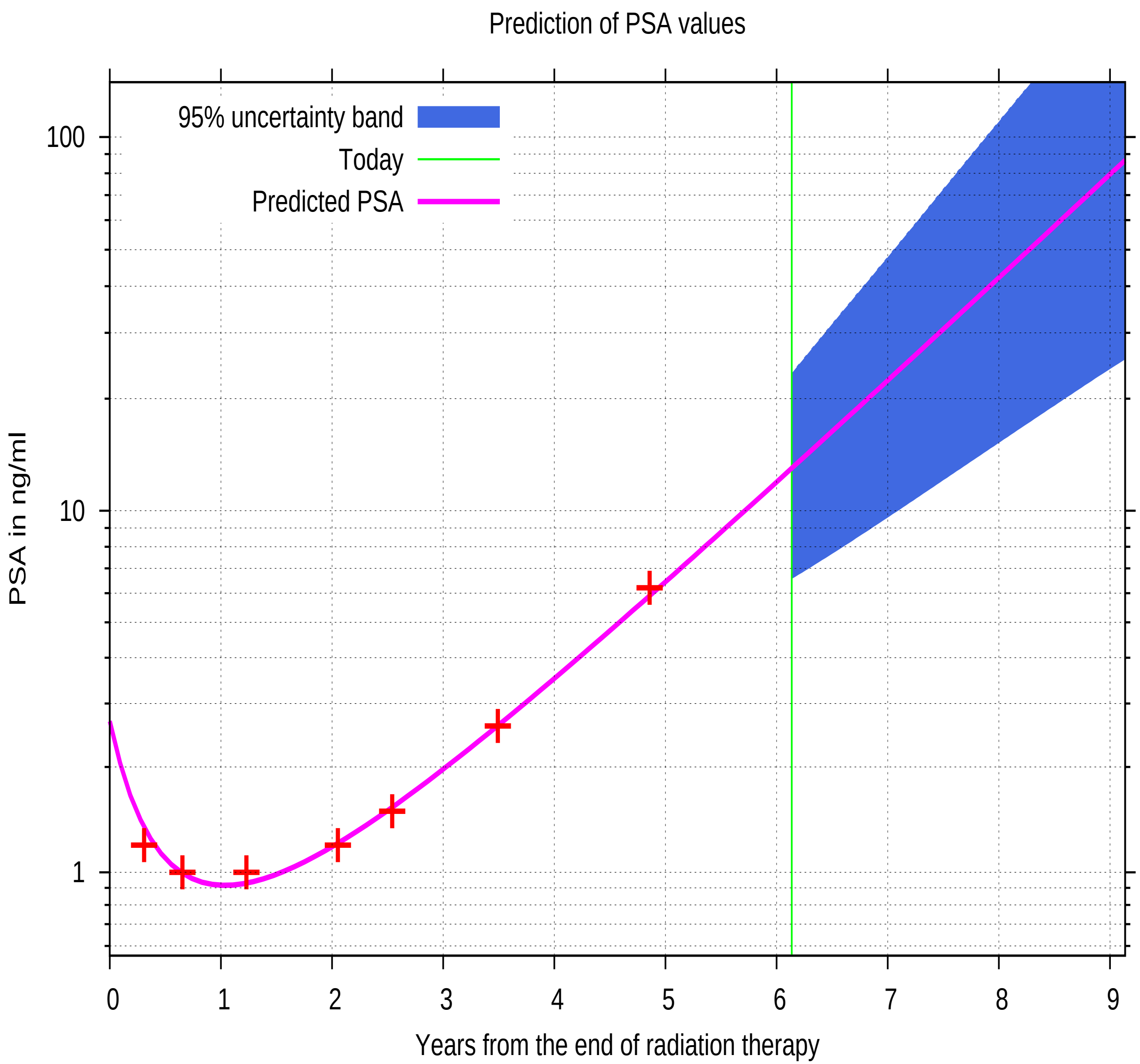
$$P[\,T > t_i + a \,|\, T > t_i, X_i, \theta, R_i]$$

$$= exp[-\int_{u=t}^{t+a} \lambda_i(u|X_i, \theta, Z_i(u), sl_i(u))du]$$

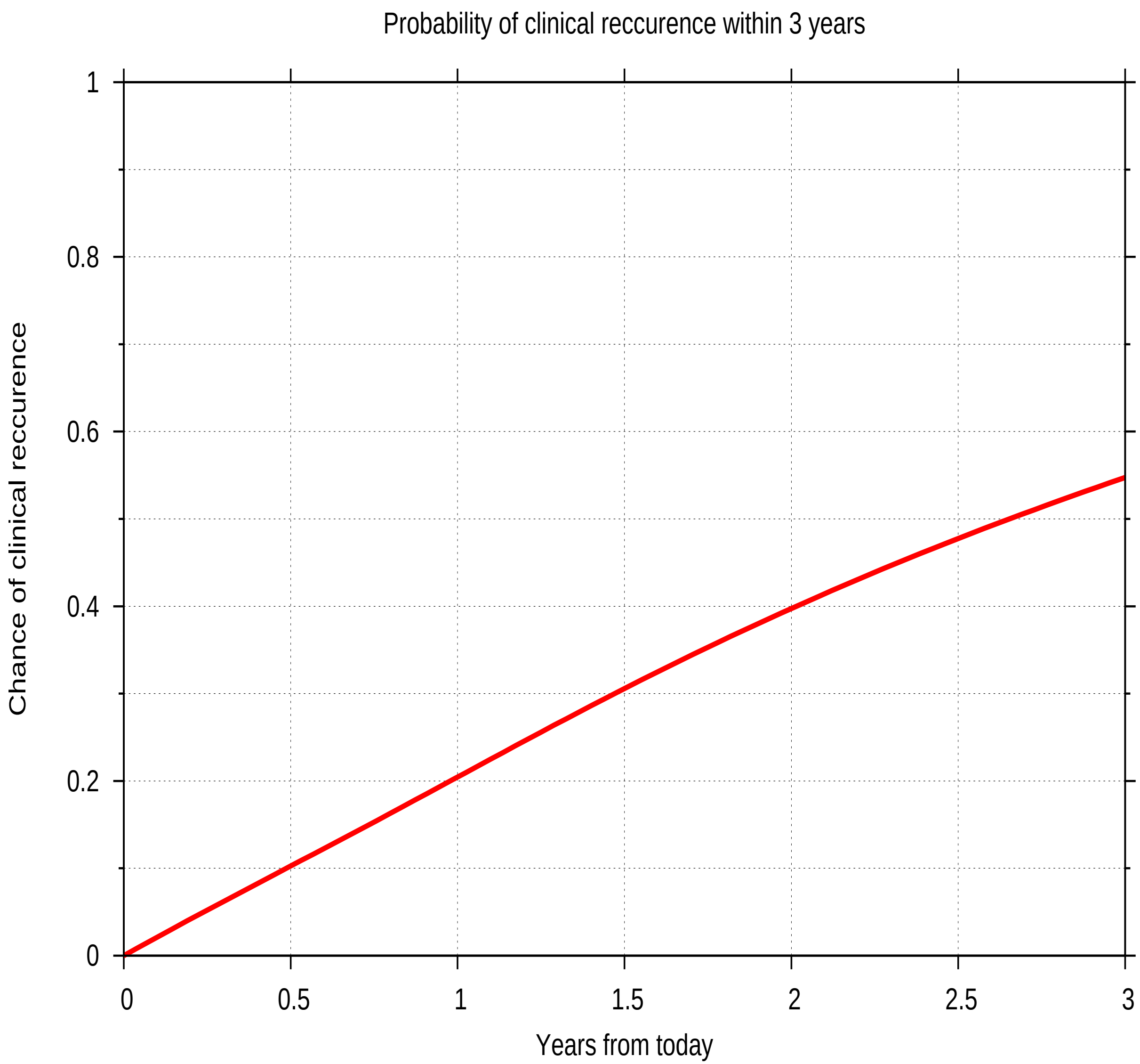
$$\lambda_i(u \mid X_i, \theta, Z_i(u), sl_i(u))$$

$$= \lambda_0(t)\exp[\eta g(Z_i(u)) + \omega sl_i(u) + \gamma X_i]$$

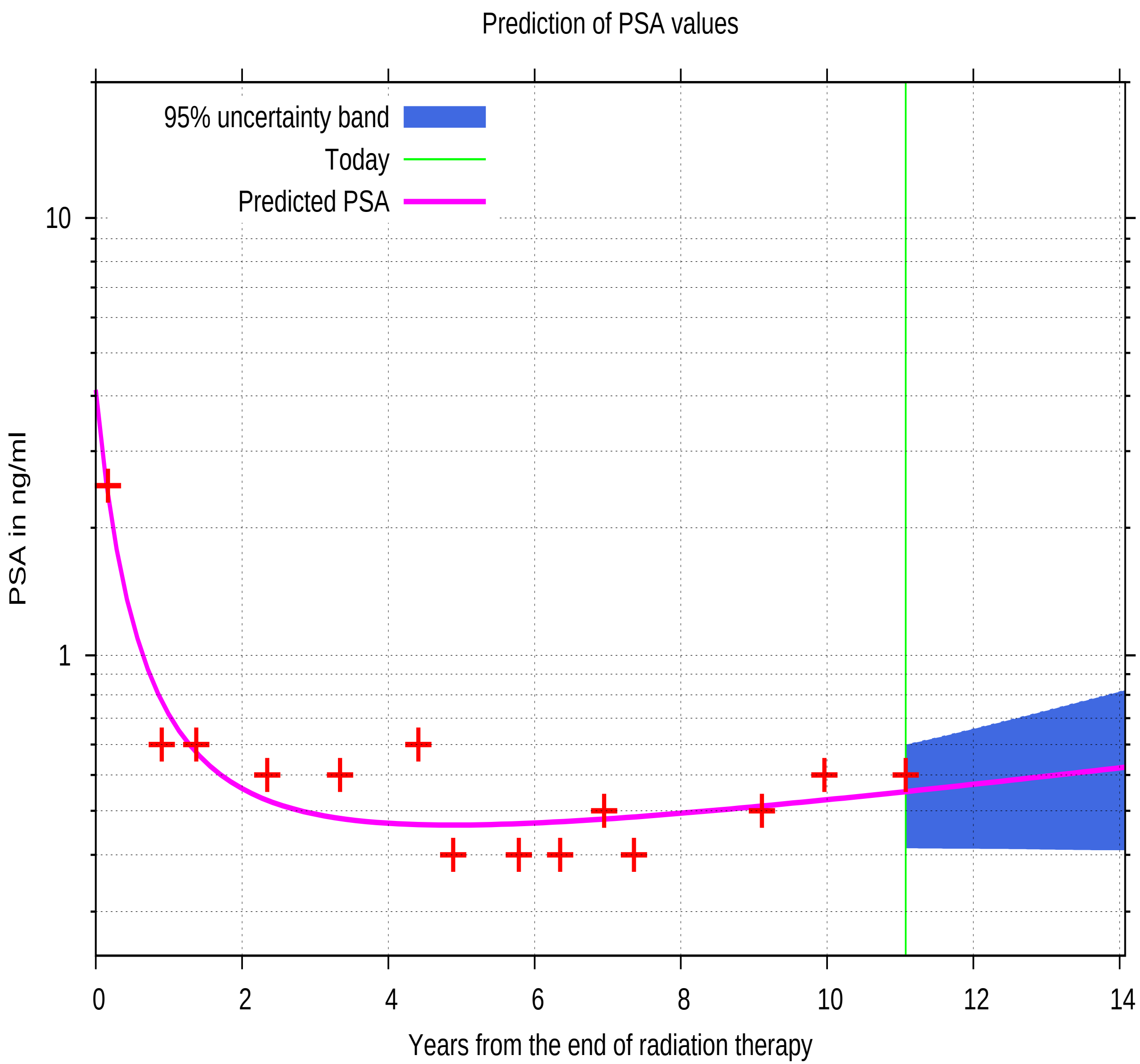
Predictions of PSA for Censored Subject



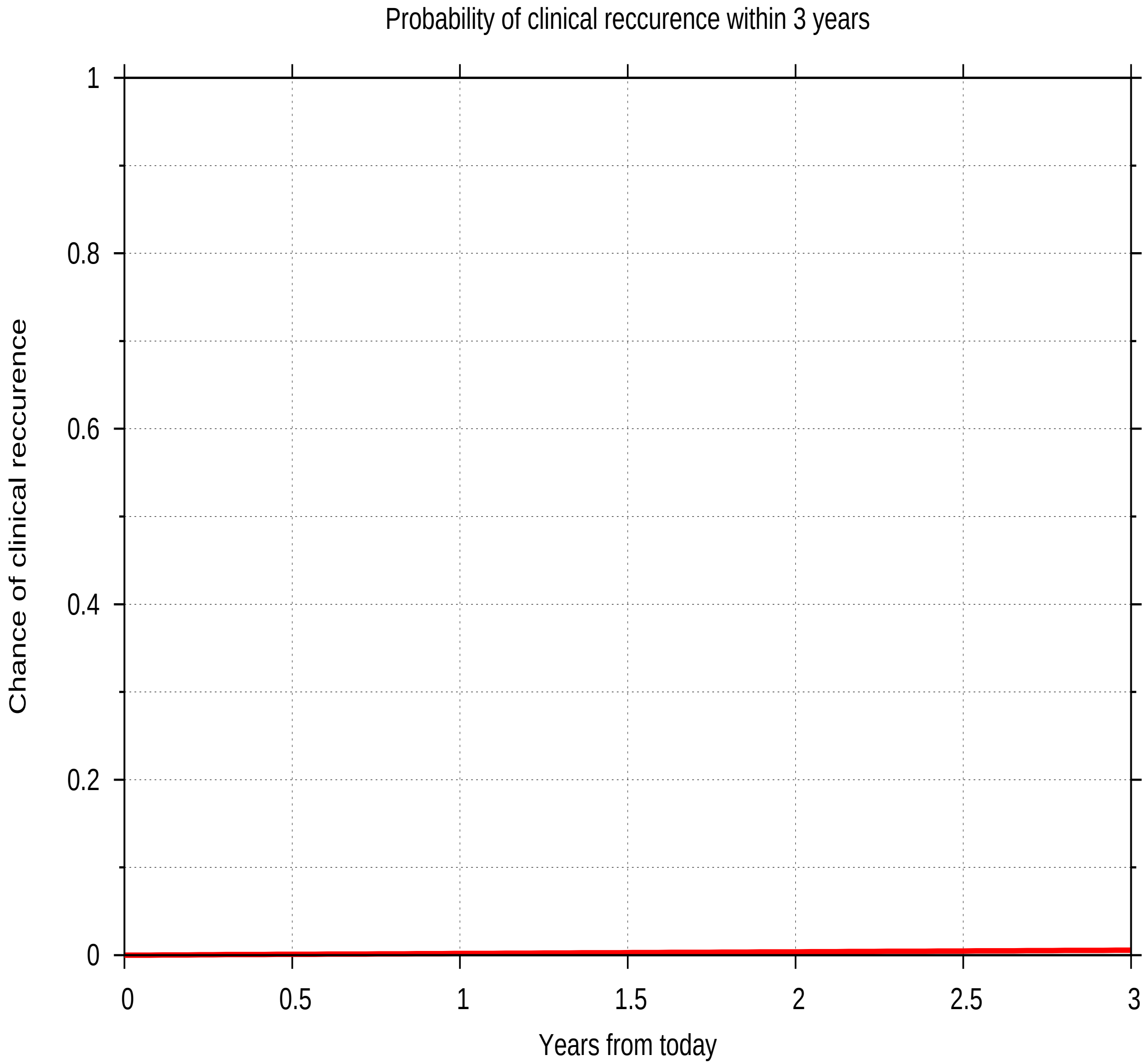
Residual Time Distribution for Censored Subject



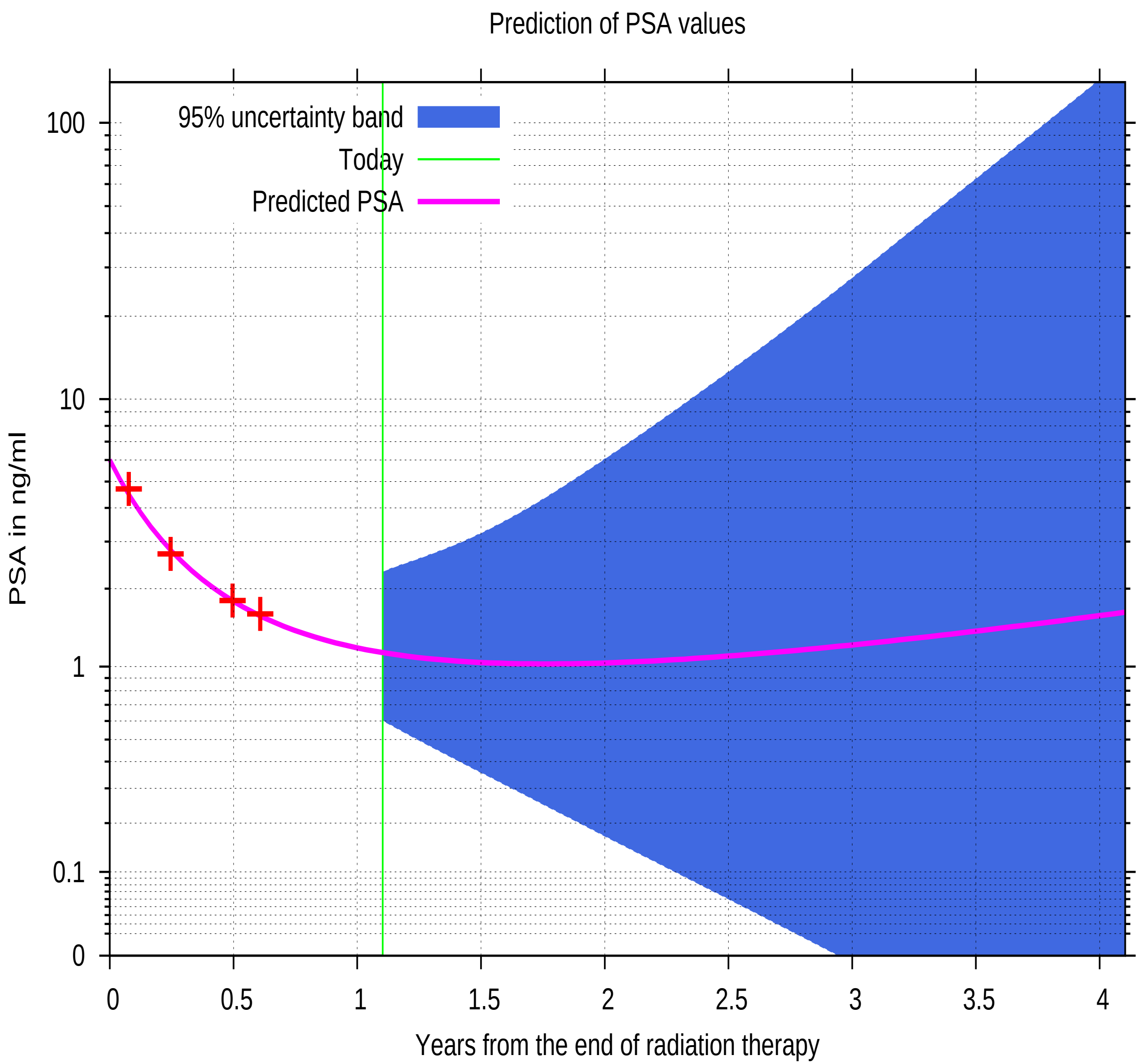
Predictions of PSA for Censored Subjects



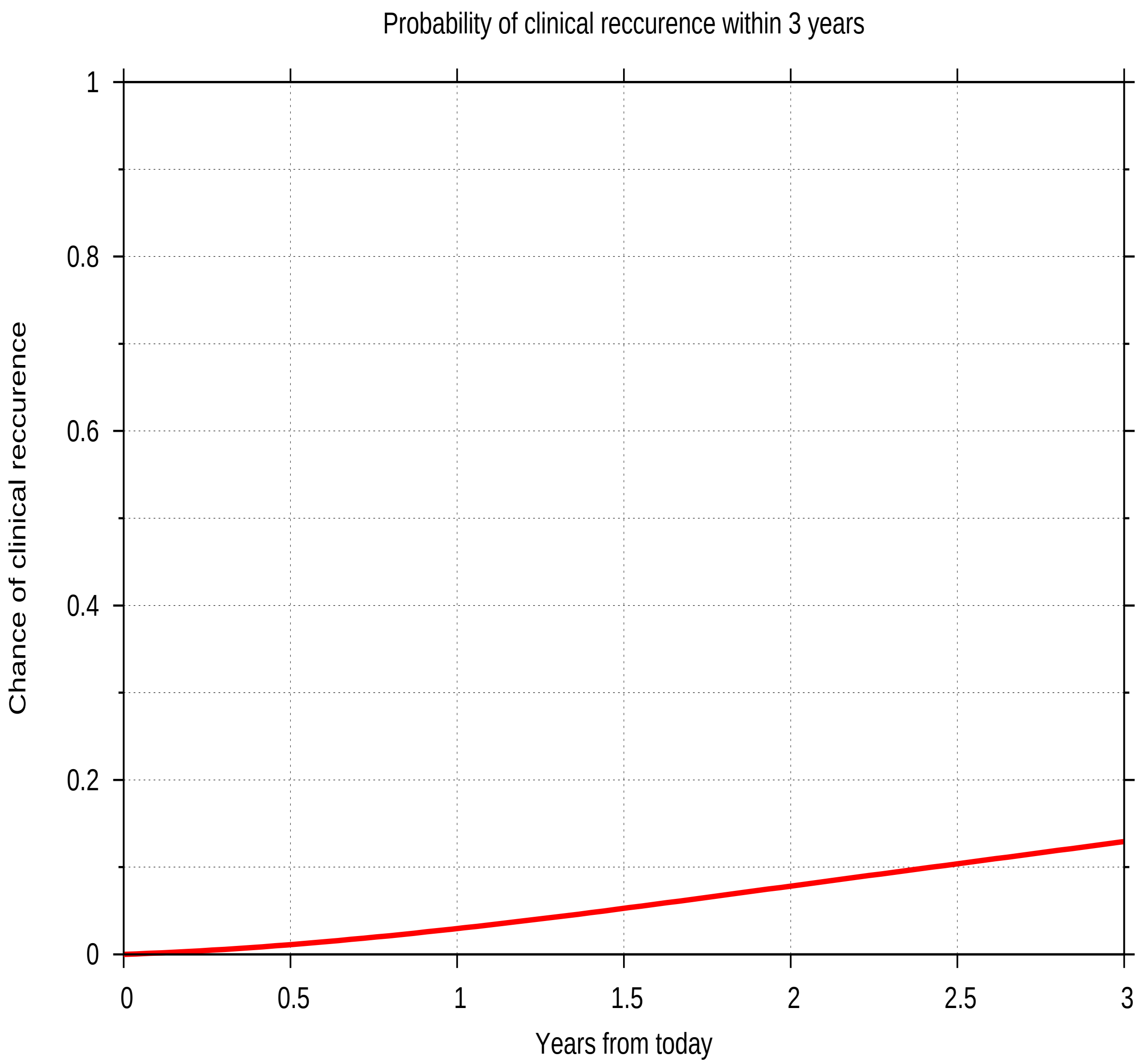
Residual Time Distribution for Censored Subject



Predictions of PSA for Censored Subjects



Residual Time Distribution for Censored Subjects



Predict Recurrence for New Censored Patient (m)

- Want $P[T < t_m + a \,|\, T > t_m, Y_m, X_m]$
- Obtain by averaging

$$P[T < t_m + a \,|\, T > t_m, X_m, \theta^k, R_m^k]$$

- Don't want to add new subject to dataset
- Have draws of θ from converged chain, needs draws of R_m
- For each θ^k run quick MCMC to get a draw of R_m^k.
- Draw R_m from $P(R_m|\theta^k, T_m > t_m, Y_m, X_m)$
- Use likelihood contribution from subject m
  $P(Y_m|\theta^k, X_m, R_m)P(T_m, \delta_m = 0|\theta^k, X_m, R_m)$

Website for the public to use

psacalc.sph.umich.edu.

Public = cancer patients and their doctors

Issues

- What to present?
- How to present it?
- A lot of clinical information is required as input
- Needs to run fast
- Aid in clinical decision making
- Could present predictions if salvage HT is started
- How to publicize it
- How much validation needs to be done and shown

Possible uses

- Individual patient monitoring
- Definition of an endpoint
  - Taylor definition:
    1st time Pr(Clinical Recurrence within 3 years) > 0.1
- Entry criteria for clinical study
  - eg Pr(Clinical Recurrence within 3 years) > 0.1

Statistical Issues

- Death from other causes is regarded as censored
- Gives "pure risk", assuming no deaths from other causes
- Very parametric approach
  - Use spline type of longitudinal model
- How to calculate P(Recurrence) if someone started HT today

ISSUES IN VALIDATION

- Training data, External validation data
- Prediction at time t about an event in $(t, t + a)$
- Prediction is a distribution function, data is censored

SIMPLE GRAPHICAL APPROACH FOR BINARY Y

(Hosmer-Lemeshow)

- $\hat{P}_i = \hat{P}(X_i) =$ predicted probability $P(Y = 1|X_i)$
- Create homogeneous groups of people with similar $\hat{P}_i$
- Estimate proportion for people in group g, $\hat{P}_g$
- Compare $\hat{P}_i$ with $\hat{P}_g$

ALTERNATIVE APPROACHES WHEN Y IS BINARY

- ROC curves, AUC
  - Popular and familiar
  - Change $\hat{P}_i$ to $\hat{P}_i/2$ would give same ROC curve.

SIMPLE GRAPHICAL APPROACH FOR SURVIVAL DATA (like Hosmer-Lemeshow)

- Create homogeneous groups with "similar" $\hat{S}(t + a|T_i > t, H_i)$
- Estimate empirical survival distribution for people in group g, $\hat{S}_g(a)$
- Compare $\hat{S}(t + a|T_i > t, H_i)$ with $\hat{S}_g(a)$
  - Calibration

1. Estimation on training data (3 cohort of patients):

   . 2386 patients

2. Predict on training data

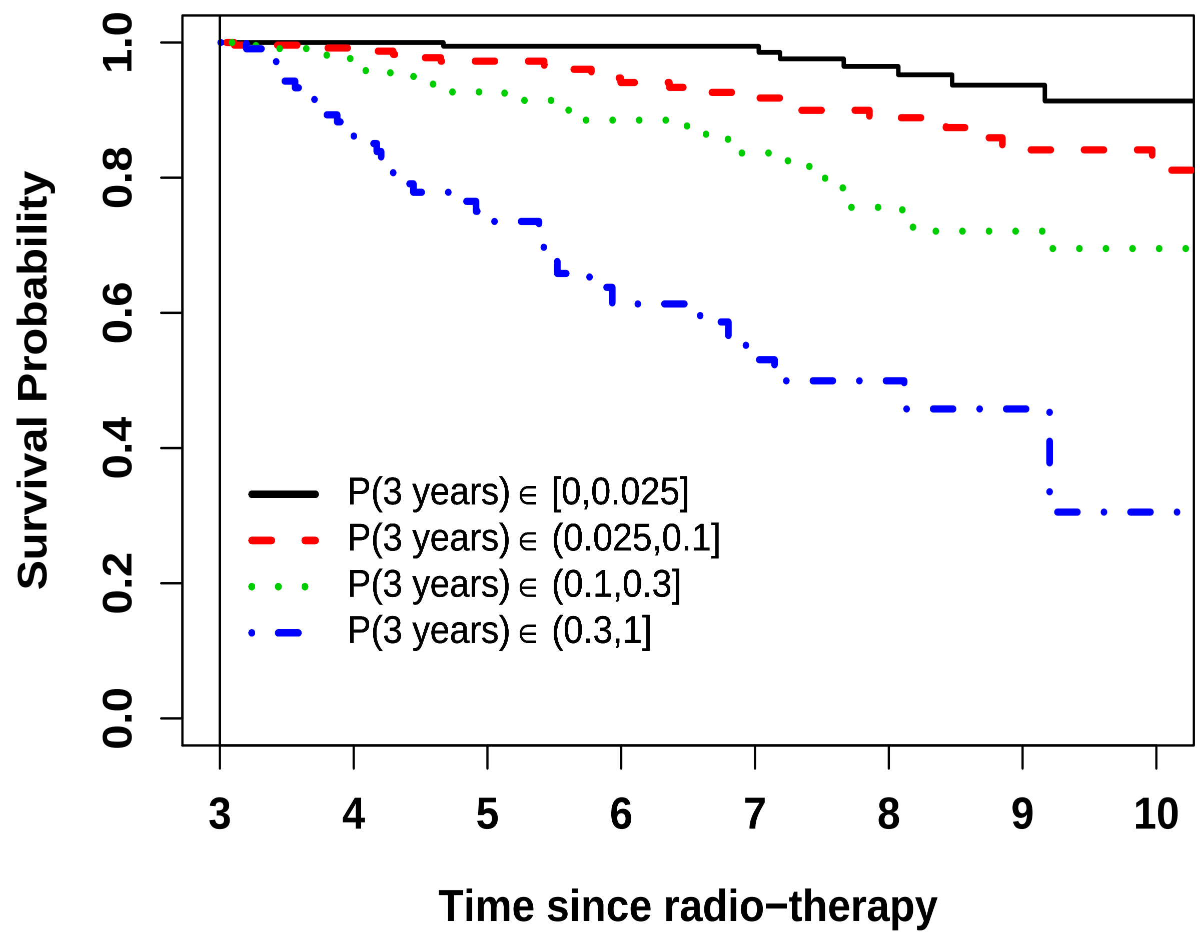
3. Prediction on 2 independent cohorts:

   . 846 patients from Vancouver cohort

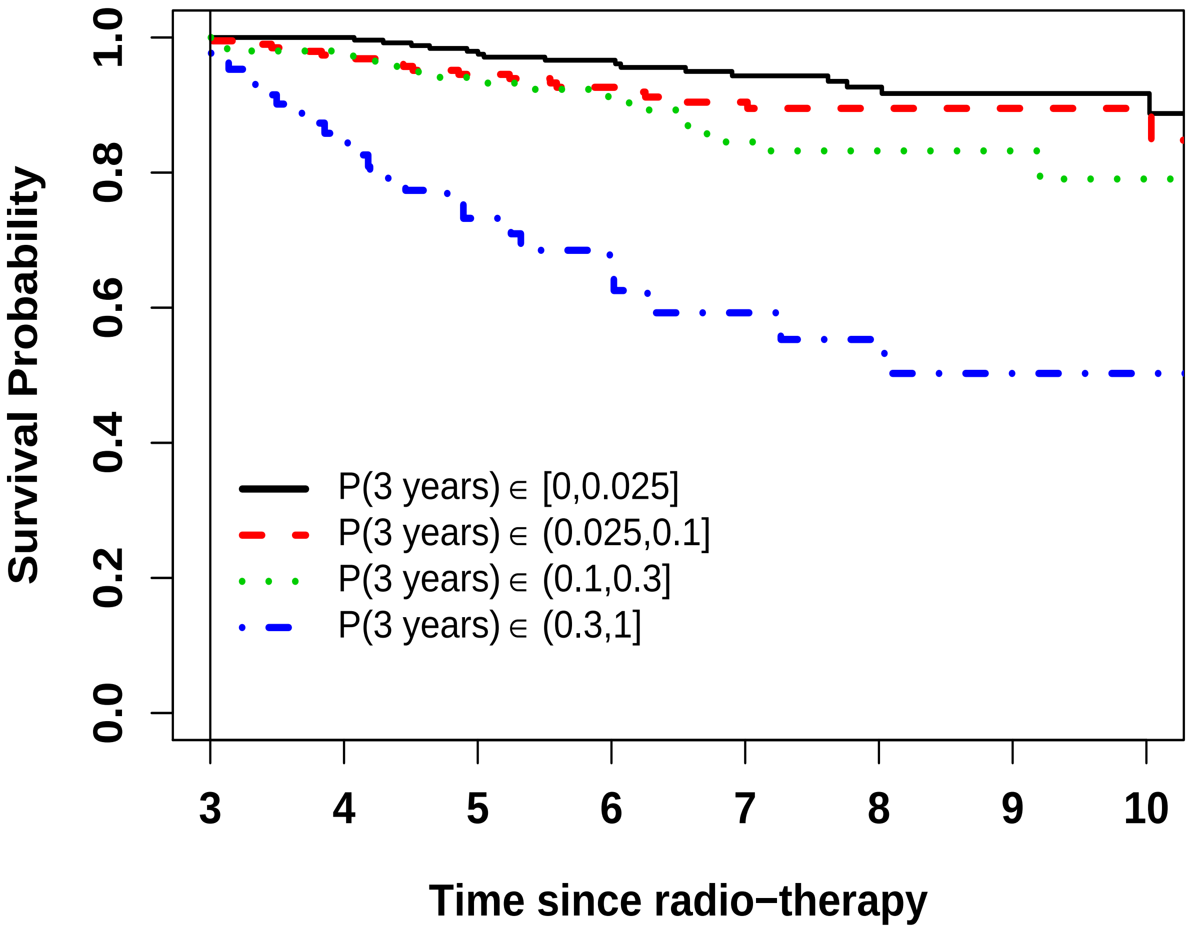
   . 395 patients from Melbourne

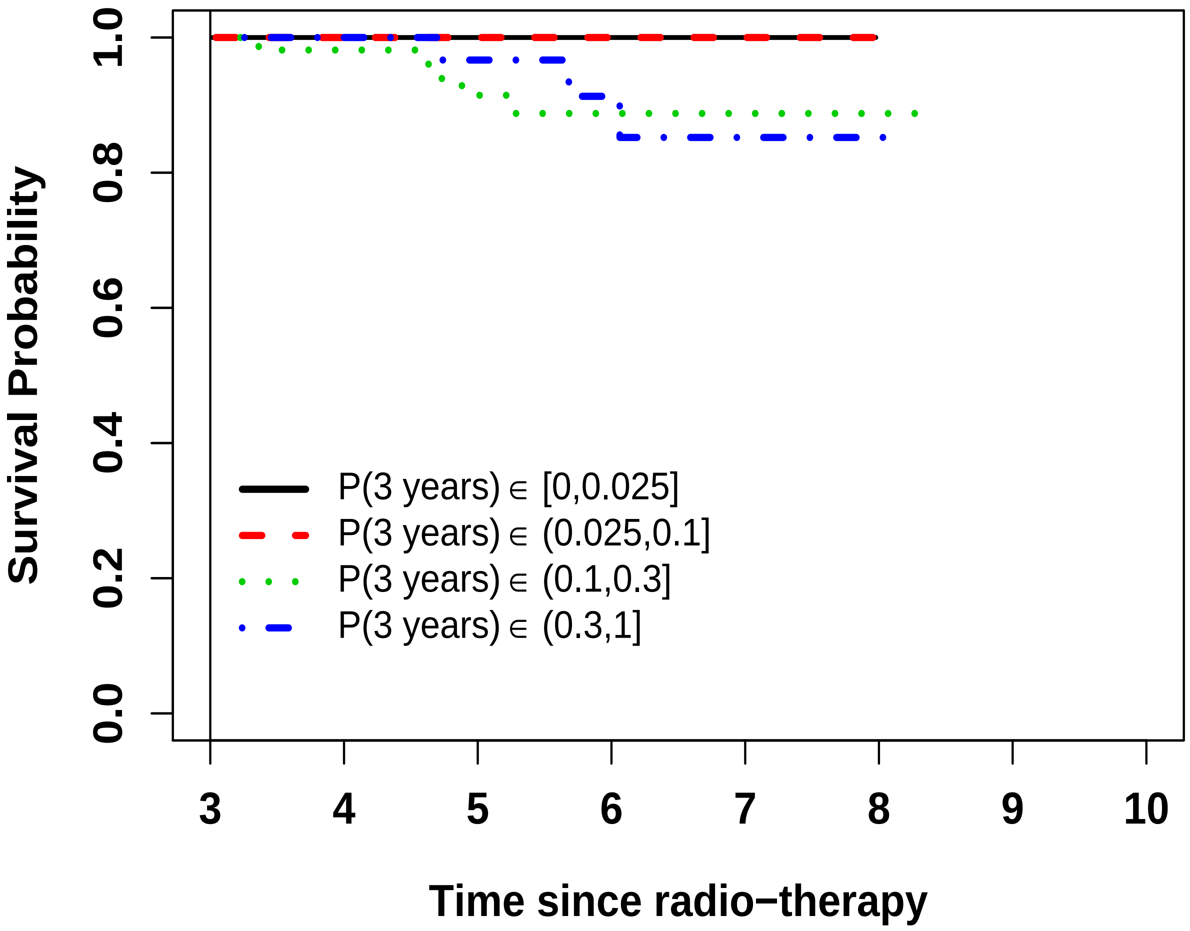
Graphical validation

- 4 groups based on $\hat{P}(T_i < 3 + a | T_i > 3, H_i(t))$
  - $\hat{P} \in$ (0.0,0.025)
  - $\hat{P} \in$ (0.025,0.10)
  - $\hat{P} \in$ (0.1,0.3)
  - $\hat{P} \in$ (0.3,1.0)
- Make predictions for everyone in testing datasets who are at risk at 3 years based on their PSA data prior to 3 years
- Place person into group
- Kaplan-Meier curve of what happened to them after 3 years



Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
P(3 years) ∈ [0,0.025]
P(3 years) ∈ (0.025,0.1]
P(3 years) ∈ (0.1,0.3]
P(3 years) ∈ (0.3,1]
3
4
5
6
7
8
9
10
Time since radio−therapy

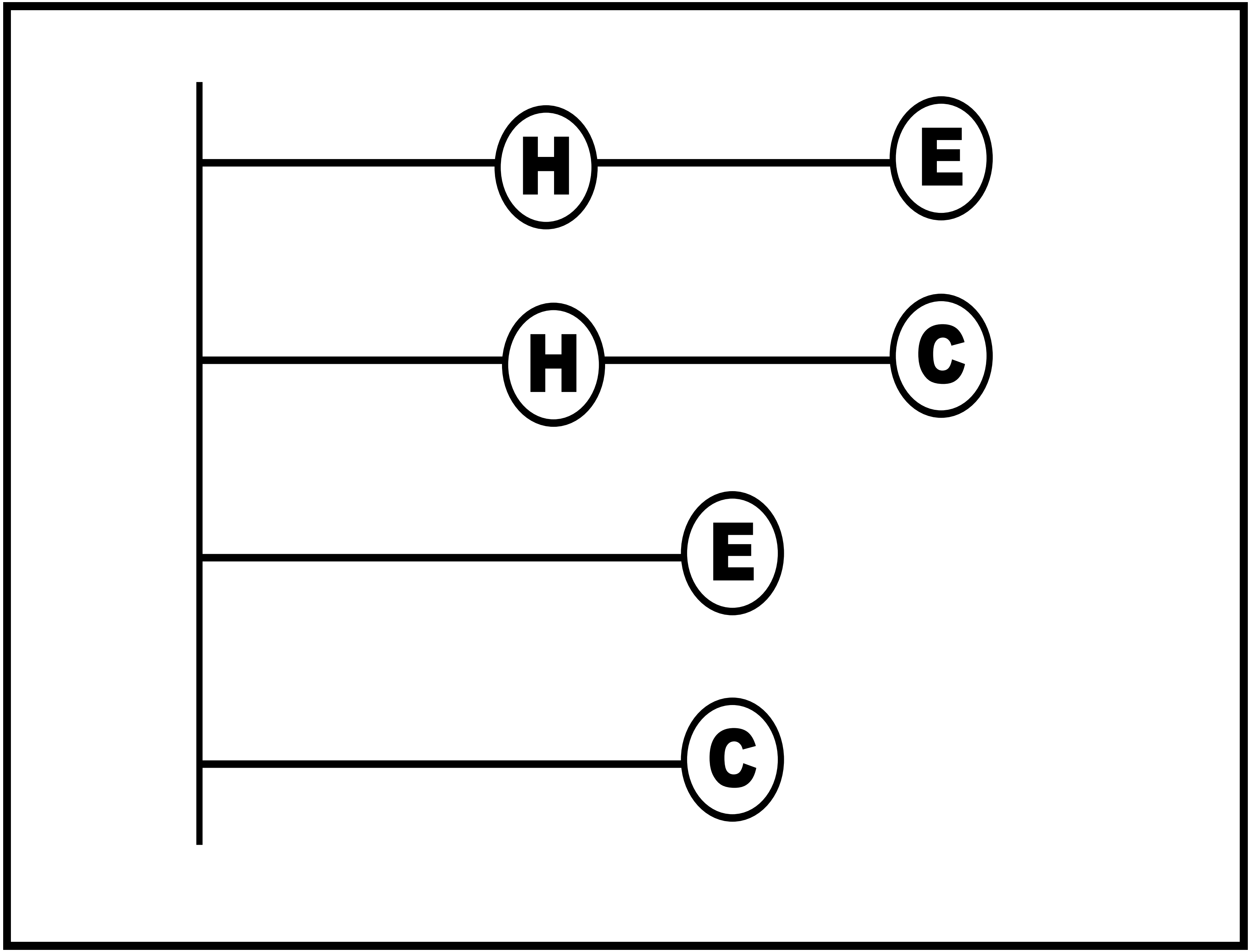

Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
3
4
5
6
7
8
9
10
Time since radio−therapy
P(3 years) ∈ [0,0.025]
P(3 years) ∈ (0.025,0.1]
P(3 years) ∈ (0.1,0.3]
P(3 years) ∈ (0.3,1]

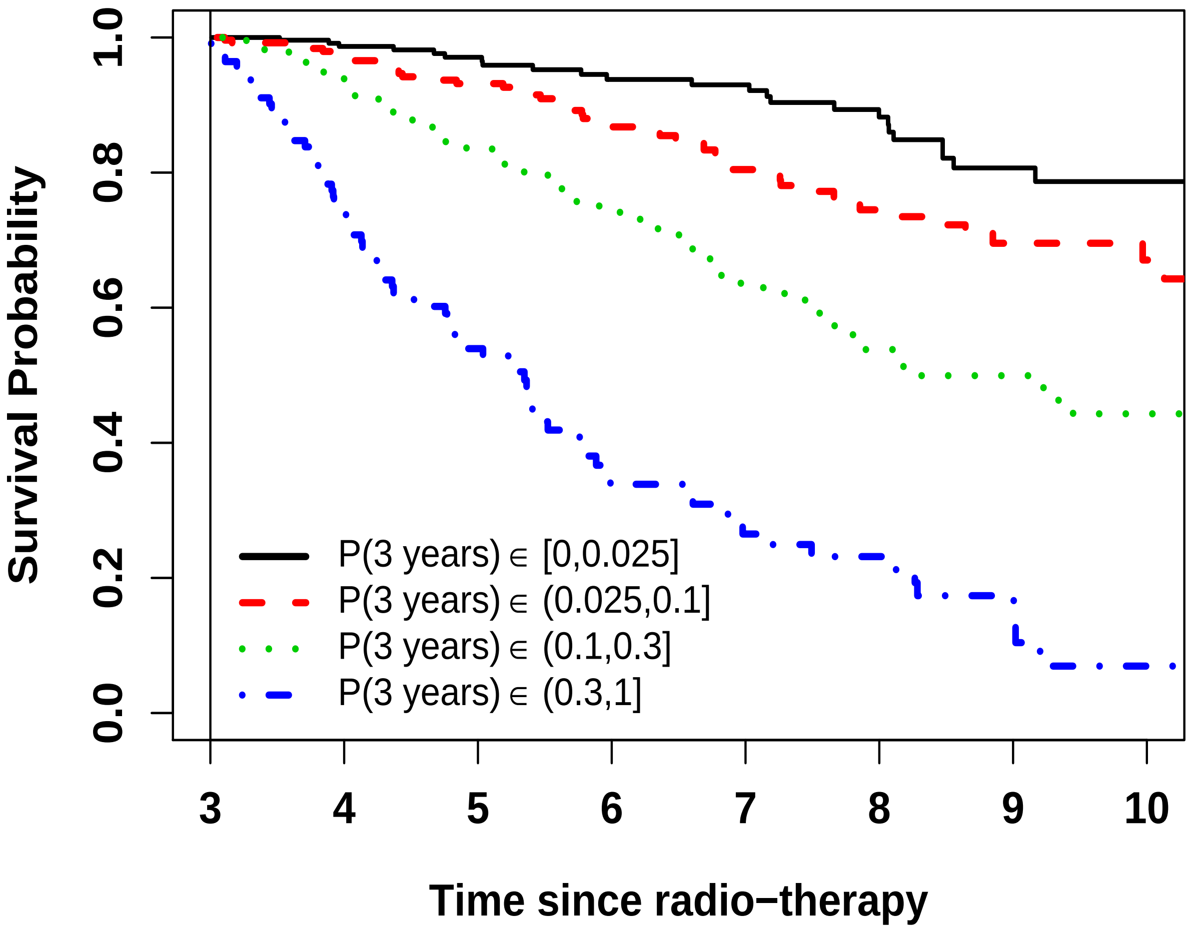


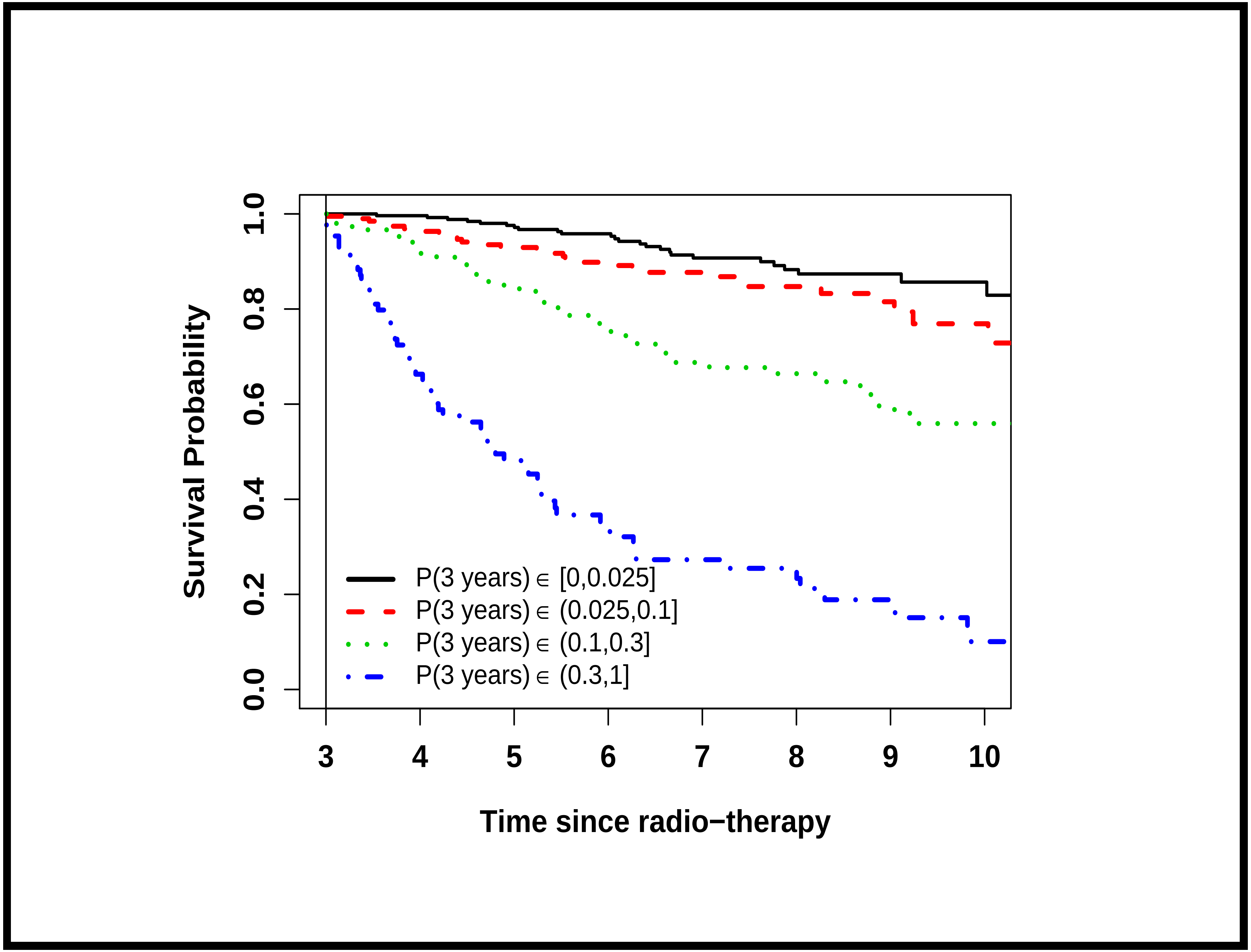
Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
3
4
5
6
7
8
9
10
Time since radio−therapy
P(3 years) ∈ [0,0.025]
P(3 years) ∈ (0.025,0.1]
P(3 years) ∈ (0.1,0.3]
P(3 years) ∈ (0.3,1]

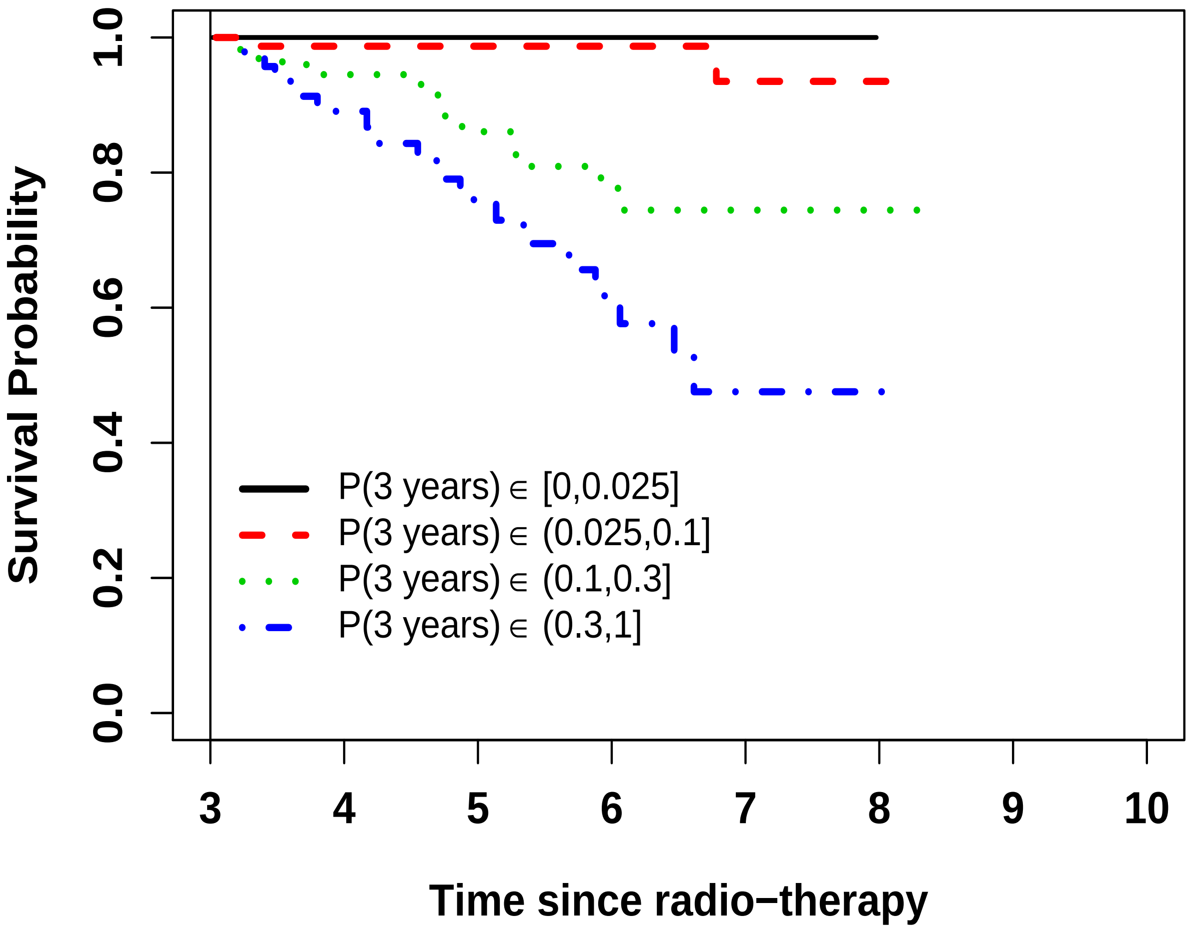
A complication with Kaplan-Meier estimate, dependent censoring

- people are given salvage hormone therapy (HT) prior to an event, because of rising PSA
- Options
  - censor at time of HT
  - call HT an event
  - ignore HT
  - something fancier


H
E
H
C
E
C


Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
3
4
5
6
7
8
9
10
Time since radio−therapy
P(3 years) ∈ [0,0.025]
P(3 years) ∈ (0.025,0.1]
P(3 years) ∈ (0.1,0.3]
P(3 years) ∈ (0.3,1]


Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
3
4
5
6
7
8
9
10
Time since radio−therapy
P(3 years) ∈ [0,0.025]
P(3 years) ∈ (0.025,0.1]
P(3 years) ∈ (0.1,0.3]
P(3 years) ∈ (0.3,1]



Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
3
4
5
6
7
8
9
10
Time since radio−therapy
P(3 years) ∈ [0,0.025]
P(3 years) ∈ (0.025,0.1]
P(3 years) ∈ (0.1,0.3]
P(3 years) ∈ (0.3,1]

Is this worth all the trouble? Simpler approaches:

- Cox model (PHM) with baseline variables: $P(T_i \leq t + a | T_i \geq t, X_i; \hat{\theta_0})$
- PHM with baseline variables & the last PSA (landmark analysis) $P(T_i \leq t + a | T_i \geq t, Y_i(t); \hat{\theta}_t)$

References

- Yu et al, 2004, Statistica Sinica
- Taylor et al, 2005, J Clin Oncol
- Yu et al, 2007, JASA
- Pauler and Finkelstein, 2002, Stat in Med
- Proust-Lima et al, 2008, Int J Rad Oncol Biol Phys
- Proust-Lima et al, 2009, Biostatistics

Collaborators and funding

Donna Ankerst

Cecile Proust-Lima

Ning Liu

Yongseok Park

Menggang Yu

Wen Ye

Howard Sandler

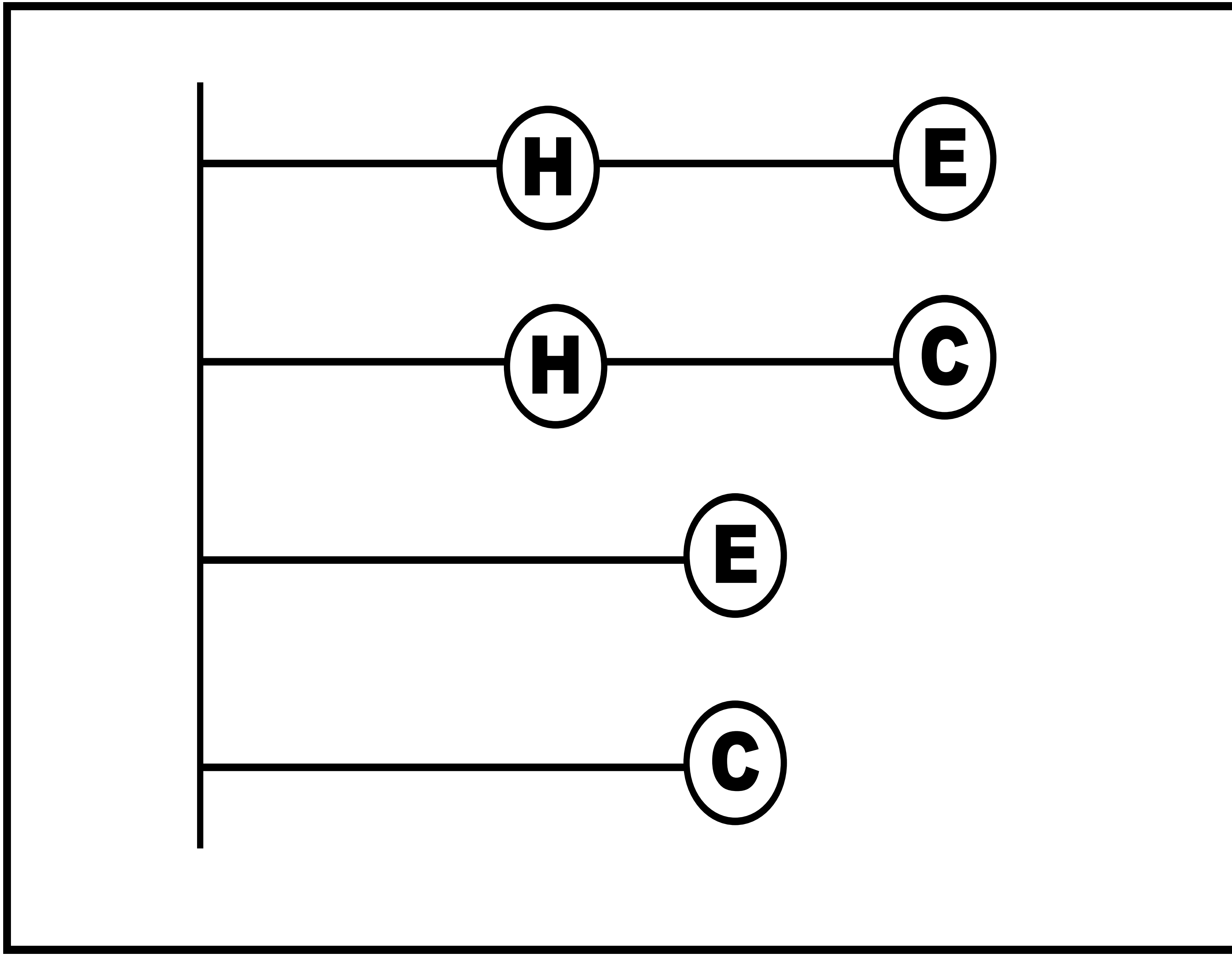
Scott Williams

Terry Gliedt

NCI grant, CA110518

Hormonal therapy and dependent censoring

- Salvage hormonal therapy (HT) is sometimes given when a patient exhibits a rising trend in PSA, before a clinical event.
- HT delays occurence of the clinical event
- HT is a nuisance factor present in the data.
- A different problem would be estimating the effect of HT.


H
E
H
C
E
C

$$\lambda_i(t \mid X_i, Z_i, sl_i, HT_i)$$
$$= \lambda_0(t)$$
$$\exp[\eta g(Z_i(t)) + \omega sl_i(t) + \gamma X_i + \phi HT_i(t)]$$

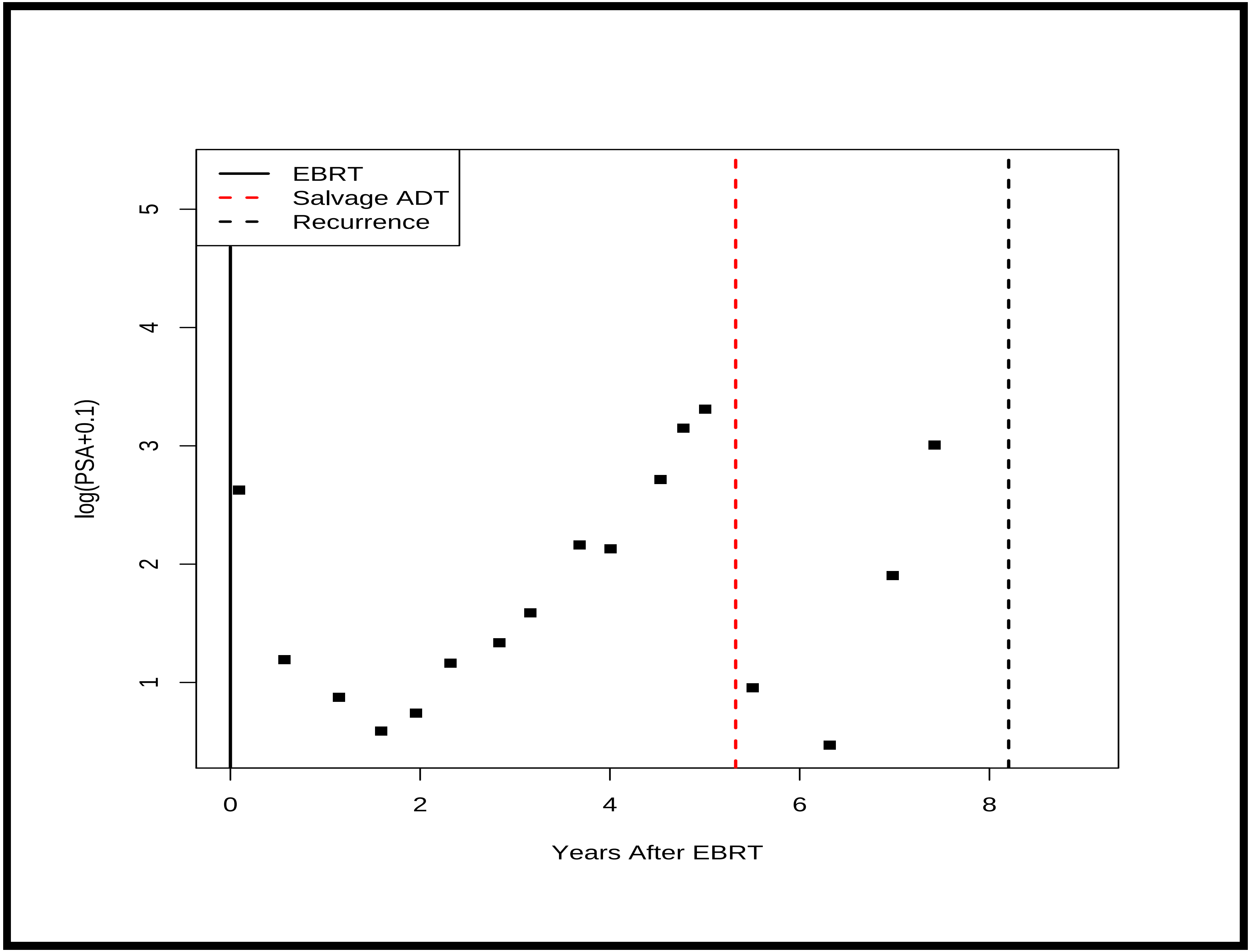
$$sl_i(t) = \text{ slope of } Z_i(t)$$

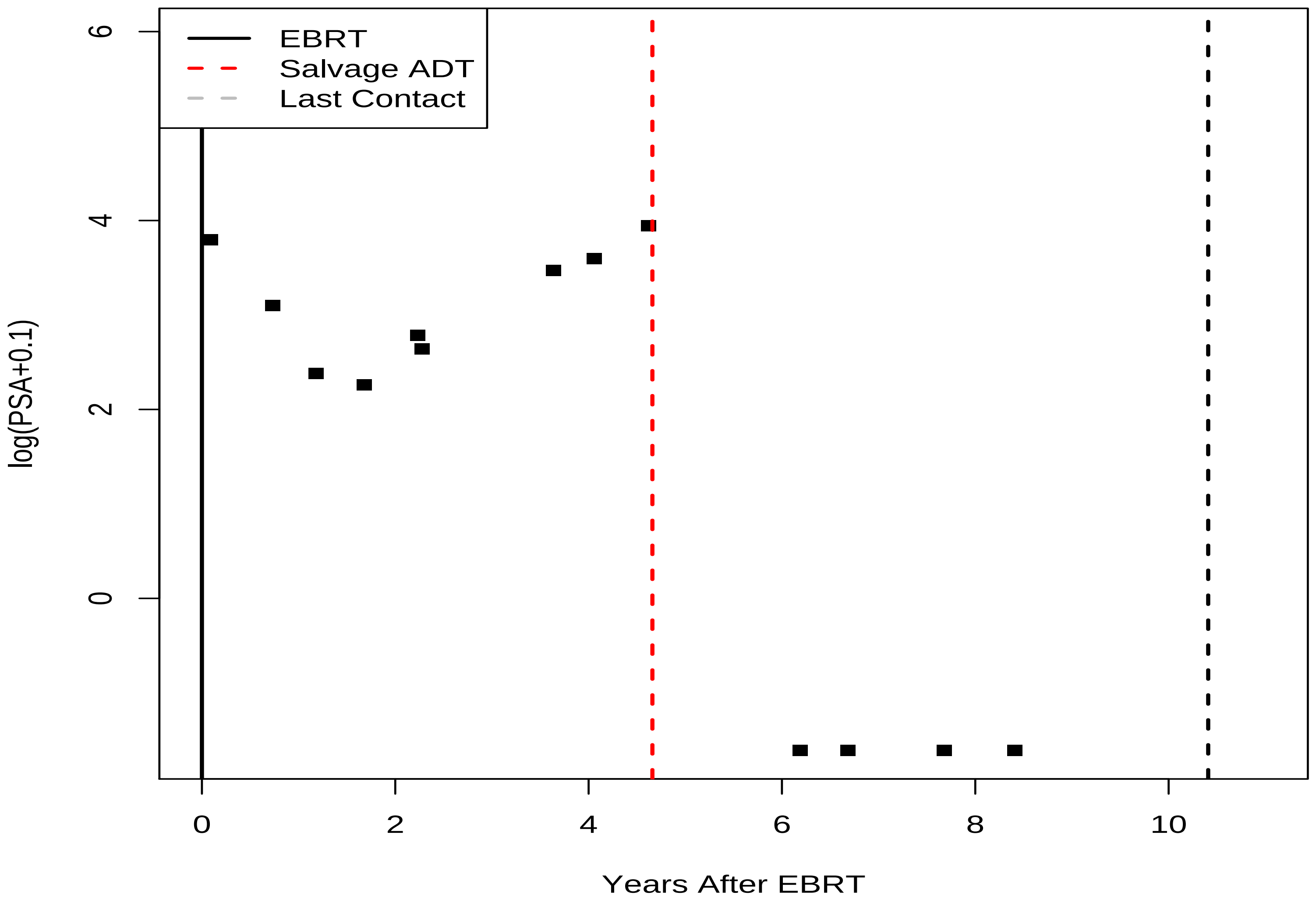
$$HT_i(t) = \begin{cases} 0 & \text{if} \quad t < S_i \\ 1 & \text{if} \quad t > S_i \end{cases}$$

Sensitivity analyses suggests estimates of η, ω, γ are stable

Using model to estimate effect of salvage hormonal therapy

- When PSA starts to rise some patients receive an intervention, which is thought to delay recurrence
- Level and slope of PSA are the important factors associated with the decision to initiate salvage therapy
- Of all patients about 10% receive salvage therapy prior to any recurrence
- Of all recurrences about 75% are before salvage therapy and 25% are after salvage therapy


EBRT
Salvage ADT
Recurrence
log(PSA+0.1)
1
2
3
4
5
0
2
4
6
8
Years After EBRT


EBRT
Salvage ADT
Last Contact
log(PSA+0.1)
0
2
4
6
0
2
4
6
8
10
Years After EBRT