A LONGITUDINAL AND SURVIVAL MODEL WITH HEALTH CARE USAGE FOR INSURED ELDERLY

Ramon Alemany
Montserrat Guillén
Xavier Piulachs Lozada

Riskcenter - IREA
Universitat de Barcelona
http://www.ub.edu/riskcenter

Funding and insurance in long-term care

Workshop

September 1st, 2014 - London

CONTENTS

Contents

1. INTRODUCTION AND GOALS

Scope of the work: Health insurance companies

The reasons leading to putting down this study are based on the following points:

- The gradual development of medical science leads to a larger number of years lived with disabilities (Robine and Michel, 2010).
- Policy holders are generally supposed to have a higher socio-economic level (Schoen et al, 2010).
- There is consequently a need of knowing how a greater elderly people cohort will evolve, as they are the principal beneficiaries of life expectancy improvements.

Proposed longitudinal variable of interest for a policy holder:

$\boldsymbol{y}$ = ANNUAL CUMULATIVE NUMBER OF REQUESTS AT SPECIFIC TIME POINTS WITHIN STUDY PERIOD

1. INTRODUCTION AND GOALS

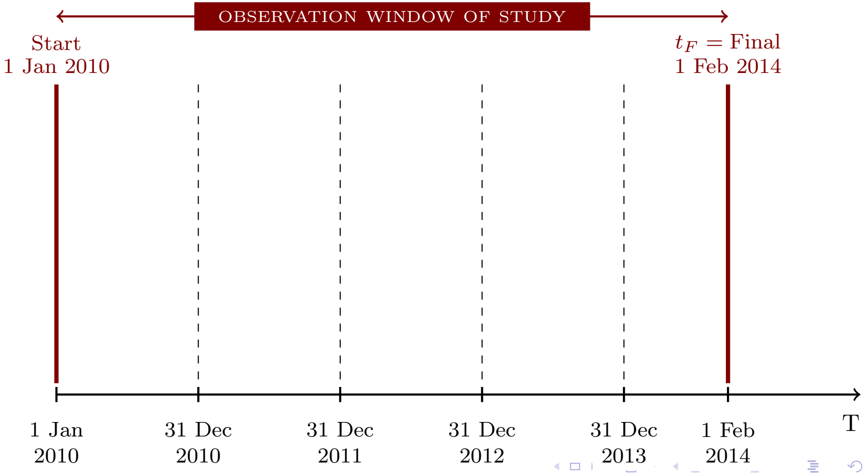
Motivation (I)

MOTIVATION: Longitudinal study with **2 outcomes**, based on **a)** Repeated measurements of response variable and **b)** Time until a particular event.

1. INTRODUCTION AND GOALS

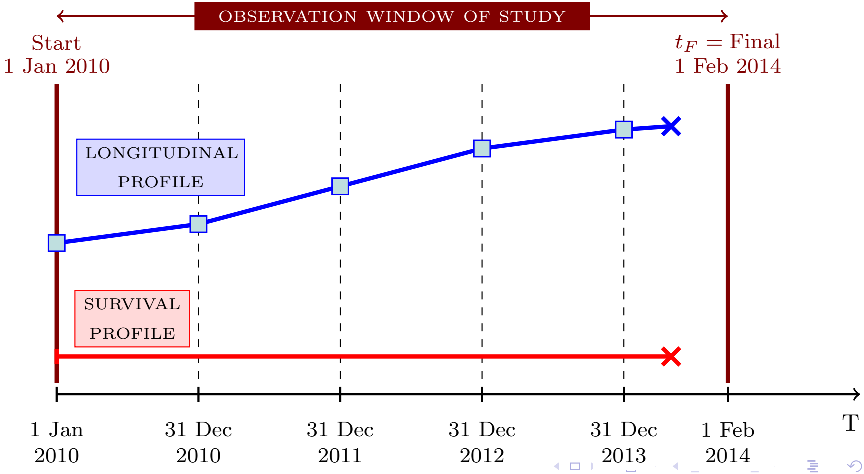
Motivation (I)

MOTIVATION: Longitudinal study with **2 outcomes**, based on **a)** Repeated measurements of response variable and **b)** Time until a particular event.



1. INTRODUCTION AND GOALS
Motivation (I)

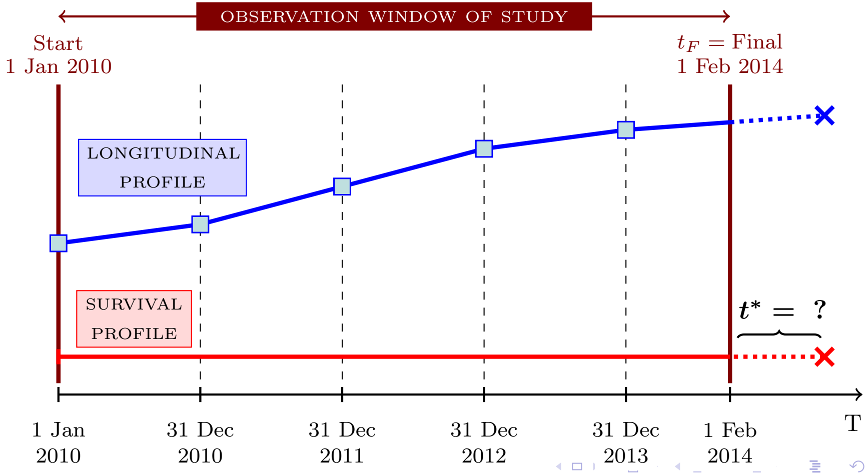
MOTIVATION: Longitudinal study with **2 outcomes**, based on **a)** Repeated measurements of response variable and **b)** Time until a particular event.



1. INTRODUCTION AND GOALS

Motivation (I)

MOTIVATION: Longitudinal study with **2 outcomes**, based on **a)** Repeated measurements of response variable and **b)** Time until a particular event.



1. INTRODUCTION AND GOALS
Motivation (III)

Therefore, it's a question of coupling longitudinal and survival information in one single model, which allows:

- To establish the degree of association between the value of the longitudinal variable with the event outcome.
- To estimate subject specific survival probabilities based on longitudinal outcome: $\mathbf{Pr}(T_i \geq t^* \mid T_i > t_F)$
- **To update personalized survival estimations as additional longitudinal information is collected.**

1. INTRODUCTION AND GOALS

Motivation (IV)

However, the coupling of longitudinal and survival information is not without difficulties ...

- If missing improperly handled, biased results (Prentice, 1982).
- The longitudinal response is often an endogenous variable.
- In some cases, the event of interest is only known to occur after a certain $t \Rightarrow$ right-censored data.

So, how to achieve a simultaneous modeling of two processes?

1. INTRODUCTION AND GOALS

Motivation (IV)

However, the coupling of longitudinal and survival information is not without difficulties ...

- If missing improperly handled, biased results (Prentice, 1982).
- The longitudinal response is often an endogenous variable.
- In some cases, the event of interest is only known to occur after a certain $t \Rightarrow$ right-censored data.

So, how to achieve a simultaneous modeling of two processes?

Joint Modeling for Longitudinal and Survival Data

Contents

2. DATABASE: SPANISH HEALTH INSURANCE COMPANY

Main characteristics of the study

- Spanish insurance company where the study period is fixed from **1 Jan 2010 and 1 Feb 2014**.
- Monitoring of **elderly 65 annual cumulative requests**.
- Subjects requests' during the four years before their study entry are treated as a baseline covariate.
- Distribution by sex:

  11912 men (39%) with 409 events (3.4%).
  18668 women (61%) with 933 events (5.0%).

- **GOAL:** To evaluate, in a personalized way, the existing degree of association between the frequency of use of medical services with the risk of mortality.

2. DATABASE: SPANISH HEALTH INSURANCE COMPANY

Database(I): Main variables

VARIABLE	DESCRIPTION
ID	Subject identifier: $i = 1, 2, \ldots, 30580$
SEX	Gender of the subject: 0 = Male, 1 = Female
OBSTIME	Age (years) over 65 at each of subject's time points
CUM0	Baseline cumulative number of requests (4 years before entry)
CUM	Cumulative number of requests at each time point
TIME	Final observation time (years), which may correspond to an event or to a right-censored data.
CENS	Censoring indicator: 0 = Right-censored, 1 = Event

Contents

3. JOINT MODELING TECHNIQUES

Longitudinal data analysis (I): Assumptions

LONGITUDINAL APPROACH

- To obtain the effect of covariates on an outcome when there is an association between outcomes.
- Association between outcomes.
- Let denote y_{ij} the response variable on the $i-$th subject, $i = 1, \ldots, n$, observed at time point t_{ij}, $j = 1, \ldots, n_i$.
- The outcome is linearly related to a set of ***p*** **explanatory covariates** and ***q*** **random effects**
- Let assume that the longitudinal outcomes for the $i-$th subject, $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in_i})^{\mathrm{T}}$, are normally distributed.

3. JOINT MODELING TECHNIQUES

Longitudinal data analysis (II): General equation

Linear mixed model equation:

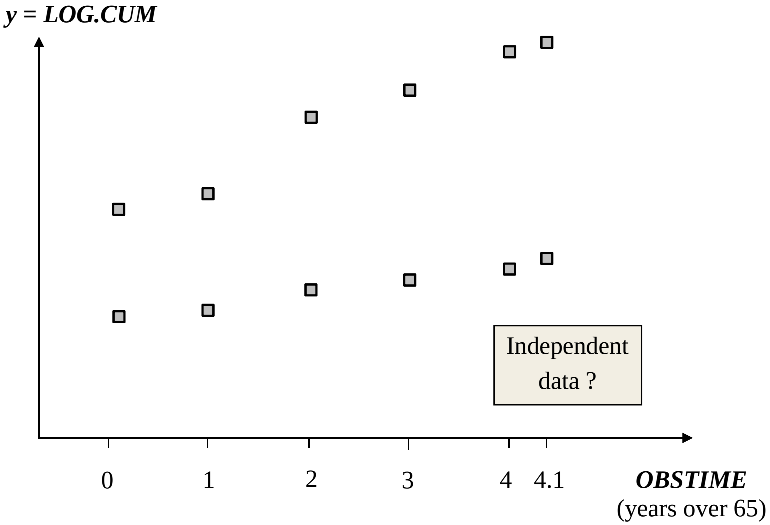
$$\begin{cases} \mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i \\ \mathbf{b}_i \sim \mathcal{N}_q(\mathbf{0}, \mathbf{D}) \\ \boldsymbol{\varepsilon}_i \sim \mathcal{N}_{n_i}(\mathbf{0}, \sigma^2\mathbf{I}_{n_i}) \end{cases}$$

- $\mathbf{X}_i$ and $\mathbf{Z}_i$ design matrices for fixed and random effects, respectively.
- $\boldsymbol{\beta}$ and $\mathbf{b}_i$ vectors for the fixed effects and random effects, respectively.
- $\{\mathbf{b}_1,\ \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ independent of $\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_n\}$.

(Laird and Ware, 1982)
(Verbeke and Molenberghs, 2000)

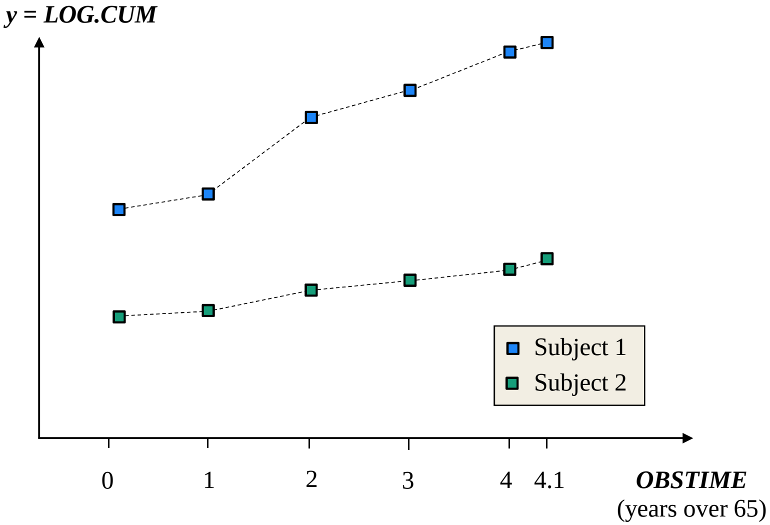
3. JOINT MODELING TECHNIQUES

Longitudinal data analysis (III): Graphic illustration



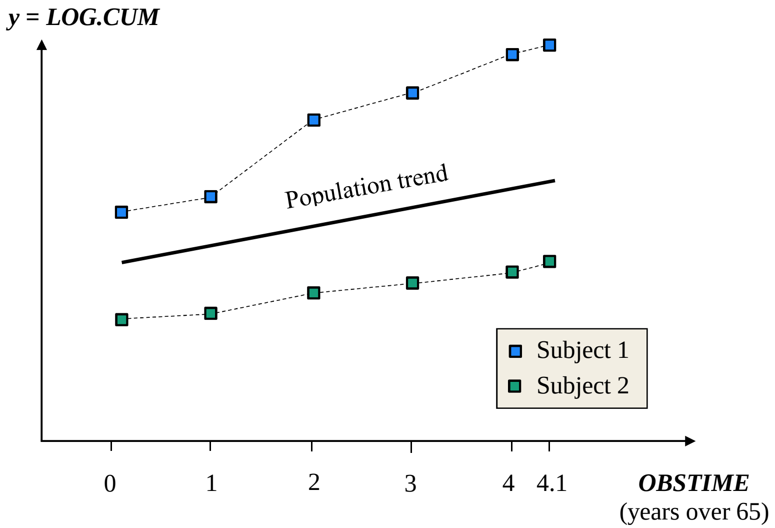
3. JOINT MODELING TECHNIQUES

Longitudinal data analysis (III): Graphic illustration



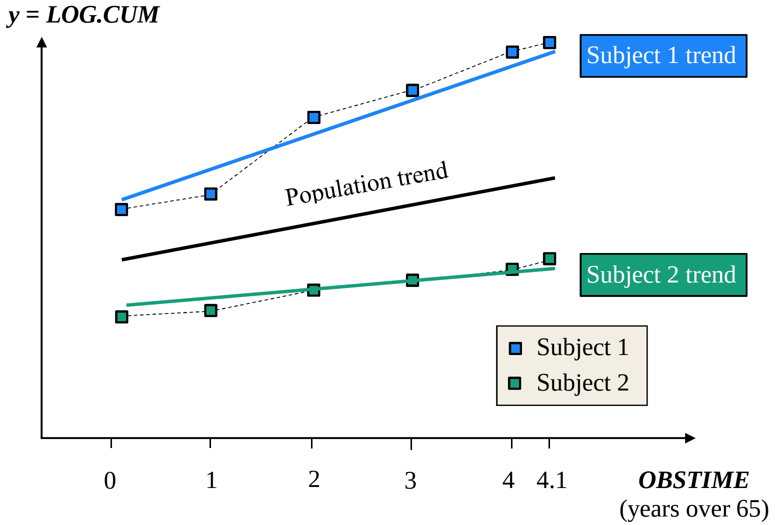
3. JOINT MODELING TECHNIQUES

Longitudinal data analysis (III): Graphic illustration



3. JOINT MODELING TECHNIQUES

Longitudinal data analysis (III): Graphic illustration



3. JOINT MODELING TECHNIQUES

Survival analysis (I): Notation and definitions

SURVIVAL APPROACH

Let consider for the i-th subject:

- T^* is a non-negative continuous random variable denoting the true survival time.
- C is the potential right-censoring time.
- For the $i-$th subject, we define the observed survival time: $Y_i = \min\{T_i, C_i\}$ and $\delta_i = I(T_i \leq C_i)$.

3. JOINT MODELING TECHNIQUES

Survival analysis (II): PH Cox Model

Semi-parametric estimation: PH Cox Model (Cox, 1972)

$$h_i(t|\mathbf{w}_i) = h_0(t)\exp(\boldsymbol{\gamma}^{\mathrm{T}}\mathbf{w}_i),$$

$\mathbf{w}_i = (w_{i1}, w_{i2}, \ldots, w_{ip})^{\mathrm{T}}$
$\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_p)^{\mathrm{T}}$

The Cox model can be extended to handle exogenous time-dependent covariates (Andersen and Gill, 1982).

But often measurements taken on the subjects are related to inherent biological changes: **endogenous covariates**.

3. JOINT MODELING TECHNIQUES

Survival analysis (II): PH Cox Model

Semi-parametric estimation: PH Cox Model (Cox, 1972)

$$h_i(t|\mathbf{w}_i) = h_0(t)\exp(\boldsymbol{\gamma}^{\mathrm{T}}\mathbf{w}_i),$$

$\mathbf{w}_i = (w_{i1}, w_{i2}, \ldots, w_{ip})^{\mathrm{T}}$
$\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_p)^{\mathrm{T}}$

The Cox model can be extended to handle exogenous time-dependent covariates (Andersen and Gill, 1982).

But often measurements taken on the subjects are related to inherent biological changes: **endogenous covariates**.

It is therefore necessary to implement:

JOINT MODELING TECHNIQUES

(Tsiatis *et al.*, 1995; Rizopoulos, 2012)

3. JOINT MODELING TECHNIQUES

Joint Modeling framework: Fitted joint model

In our particular database:

$$\begin{cases} \text{For the } i-\text{th man, } i = 1, \ldots, 11912, \text{ at time } t \\ \text{For the } i-\text{th woman, } i = 1, \ldots, 18668, \text{ at time } t \end{cases}$$

Longitudinal submodel

$$\begin{cases} LOG.CUM_i(t) = \boldsymbol{\beta_0} + \boldsymbol{b_{i0}} + \boldsymbol{\beta_1 t} + \varepsilon_i(t) \\ \boldsymbol{\beta} = (\beta_0, \beta_1)^{\mathrm{T}} \\ b_{i0} \sim \mathcal{N}(0, \sigma^2_{b_0}) \\ \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2) \end{cases}$$

Survival submodel

$$h_i(t|\mathbf{w}_i) = h_0(t) R_i(t) \exp\{\gamma LOG.CUMFENT_i\}.$$

3. JOINT MODELING TECHNIQUES

Joint Modeling framework: Fitted joint model

In our particular database:

$$\begin{cases} \text{For the } i\text{−th man, } i = 1, \ldots, 11912, \text{ at time } t \\ \text{For the } i\text{−th woman, } i = 1, \ldots, 18668, \text{ at time } t \end{cases}$$

Longitudinal submodel

$$\begin{cases} LOG.CUM_i(t) = \boldsymbol{\beta_0 + b_{i0} + \beta_1 t} + \varepsilon_i(t) \\ \boldsymbol{\beta} = (\beta_0, \beta_1)^{\mathrm{T}} \\ b_{i0} \sim \mathcal{N}(0, \sigma^2_{b_0}) \\ \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2) \end{cases}$$

Survival submodel

$$h_i(t|\mathbf{w}_i) = h_0(t) R_i(t) \exp\{\gamma LOG.CUMFENT_i\}.$$

JOINT MODEL

$$h_i(t|\mathcal{M}_i(t), \mathbf{w}_i) = h_0(t) R_i(t) \exp\{\gamma LOG.CUMFENT_i + \alpha(\boldsymbol{\beta_0 + b_{i0} + \beta_1 t})\}$$

Contents

4. RESULTS FOR THE DATABASE

Joint modeling results (I): Estimation by sex

$$h_i(t|\mathcal{M}_i(t), \mathbf{w}_i) = h_0(t)R_i(t)\exp\{\gamma LOG.CUMFENT_i + \alpha(\boldsymbol{\beta_0 + b_{i0} + \beta_1 t})\}$$

	JM for men		JM for women	
Parameters	Estimate	95% CI	Estimate	95% CI
Longitudinal				
β_0	2.140	(2.100, 2.180)	2.158	(2.123, 2.192)
β_1	0.170	(0.167, 0.172)	0.157	(0.155, 0.159)
σ	0.332	(0.329, 0.334)	0.314	(0.312, 0.316)
σ_{b_0}	1.648	(1.597, 1.698)	1.791	(1.748, 1.834)
Survival				
γ	-1.174	(-1.306,-1.042)	-0.964	(-1.042, -0.887)
Association				
α	1.437	(1.275, 1.598)	1.273	(1.179, 1.367)

4. RESULTS FOR THE DATABASE

Joint modeling results (I): Estimation by sex

$$h_i(t|\mathcal{M}_i(t), \mathbf{w}_i) = h_0(t)R_i(t)\exp\{\gamma LOG.CUMFENT_i + \alpha(\boldsymbol{\beta_0 + b_{i0} + \beta_1 t})\}$$

	JM for men		JM for women	
Parameters	Estimate	95% CI	Estimate	95% CI
Longitudinal				
β_0	2.140	(2.100, 2.180)	2.158	(2.123, 2.192)
β_1	0.170	(0.167, 0.172)	0.157	(0.155, 0.159)
σ	0.332	(0.329, 0.334)	0.314	(0.312, 0.316)
σ_{b_0}	1.648	(1.597, 1.698)	1.791	(1.748, 1.834)
Survival				
γ	-1.174	(-1.306,-1.042)	-0.964	(-1.042, -0.887)
Association				
α	1.437	(1.275, 1.598)	1.273	(1.179, 1.367)

4. RESULTS FOR THE DATABASE

Joint modeling results (I): Estimation by sex

$$h_i(t|\mathcal{M}_i(t), \mathbf{w}_i) = h_0(t)R_i(t)\exp\{\gamma LOG.CUMFENT_i + \alpha(\boldsymbol{\beta_0 + b_{i0} + \beta_1 t})\}$$

	JM for men		JM for women	
Parameters	Estimate	95% CI	Estimate	95% CI
Longitudinal				
β_0	2.140	(2.100, 2.180)	2.158	(2.123, 2.192)
β_1	0.170	(0.167, 0.172)	0.157	(0.155, 0.159)
σ	0.332	(0.329, 0.334)	0.314	(0.312, 0.316)
σ_{b_0}	1.648	(1.597, 1.698)	1.791	(1.748, 1.834)
Survival				
γ	-1.174	(-1.306,-1.042)	-0.964	(-1.042, -0.887)
Association				
α	1.437	(1.275, 1.598)	1.273	(1.179, 1.367)

4. RESULTS FOR THE DATABASE

Joint modeling results (I): Estimation by sex

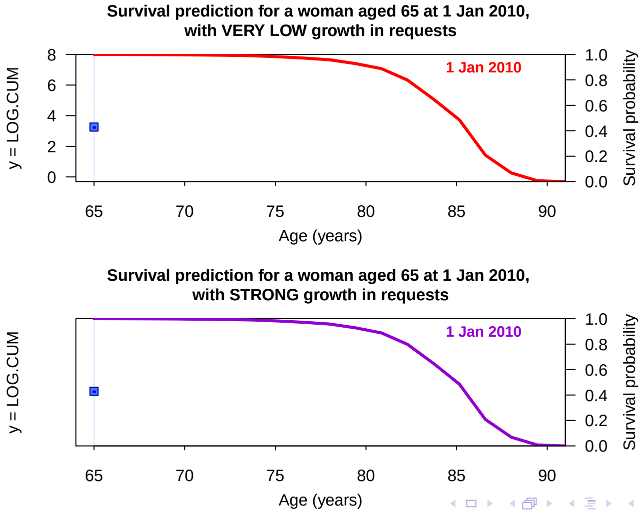
$$h_i(t|\mathcal{M}_i(t), \mathbf{w}_i) = h_0(t)R_i(t)\exp\{\gamma LOG.CUMFENT_i + \alpha(\boldsymbol{\beta_0 + b_{i0} + \beta_1 t})\}$$

	JM for men		JM for women	
Parameters	Estimate	95% CI	Estimate	95% CI
Longitudinal				
β_0	2.140	(2.100, 2.180)	2.158	(2.123, 2.192)
β_1	0.170	(0.167, 0.172)	0.157	(0.155, 0.159)
σ	0.332	(0.329, 0.334)	0.314	(0.312, 0.316)
σ_{b_0}	1.648	(1.597, 1.698)	1.791	(1.748, 1.834)
Survival				
γ	-1.174	(-1.306,-1.042)	-0.964	(-1.042, -0.887)
Association				
α	1.437	(1.275, 1.598)	1.273	(1.179, 1.367)

4. RESULTS FOR THE DATABASE

Joint modeling results (II): Dynamic predictions

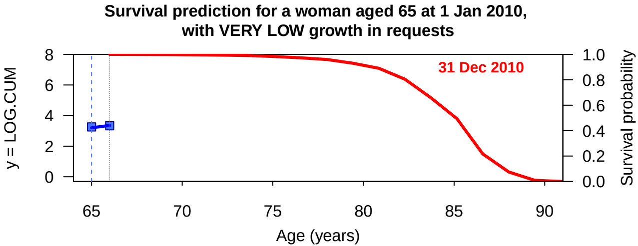
Comparison: Two women aged 65 at 1st Jan 2010 with 25 baseline requests

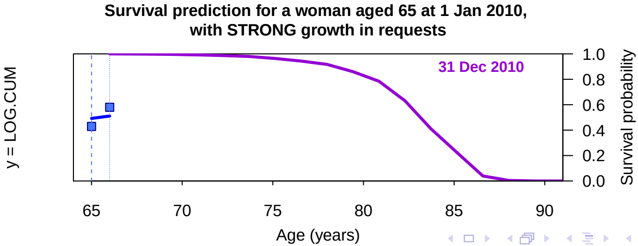


4. RESULTS FOR THE DATABASE

Joint modeling results (II): Dynamic predictions

Comparison: Two women aged 65 at 1st Jan 2010 with 25 baseline requests

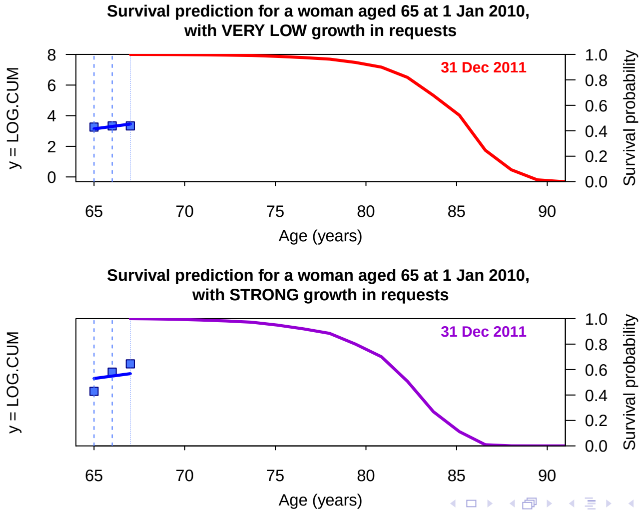




4. RESULTS FOR THE DATABASE

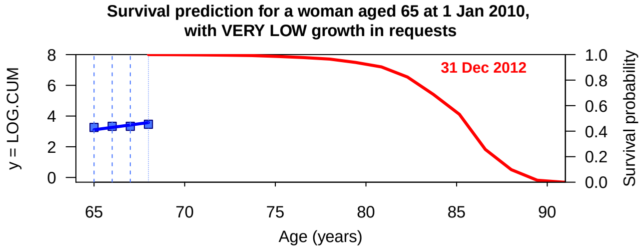
Joint modeling results (II): Dynamic predictions

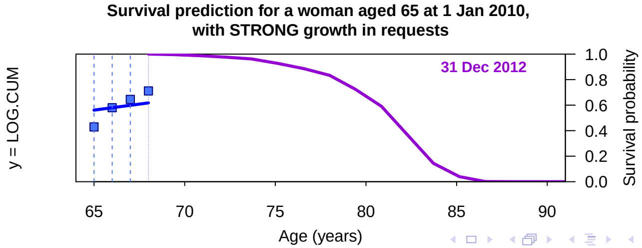
Comparison: Two women aged 65 at 1st Jan 2010 with 25 baseline requests



4. RESULTS FOR THE DATABASE
Joint modeling results (II): Dynamic predictions

Comparison: Two women aged 65 at 1st Jan 2010 with 25 baseline requests

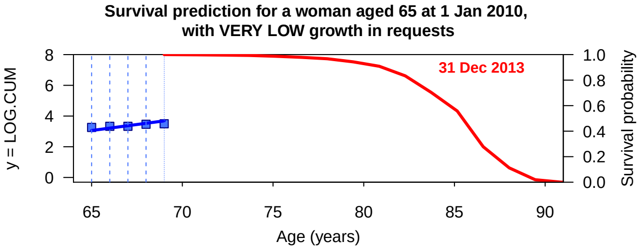


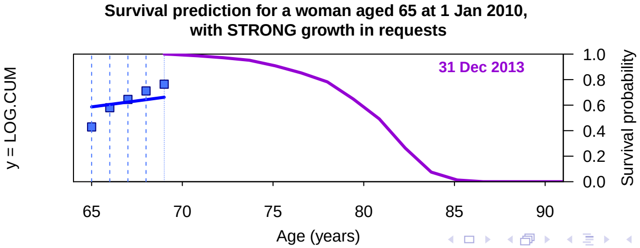


4. RESULTS FOR THE DATABASE

Joint modeling results (II): Dynamic predictions

Comparison: Two women aged 65 at 1st Jan 2010 with 25 baseline requests

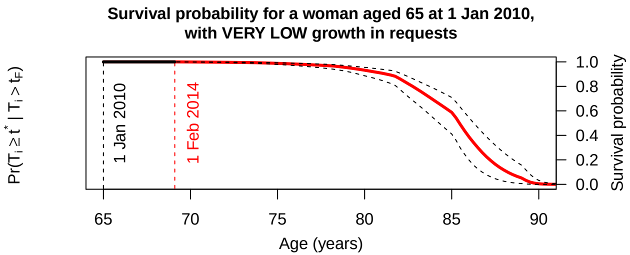


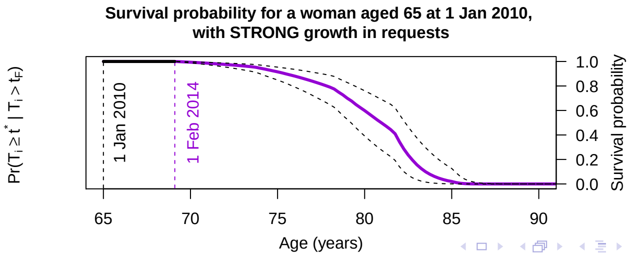


4. RESULTS FOR THE DATABASE

Joint modeling results (II): Dynamic predictions

Comparison: Two women aged 65 at 1st Jan 2010 with 25 baseline requests





Contents

5. DISCUSSION AND FUTURE RESEARCH

Discussion and Conclusions

CONCLUSIONS

- The fitted joint model indicates that the observed number of cumulated requests is highly associated with the risk of death (event of interest).
- The baseline cumulated acts has a protective effect.
- The joint modeling techniques allow to obtain an unbiased and personalized estimate of the the impact of $y = LOG.CUM$ trajectories on time to mortality event. As longitudinal information was collected for all subjects, the joint modeling methodology has allowed to continuously update the predictions of their survival probabilities.

5. DISCUSSION AND FUTURE RESEARCH

Future Research

TO GO FURTHER ...

- Different types of requests need to be distinguished
- To relate the subject specific profile to an estimated cost.
- To consider extensions of the standard joint modeling approach (*e.g.* the subject-specific slope $m_i'(t)$).

REFERENCES

MAIN REFERENCES:

- *Modeling of Policyholder Behavior for Life Insurance and Annuity Products*, (Campbell *et al.*, 2014)
- *Regression models and life-tables*, (Cox *et al.*, 2014)
- *Joint Models for Longitudinal and Time-to-Event Data*, (Rizopoulos, 2012)
- *Looking Forward to a General Theory on Population Aging*, (Robine and Michel, 2004)
- *How Health Insurance Design Affects Access To Care And Costs, By Income, In Eleven Countries*, (Schoen *et al.*, 2010)
- *Joint modeling of longitudinal and time-to-event data: An overview* , (Tsiatis *et al.*, 2004)
- *Linear Mixed Models for Longitudinal Data*, (Verbeke and Mohlenbergs, 2010)

Draft working paper:
http://www.ub.edu/riskcenter/research/WP/UBriskcenterWP201407.pdf

ACKNOWLEDGEMENTS

THANK YOU VERY MUCH
FOR YOUR ATTENTION!