Improving long-term survival estimation through flexible models, combining evidence and accessible software

Chris Jackson
MRC Biostatistics Unit, Cambridge, U.K.

UCL, 7 July 2015

Overview

"Improving long-term survival estimation through

1. flexible models
2. combining evidence
3. accessible software...

... but you can't currently have all three at the same time!"

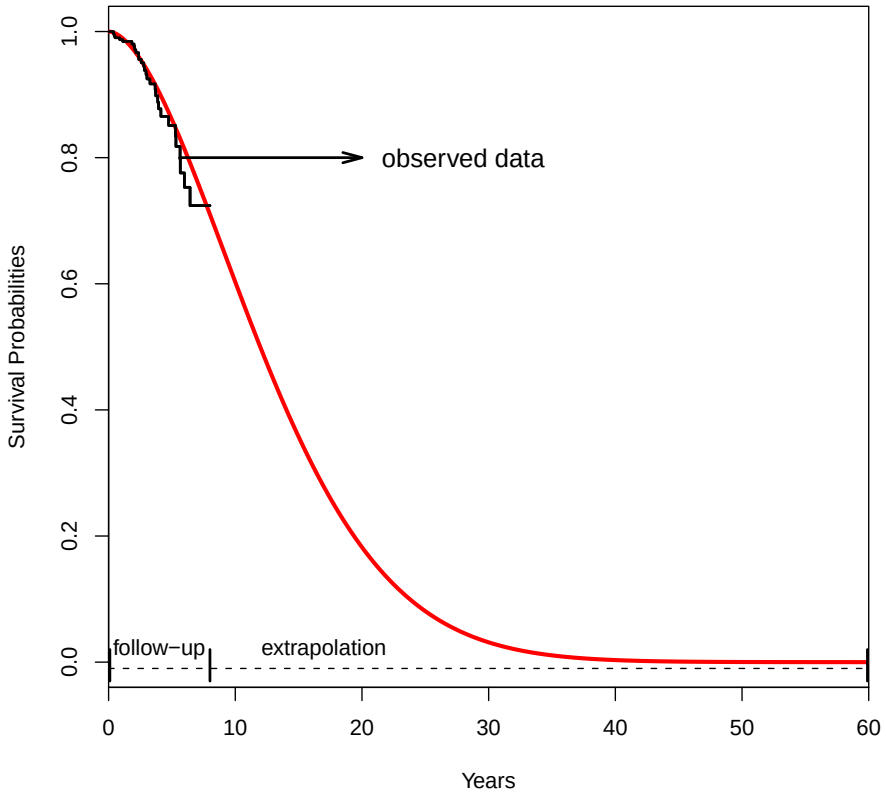
Overview

"Improving long-term survival estimation through

1. flexible models
2. combining evidence
3. accessible software...

... but you can't currently have all three at the same time!"

Premise of the talk



Long-term survival modelling needs

- Sufficient long-term data:
  - individual data with long follow up, or
  - individual data with short follow up + other data on the long-term period
- Models flexible enough to capture the data
- Usable software, skills.

Overview

1. Using long-term population data to extrapolate RCT evidence over time (slides 6–21)

   (Benaglia, Jackson & Sharples, Stat. Med (2015) 34(5):796–811)

flexible modelling | combining evidence | accessible software

2. Flexible parametric survival models for one individual dataset (slides 23–37)
   - The `flexsurv` R package for parametric survival modelling

flexible modelling | | accessible software

Part I

Using long-term population data to extrapolate short-term survival data

Motivating example: ICDs

ICD (Implantable Cardioverter Defibrillators) compared to anti-arrhythmic drugs (AAD) for prevention of sudden cardiac death in people with cardiac arrhythmia.

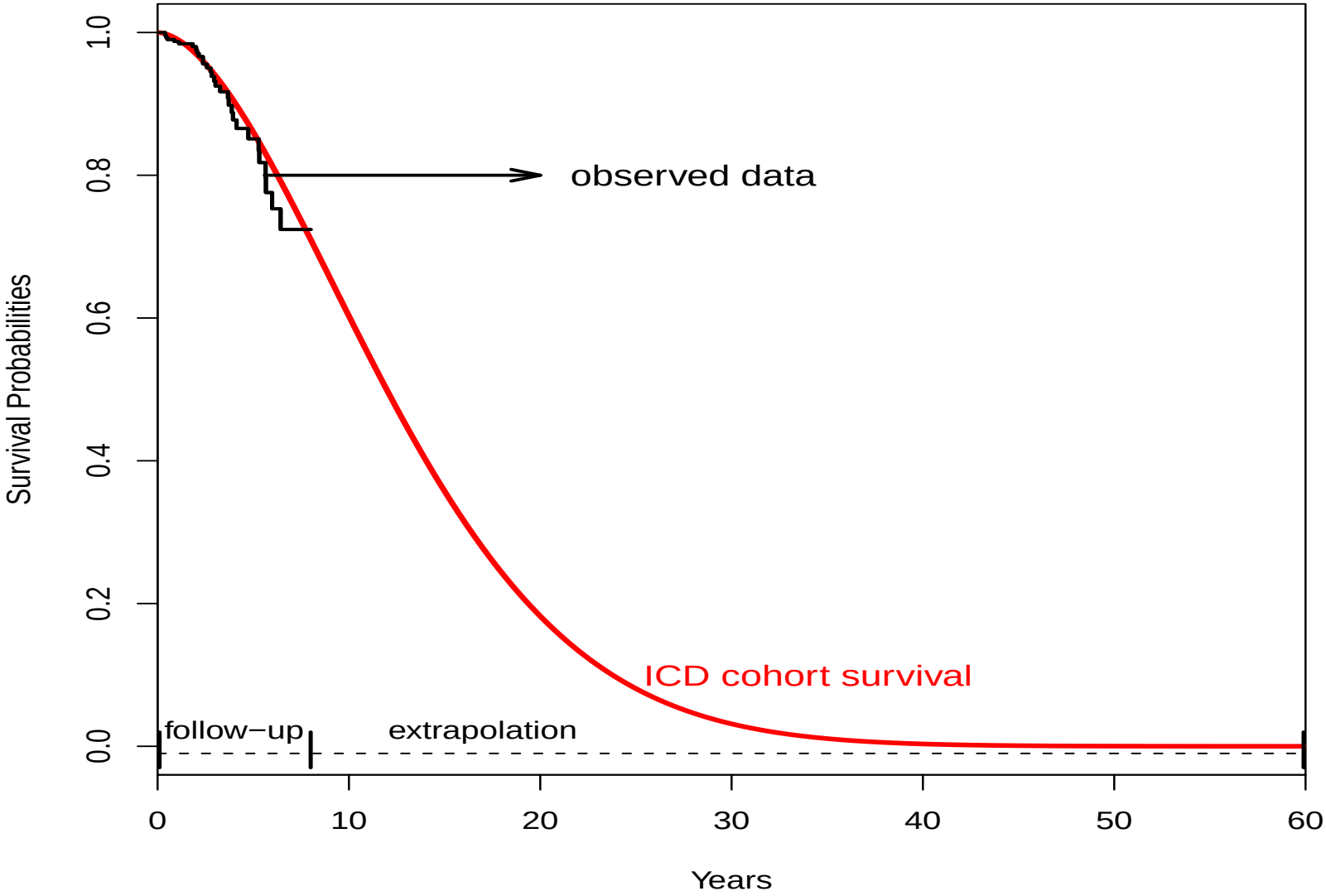
Data:

- Individual data from cohort of 535 UK cardiac arrhythmia patients implanted with ICDs between 1991 and 2002.
- Meta-analysis of three (non-UK) RCTs (published HRs).
  - Relatively short-term follow-up: approximately 75% people followed for less than 5 years, maximum 10 years
- UK population mortality statistics by age, sex, cause of death.

Estimate the survival curve over the lifetime of ICD and AAD patients in UK

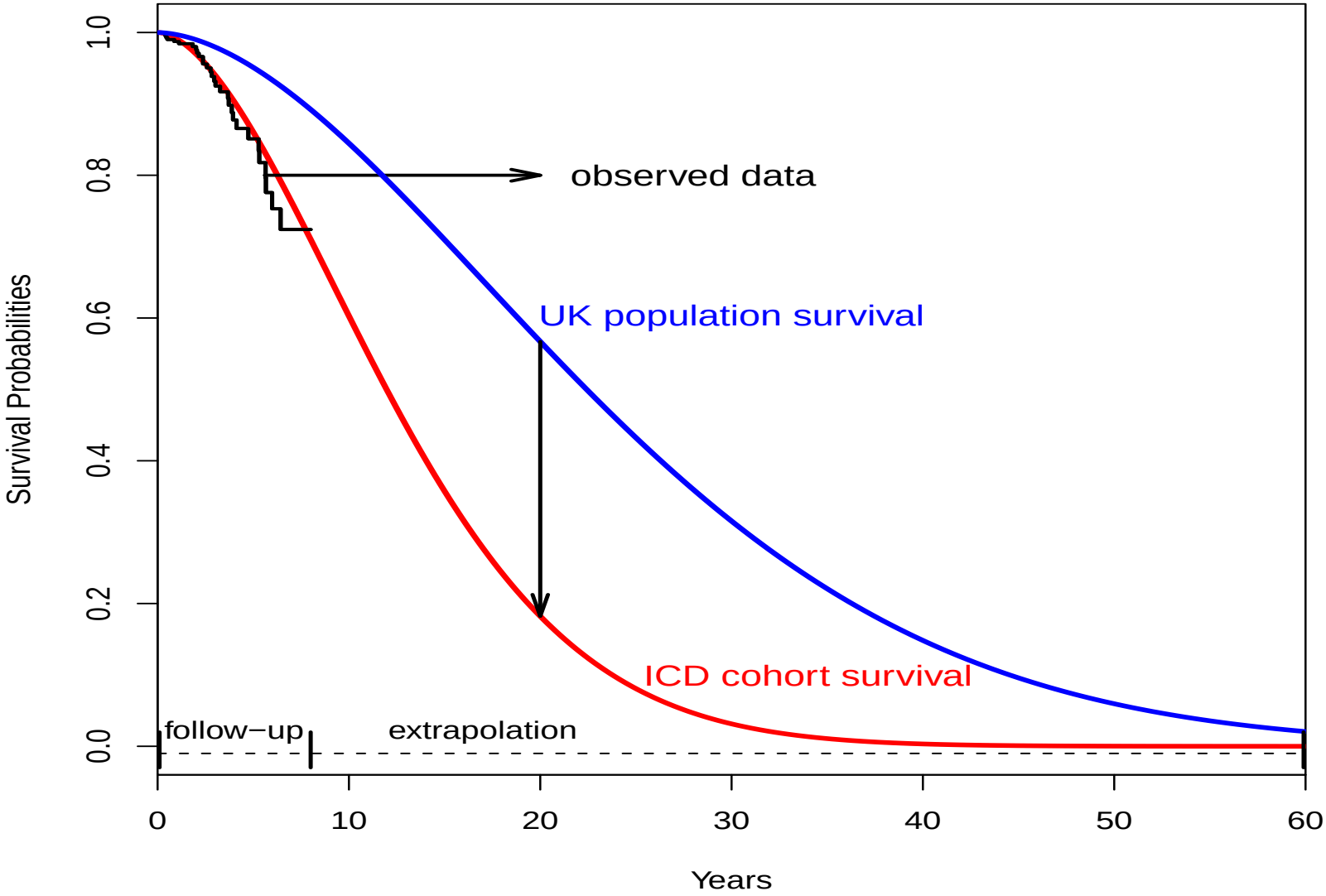
(Benaglia, Jackson & Sharples, Stat. Med (2015) 34(5):796–811)

Previous Work: central idea (Demiris & Sharples, Stat. Med. 2006)



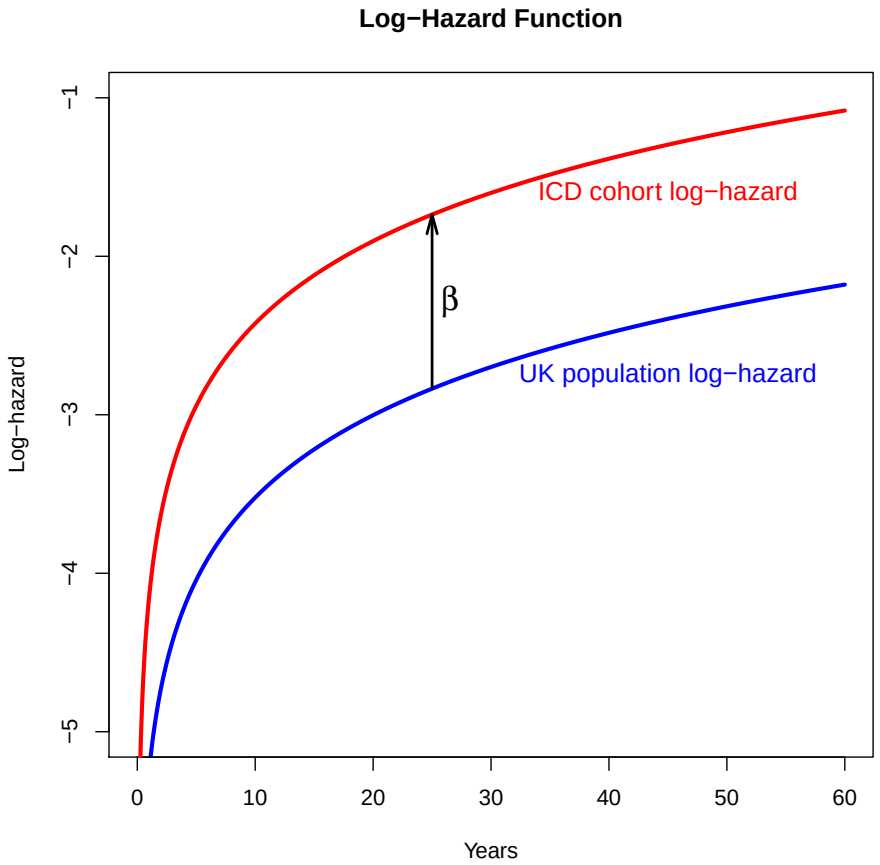
Use UK population data with same age/sex distribution to anchor the ICD population risk

Previous Work: central idea (Demiris & Sharples, Stat. Med. 2006)



Use UK population data with same age/sex distribution to anchor the ICD population risk

Previous Work: key assumption



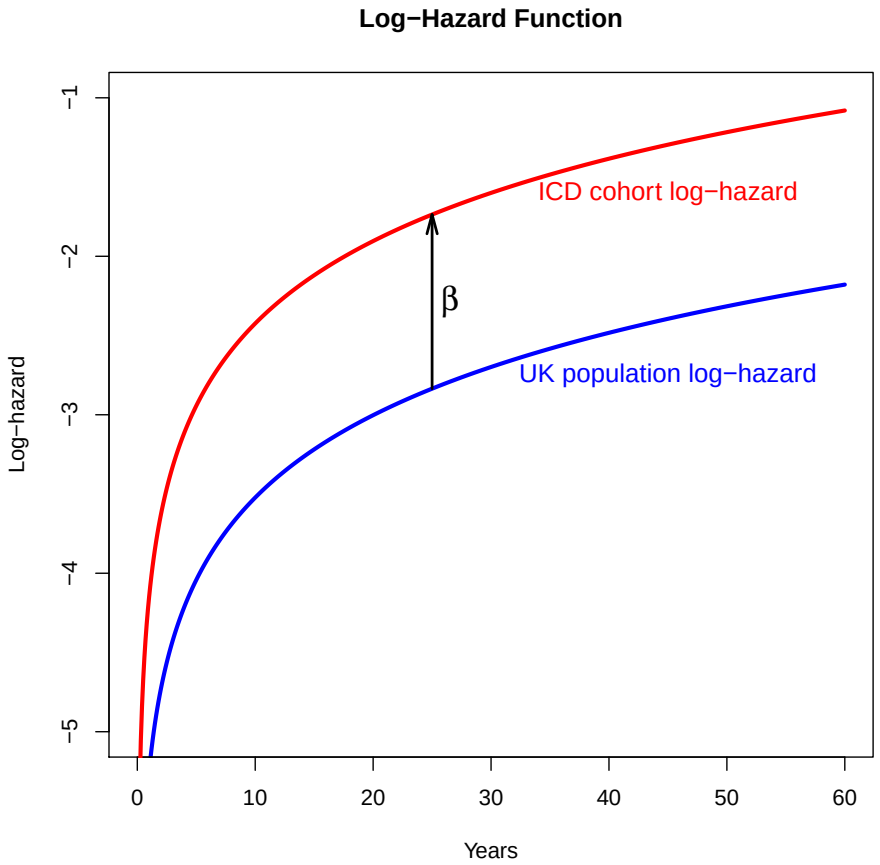
$$h_{ICD}(t) = e^{\beta} h_{UK}(t), \text{ for } t > 0$$

Constant (multiplicative) hazard ratio between ICD and UK population

This seems a strong assumption:

1. ICD patients at greater risk of arrhythmia death
2. If proportion of deaths caused by arrhythmia changes over time, then extrapolating constant HR for all causes may be inaccurate

Previous Work: key assumption



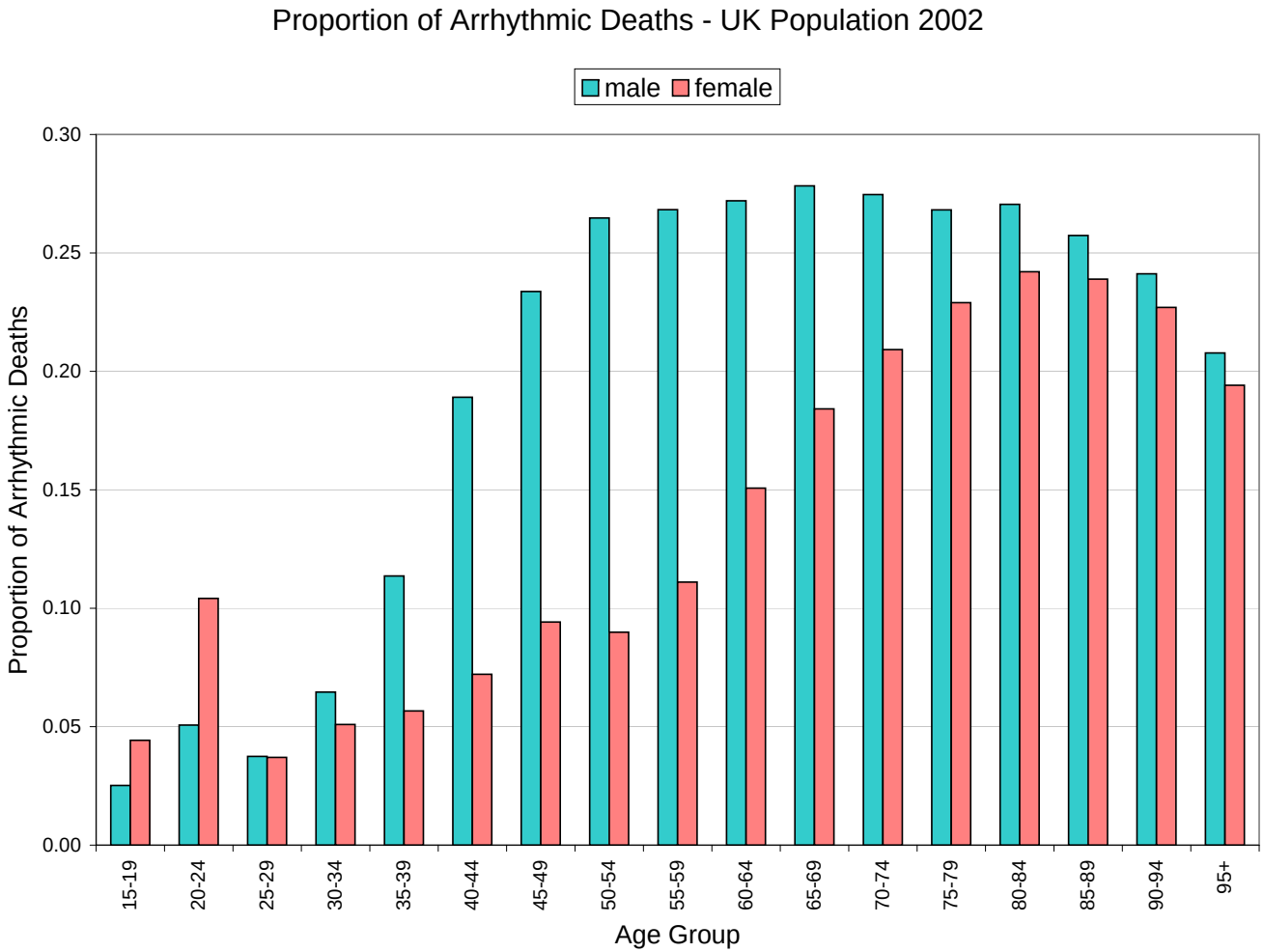
$$h_{ICD}(t) = e^{\beta} h_{UK}(t), \text{ for } t > 0$$

Constant (multiplicative) hazard ratio between ICD and UK population

This seems a strong assumption:

1. ICD patients at greater risk of arrhythmia death
2. If proportion of deaths caused by arrhythmia changes over time, then extrapolating constant HR for all causes may be inaccurate

Proportion of UK deaths which are due to arrhythmia



Account for multiple causes of death when modelling

Extrapolate constant cause-specific instead of all-cause hazard ratio

Simulation study: ignoring cause-specific nature of hazard gives bias in mean survival

- particularly when hazard increases much quicker for other cause
- See paper for full details

  (Benaglia, Jackson & Sharples, Stat. Med (2015) 34(5):796–811)

Application to ICD example...

Model to extrapolate survival for ICD patients

(not considering AAD control group, RCT data for the moment...)

- General population data: cause of death (k =arrhythmic, non-arrhythmic) known.
  Cause-specific survival is Weibull with hazard:

$$h_{UK}^{(k)}(t) = \alpha_k \lambda_k t^{\alpha_k - 1}$$

- ICD cohort: cause of death unknown
  Overall survival follows a polyhazard model (Louzada-Neto, Biometrics 1999):

$$h_{ICD}(t) = h_{ICD}^{arr}(t) + h_{ICD}^{other}(t)$$

  - t: minimum time to one of 2 possible causes of death
  - Hazard is the sum of 2 cause-specific hazards

Cause-specific proportional hazards assumption

ICD cohort hazard is related to the general population hazard as:

$$\begin{aligned} h_{ICD}(t) &= h^{arr}_{ICD}(t) + h^{other}_{ICD}(t) \\ &= e^{\beta} h^{arr}_{UK}(t) + h^{other}_{UK}(t) \\ &= e^{\beta} \alpha_1 \lambda_1 t^{\alpha_1 - 1} + \alpha_2 \lambda_2 t^{\alpha_2 - 1} \text{ (poly-Weibull)} \end{aligned}$$

Arrhythmia hazard is proportional
Other-cause hazard is identical
} to UK matched population.

- Joint Bayesian model for ICD cohort + UK population data
- Estimate joint posterior of parameters $\alpha_1, \alpha_2, \lambda_1, \lambda_2, \beta$ by MCMC (using WinBUGS).
- WBDev add-on needed to implement the poly-Weibull distribution for the cohort data

Weakly informative prior distributions

Express beliefs on an intuitive scale — exact choice may make a difference for small populations

Weibull rate λ:

- Age around 60 on study entry: cannot survive more than 60 additional years. Mean survival $\sim U(0, 60)$.
- $1/\lambda \sim U(0, 100)$, gives a mean $1/\lambda\Gamma(1 + 1/\alpha)$ of < 60, for all plausible α.

Weibull shape α: controls hazard vs. time: $h(t) = \alpha\lambda(\lambda t)^{\alpha-1}$

- Hazard ratio for doubled time t is $2^{\alpha-1}$.
- Prior mean of 1.5 for this, with 95% CI about (0.64, 100)
- implies $\log(\alpha) \sim \mathrm{N}(0.5, \sigma = 0.78)$

Log HR β between ICD patients and general population: 95% CI for HR (1/150,150) $\rightarrow \beta \sim N(0, \sigma = 2.5)$

Weakly informative prior distributions

Express beliefs on an intuitive scale — exact choice may make a difference for small populations

Weibull rate λ:

- Age around 60 on study entry: cannot survive more than 60 additional years. Mean survival $\sim U(0, 60)$.
- $1/\lambda \sim U(0, 100)$, gives a mean $1/\lambda\Gamma(1 + 1/\alpha)$ of < 60, for all plausible α.

Weibull shape α: controls hazard vs. time: $h(t) = \alpha\lambda(\lambda t)^{\alpha-1}$

- Hazard ratio for doubled time t is $2^{\alpha-1}$.
- Prior mean of 1.5 for this, with 95% CI about (0.64, 100)
- implies $\log(\alpha) \sim \mathrm{N}(0.5, \sigma = 0.78)$

Log HR β between ICD patients and general population: 95% CI for HR (1/150,150) $\rightarrow \beta \sim N(0, \sigma = 2.5)$

Weakly informative prior distributions

Express beliefs on an intuitive scale — exact choice may make a difference for small populations

Weibull rate λ:

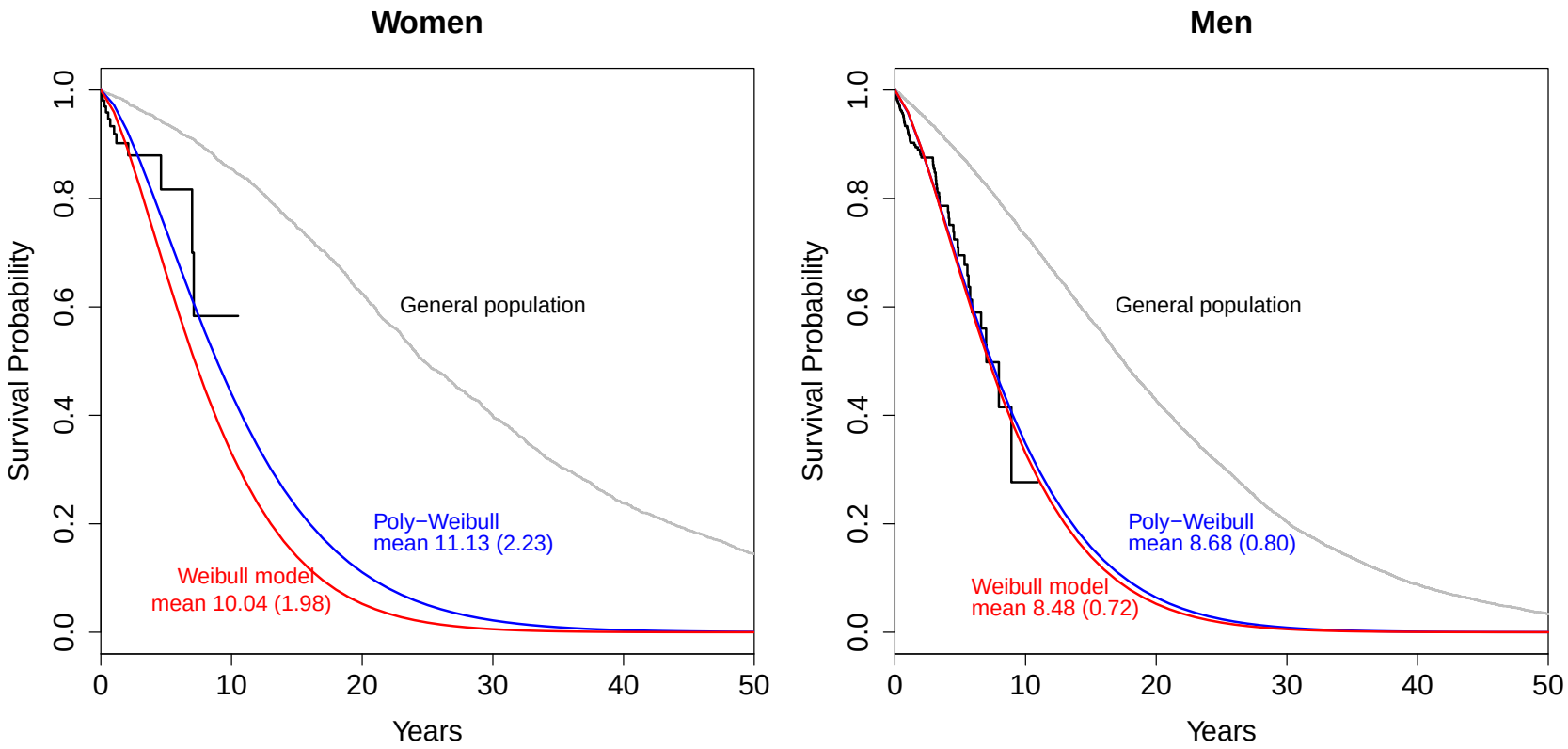
- Age around 60 on study entry: cannot survive more than 60 additional years. Mean survival $\sim U(0, 60)$.
- $1/\lambda \sim U(0, 100)$, gives a mean $1/\lambda\Gamma(1 + 1/\alpha)$ of < 60, for all plausible α.

Weibull shape α: controls hazard vs. time: $h(t) = \alpha\lambda(\lambda t)^{\alpha-1}$

- Hazard ratio for doubled time t is $2^{\alpha-1}$.
- Prior mean of 1.5 for this, with 95% CI about (0.64, 100)
- implies $\log(\alpha) \sim \mathrm{N}(0.5, \sigma = 0.78)$

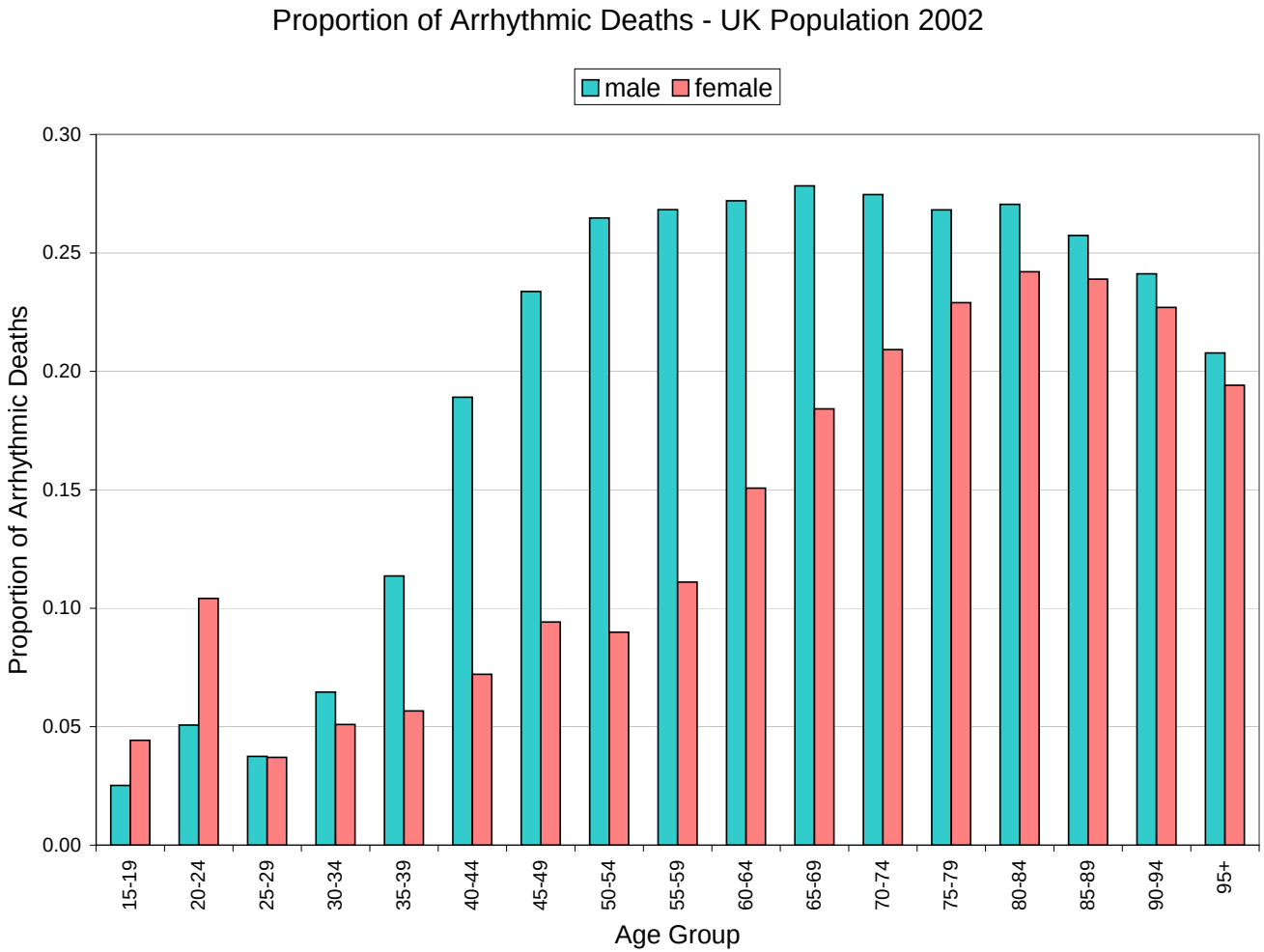
Log HR β between ICD patients and general population: 95% CI for HR (1/150,150) $\rightarrow \beta \sim N(0, \sigma = 2.5)$

Extrapolating real ICD cohort data



- Ignore the cause-specific hazard (Weibull) or account for it (poly-Weibull)
- More bias for women when ignoring it
  - due to time-varying proportion of deaths due to arrhythmia.

Proportion of UK deaths which are due to arrhythmia



Including an intervention effect from literature

Hazards for three groups under Poly-Weibull model:

$$\begin{aligned} h_{\mathrm{UK}}(t) &= h^{arr}_{\mathrm{UK}}(t) + h^{other}_{\mathrm{UK}}(t) \\ h_{\mathrm{ICD}}(t) &= e^{\beta} h^{arr}_{\mathrm{UK}}(t) + h^{other}_{\mathrm{UK}}(t) \\ h_{\mathrm{AAD}}(t) &= e^{\gamma_a + \beta} h^{arr}_{\mathrm{UK}}(t) + h^{other}_{\mathrm{UK}}(t), \end{aligned}$$

Meta-analysis of ICD vs AAD trials, published HR for arrhythmia mortality, gives a prior for γ_a.

For the (probably biased) Weibull model we have:

$$\begin{aligned} h_{\mathrm{UK}}(t) &= \mu_1 \alpha t^{\alpha-1} = e^{\beta_0} \alpha t^{\alpha-1} \\ h_{\mathrm{ICD}}(t) &= e^{\beta_1} h_{\mathrm{UK}}(t) = e^{\beta_0 + \beta_1} \alpha t^{\alpha-1} \\ h_{\mathrm{AAD}}(t) &= e^{\gamma} h_{\mathrm{ICD}}(t) = e^{\beta_0 + \beta_1 + \gamma} \alpha t^{\alpha-1}, \end{aligned}$$

Prior for γ from published meta-analysis HR for all-cause mortality.

Outcome of interest $\rightarrow$ life years gained (LYG) by ICDs vs AADs.

Including an intervention effect from literature

Hazards for three groups under Poly-Weibull model:

$$\begin{aligned} h_{\text{UK}}(t) &= h^{arr}_{\text{UK}}(t) + h^{other}_{\text{UK}}(t) \\ h_{\text{ICD}}(t) &= e^{\beta} h^{arr}_{\text{UK}}(t) + h^{other}_{\text{UK}}(t) \\ h_{\text{AAD}}(t) &= e^{\gamma_a + \beta} h^{arr}_{\text{UK}}(t) + h^{other}_{\text{UK}}(t), \end{aligned}$$

Meta-analysis of ICD vs AAD trials, published HR for arrhythmia mortality, gives a prior for γ_a.

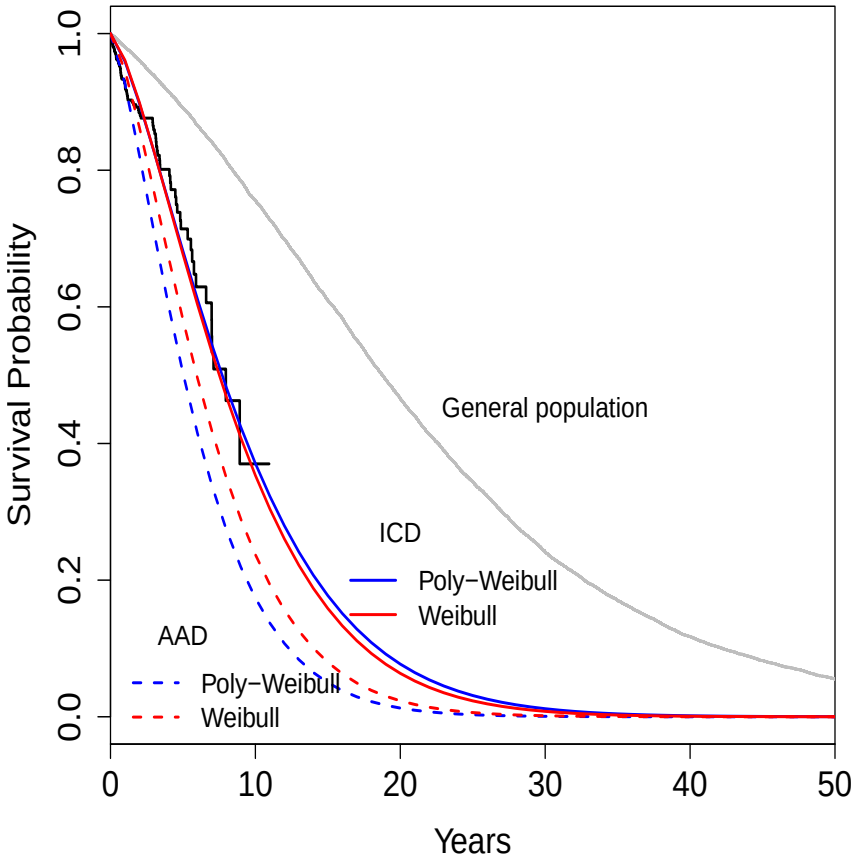
For the (probably biased) Weibull model we have:

$$\begin{aligned} h_{\text{UK}}(t) &= \mu_1 \alpha t^{\alpha-1} = e^{\beta_0} \alpha t^{\alpha-1} \\ h_{\text{ICD}}(t) &= e^{\beta_1} h_{\text{UK}}(t) = e^{\beta_0 + \beta_1} \alpha t^{\alpha-1} \\ h_{\text{AAD}}(t) &= e^{\gamma} h_{\text{ICD}}(t) = e^{\beta_0 + \beta_1 + \gamma} \alpha t^{\alpha-1}, \end{aligned}$$

Prior for γ from published meta-analysis HR for all-cause mortality.

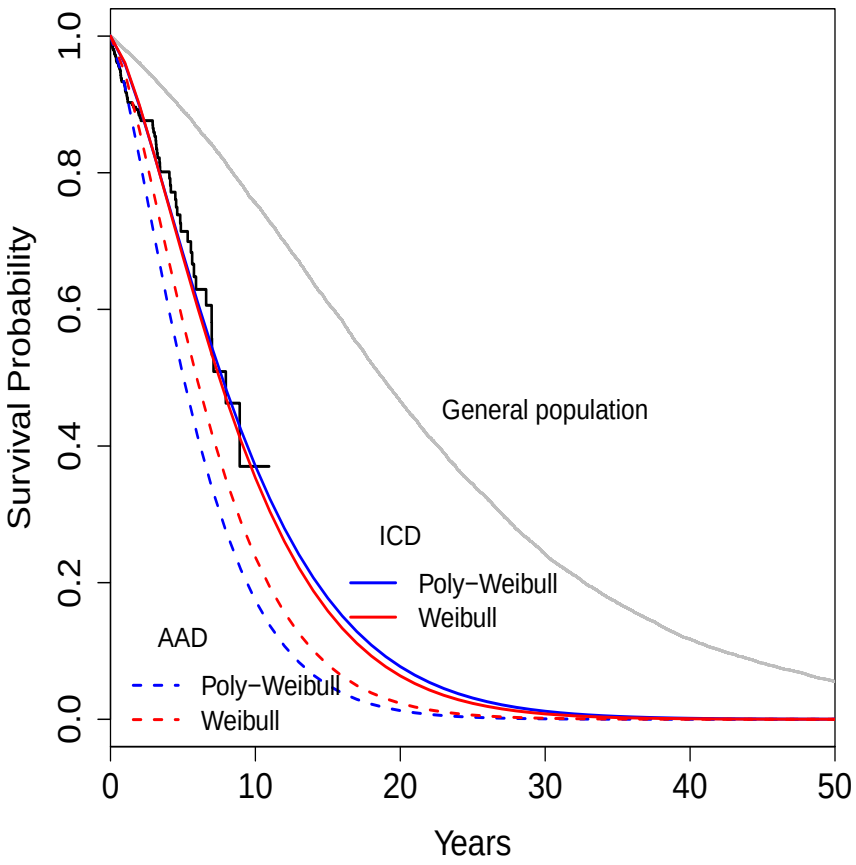
Outcome of interest $\rightarrow$ life years gained (LYG) by ICDs vs AADs.

Extrapolating incremental survival between interventions



- ICD cohort extrapolated using population data
- AAD survival generated with aid of meta-analysis.
- Life-years gained from ICD appears biased if use Weibull

Extrapolating incremental survival between interventions

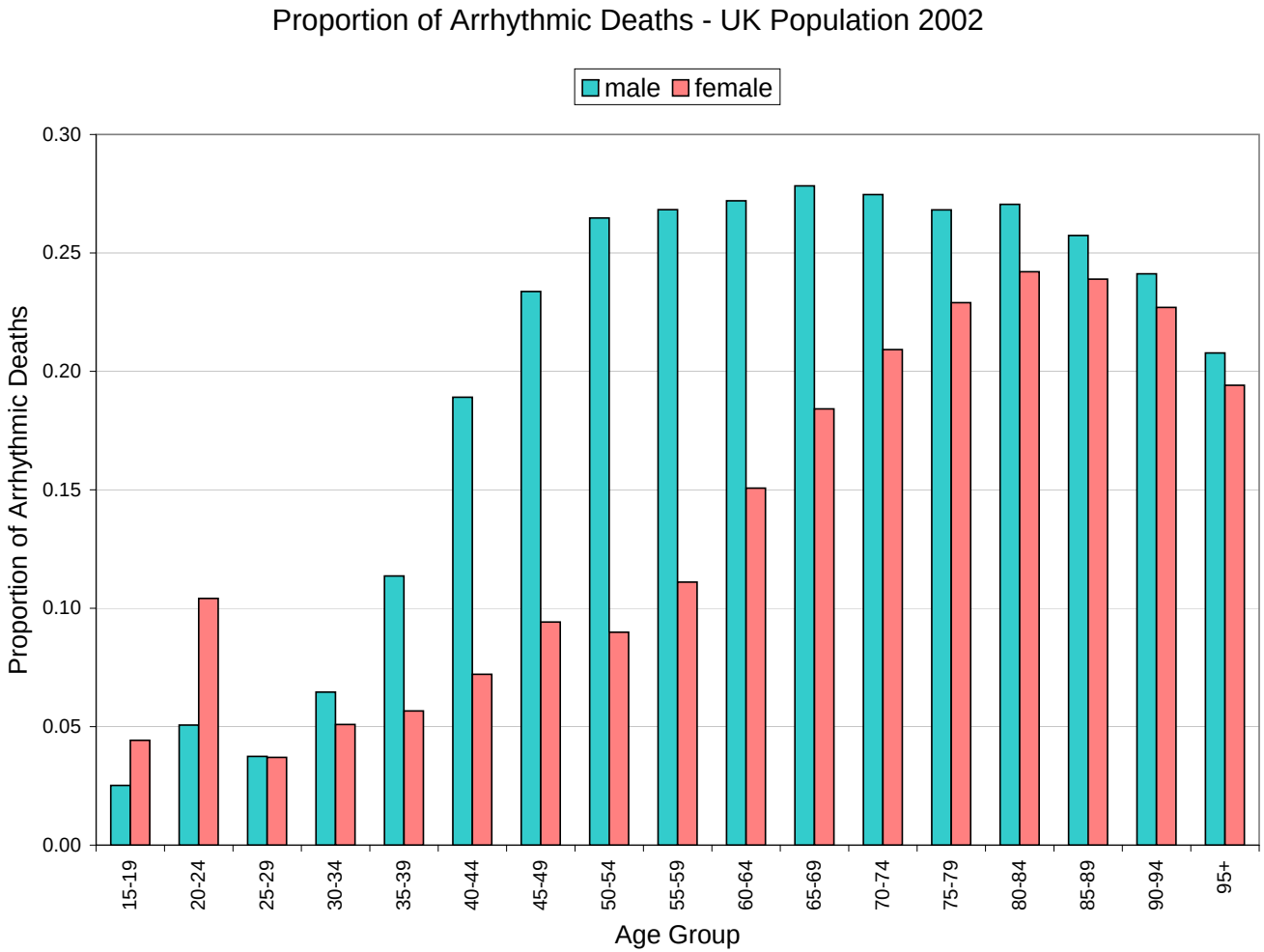


- ICD cohort extrapolated using population data
- AAD survival generated with aid of meta-analysis.
- Life-years gained from ICD appears biased if use Weibull
- Slightly more apparent bias for women

Life-years gained from ICD	Weibull	Poly-Weibull
Overall	1.82 (0.49)	3.12 (0.61)
Women	1.89 (0.62)	3.11 (0.76)
Men	1.73 (0.47)	2.91 (0.58)

. . . Still bias for men

Proportion of UK deaths which are due to arrhythmia



Risks of combining data (1) – data inconsistency

Causes of death may be recorded inconsistently between

- meta-analysis of ICD vs drug trials "HR for arrhythmia deaths"
- population mortality data

Sensitivity analysis — assume 10%-20% "arrhythmia"/"non-arrhythmia" deaths are misclassified.

- e.g. if fewer deaths actually affected by treatment, expected survival gains from treatment lower
- Still doesn't explain the discrepancy between models for men.

Risks of combining data (1) – data inconsistency

Causes of death may be recorded inconsistently between

- meta-analysis of ICD vs drug trials "HR for arrhythmia deaths"
- population mortality data

Sensitivity analysis — assume 10%-20% "arrhythmia"/"non-arrhythmia" deaths are misclassified.

- e.g. if fewer deaths actually affected by treatment, expected survival gains from treatment lower
- Still doesn't explain the discrepancy between models for men.

Risks of combining data (2) — assumptions about hazards

We assumed ICD patients had

arrhythmia hazard proportional(=greater)
other-cause hazard identical
} to general population.

- What if ICD patients at greater risk from some other causes (other heart disease), as well as arrhythmia?
- May have led to biases in survival
  (underestimation of AAD-specific survival in poly-Weibull model. . .reasoning for this in paper)

Conclusions: survival extrapolation example

- Bayesian models useful for combining short-term RCT / cohort and longer-term survival data.
- Ignoring cause-specific hazard, thus misspecifying the underlying model, introduces bias in survival estimates.
  - may underestimate or overestimate overall survival.
- Bias can be alleviated by modelling cause-specific hazards
  - but requires cause-specific survival data / treatment effects
  - and information about which causes will be affected by disease status and / or treatment
- Bias for treatment comparisons may be less if bias acts in the same way in all treatment groups.
- Sensitivity analysis to model / data assumptions important
- Requires model-specific BUGS code...

(Benaglia, Jackson & Sharples, Stat. Med (2015) 34(5):796–811)

Conclusions: survival extrapolation example

- Bayesian models useful for combining short-term RCT / cohort and longer-term survival data.
- Ignoring cause-specific hazard, thus misspecifying the underlying model, introduces bias in survival estimates.
  - may underestimate or overestimate overall survival.
- Bias can be alleviated by modelling cause-specific hazards
  - but requires cause-specific survival data / treatment effects
  - and information about which causes will be affected by disease status and / or treatment
- Bias for treatment comparisons may be less if bias acts in the same way in all treatment groups.
- Sensitivity analysis to model / data assumptions important
- Requires model-specific BUGS code...

(Benaglia, Jackson & Sharples, Stat. Med (2015) 34(5):796–811)

Conclusions: survival extrapolation example

- Bayesian models useful for combining short-term RCT / cohort and longer-term survival data.
- Ignoring cause-specific hazard, thus misspecifying the underlying model, introduces bias in survival estimates.
  - may underestimate or overestimate overall survival.
- Bias can be alleviated by modelling cause-specific hazards
  - but requires cause-specific survival data / treatment effects
  - and information about which causes will be affected by disease status and / or treatment
- Bias for treatment comparisons may be less if bias acts in the same way in all treatment groups.
- Sensitivity analysis to model / data assumptions important
- Requires model-specific BUGS code...

(Benaglia, Jackson & Sharples, Stat. Med (2015) 34(5):796–811)

Conclusions: survival extrapolation example

- Bayesian models useful for combining short-term RCT / cohort and longer-term survival data.
- Ignoring cause-specific hazard, thus misspecifying the underlying model, introduces bias in survival estimates.
  - may underestimate or overestimate overall survival.
- Bias can be alleviated by modelling cause-specific hazards
  - but requires cause-specific survival data / treatment effects
  - and information about which causes will be affected by disease status and / or treatment
- Bias for treatment comparisons may be less if bias acts in the same way in all treatment groups.
- Sensitivity analysis to model / data assumptions important
- Requires model-specific BUGS code...

(Benaglia, Jackson & Sharples, Stat. Med (2015) 34(5):796–811)

Conclusions: survival extrapolation example

- Bayesian models useful for combining short-term RCT / cohort and longer-term survival data.
- Ignoring cause-specific hazard, thus misspecifying the underlying model, introduces bias in survival estimates.
  - may underestimate or overestimate overall survival.
- Bias can be alleviated by modelling cause-specific hazards
  - but requires cause-specific survival data / treatment effects
  - and information about which causes will be affected by disease status and / or treatment
- Bias for treatment comparisons may be less if bias acts in the same way in all treatment groups.
- Sensitivity analysis to model / data assumptions important
- Requires model-specific BUGS code...

(Benaglia, Jackson & Sharples, Stat. Med (2015) 34(5):796–811)

Part II

Flexible parametric survival models and software

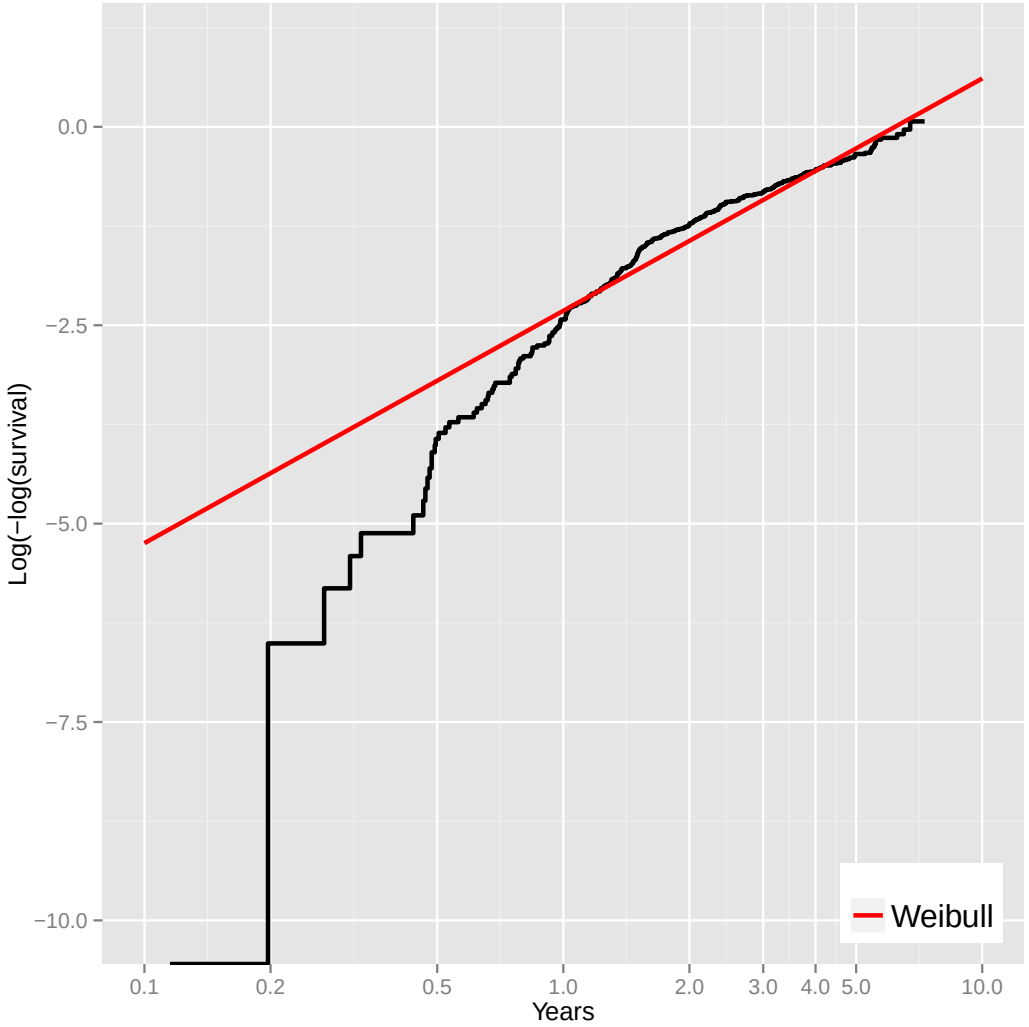
Flexible parametric models

- 2-parameter Weibull model judged adequate in ICD application — but this is not always the case.
- 3–4 parameter models (generalised gamma / F) (see e.g. Jackson et al Int J Biostat) sometimes better

More flexible is a spline-based model (Royston & Parmar, Stat. Med. 2002)

- can model both baseline hazard and non-proportional hazards between groups with any number of parameters
  - can fit as well as needed
- can be implemented in Stata (`stpm2`) and R (`flexsurv`)

Spline survival model



(German breast cancer data, see Sauerbrei & Royston, J.Roy.Stat.Soc A 1999)

Log cumulative hazard (log(-log survival))

- Weibull: linear function of log time t

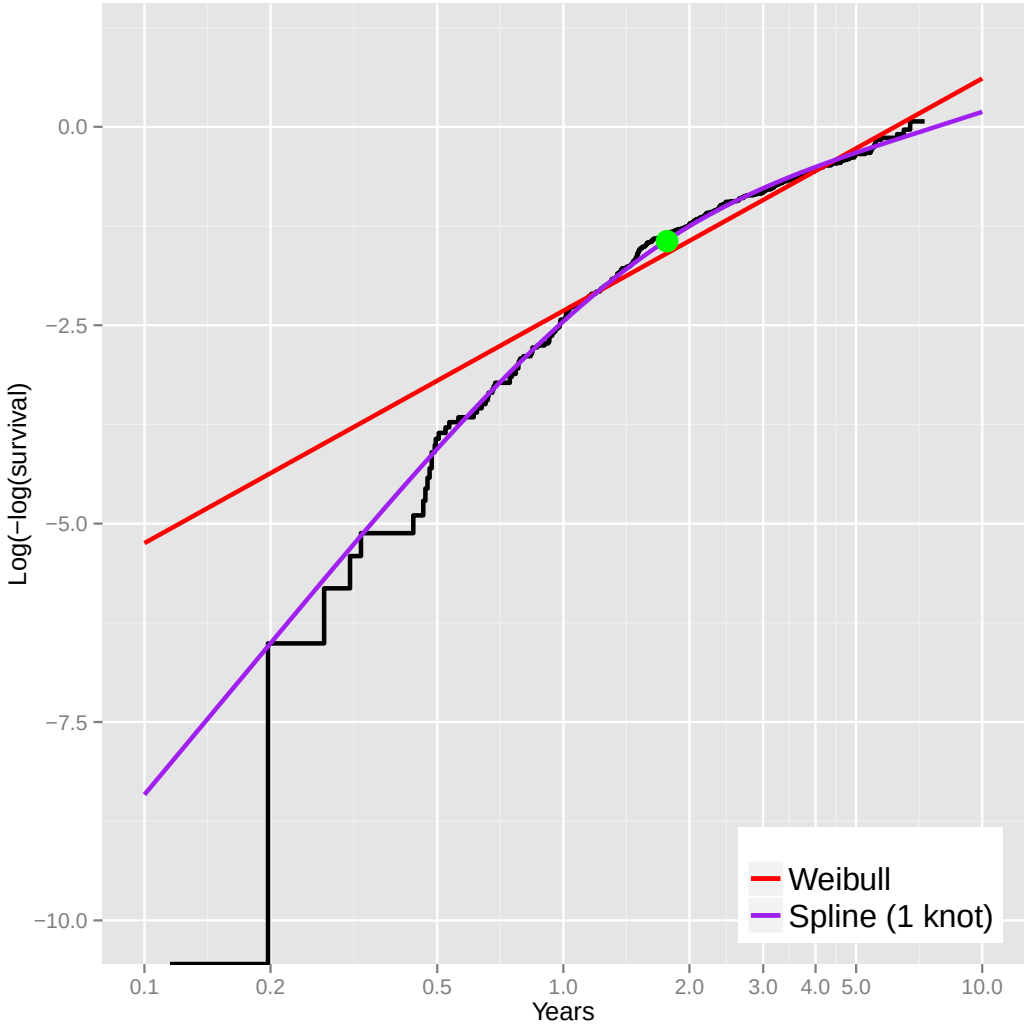
Spline: piecewise cubic function of $\log(t)$

- Pieces separated by knots to span the data

0 knots: Weibull

n knots: $n + 2$ parameters

Spline survival model



(German breast cancer data, see Sauerbrei & Royston, J.Roy.Stat.Soc A 1999)

Log cumulative hazard (log(-log survival))

- Weibull: linear function of log time t

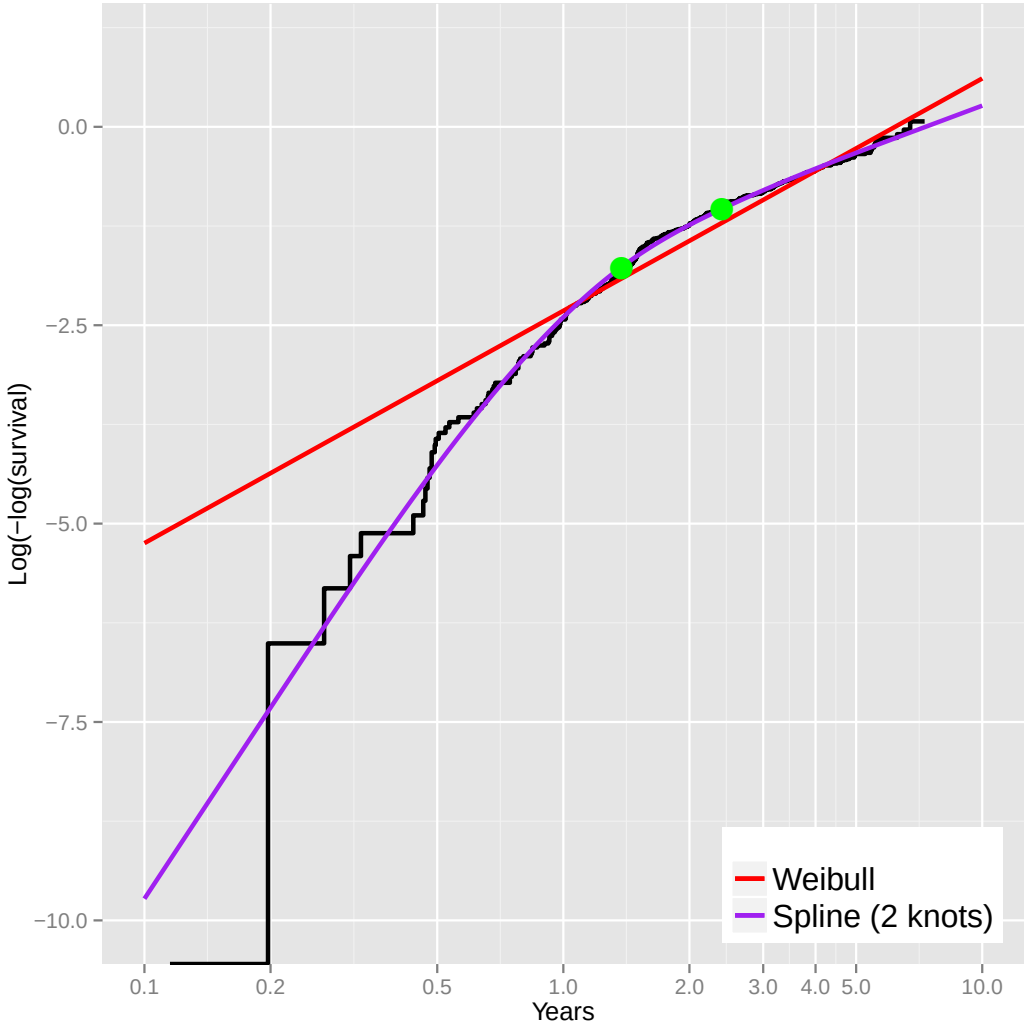
Spline: piecewise cubic function of $\log(t)$

- Pieces separated by knots to span the data

0 knots: Weibull

n knots: $n + 2$ parameters

Spline survival model



(German breast cancer data, see Sauerbrei & Royston, J.Roy.Stat.Soc A 1999)

Log cumulative hazard (log(-log survival))

- Weibull: linear function of log time t

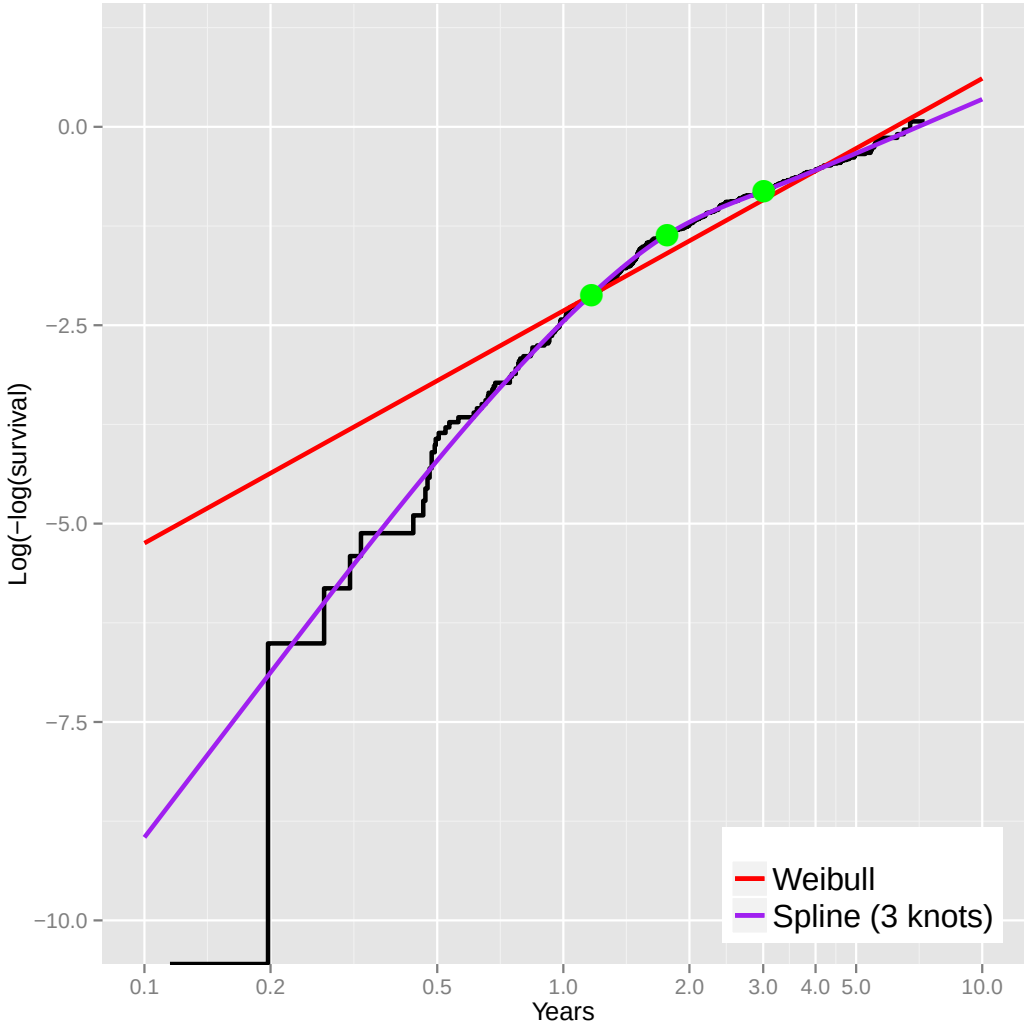
Spline: piecewise cubic function of $\log(t)$

- Pieces separated by knots to span the data

0 knots: Weibull

n knots: $n + 2$ parameters

Spline survival model



(German breast cancer data, see Sauerbrei & Royston, J.Roy.Stat.Soc A 1999)

Log cumulative hazard (log(-log survival))

- Weibull: linear function of log time t

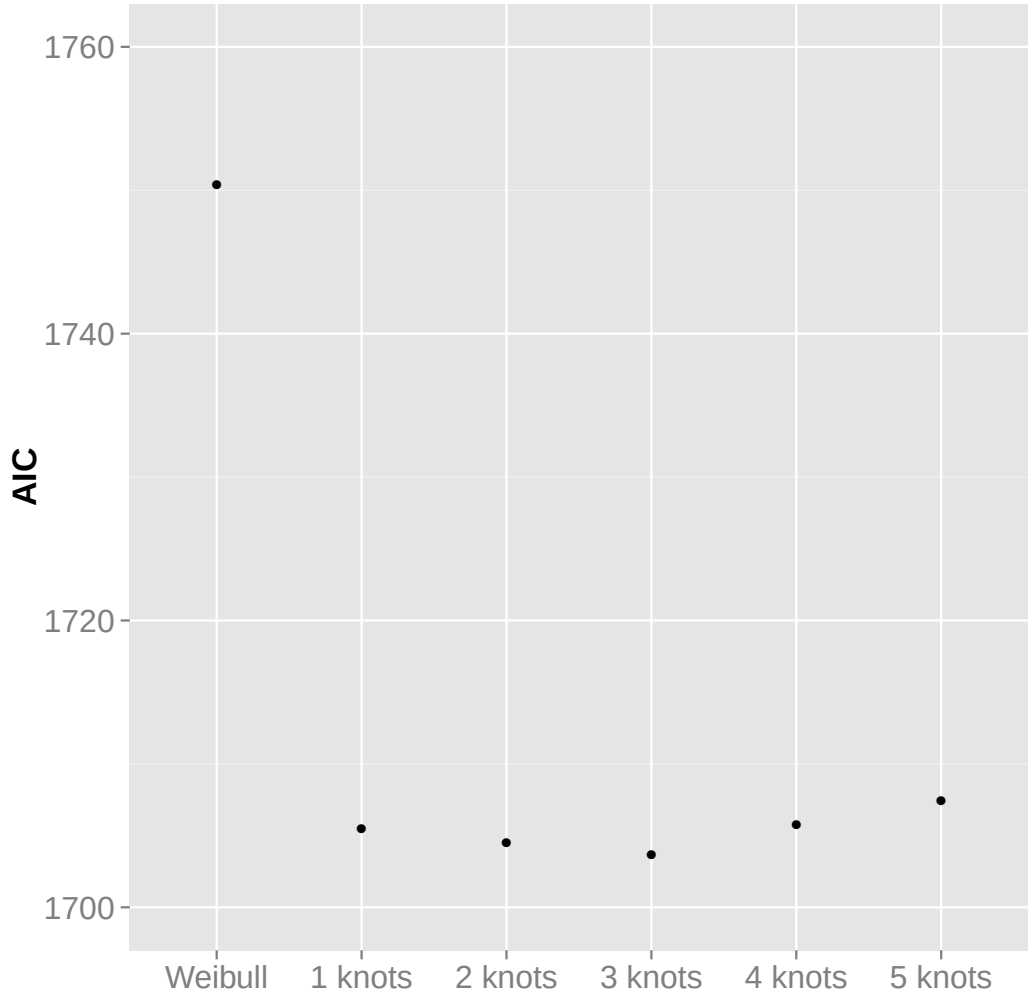
Spline: piecewise cubic function of $\log(t)$

- Pieces separated by knots to span the data

0 knots: Weibull

n knots: $n + 2$ parameters

Spline survival model



(German breast cancer data, see Sauerbrei & Royston, J.Roy.Stat.Soc A 1999)

Log cumulative hazard (log(-log survival))

- Weibull: linear function of log time t

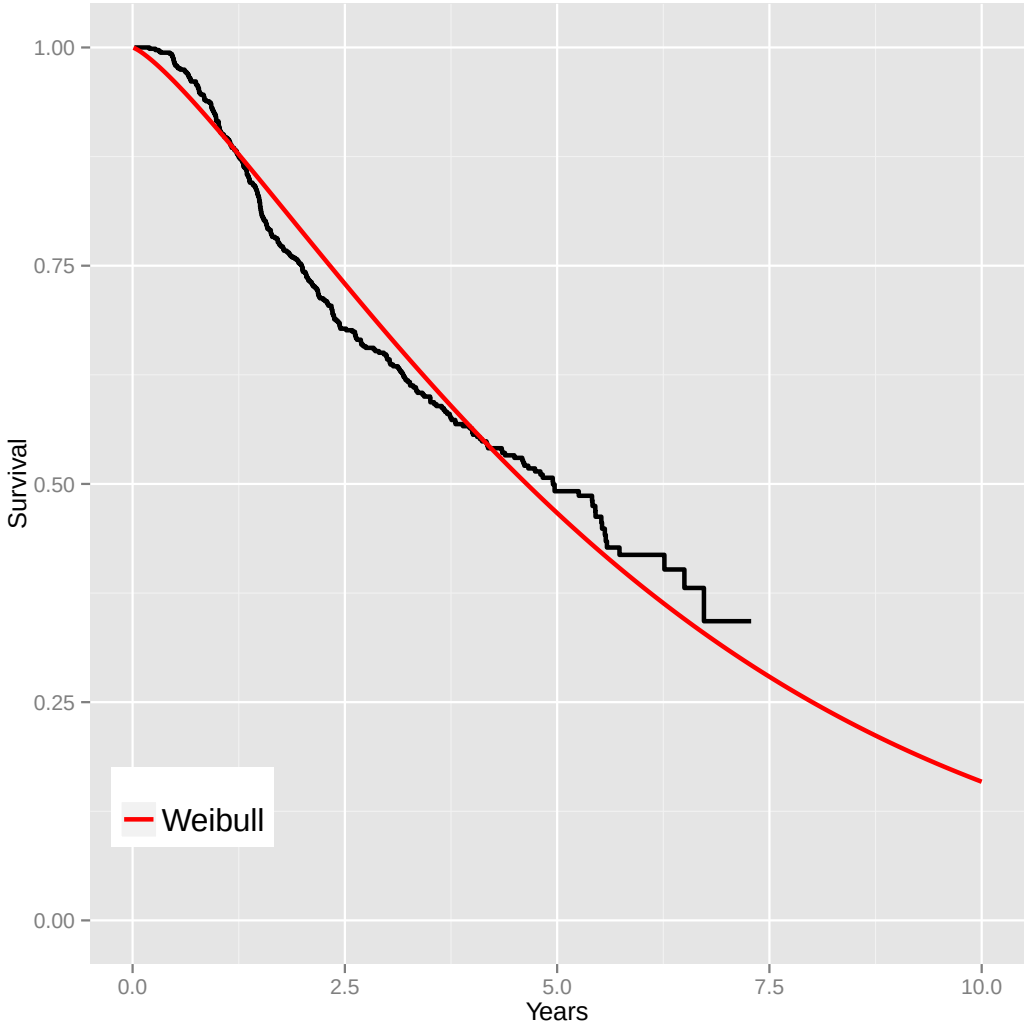
Spline: piecewise cubic function of $\log(t)$

- Pieces separated by knots to span the data

0 knots: Weibull

n knots: $n + 2$ parameters

Spline survival model



(German breast cancer data, see Sauerbrei & Royston, J.Roy.Stat.Soc A 1999)

Log cumulative hazard (log(-log survival))

- Weibull: linear function of log time t

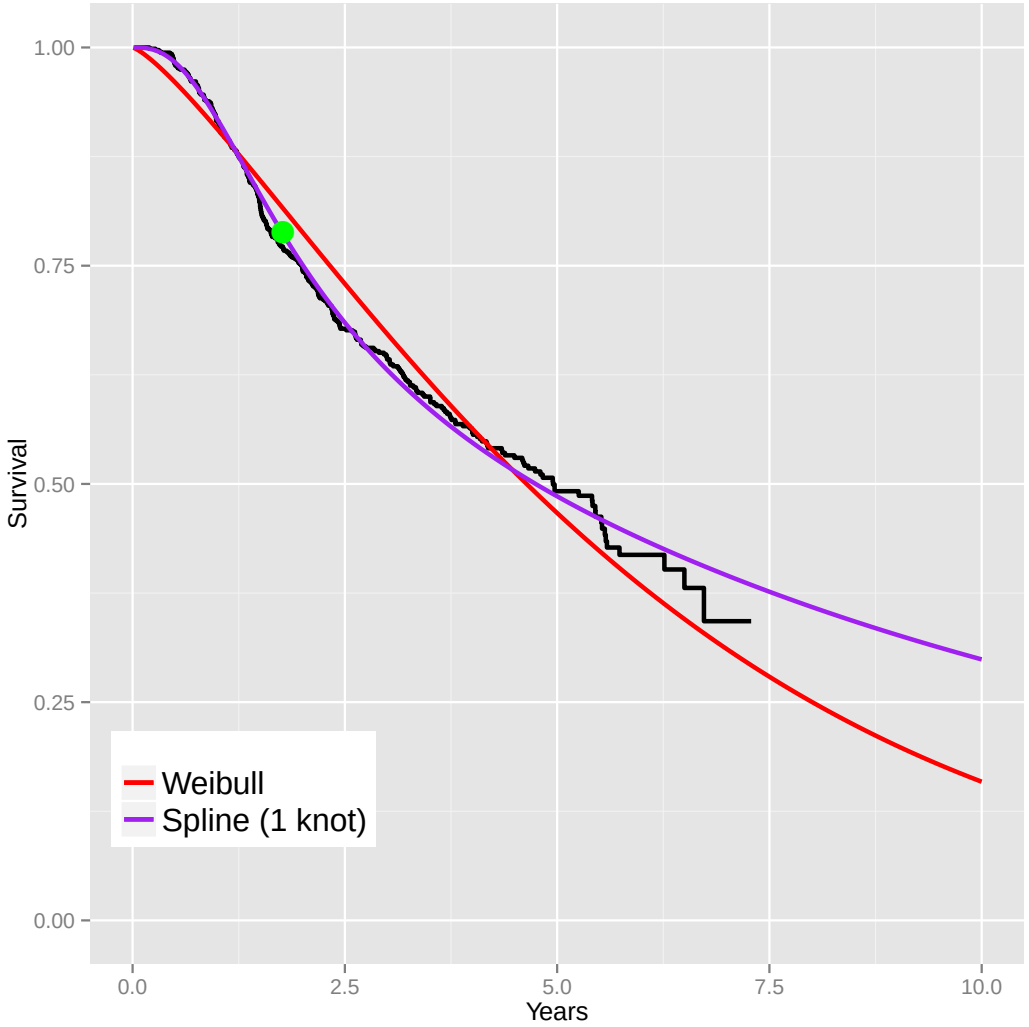
Spline: piecewise cubic function of $\log(t)$

- Pieces separated by knots to span the data

0 knots: Weibull

n knots: $n + 2$ parameters

Spline survival model



(German breast cancer data, see Sauerbrei & Royston, J.Roy.Stat.Soc A 1999)

Log cumulative hazard (log(-log survival))

- Weibull: linear function of log time t

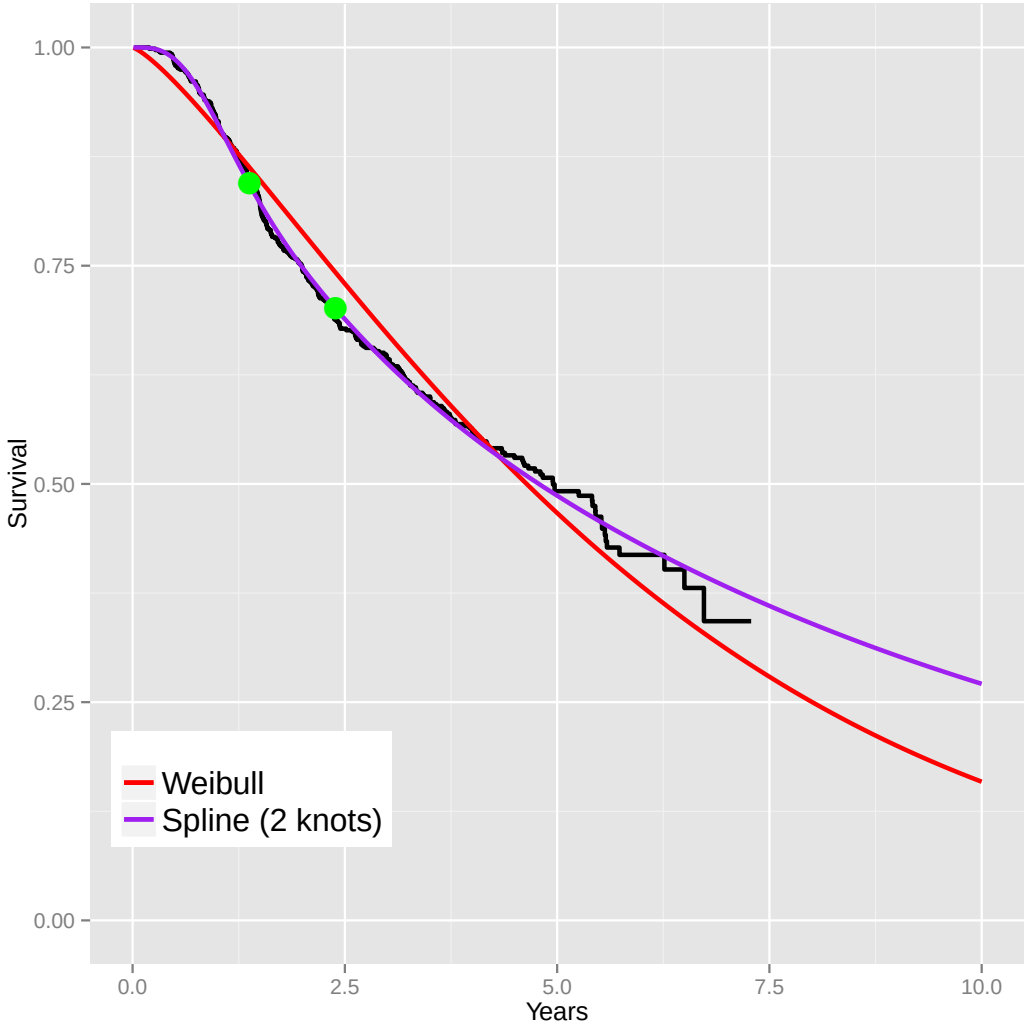
Spline: piecewise cubic function of $\log(t)$

- Pieces separated by knots to span the data

0 knots: Weibull

n knots: $n + 2$ parameters

Spline survival model



(German breast cancer data, see Sauerbrei & Royston, J.Roy.Stat.Soc A 1999)

Log cumulative hazard (log(-log survival))

- Weibull: linear function of log time t

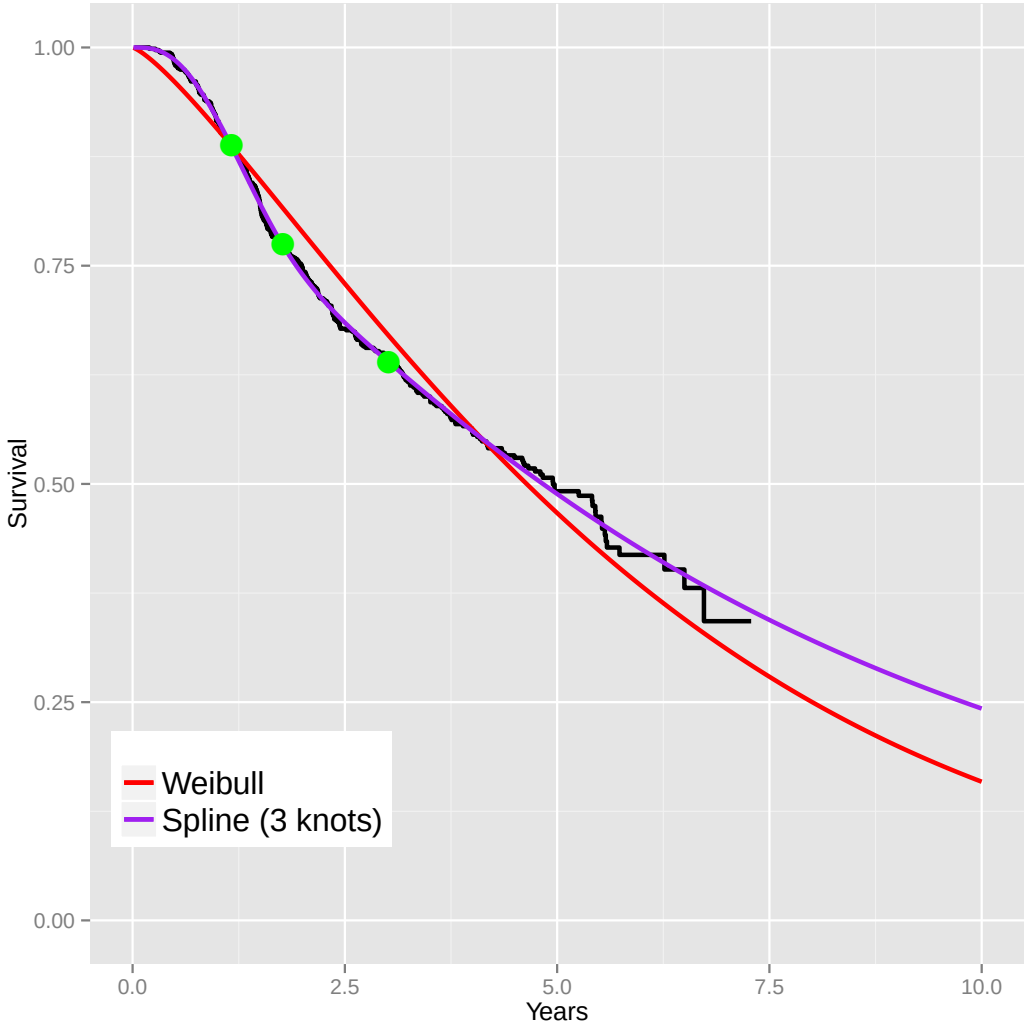
Spline: piecewise cubic function of $\log(t)$

- Pieces separated by knots to span the data

0 knots: Weibull

n knots: $n + 2$ parameters

Spline survival model



(German breast cancer data, see Sauerbrei & Royston, J.Roy.Stat.Soc A 1999)

Log cumulative hazard (log(-log survival))

- Weibull: linear function of log time t

Spline: piecewise cubic function of $\log(t)$

- Pieces separated by knots to span the data

0 knots: Weibull

n knots: $n + 2$ parameters

Spline equations

Log cumulative hazard in terms of $x = log(t)$, parameters γ

$$log(-log(S(x, \gamma))) = \gamma_0 + \gamma_1 x$$

for Weibull

Spline equations

Log cumulative hazard in terms of $x = log(t)$, parameters γ

$$log(-log(S(x, \gamma))) = \gamma_0 + \gamma_1 x + \gamma_2 v_1(x) + \ldots + \gamma_{m+1} v_m(x)$$

for spline with m knots. $v_j(x)$ is jth *basis* function

Spline equations

Log cumulative hazard in terms of $x = log(t)$, parameters γ

$$log(-log(S(x, \gamma))) = \gamma_0 + \gamma_1 x + \gamma_2 v_1(x) + \ldots + \gamma_{m+1} v_m(x)$$

for spline with m knots. $v_j(x)$ is jth *basis* function

$$v_j(x) = (x - k_j)_+^3 - \lambda_j (x - k_{min})_+^3 - (1 - \lambda_j)(x - k_{max})_+^3, \qquad \lambda_j = \frac{k_{max} - k_j}{k_{max} - k_{min}}$$

and $(x - a)_+ = max(0, x - a)$.

a cubic specially constructed to give a smooth function at the knots $k_{min}, \ldots, k_{max}$

Modelling covariates, and non-proportional hazards

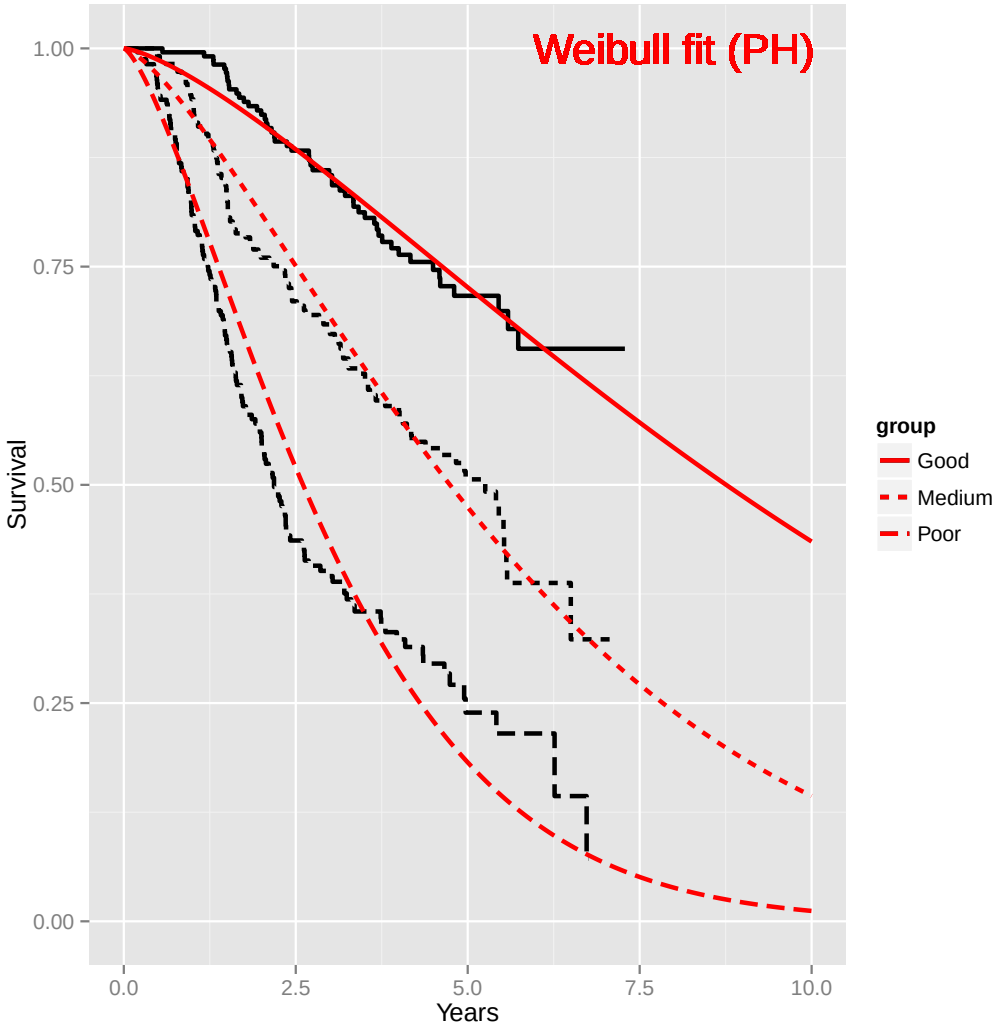
Log cumulative hazard defined by

$$log(-log(S(x, \boldsymbol{\gamma}))) = \gamma_0 + \gamma_1 x + \gamma_2 v_1(x) + \ldots + \gamma_{m+1} v_m(x)$$

Covariates $\mathbf{z}$ can be placed on any parameter γ

- $\gamma_0 = \boldsymbol{\beta}^\top \mathbf{z}$ gives a proportional hazards model
  - Sufficiently flexible splines give similar answers to Cox models
- Modelling any other γ_r as linear in $\mathbf{z}$ gives non-proportional hazards:
  - hazard ratio can be an arbitrarily flexible function of time

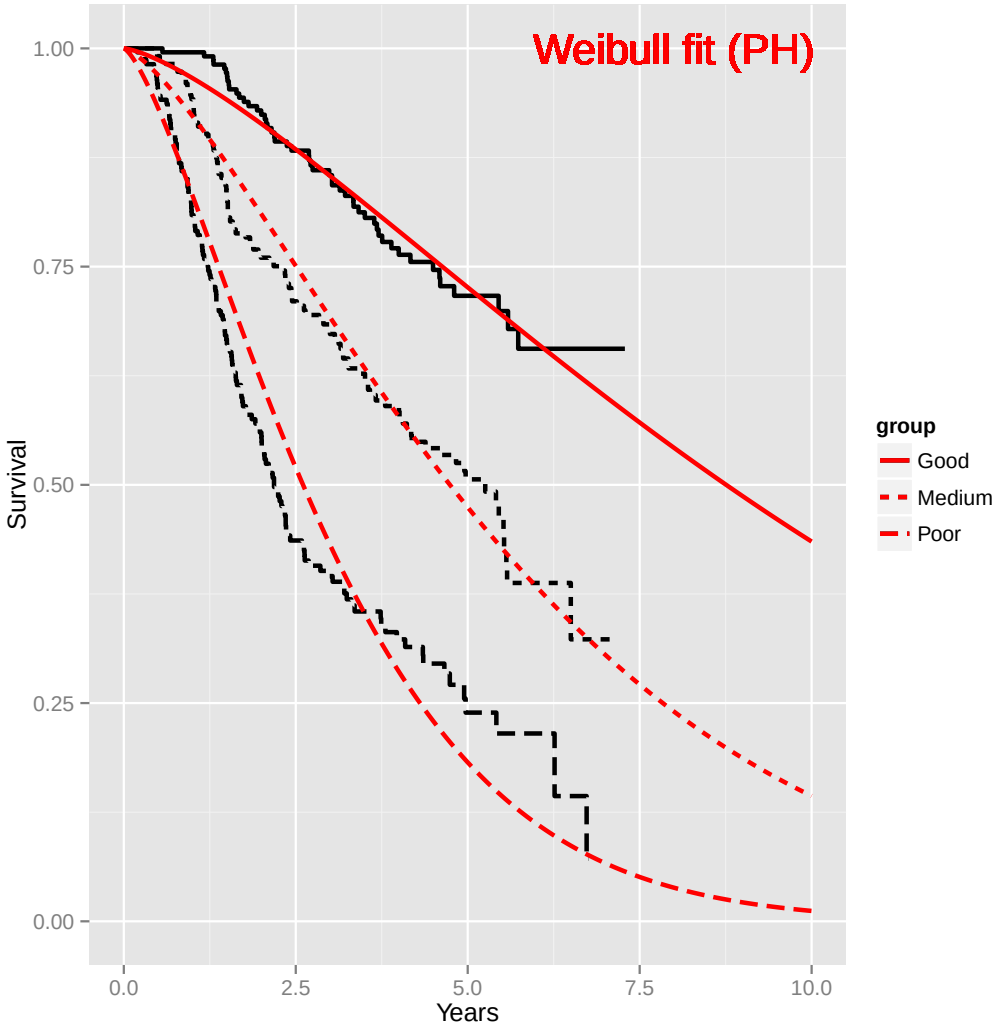
Modelling covariates, and non-proportional hazards



Hazard ratios (vs. Good) under proportional hazards:

Medium vs good prognosis	
Weibull	2.33 (1.67, 3.26)
Poor vs good prognosis	
Weibull	5.33 (3.86, 7.35)

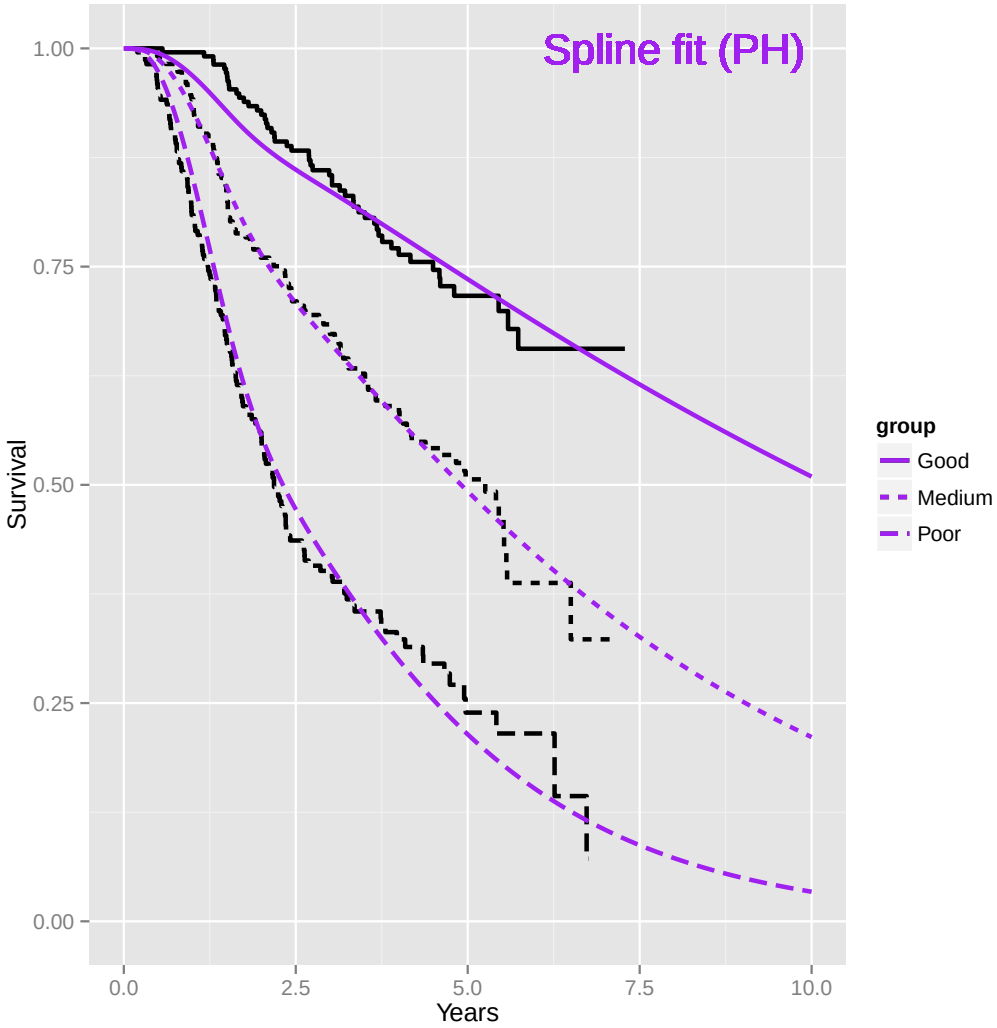
Modelling covariates, and non-proportional hazards



Hazard ratios (vs. Good) under proportional hazards:

Medium vs good prognosis	
Weibull	2.33 (1.67, 3.26)
Cox	2.32 (1.66, 3.24)
Poor vs good prognosis	
Weibull	5.33 (3.86, 7.35)
Cox	5.04 (3.65, 6.96)

Modelling covariates, and non-proportional hazards



Hazard ratios (vs. Good) under proportional hazards:

Medium vs good prognosis	
Weibull	2.33 (1.67, 3.26)
Cox	2.32 (1.66, 3.24)
Spline	2.31 (1.65, 3.23)
Poor vs good prognosis	
Weibull	5.33 (3.86, 7.35)
Cox	5.04 (3.65, 6.96)
Spline	5.02 (3.64, 6.93)

Modelling covariates, and non-proportional hazards

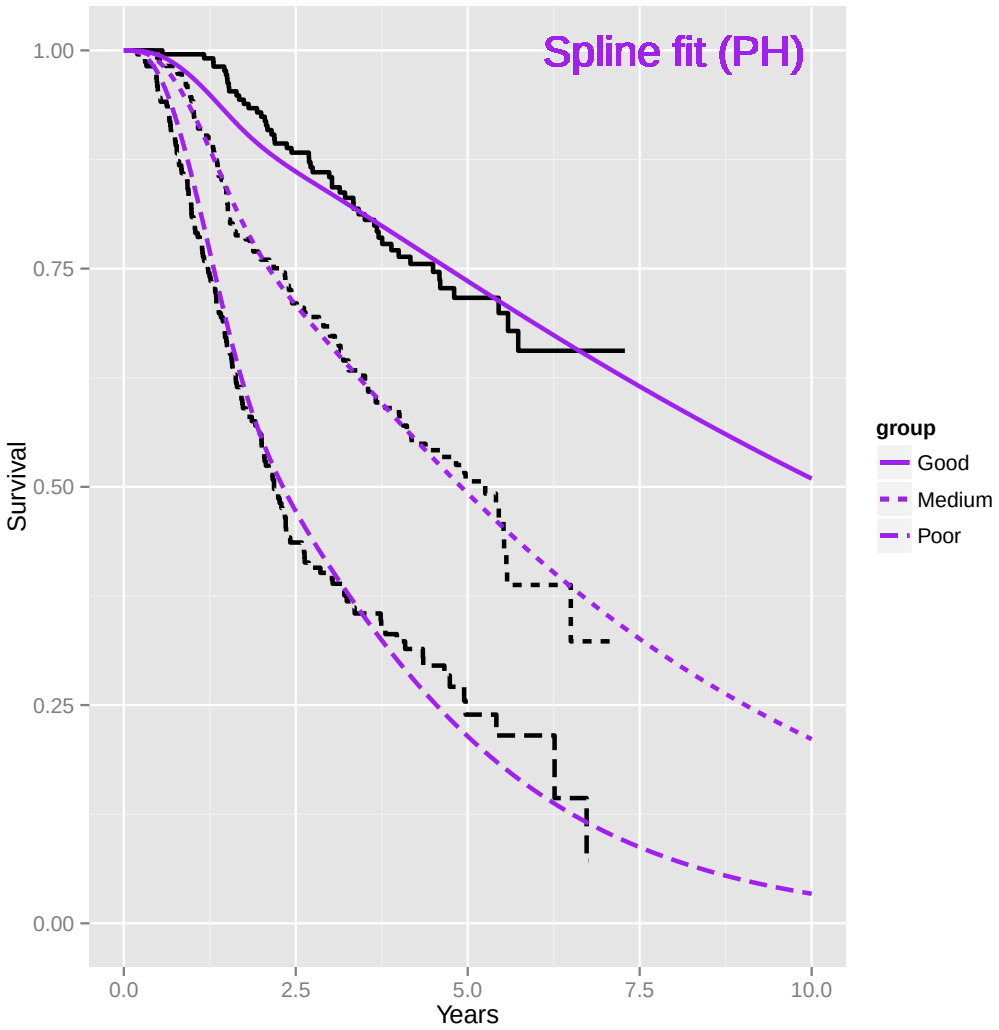
Log cumulative hazard defined by

$$log(-log(S(x, \boldsymbol{\gamma}))) = \gamma_0 + \gamma_1 x + \gamma_2 v_1(x) + \ldots + \gamma_{m+1} v_m(x)$$

Covariates $\mathbf{z}$ can be placed on any parameter γ

- $\gamma_0 = \boldsymbol{\beta}^\top \mathbf{z}$ gives a proportional hazards model
  - Sufficiently flexible splines give similar answers to Cox models
- Modelling any other γ_r as linear in $\mathbf{z}$ gives non-proportional hazards:
  - hazard ratio can be an arbitrarily flexible function of time

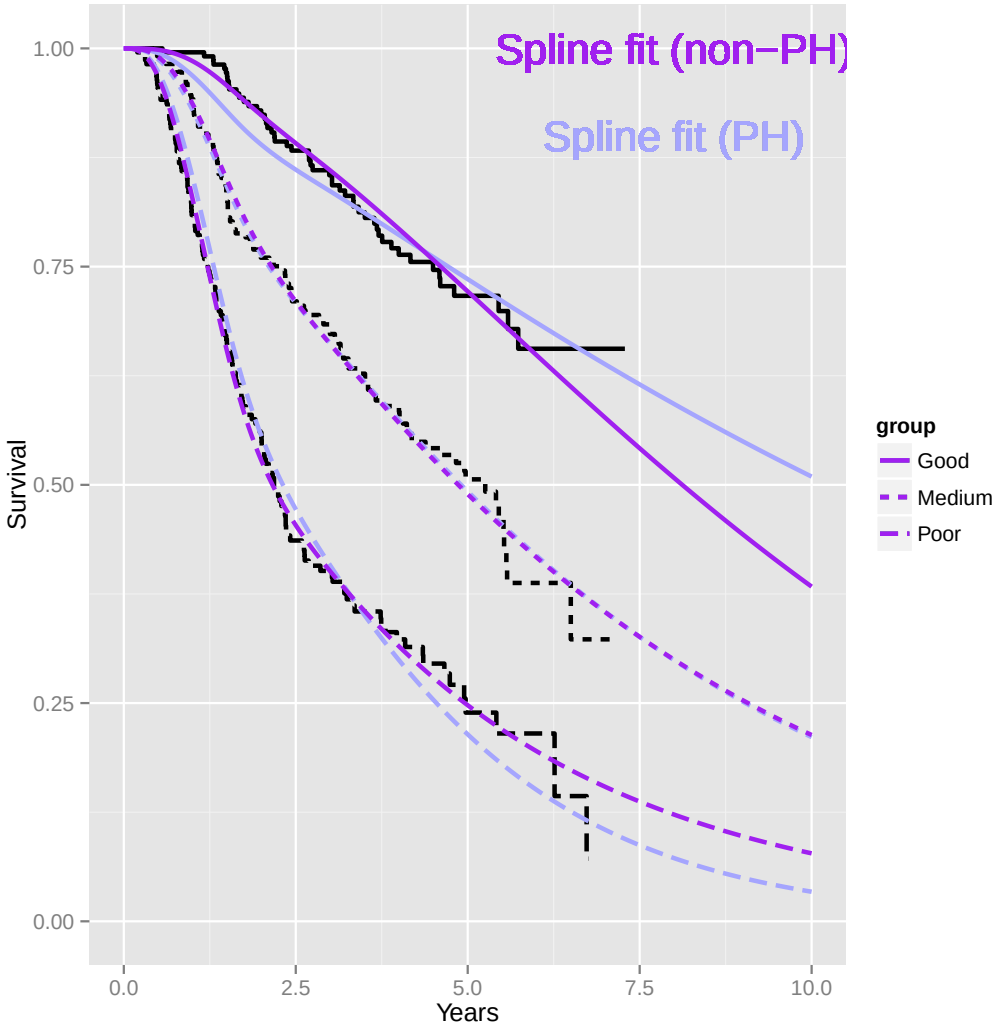
Modelling covariates, and non-proportional hazards



Relaxing the proportional hazards assumption

- improves short-term fit
- changes extrapolation

Modelling covariates, and non-proportional hazards



Relaxing the proportional hazards assumption

- improves short-term fit
- changes extrapolation

Ongoing challenges with spline models

Interconnected issues:

- Hard to combine with evidence synthesis...
- ...No easy Bayesian implementation (tricky in BUGS)...
- ...Priors for / interpretation of spline coefficients γ.

`flexsurv` R package

`flexsurv` (Jackson, CRAN (2011–2015))

- Fit any parametric survival model to individual data by maximum likelihood
  - right/interval censoring, left-truncation, multi-state models
- Standard models built in, plus spline, generalized gamma/F
- Can program your own novel parametric distributions, given the hazard or probability density function
- Any parameter of any distribution can depend on covariates
  - proportional hazards / accelerated failure time typical, but not necessary
- Customisable outputs / summary presentation

Similar Stata packages

- `stpm2` (Royston & Lambert) for spline model
- `stgenreg` (Crowther & Lambert) for user-defined parametric survival models

`flexsurv` R package

`flexsurv` (Jackson, CRAN (2011–2015))

- Fit any parametric survival model to individual data by maximum likelihood
  - right/interval censoring, left-truncation, multi-state models
- Standard models built in, plus spline, generalized gamma/F
- Can program your own novel parametric distributions, given the hazard or probability density function
- Any parameter of any distribution can depend on covariates
  - proportional hazards / accelerated failure time typical, but not necessary
- Customisable outputs / summary presentation

Similar Stata packages

- `stpm2` (Royston & Lambert) for spline model
- `stgenreg` (Crowther & Lambert) for user-defined parametric survival models

Fitting models, and default output

Generalized gamma 3-parameter model.
Two out of the three parameters depend on a covariate (prognostic group)

```
fs2 <- flexsurvreg(Surv(recyrs, censrec) ~ group + sigma(group),
                   data=bc, dist="gengamma")
fs2
```

```
  Call:
flexsurvreg(formula = Surv(recyrs, censrec) ~ group + sigma(group),
    data = bc, dist = "gengamma")

Estimates:
                   data mean  est      L95%     U95%     se       exp(est)  L95%     U95%
mu                      NA     1.9840   1.6822   2.2859   0.1540       NA        NA       NA
sigma                   NA     1.1765   0.9392   1.4738   0.1352       NA        NA       NA
Q                       NA    -0.7074  -1.1795  -0.2353   0.2409       NA        NA       NA
groupMedium         0.3338    -0.7105  -1.0346  -0.3863   0.1654   0.4914    0.3554   0.6795
groupPoor           0.3324    -1.4043  -1.7102  -1.0983   0.1561   0.2455    0.1808   0.3334
sigma(groupMedium)  0.3338    -0.0494  -0.3105   0.2118   0.1332   0.9518    0.7331   1.2358
sigma(groupPoor)    0.3324    -0.1768  -0.4293   0.0757   0.1288   0.8380    0.6510   1.0787

N = 686,  Events: 299,  Censored: 387
Total time at risk: 2113.425
Log-likelihood = -786.2172, df = 7
AIC = 1586.434
```

Fitting models, and default output

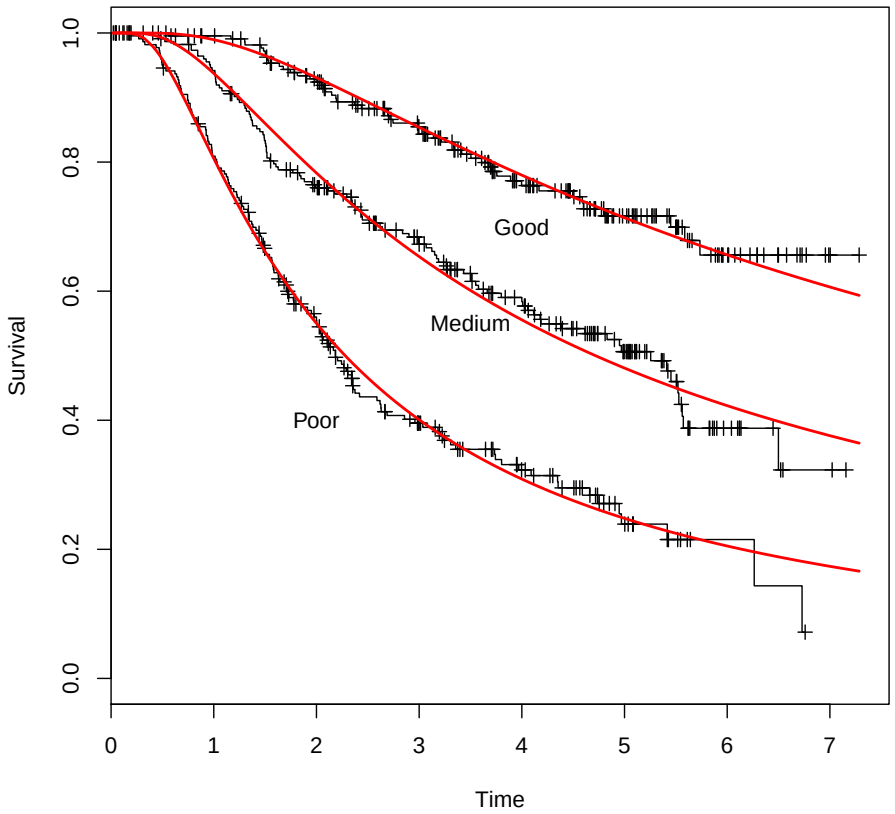
Generalized gamma 3-parameter model.
Two out of the three parameters depend on a covariate (prognostic group)

```
fs2 <- flexsurvreg(Surv(recyrs, censrec) ~ group + sigma(group),
                   data=bc, dist="gengamma")
fs2
```

```
  Call:
flexsurvreg(formula = Surv(recyrs, censrec) ~ group + sigma(group),
    data = bc, dist = "gengamma")

Estimates:
                    data mean  est      L95%     U95%     se       exp(est)  L95%     U95%
mu                       NA     1.9840   1.6822   2.2859   0.1540       NA        NA       NA
sigma                    NA     1.1765   0.9392   1.4738   0.1352       NA        NA       NA
Q                        NA    -0.7074  -1.1795  -0.2353   0.2409       NA        NA       NA
groupMedium          0.3338    -0.7105  -1.0346  -0.3863   0.1654   0.4914    0.3554   0.6795
groupPoor            0.3324    -1.4043  -1.7102  -1.0983   0.1561   0.2455    0.1808   0.3334
sigma(groupMedium)   0.3338    -0.0494  -0.3105   0.2118   0.1332   0.9518    0.7331   1.2358
sigma(groupPoor)     0.3324    -0.1768  -0.4293   0.0757   0.1288   0.8380    0.6510   1.0787

N = 686,  Events: 299,  Censored: 387
Total time at risk: 2113.425
Log-likelihood = -786.2172, df = 7
AIC = 1586.434
```

Fitting models, plot output

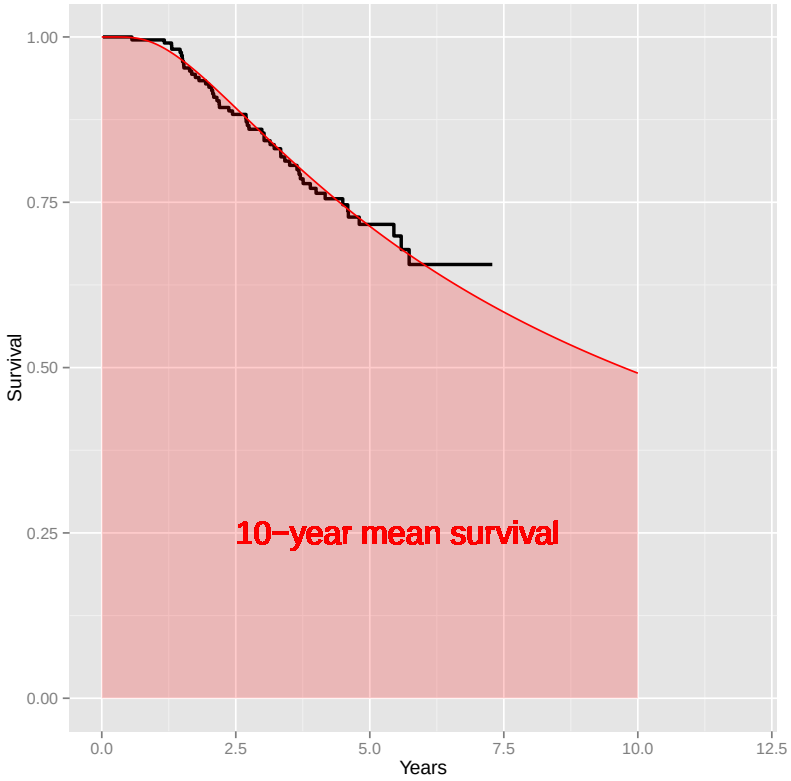
```
plot(fs2, xlim=c(0, 10))
text(x=c(2, 3.5, 4), y=c(0.4, 0.55, 0.7),
     c("Poor","Medium","Good"))
```



Summarise user-defined functions of the parameters

e.g. restricted mean survival for a model with parameters $\boldsymbol{\alpha}$

$$E(T|T < t, \boldsymbol{\alpha}) = \int_0^t S(u|\boldsymbol{\alpha})du$$

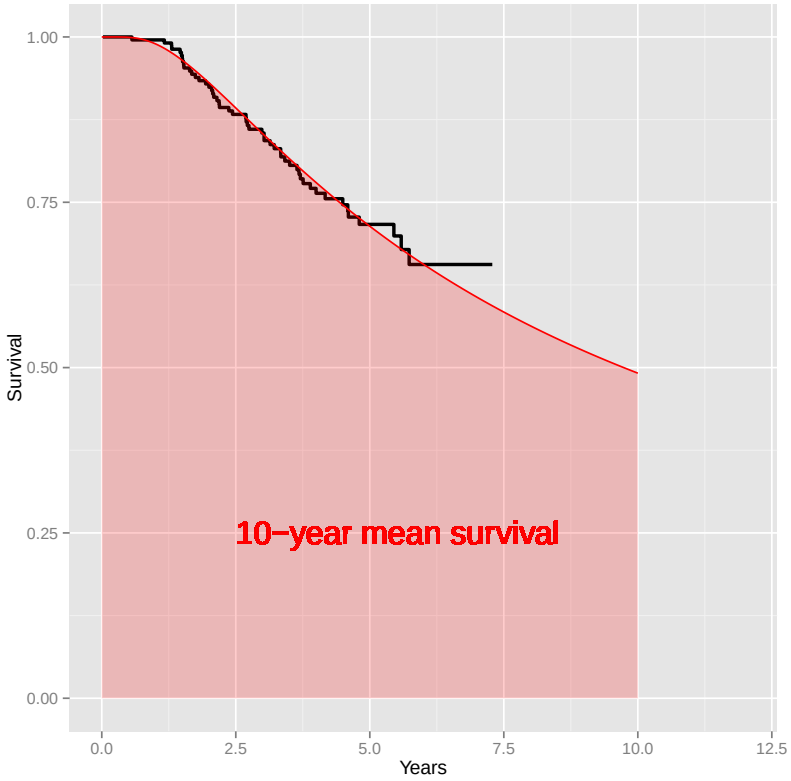



```
mean.gengamma <- function(mu, sigma, Q,
                          horizon=10, ...){
    surv <- function(t, ...) {
        1 - pgengamma(q=t, mu=mu,
                      sigma=sigma, Q=Q, ...)
    }
    integrate(surv, 0, horizon, ...)$value
}

summary(fs2, newdata=list(group="Good"),
        t=1, fn=mean.gengamma)

group=Good
  time     est      lcl     ucl
1    1 7.36671 6.725084 7.85355
```

Summarise user-defined functions of the parameters

e.g. restricted mean survival for a model with parameters $\boldsymbol{\alpha}$

$$E(T|T < t, \boldsymbol{\alpha}) = \int_0^t S(u|\boldsymbol{\alpha})du$$




```
mean.gengamma <- function(mu, sigma, Q,
                          horizon=10, ...){
    surv <- function(t, ...) {
        1 - pgengamma(q=t, mu=mu,
                      sigma=sigma, Q=Q, ...)
    }
    integrate(surv, 0, horizon, ...)$value
}

summary(fs2, newdata=list(group="Good"),
        t=1, fn=mean.gengamma)
```

```
group=Good
  time     est      lcl     ucl
1    1 7.36671 6.725084 7.85355
```

`flexsurv` is user-extensible

- Similar facility for people to fit their own parametric models
  - given definitions of the hazard or probability density as R functions
- User guide and Supplementary examples manuals give lots of worked examples
  - User guide to be published in Journal of Statistical Software
- Suggestions for additions welcome
  - ideally build a user-contributed library of examples

Limitations of `flexsurv`

- Restricted to parametric modelling of one individual-level dataset
- No random effects / frailty, though could be extended in theory

For synthesis of multiple related datasets, Bayesian probabilistic modelling languages (BUGS / JAGS / Stan) preferred

- though higher barrier to entry at the moment

In summary

Modelling long term survival may need

- a combination of data sources...
  - but modelling/software tools are tricky
  - can be hard work to assess assumptions
- flexible models ...
  - hard to combine with synthesis of multiple datasets
- accessible software

In summary

Modelling long term survival may need

- a combination of data sources...
  - but modelling/software tools are tricky
  - can be hard work to assess assumptions
- flexible models ...
  - hard to combine with synthesis of multiple datasets
- accessible software

In summary

Modelling long term survival may need

- a combination of data sources...
  - but modelling/software tools are tricky
  - can be hard work to assess assumptions
- flexible models ...
  - hard to combine with synthesis of multiple datasets
- accessible software