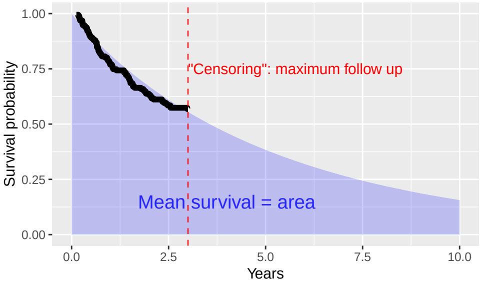
survextrap: a new tool for flexible and transparent extrapolation of survival data to inform health policy

Christopher Jackson

Royal Statistical Society conference, Harrogate, 6 September 2023

MRC Biostatistics Unit
UNIVERSITY OF CAMBRIDGE

Survival extrapolation: long-term decisions from short-term, time-to-event data



Examples:

Should health service adopt a new treatment, given 3 years follow up data from a trial?

Predicted burden on hospitals in an epidemic, given data on hospital stays, where many are people still in hospital?

Consequences of policy decisions will last longer than the end of the data

Estimating expected survival over the long term

Policy-maker typically wants to know the expected time to event

- equivalent to knowing the total outcome (e.g. survival, hospital length of stay) over the population
- Not provided by most common survival analysis tools e.g. Kaplan-Meier estimators, Cox models.

Provided by a fully-parametric distribution for the time T to the event.

- Many choices for how to specify this

Estimating expected survival over the long term

Policy-maker typically wants to know the expected time to event

- equivalent to knowing the total outcome (e.g. survival, hospital length of stay) over the population
- Not provided by most common survival analysis tools e.g. Kaplan-Meier estimators, Cox models.

Provided by a fully-parametric distribution for the time T to the event.

- Many choices for how to specify this

Estimating expected survival over the long term

Policy-maker typically wants to know the expected time to event

- equivalent to knowing the total outcome (e.g. survival, hospital length of stay) over the population
- Not provided by most common survival analysis tools e.g. Kaplan-Meier estimators, Cox models.

Provided by a fully-parametric distribution for the time T to the event.

- Many choices for how to specify this

Estimating expected survival over the long term

Policy-maker typically wants to know the expected time to event

- equivalent to knowing the total outcome (e.g. survival, hospital length of stay) over the population
- Not provided by most common survival analysis tools e.g. Kaplan-Meier estimators, Cox models.

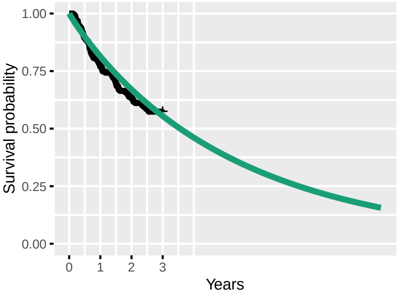
Provided by a fully-parametric distribution for the time T to the event.

- Many choices for how to specify this

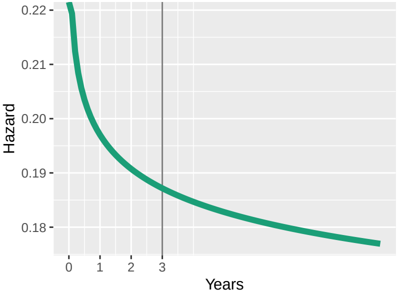
Parametric survival models: examples

Weibull distribution

Survivor function
$S(t|\lambda, \alpha) = \exp(-\lambda t^{\alpha})$



Hazard function
$h(t|\lambda, \alpha) = \lambda \alpha t^{\alpha - 1}$

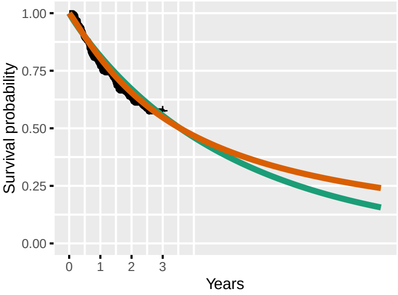


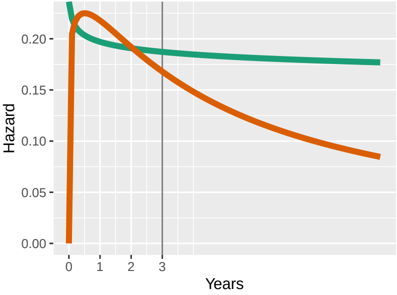
Parametric survival models: examples

Log-logistic distribution

Survivor function
$S(t|\lambda, \alpha) = 1/(1 + (t/b)^a)$

Hazard function $h(t|\lambda, \alpha) =$

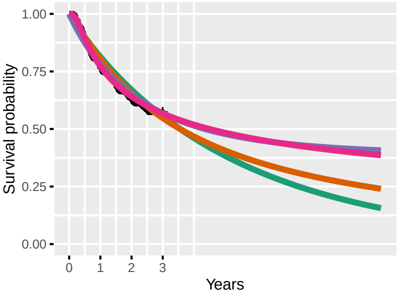


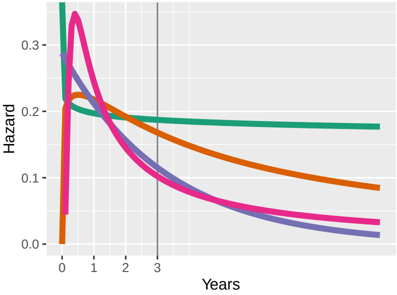


Parametric survival models: examples

Typical set of models in standard software

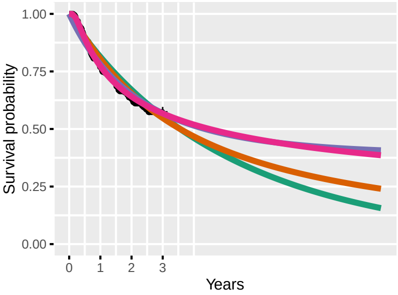
Give different extrapolations from short-term data

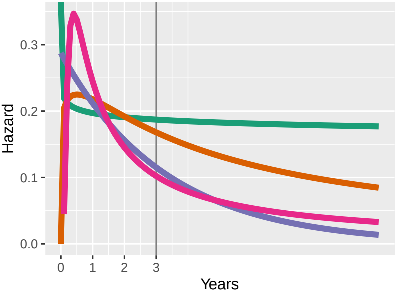




Parametric survival models: examples

Typical set of models in standard software
Give different extrapolations from short-term data





Fit a set of models and judge which :

- fits the data best in the short term? (easy, e.g. AIC)
- gives more plausible risk changes in the long term? harder

Where might information about the long term come from?

Data on general population, or disease registry

- survival of people with a specific disease cannot be better than a comparable set of people in the general population?

Clinical judgements about the mechanism

- e.g. some people get cured of the disease, so disease-specific hazard of death reduces to zero

Formally-elicited judgements about, e.g. 5-year, 10-year survival?

Want to be able to use this kind of information in a transparent and statistically-principled way

Reviews for methods for extrapolation with external data

Bullement et al. (Medical Decision Making 2023)

Medical Decision Making
1–11

DOI: 10.1177/0272989X231168618
journals.sagepub.com/home/mdm
SAGE

A Systematic Review of Methods to Incorporate External Evidence into Trial-Based Survival Extrapolations for Health Technology Assessment

Ash Bullement, Matthew D. Stevenson, Gianluca Baio, Gemma E. Shields, and Nicholas R. Latimer

Jackson et al. (Medical Decision Making 2017)

Extrapolating Survival from Randomized Trials Using External Data: A Review of Methods

Christopher Jackson, PhD, John Stevens, PhD, Shijie Ren, MPhil, PhD, Nick Latimer, PhD, MSc, Laura Bojke, PhD, MSc, Andrea Manca, PhD, MSc, Linda Sharples, PhD Show less

First Published July 10, 2016 | Research Article | Find in PubMed | Check for updates
https://doi-org.ezp.lib.cam.ac.uk/10.1177/0272989X16639900

NICE DSU (2021) `https://www.sheffield.ac.uk/nice-dsu/tsds/flexible-methods-survival`

...piecewise models, spline models, cure, relative survival, proportional and additive hazards, converging hazards, diminishing treatment effects, Bayesian methods ...

Ideal characteristics of a method / tool

1. Incorporate all available data
2. Fit the data as well as possible
3. Make any assumptions transparent
4. Quantify uncertainty
5. Be easy to use!

In particular, Bayesian evidence synthesis methods

- are comprehensive, flexible, principled
- but have needed specialised programming (BUGS / JAGS), advanced statistical expertise

Ideal characteristics of a method / tool

1. Incorporate all available data
2. Fit the data as well as possible
3. Make any assumptions transparent
4. Quantify uncertainty
5. Be easy to use!

In particular, Bayesian evidence synthesis methods

- are comprehensive, flexible, principled
- but have needed specialised programming (BUGS / JAGS), advanced statistical expertise

survextrap R package

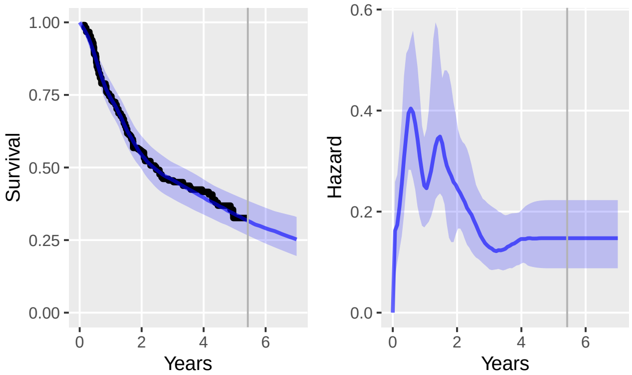
https://chjackson.github.io/survextrap
Paper at https://arxiv.org/abs/2306.03957

Package for Bayesian survival modelling

- Multiple sources of external data to aid extrapolation
- Flexible parametric spline model for the hazard
- Multiparameter evidence synthesis, MCMC estimation (Stan)
- Principle: data and influential assumptions made as transparent as possible.
  - "you say what you know, then the computer does the hard work of converting that to answers"!

Rapid tour...

Spline-based, flexible model for the hazard



Running example: head and neck cancer trial data from Guyot et al. 2017

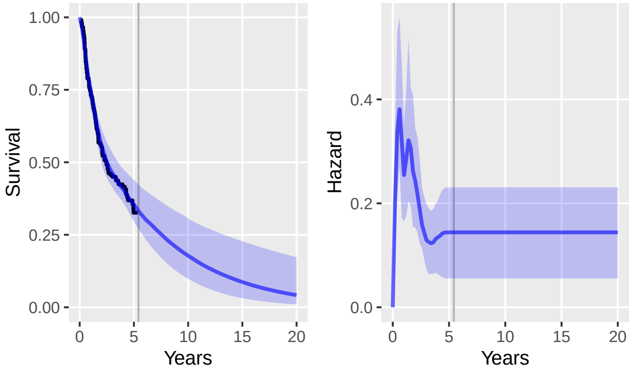
Default model: M-spline with 11 parameters

- penalized through a hierachical prior to control overfitting

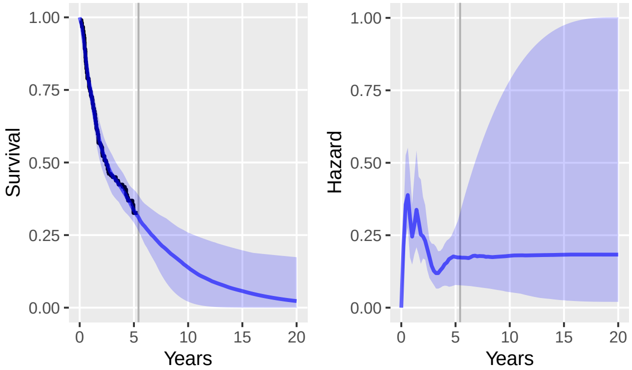
Flexible hazard within data, constant beyond data

```
mod1 <- survextrap(Surv(years, d) ~ 1, data=dat)
```

Spline-based, flexible model for the hazard



Spline-based, flexible model for the hazard



Allow hazard to vary up to 20 years, by adding a spline "knot".
Assumes only that the hazard is smooth

- short-term extrapolation influenced weakly by latest data.

Posterior distribution represents uncertainty due to lack of data

```
mod2 <- survextrap(Surv(years, d) ~ 1, add_knots=20)
```

Use of external data

Supplied as data frame of aggregate counts of survivors over a series of intervals. Example:

Follow-up period		Number		Treatment
From t	To u	Alive at t	Still alive at u	
5	6	358	325	Control
6	7	308	285	Control
8	9	221	198	Control
etc.				

- Registry data, population data, and elicited judgements can often be expressed in this form
- Each count of survivors is a binomal outcome, with probability defined by the spline model and covariates
- Bayesian evidence synthesis: posterior for parameters determined given individual and external data together

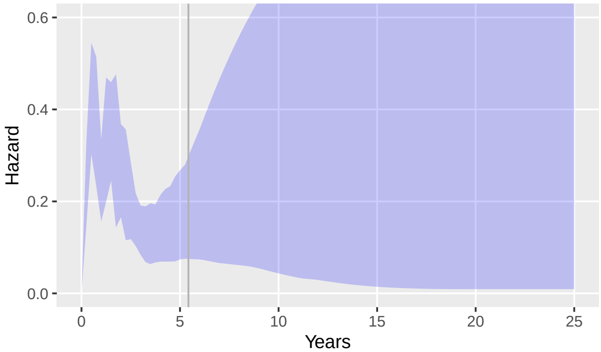
Use of external data

Supplied as data frame of aggregate counts of survivors over a series of intervals. Example:

Follow-up period		Number		Treatment
From t	To u	Alive at t	Still alive at u	
5	6	358	325	Control
6	7	308	285	Control
8	9	221	198	Control
etc.				

- Registry data, population data, and elicited judgements can often be expressed in this form
- Each count of survivors is a binomal outcome, with probability defined by the spline model and covariates
- Bayesian evidence synthesis: posterior for parameters determined given individual and external data together

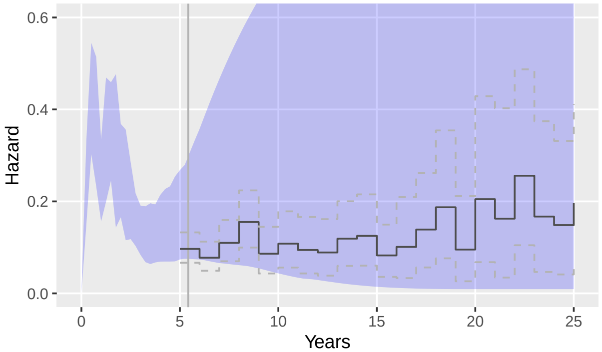
Survival modelling with external data: example



With no external data, and just a smoothness assumption, estimated long-term hazard is extremely uncertain

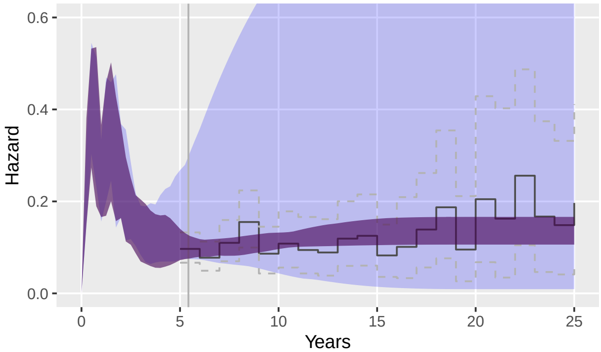
```
mod_seer <- survextrap(Surv(years, d) ~ 1, data=dat,
                       add_knots=20)
```

Survival modelling with external data: example



Registry data give annual survival rates from 5 to 25 years.

Survival modelling with external data: example

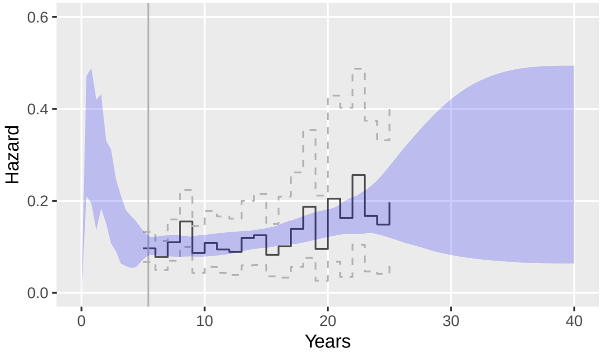


Registry data give annual survival rates from 5 to 25 years.

- Enables a confident "extrapolation" of the hazard for control group survival

```
mod_seer <- survextrap(Surv(years, d) ~ 1, data=dat,
                         external=ext_dat, add_knots=20)
```

Survival modelling with external data: example



Extrapolating up to 40 years now...

```
mod_seer <- survextrap(Surv(years, d) ~ 1, data=dat,
                       external=ext_dat,
                       add_knots=c(20,30,40))
```

Additive hazards model

Can supply a known background hazard $h_b(t)$

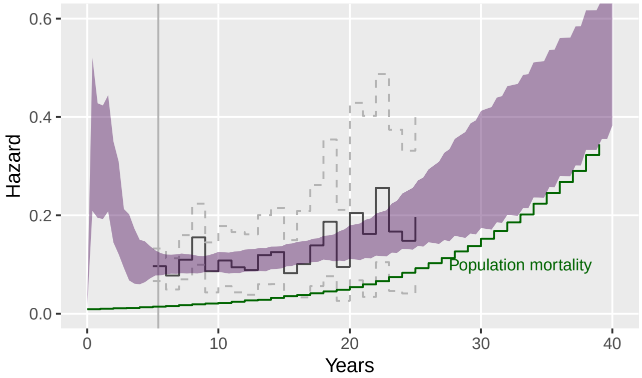
- e.g. national mortality statistics for people of comparable age/sex

Hazard for trial patients modelled as additive:

$$h(t) = h_b(t) + h_e(t)$$

i.e. no lower than population hazard

Additive hazards model




```
mod_seer <- survextrap(Surv(years, d) ~ 1, data=dat,
                       external=ext_dat, bh=back_dat,
                       add_knots=c(20,30,40))
```

Including expert judgements on long-term survival

Suppose we judge by 40 years, survival will be similar to general population (though no evidence of this here!).

We elicit a conditional annual survival of 0.72

- taken from population data
- with 95% credible interval of 0.69–0.75

Equivalent information to having observed 724 survivors out of 1000 people

- Bayesian principles: "conjugacy" of Beta and Binomial

Implement by appending these counts to the "external" data

Including expert judgements on long-term survival

Suppose we judge by 40 years, survival will be similar to general population (though no evidence of this here!).

We elicit a conditional annual survival of 0.72

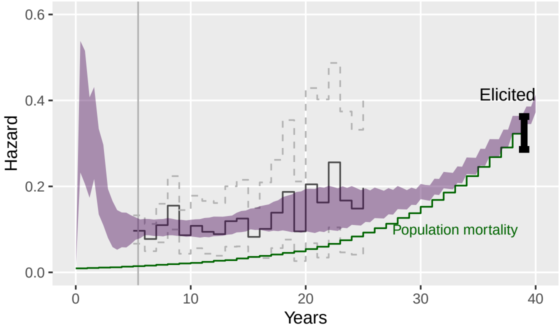
- taken from population data
- with 95% credible interval of 0.69–0.75

Equivalent information to having observed 724 survivors out of 1000 people

- Bayesian principles: "conjugacy" of Beta and Binomial

Implement by appending these counts to the "external" data

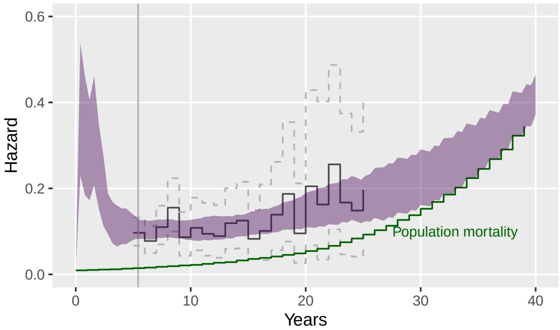
Model including long term expert judgement



Full evidence synthesis: trial, registry, population and elicited data all combined

```
mod_seer <- survextrap(Surv(years, d) ~ 1, data=dat,
                       external=ext_elic_dat,
                       add_knots=c(20,30,40))
```

Model including long term expert judgement



Mixture cure: parametric model decreasing to zero excess hazard
Learnt from data up to 25 years, extrapolated to 40

```
mod_seer <- survextrap(Surv(years, d) ~ 1, data=dat,
                       external=ext_dat, cure=TRUE,
                       add_knots=c(20,30,40))
```

Treatment or covariate effects

Proportional hazards model

```
mod_seer <- survextrap(Surv(years, d) ~ treat + age,
                       data=dat, external=ext_dat)
```

Flexible non-proportional hazards model

```
mod_seer <- survextrap(Surv(years, d) ~ treat + age,
                       data=dat, external=ext_dat,
                       nonprop=TRUE)
```

Additionally can impose a waning effect when predicting from models

- log hazard ratio for a treatment decreases from modelled value to zero over some time period

Treatment or covariate effects

Proportional hazards model

```
mod_seer <- survextrap(Surv(years, d) ~ treat + age,
                       data=dat, external=ext_dat)
```

Flexible non-proportional hazards model

```
mod_seer <- survextrap(Surv(years, d) ~ treat + age,
                       data=dat, external=ext_dat,
                       nonprop=TRUE)
```

Additionally can impose a waning effect when predicting from models

- log hazard ratio for a treatment decreases from modelled value to zero over some time period

Treatment or covariate effects

Proportional hazards model

```
mod_seer <- survextrap(Surv(years, d) ~ treat + age,
                       data=dat, external=ext_dat)
```

Flexible non-proportional hazards model

```
mod_seer <- survextrap(Surv(years, d) ~ treat + age,
                       data=dat, external=ext_dat,
                       nonprop=TRUE)
```

Additionally can impose a waning effect when predicting from models

- log hazard ratio for a treatment decreases from modelled value to zero over some time period

Final discussion

Further info in the paper, full doc and worked examples on
https://chjackson.github.io/survextrap

Software status: "beta": not expected to change drastically before "full release" (CRAN), but more testing planned

Many details I haven't mentioned today

- choice of prior distributions
- how the spline is built: knots, model comparison
- "post-estimation" outputs e.g. restricted mean survival

Ongoing further work

- Simulation studies to assess if defaults are sensible
- Listen to users. Usable? Understandable? What is most challenging?

Final discussion

Further info in the paper, full doc and worked examples on
https://chjackson.github.io/survextrap

Software status: "beta": not expected to change drastically before "full release" (CRAN), but more testing planned

Many details I haven't mentioned today

- choice of prior distributions
- how the spline is built: knots, model comparison
- "post-estimation" outputs e.g. restricted mean survival

Ongoing further work

- Simulation studies to assess if defaults are sensible
- Listen to users. Usable? Understandable? What is most challenging?

Final discussion

Further info in the paper, full doc and worked examples on
https://chjackson.github.io/survextrap

Software status: "beta": not expected to change drastically before "full release" (CRAN), but more testing planned

Many details I haven't mentioned today

- choice of prior distributions
- how the spline is built: knots, model comparison
- "post-estimation" outputs e.g. restricted mean survival

Ongoing further work

- Simulation studies to assess if defaults are sensible
- Listen to users. Usable? Understandable? What is most challenging?

MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Thank you