Standardised survival probabilities: a useful and informative tool for reporting regression models for survival data

Elisavet Syriopoulou

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet

Email: elisavet.syriopoulou@ki.se
26th April 2024

What to report after fitting a survival model?

- Survival probabilities under different exposure groups are frequently compared using the Kaplan–Meier estimator.
- To adjust for confounding, we often fit survival models e.g Cox model, flexible parametric survival models.
- After fitting a survival model the analysis is often summarised using the hazard ratio (HR).
- Many authors have previously argued about limitations related to the use of HRs.
- I will focus on an example of relapse-free survival (measured as time from primary surgery to relapse or death, whichever occurred first) of breast cancer patients.

Useful references

- My presentation today is a summary of a tutorial paper:

  Syriopoulou E, Wästerlid T, Lambert PC, Andersson TM-L. Standardised survival probabilities: a useful and informative tool for reporting regression models for survival data. Br J Cancer 127, 1808–1815 (2022), 10.1038/s41416-022-01949-6

- The published paper includes code in Stata.
- Recently, Alessandro Gasparini translated the code into R:

  https://github.com/ellessenne/standsurv-tutorial-r

- I will focus on flexible parametric survival models but it is also possible to obtain standardised survival curves after Cox models with the {stdReg} package.
- There is also a `standsurv` function in flexsurv:

  https://CRAN.R-project.org/package=flexsurv

Example - Kaplan Meier survival curves
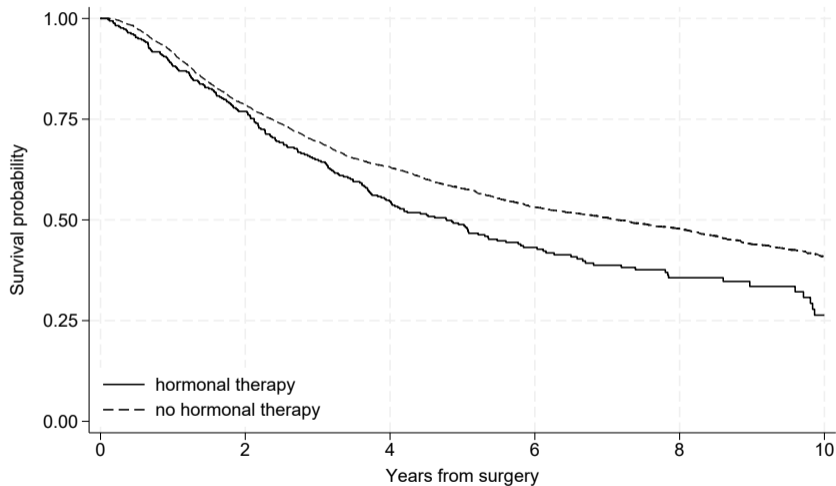
Example - descriptives

Table 1. Summary characteristics of breast cancer patients by treatment arm.

Variables	Hormonal therapy: no N = 2643	Hormonal therapy: yes N = 339
Age at surgery (years)	53 (44–64)	62 (57–69)
Number of positive nodes	0 (0–3)	4 (2–9)
Progesterone level (fmol/l)	46 (5–208)	19 (1–117)
Differentiation grade		
2	735 (28%)	59 (17%)
3	1908 (72%)	280 (83%)
Tumour size		
<=20 mm	1283 (49%)	104 (31%)
>20–50 mm	1119 (42%)	172 (51%)
>50 mm	241 (9%)	63 (18%)

For categorical variables, the number of individuals with relevant proportions is given.
For continuous variables, median with 25th and 75th percentiles are given.

Patients who received hormonal therapy were older, had a higher number of positive nodes and there was a larger proportion of patients with a tumour above 50 mm.

Flexible parametric models (Royston-Parmar)

- Explicitly estimate the baseline log-cumulative hazard using restricted cubic splines for the logarithm of time rather than assuming linearity with time.
- It allows a wide range of hazard functions to be captured.
- It can be written mathematically as:

$$\ln(\Lambda(t)) = s(\ln(t)|\boldsymbol{\gamma}, \boldsymbol{k_0}) + \mathsf{X}\beta$$

- Need to specify the number of knots used to create the spline function (or number of degrees of freedom, df, which is equal to the number of knots minus one).
- R packages: {rstpm2} and {flexsurv} (among others).

Royston P, Lambert P. Flexible parametric survival analysis in stata: beyond the Cox model. College Station, TX: Stata Press; 2011.

Example - Fitting a survival model

```
stpm2(Surv(exit, failure == 1) ~ hormon + size2 + size3 + grade + enodes + pr_1 + age,
    df = 4, data = rott2)

# Call:
# stpm2(formula = Surv(exit, failure == 1) ~ hormon + size2 + size3 +
#     grade + enodes + pr_1 + age, data = rott2, df = 4)
#
# Coefficients:
#                          Estimate Std. Error  z value     Pr(z)
# (Intercept)            -7.1206808  0.4142418 -17.1897 < 2.2e-16 ***
# hormon                 -0.2745180  0.0789263  -3.4782 0.0005049 ***
# size2                   0.2995659  0.0565496   5.2974 1.175e-07 ***
# size3                   0.5031879  0.0845688   5.9500 2.681e-09 ***
# grade                   0.3086666  0.0619439   4.9830 6.260e-07 ***
# enodes                 -1.7531807  0.0934000 -18.7707 < 2.2e-16 ***
# pr_1                   -0.0280997  0.0113472  -2.4764 0.0132728 *
# age                     0.0033642  0.0019880   1.6923 0.0905919 .
# nsx(log(exit), df = 4)1  6.2255334  0.3410950  18.2516 < 2.2e-16 ***
# nsx(log(exit), df = 4)2  5.3544108  0.2302276  23.2570 < 2.2e-16 ***
# nsx(log(exit), df = 4)3 11.2727237  0.6808929  16.5558 < 2.2e-16 ***
# nsx(log(exit), df = 4)4  4.7519474  0.1464113  32.4562 < 2.2e-16 ***
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
# -2 log L: 10024.43
```

Hazard ratios

- The interpretation of HRs remains challenging as HRs are often misinterpreted as relative risks.
  - The relative risk is the ratio of the probability of experiencing the event by a specific time for the exposed to the probability for the unexposed.
  - The relative risk is always a function of time.
  - HRs should be interpreted as relative rates and not relative risks!
- Studies often report a single HR estimate for the whole study follow-up (i.e. assuming proportional hazards).
  - Often an unrealistic assumption and the HR will vary over time.

Hazard ratios

- The HR is a relative measure and provides no information on whether this effect is clinically meaningful.
  - The corresponding difference in survival probabilities might be very small and not important from a clinical point of view.
  - Absolute measures such as the difference in survival probabilities can be more informative than relative measures.
- HRs are estimated based on individuals who have survived up to a particular time.
  - As time increases, the characteristics of individuals who are still in follow-up in each exposure group might differ, resulting in an imbalanced comparison between exposures.
  - This is true even if we have sufficiently adjusted for confounding at the start of follow-up.
  - This is often referred to as built-in selection bias of HRs.

Adjusted survival curves

- A more informative way to summarise the exposure effect is to use adjusted survival probabilities.
- It can be obtained using standard statistical software.
- After fitting a model, adjusted survival probabilities are often estimated using the average covariate value for the adjusting covariates.
- Only one survival curve is estimated for each exposure level based on the average values of the adjusting covariates.

Adjusted survival curves in R

First, we create a new dataset with average values for the model covariates:

```
nd <- summarise(rott2,
         across(.cols = c("age", "grade", "size2", "size3", "enodes", "pr_1"),
         .fns = mean))
```

Then, add 100 time points between 0 and 10 years

```
nd <- crossing(nd, exit = seq(0, 10, length.out = 100))
```

Estimate adjusted survival probability for the event of relapse/death:

```
nd <- crossing(nd, hormon = c(0, 1))
nd <- cbind(nd, predict(fpm, newdata = nd, type = "surv", se.fit = TRUE))
# Adjusted relapse-free survival probability estimates for breast cancer patients at 10 years
filter(nd, exit == 10)
```

Example - adjusted survival curves plot
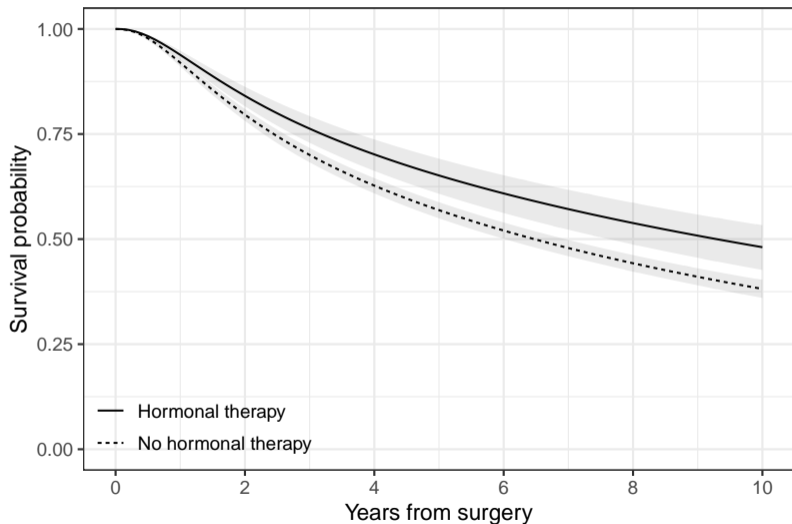
Adjusted survival curves

- The survival probability of an "average" individual if this "average" individual received hormonal therapy and an "average" individual who did not receive it.
- A caveat with "naively" adjusted survival curves is the need to calculate an "average" for included variables.
- For continuous variables, the "average" individual (mean value) might be easy to interpret.
- For categorical variables, such as sex, it is not clear what an average individual is. (e.g. if 40% of the individual were females it will be equal to 0.4) and has no meaning on an individual level (as it does not correspond to either females or males).
- Alternatively, obtain survival curves at fixed values for the adjusting variables - restricted to a specific covariate pattern.

Standardised survival curves

- Another way to overcome the need to estimate adjusted survival probabilities for an "average" individual is to obtain so-called standardised survival probabilities.

- This is done as follows:

    1. Fit a survival model (like the flexible parametric model above)
    2. Estimate individual-specific survival probabilities for each individual given the individual's covariate pattern and if they were exposed (e.g. received hormonal treatment).
    3. Then, the individual-specific survival probabilities are averaged to obtain the standardised survival probability under exposure.
    4. Repeat steps 2-3, if each individual was unexposed (e.g. had received no hormonal treatment).

Standardised survival curves

- We do not calculate an average over those who were exposed and an average over those who were unexposed.
- This would result on comparing two groups with very different covariate distributions.
- For a study population of N individuals, N estimates of individual-specific survival probabilities are obtained and then averaged to obtain the standardised survival curve in the whole population under each exposure level.

Standardised survival curves in R

We need to predict for hormon=0 and hormon=1, for all time points.
We start by predicting at 10 years, which we later generalise.

```
pred_h1 <- mean(predict(fpm, newdata = mutate(rott2, hormon = 1, exit = 10), type = "surv"))
pred_h0 <- mean(predict(fpm, newdata = mutate(rott2, hormon = 0, exit = 10), type = "surv"))
cbind(pred_h1 = pred_h1, pred_h0 = pred_h0, diff = pred_h1 - pred_h0)
```

Now, we define a function that can do that for us, for any number of
timepoints. This is specific to this application, so will have to be
adjusted/redefined for other prediction settings.

Standardised survival curves in R - cont

```
standsurv <- function(fit, newdata, timevar, hormon) {
  nd <- newdata
  nd$hormon <- hormon
  nd <- lapply(X = timevar, FUN = function(x) {
    this <- nd
    this$exit <- x
    this
  })
  nd <- do.call(rbind.data.frame, nd)
  nd$s <- predict(fit, newdata = nd, type = "surv")
  nd <- group_by(nd, exit) |>
    summarise(s = mean(s)) |>
    pull(s)
  return(nd)
}
```

Then, we use the function above to create predictions for time 0 to 10:

```
timevar <- seq(0, 10, length.out = 100)
pred_h0 <- standsurv(fit = fpm, newdata = rott2, timevar = timevar, hormon = 0)
pred_h1 <- standsurv(fit = fpm, newdata = rott2, timevar = timevar, hormon = 1)
```

Example - Standardised survival curves plot
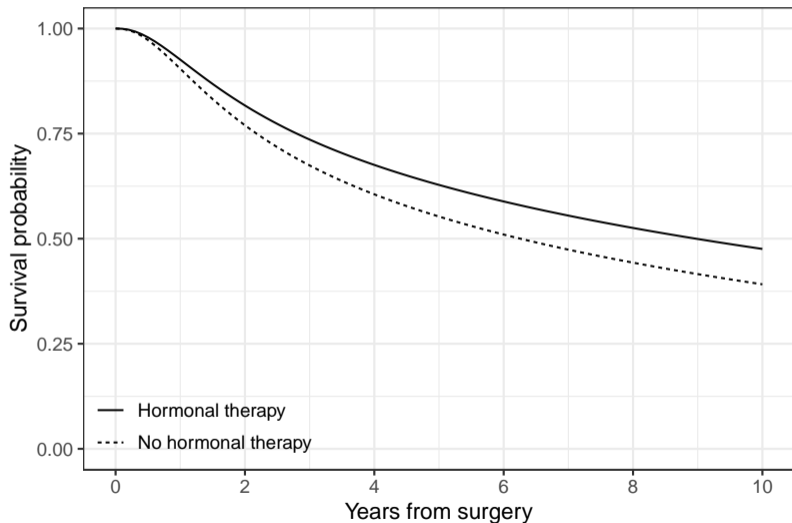
Standardised survival curves

- Interpretation: the average survival probabilities if everyone in the study population had received hormonal treatment or if everyone in the study population had not received treatment.

- Since the distribution of all other adjusting covariates are the same in the two standardised probabilities, fairer comparisons between exposed and unexposed can be made.

- Differences in the two survival curves are not due to differences in e.g. differences in tumour size.

Difference in standardised survival

Define a function for predicting the difference:

```
standsurv_diff <- function(value, ref, ...) {
  sv <- standsurv(..., hormon = value)
  sr <- standsurv(..., hormon = ref)
  return(sv - sr)
}

pred_diff <- standsurv_diff(value = 1, ref = 0, fit = fpm, newdata = rott2, timevar = timevar)
```

Example - Difference in standardised survival curves
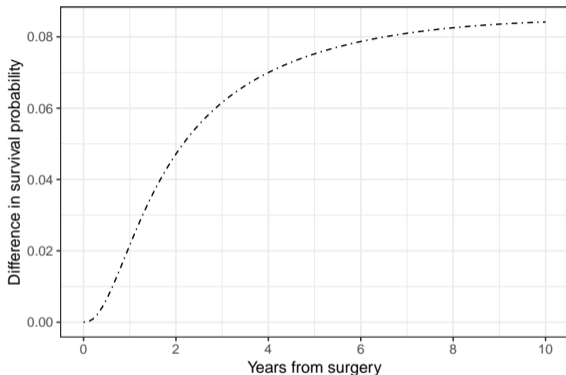
This can now be interpreted as risk (e.g., difference in risk of experiencing the event by a specific time under treatment in comparison to no treatment).

Estimating confidence intervals

The predictions above do not include confidence intervals, as we did not calculate them. For this, we can use the predictnl() function from the rstpm2 package. This uses the numerical delta method:

```
pred_h0_se <- predictnl(fpm, newdata = rott2, fun = standsurv, timevar = timevar, hormon = 0)
pred_h0_se <- cbind(pred_h0_se, confint(pred_h0_se))

pred_h1_se <- predictnl(fpm, newdata = rott2, fun = standsurv, timevar = timevar, hormon = 1)
pred_h1_se <- cbind(pred_h1_se, confint(pred_h1_se))

pred_diff_se <- predictnl(fpm, newdata = rott2, fun = standsurv_diff, value = 1, ref = 0,
    timevar = timevar)
pred_diff_se <- cbind(pred_diff_se, confint(pred_diff_se))
```

Example - with confidence intervals
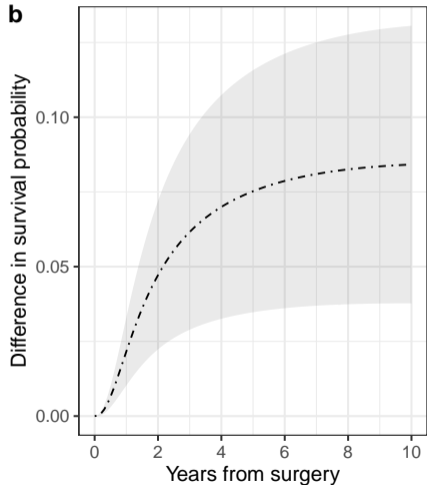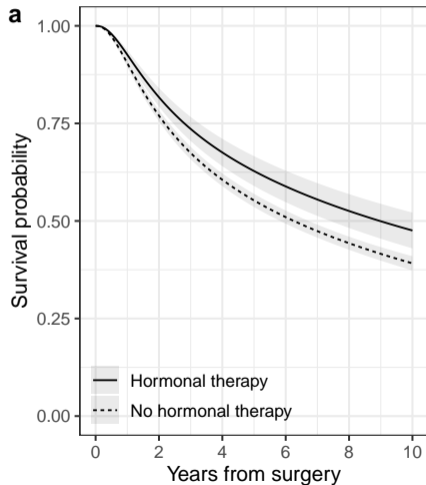
Ratio of standardised survival probabilities

```
standsurv_ratio <- function(value, ref, ...) {
  sv <- standsurv(..., hormon = value)
  sr <- standsurv(..., hormon = ref)
  return(sv / sr)
}
pred_ratio_se <- predictnl(
    fpm,
    newdata = rott2,
    fun = standsurv_ratio,
    value = 1,
    ref = 0,
    timevar = timevar
)
pred_ratio_se <- cbind(pred_ratio_se, confint(pred_ratio_se))
```

Example - Ratio of standardised survival probabilities
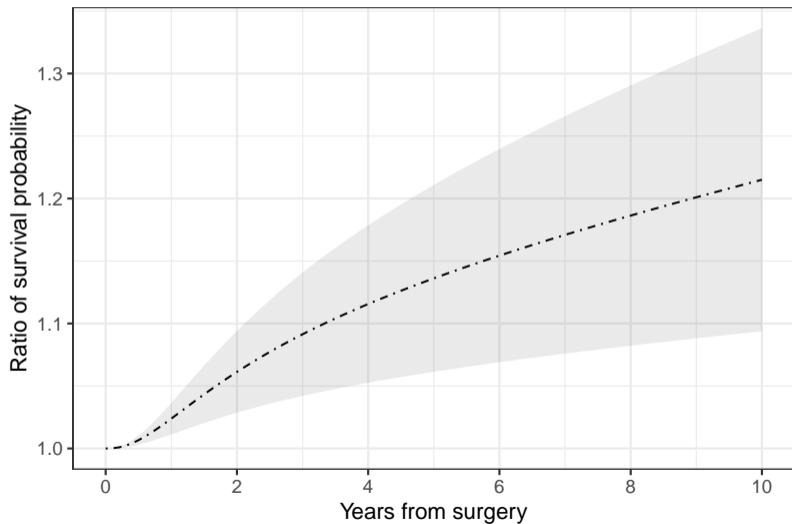
# Standardising within a subset of the population

- Before we used the empirical covariate distribution in the whole population, i.e., we estimated the average survival probability for the whole population if everyone was compared to if no one was treated.

- It may be more relevant to apply the empirical covariate distribution of a subset of the total study population, such as the covariate distribution among the treated.

- For instance, how large was the improvement in the probability of being alive with no relapse for the breast cancer patients who actually received hormonal therapy?

Standardising within subsets in R

We standardise to the subset of study subjects with hormon=1:

```
predsub_h0_se <- predictnl(fpm, newdata = filter(rott2, hormon == 1), fun = standsurv,
    timevar = timevar, hormon = 0)
predsub_h0_se <- cbind(predsub_h0_se, confint(predsub_h0_se))

predsub_h1_se <- predictnl(fpm, newdata = filter(rott2, hormon == 1), fun = standsurv,
    timevar = timevar, hormon = 1)
predsub_h1_se <- cbind(predsub_h1_se, confint(predsub_h1_se))

predsub_diff_se <- predictnl(fpm, newdata = filter(rott2, hormon == 1), fun = standsurv_diff,
    value = 1, ref = 0, timevar = timevar)
predsub_diff_se <- cbind(predsub_diff_se, confint(predsub_diff_se))
```

Example - standardised among those treated
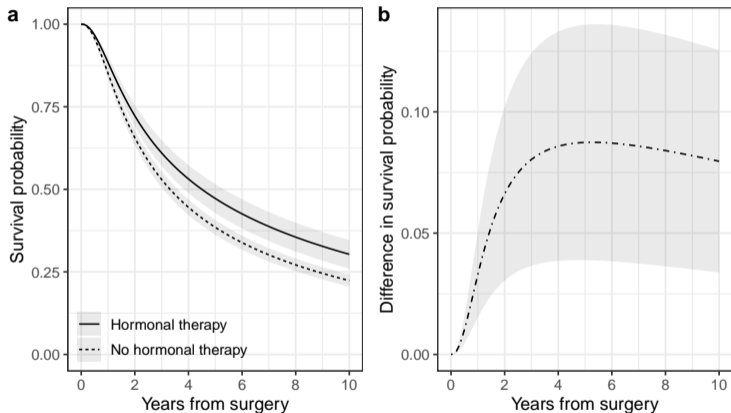
Example - over the whole population vs only treated
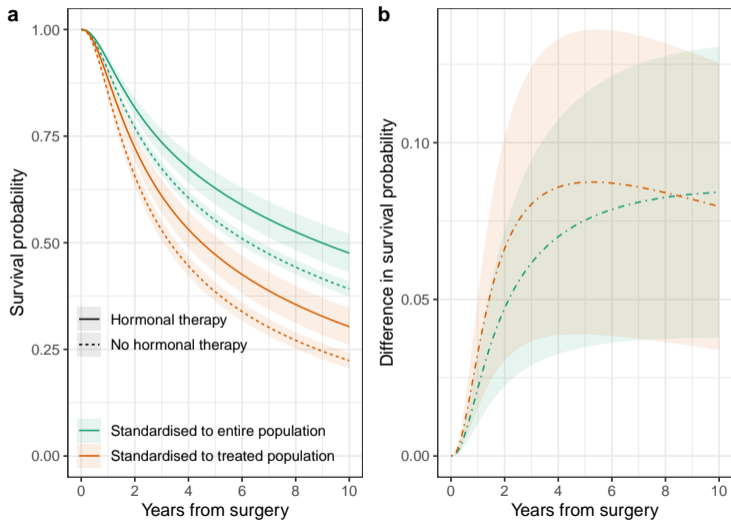
Comments

- For the standardised survival curves (among treated) we used the same population to obtain estimates under treatment and under no treatment.

- However, this time only the covariate distribution of a specific subset (treated) is used for the standardisation.

- The standardised survival probabilities within the treated group is lower than the survival within the total population.
  - This is expected as the patients who had hormonal therapy were older, had a higher number of positive nodes, and there was a larger proportion of tumours above 50 mm in comparison to patients who did not receive hormonal therapy.

Conclusions

- Survival probabilities have a less challenging interpretation in comparison to hazards and hazard ratios.
- The difference in standardised survival probabilities provide an estimate of the absolute *risk* with/without treatment.
- Comparing the effect of treatment on survival using absolute rather than relative measures makes it easier to understand whether treatment results in clinically meaningful improvements in survival.

Conclusions

- In some situations, it may be more relevant to apply the empirical covariate distribution of a subset of the population:
    - If the interest is to assess the potential impact of an intervention on a population who have not yet received it, it would be more relevant to apply the empirical distribution among the untreated.
    - For instance, what would the improvement be in the probability of being alive and having no relapse if the untreated group had hormonal therapy?