Survival Analysis
Basic concepts

Statistisk analyse af overlevelsesdata

Maj-Britt Jensen, DBCG

In many studies of breast cancer, the main outcome is the time to an event of interest,

e.g. time to death

time to recurrence (disease-free survival)

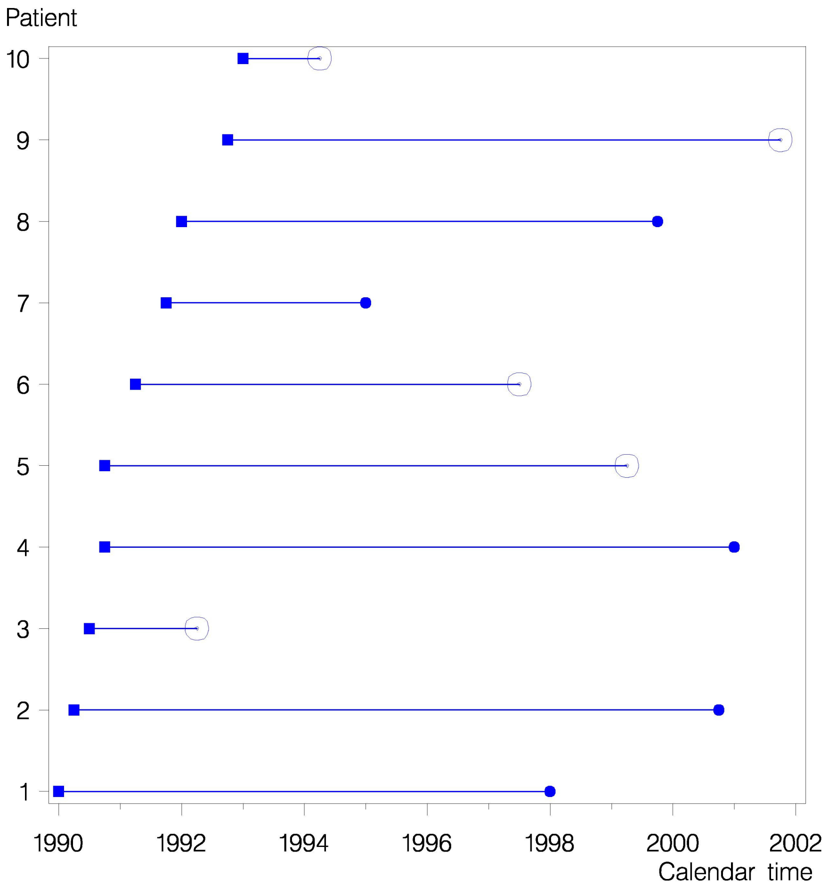
Usually for some individuals true time to event is unknown, and also survival data rarely normally distributed, therefore survival analysis necessary

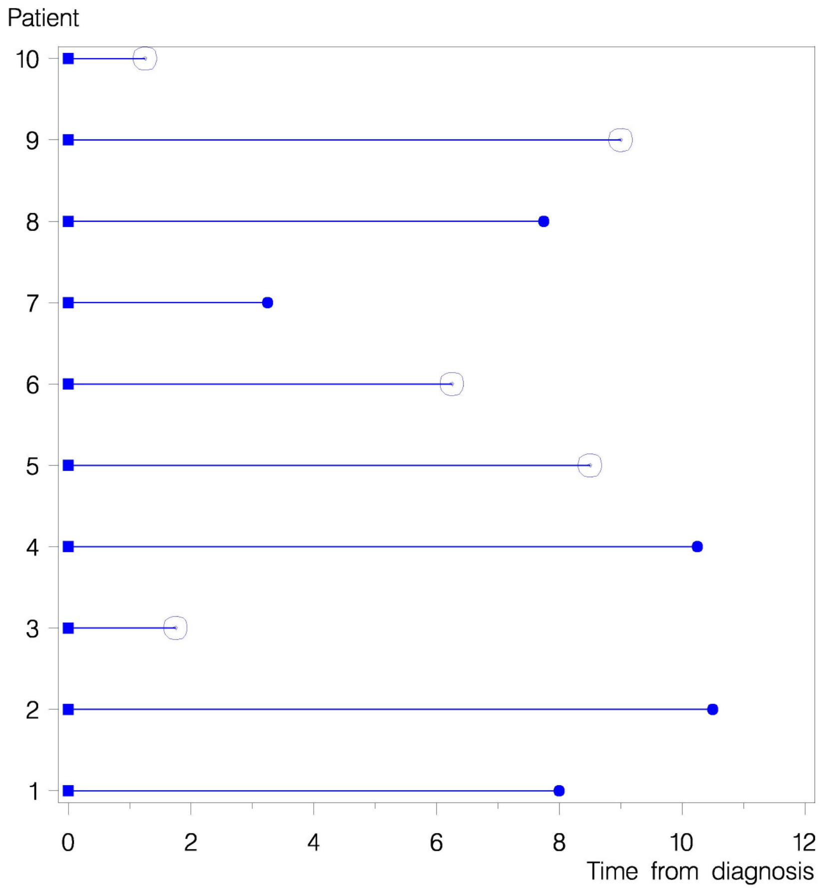
Survival analysis is a collection of statistical procedures for data analysis where the outcome variable of interest is time until an event occurs

Survival times unknown for a subset of the study group: censoring

e.g. a patient has not (yet) experienced the relevant outcome

a patient is lost to follow-up


Patient
10
9
8
7
6
5
4
3
2
1
1990
1992
1994
1996
1998
2000
2002
Calendar time


Patient
10
9
8
7
6
5
4
3
2
1
0
2
4
6
8
10
12
Time from diagnosis

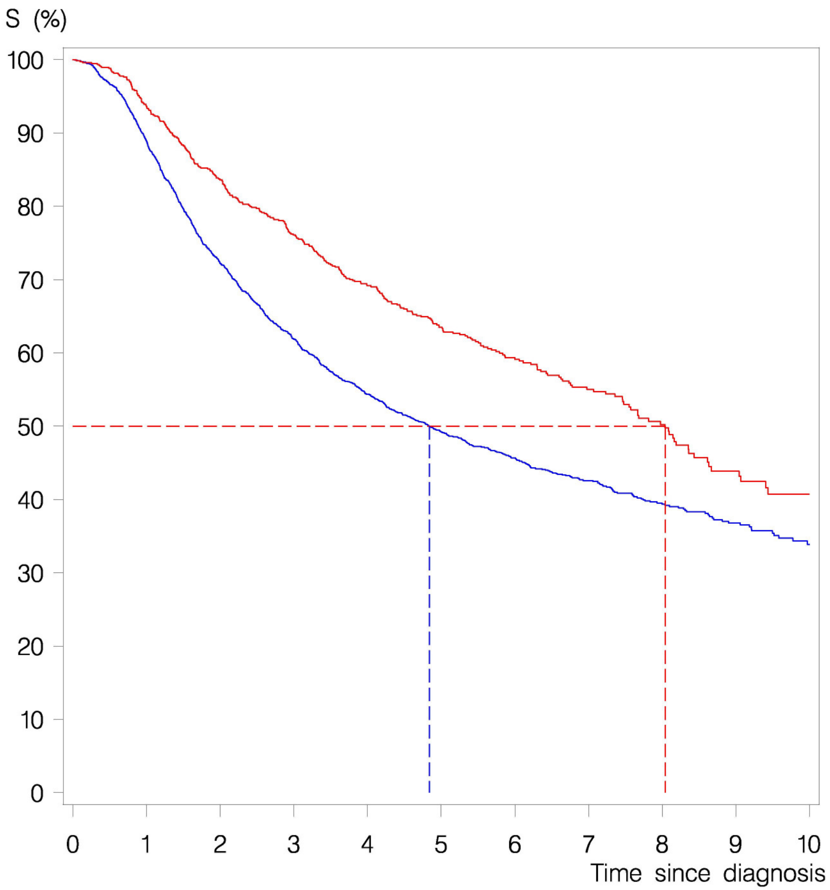
Survival data are generally described in terms of two related probabilities; *survival* $S(t)$ and *hazard* $h(t)$ or $\lambda(t)$

Hazard is the probability that an individual who is under observation at time t has an event at that time

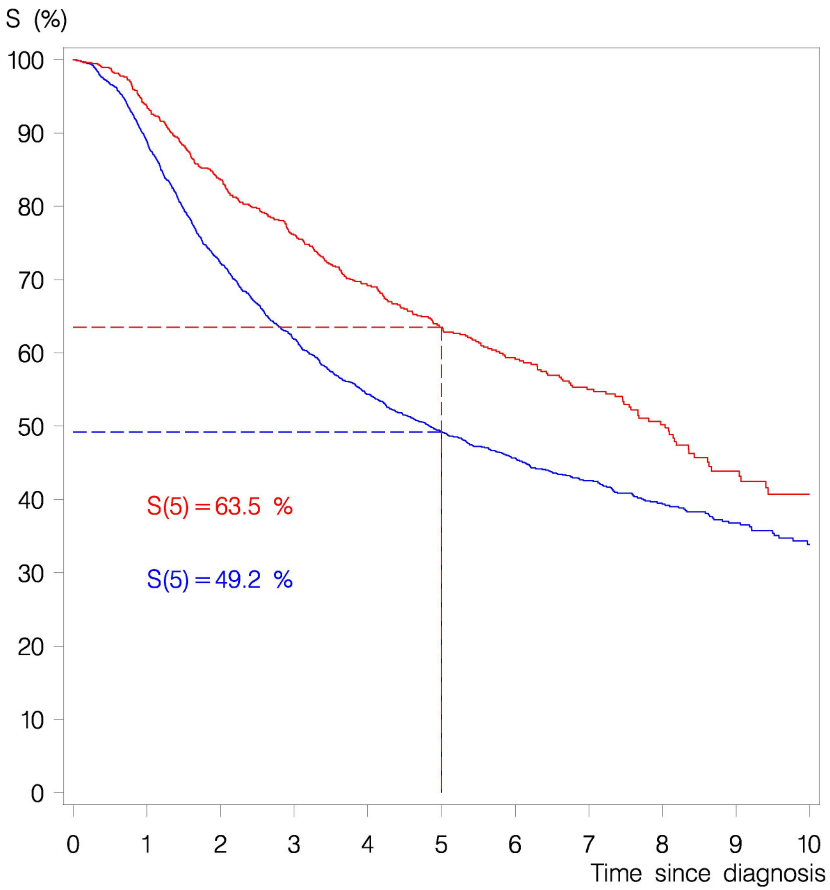
Survival is the probability that an individual survives from the time origin to a specified time t.

Kaplan-Meier survival curve

- a useful summary



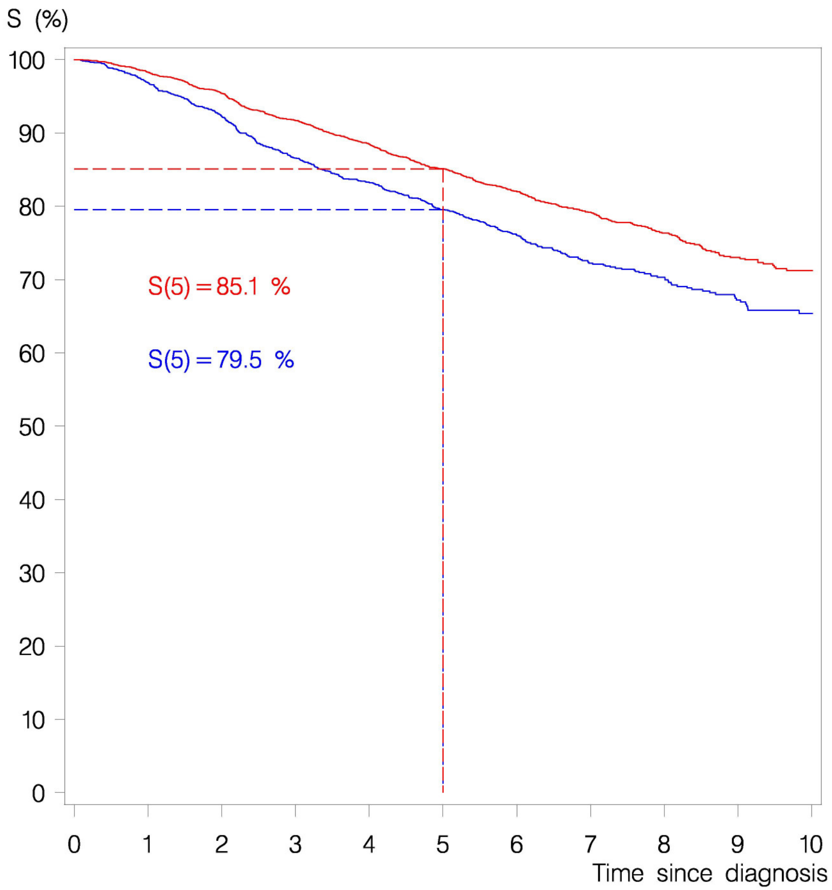
Median survival time
4.8 yrs vs 8.0 yrs



ARR = Absolute Risk Reduction
(100%- 49.2%)-(100%-63.5%)
= 14.3%

RR = Relative Risk at 5 yrs
(100-63.5)/(100-49.2) = 0.72

i.e.
the relative risk reduction is 28%

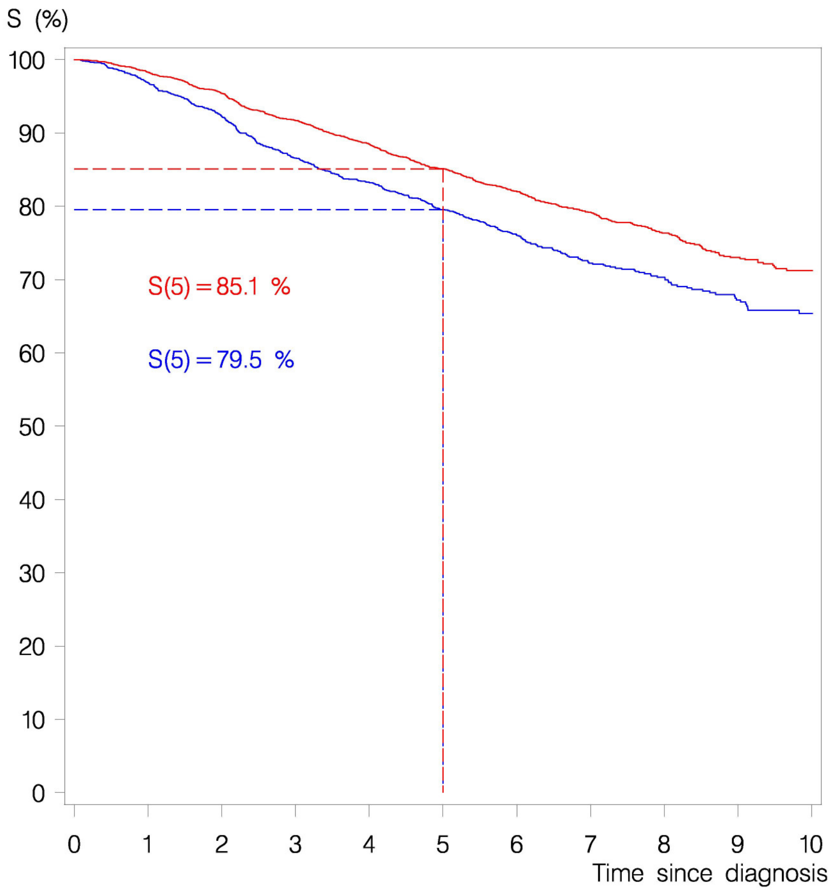


Previous:
ARR = 14.3%
RR = 0.72

Absolute risk reduction
85.1%-79.5% = 5.6%

Relative risk at 5 yrs
(100-85.1)/(100-79.5) = 0.73

i.e.
the relative risk reduction is 27%



Absolute risk reduction
85.1%-79.5% = 5.6%

Relative risk reduction 27%

Number needed to treat

The number of pts who need to be treated to prevent one additional event

NNT = 1/ARR = 1/0.056 = 18

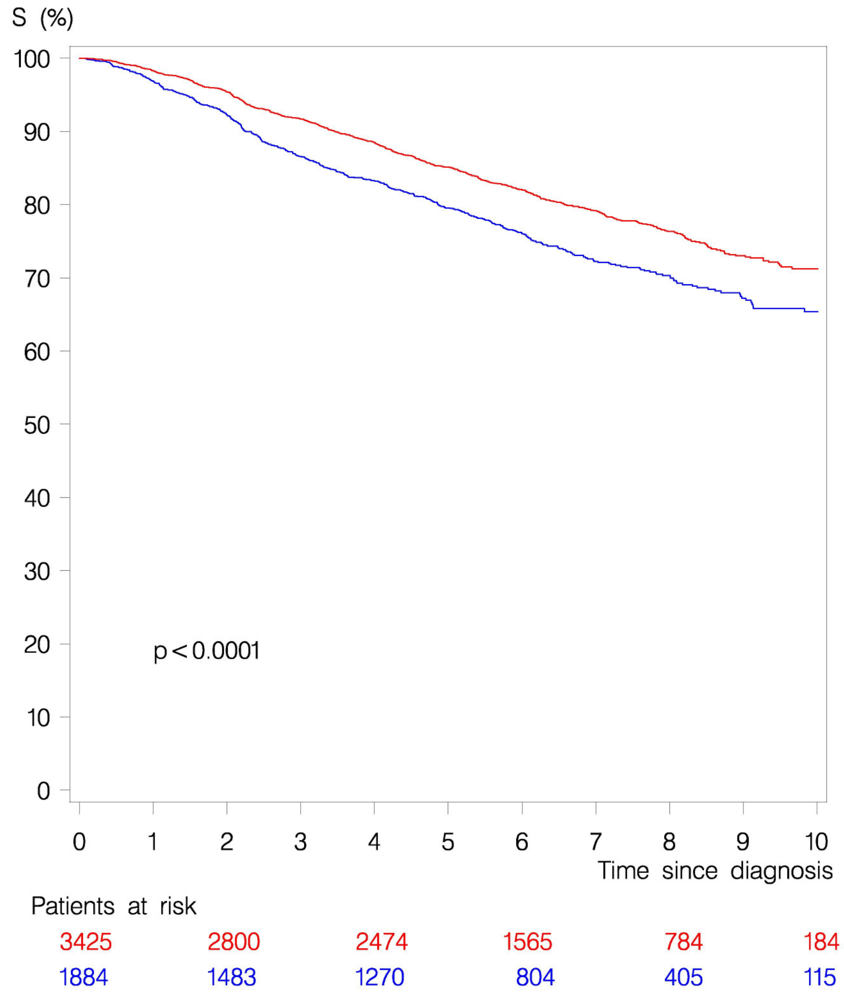
Clinical vs statistical assessment

Patients at risk

The logrank test
- the most widely used method of comparing two or more survival curves

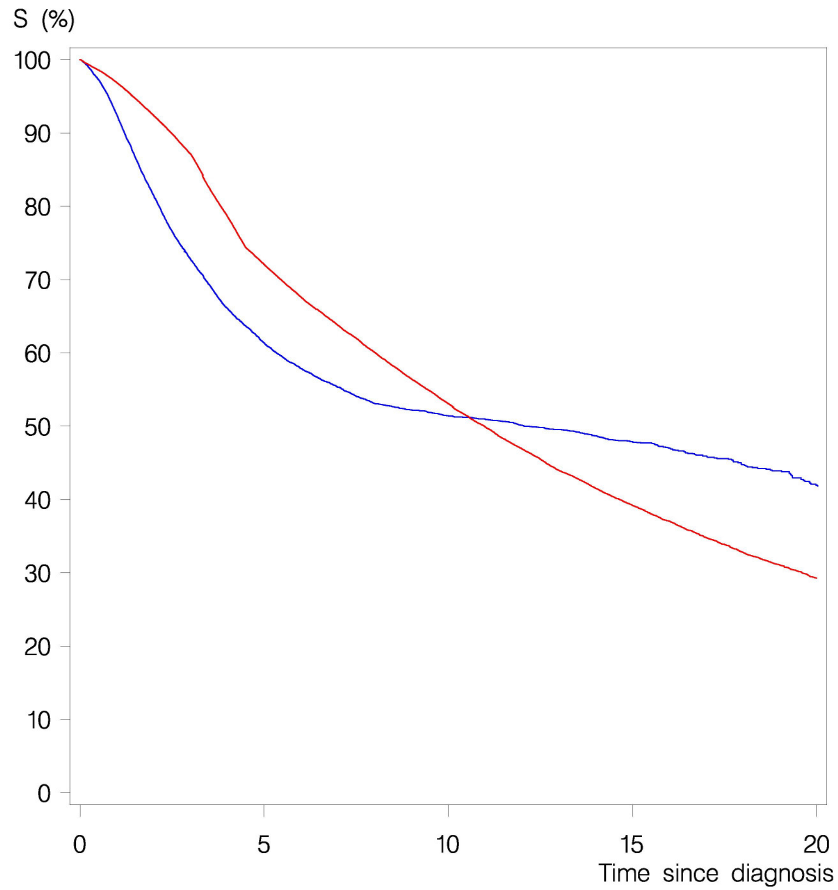
The Cox proportional hazards model
A survival analysis regression model which describes the relation between the event incidence and a set of covariates

HR = 0.75 (95% CI: 0.66-0.84), $p<0.0001$



In some cases one estimate is not representative

Examples:
Hormone receptor status
Grade of malignancy
Treatment

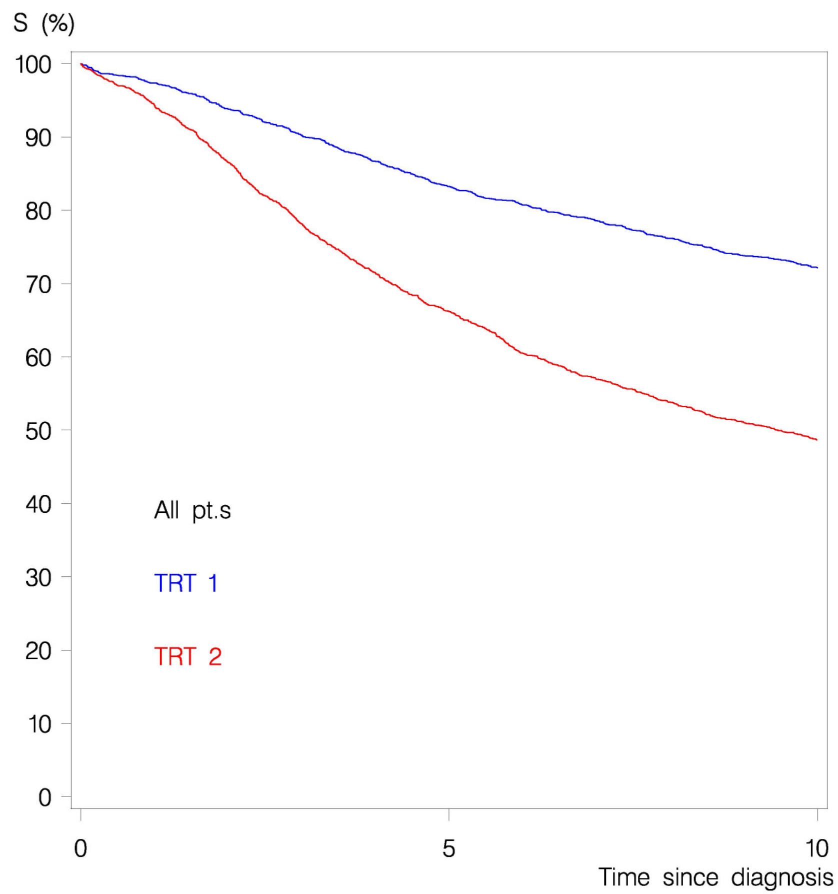


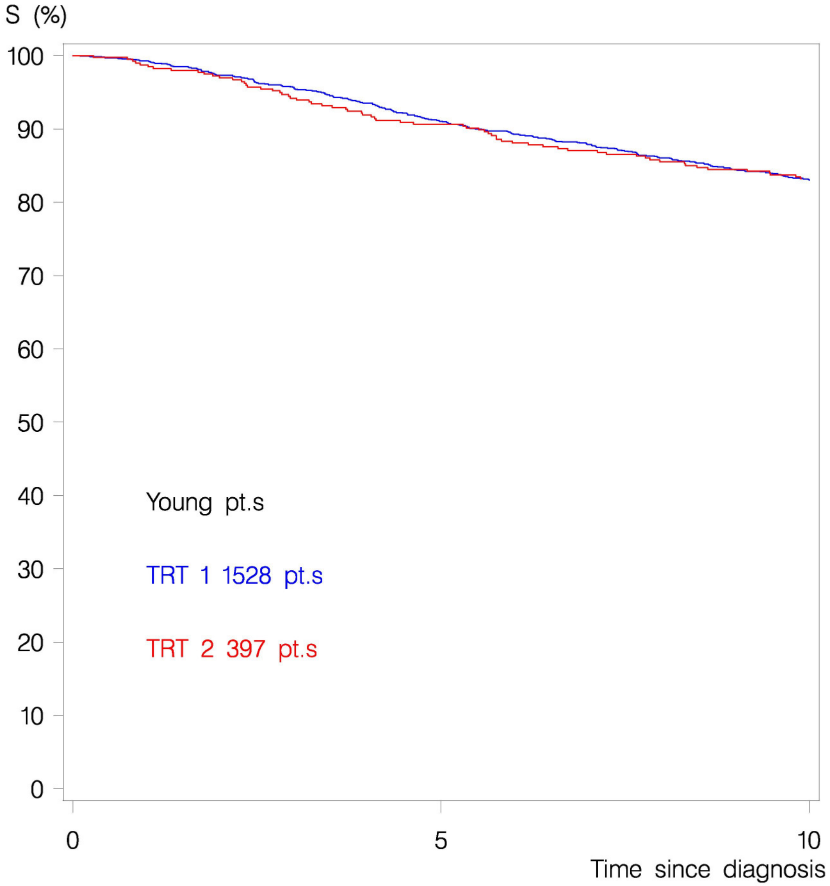
Adjustment

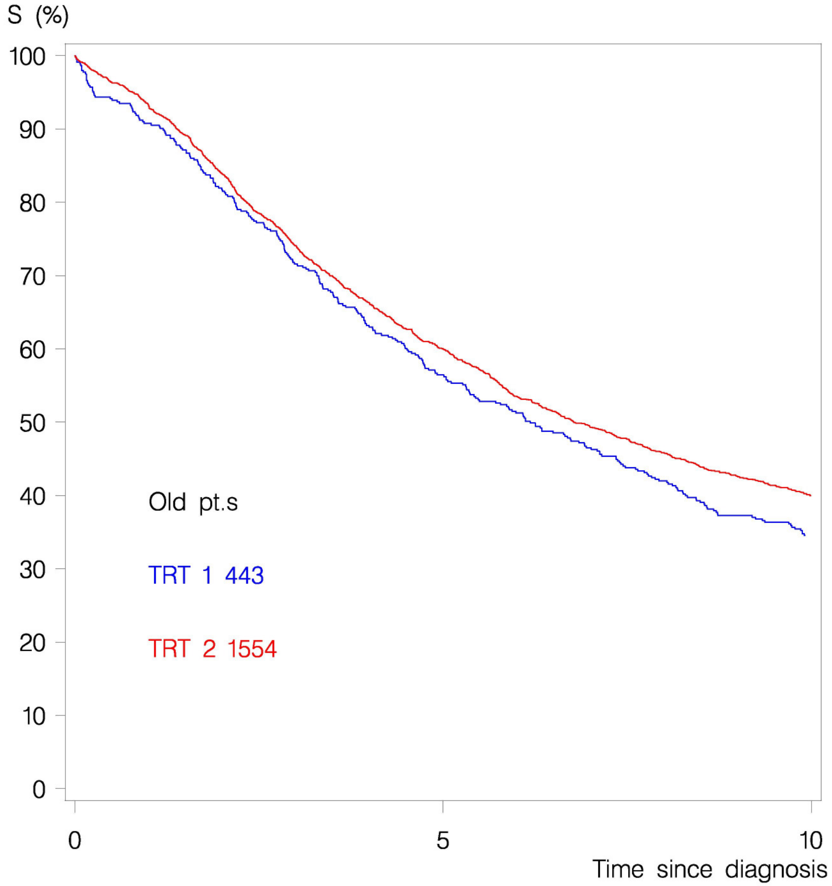
Non-randomised studies
(confounding/covariates)

Multiple prognostic factors can be adjusted for using multivariate modelling

Rule of thumb
Minimum 10 events pr. variable




S (%)
100
90
80
70
60
50
40
30
20
10
0
0
5
10
Time since diagnosis
Young pt.s
TRT 1 1528 pt.s
TRT 2 397 pt.s

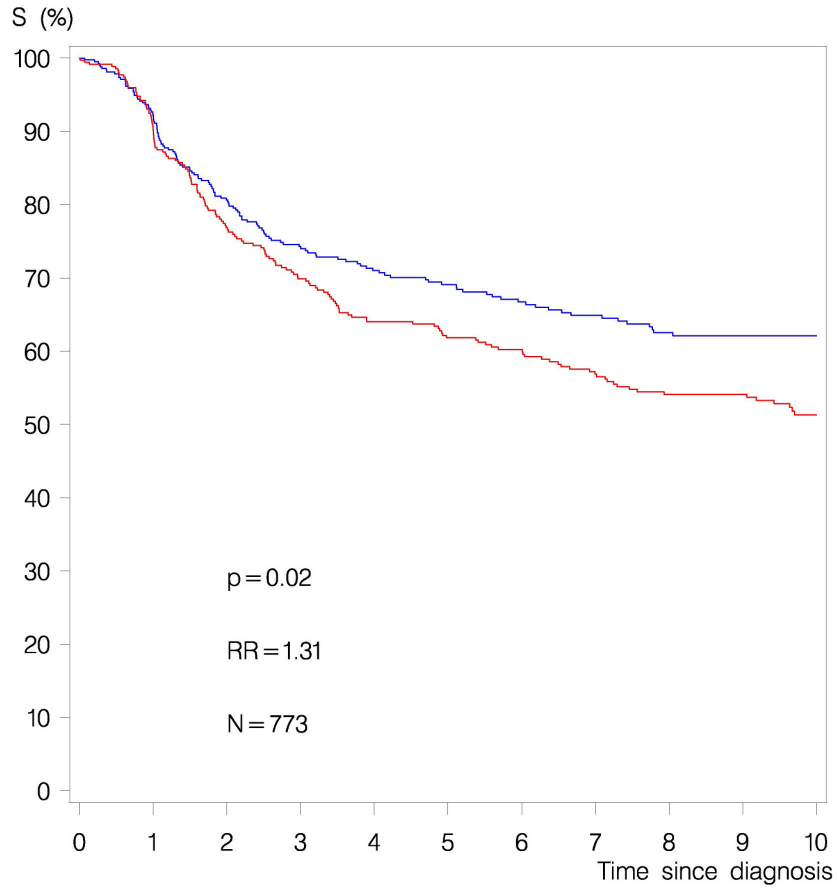

S (%)
100
90
80
70
60
50
40
30
20
10
0
0
5
10
Time since diagnosis
Old pt.s
TRT 1 443
TRT 2 1554

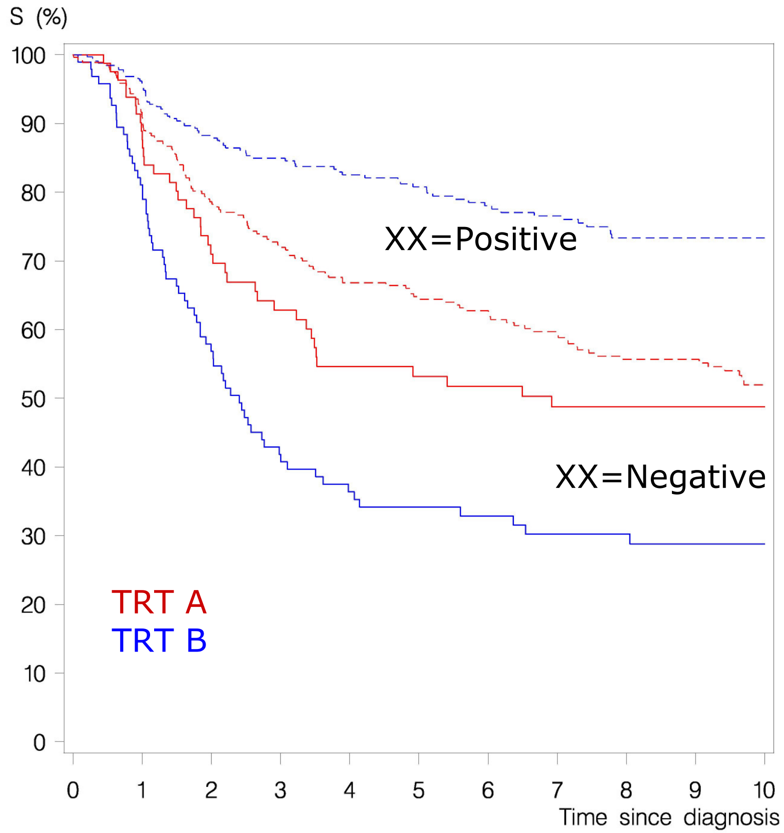
Interaction

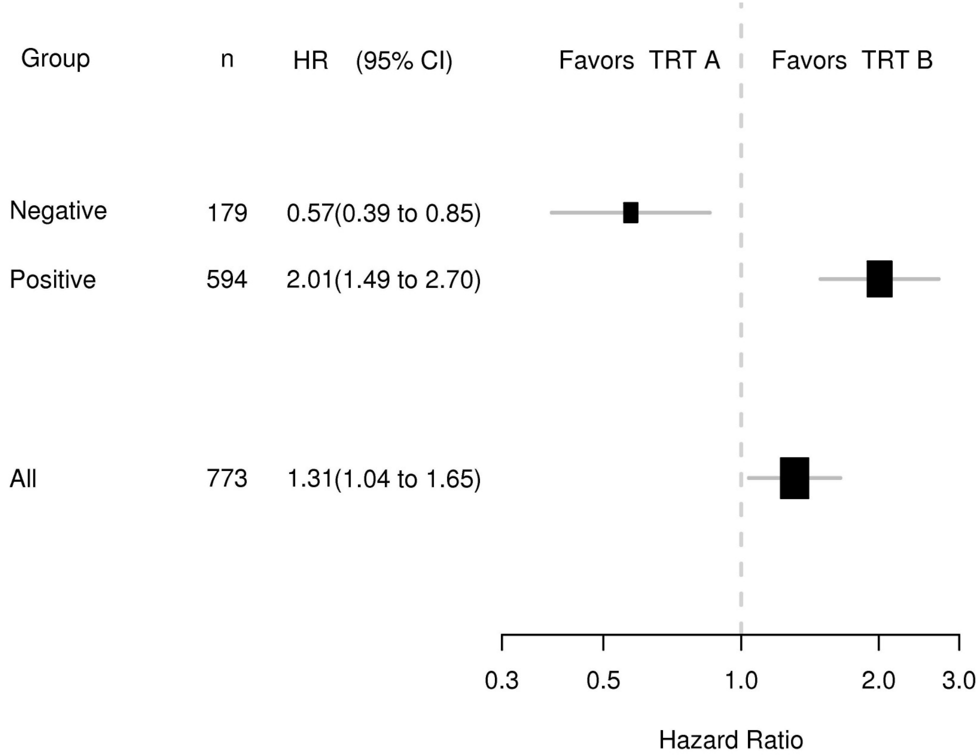
compare the treatment effect in subgroups

A specific prior suspicion

Test of interaction




S (%)
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
6
7
8
9
10
Time since diagnosis
XX=Positive
XX=Negative
TRT A
TRT B


Group
n
HR (95% CI)
Favors TRT A
Favors TRT B
Negative 179 0.57(0.39 to 0.85)
Positive 594 2.01(1.49 to 2.70)
All 773 1.31(1.04 to 1.65)
0.3
0.5
1.0
2.0
3.0
Hazard Ratio

Statistical models may give rise to misleading conclusions

Checking that a given model is an appropriate representation of the data is important, but can be complicated

Advanced survival analysis

Competing risks

Multistate models

Time-dependent covariates

Relative survival

References available on request

mj@dbcg.dk