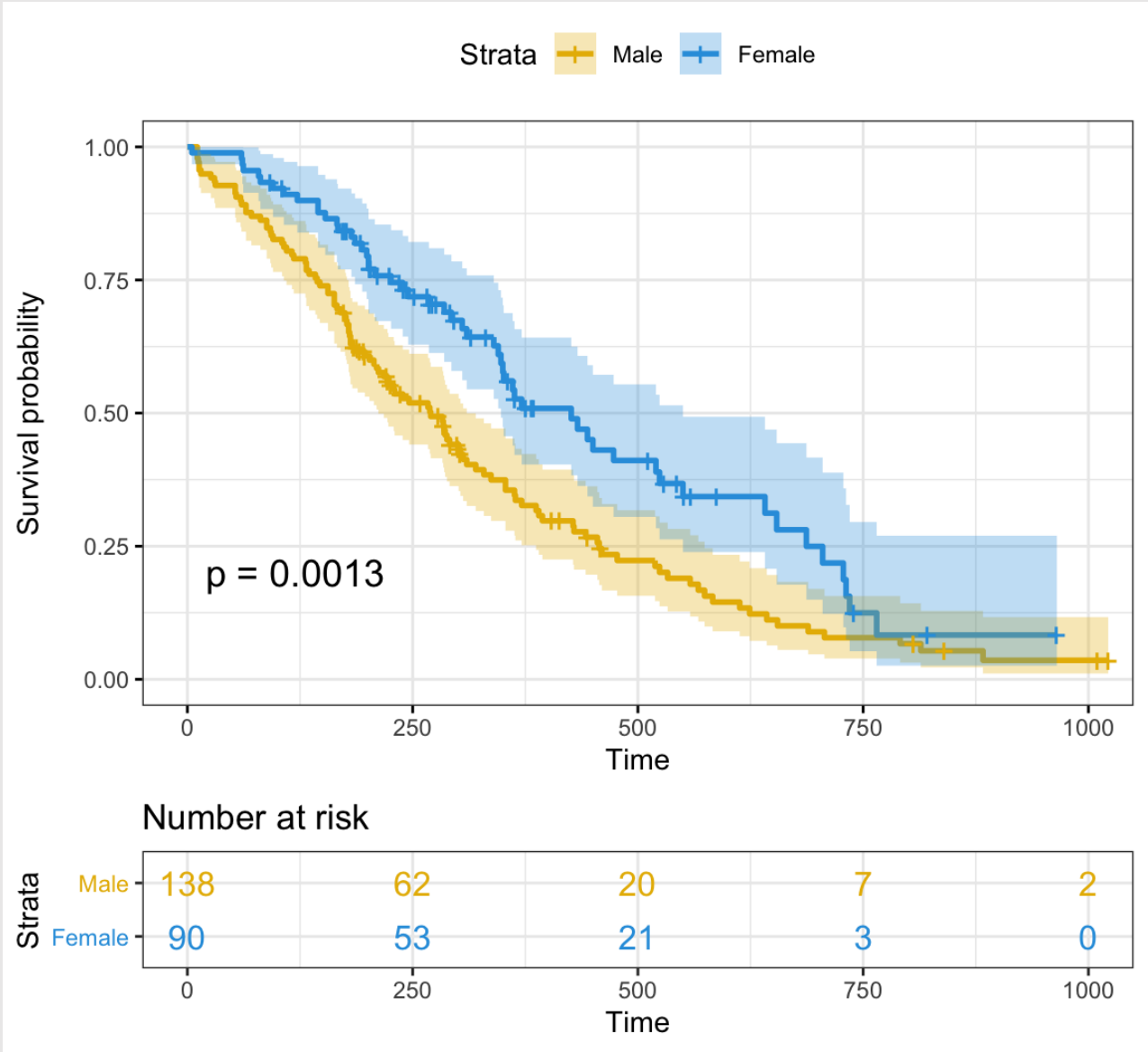
- Comparing survival curves
- Cox Regression



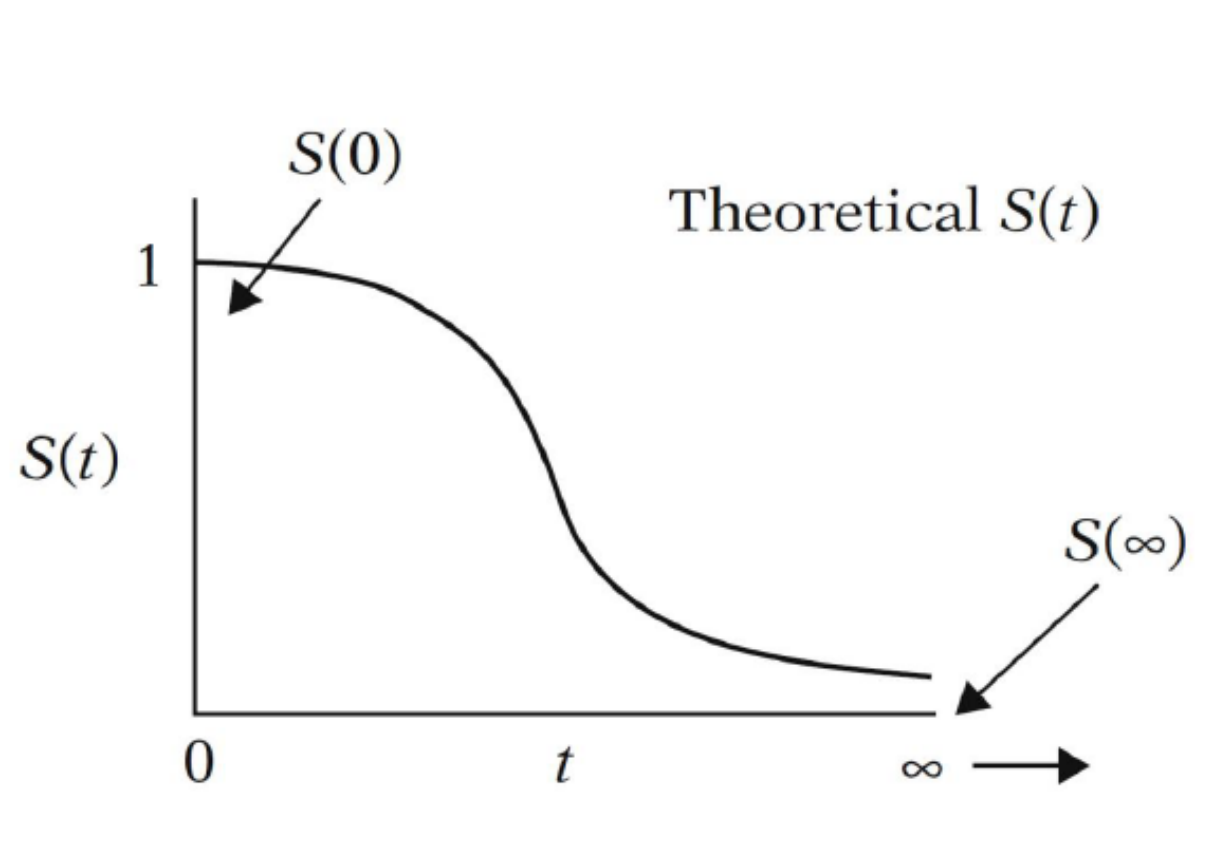
Aims of Survival Analysis

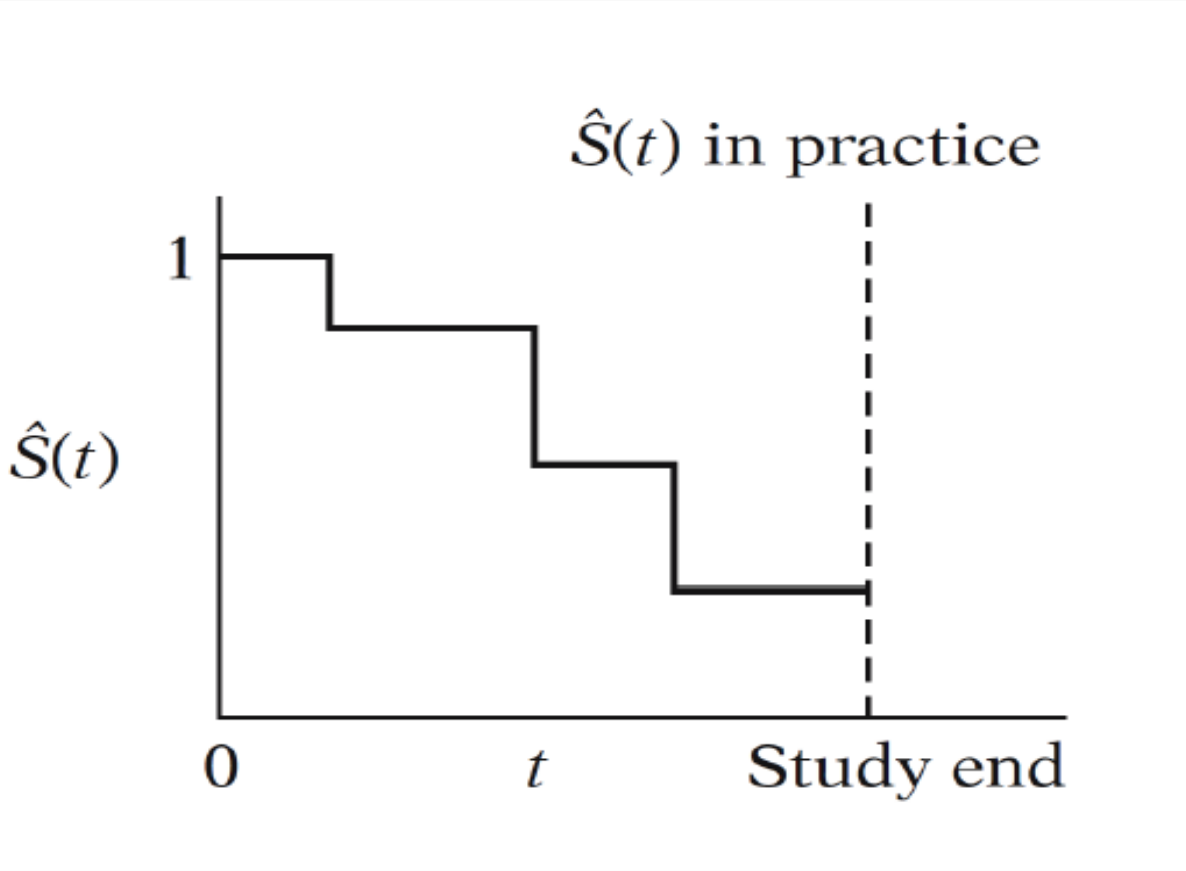
- **Estimate time-to-event for a group of individuals**, such as time until hospitalization or death for a group of patients.

- **To compare time-to-event between two or more groups**, such as treated vs. placebo patients in a randomized controlled trial.

- **To assess the relationship of co-variables to time-to-event**, such as: does weight, insulin resistance, or cholesterol influence survival time of CV patients?

Survival function: $S(t) = P(T > t)$





UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Comparison of two groups of survival data

The aim is to compare survival times of two (or more) groups of patients: one **exposed** to a certain *treatment/risk factor* another **not exposed**.

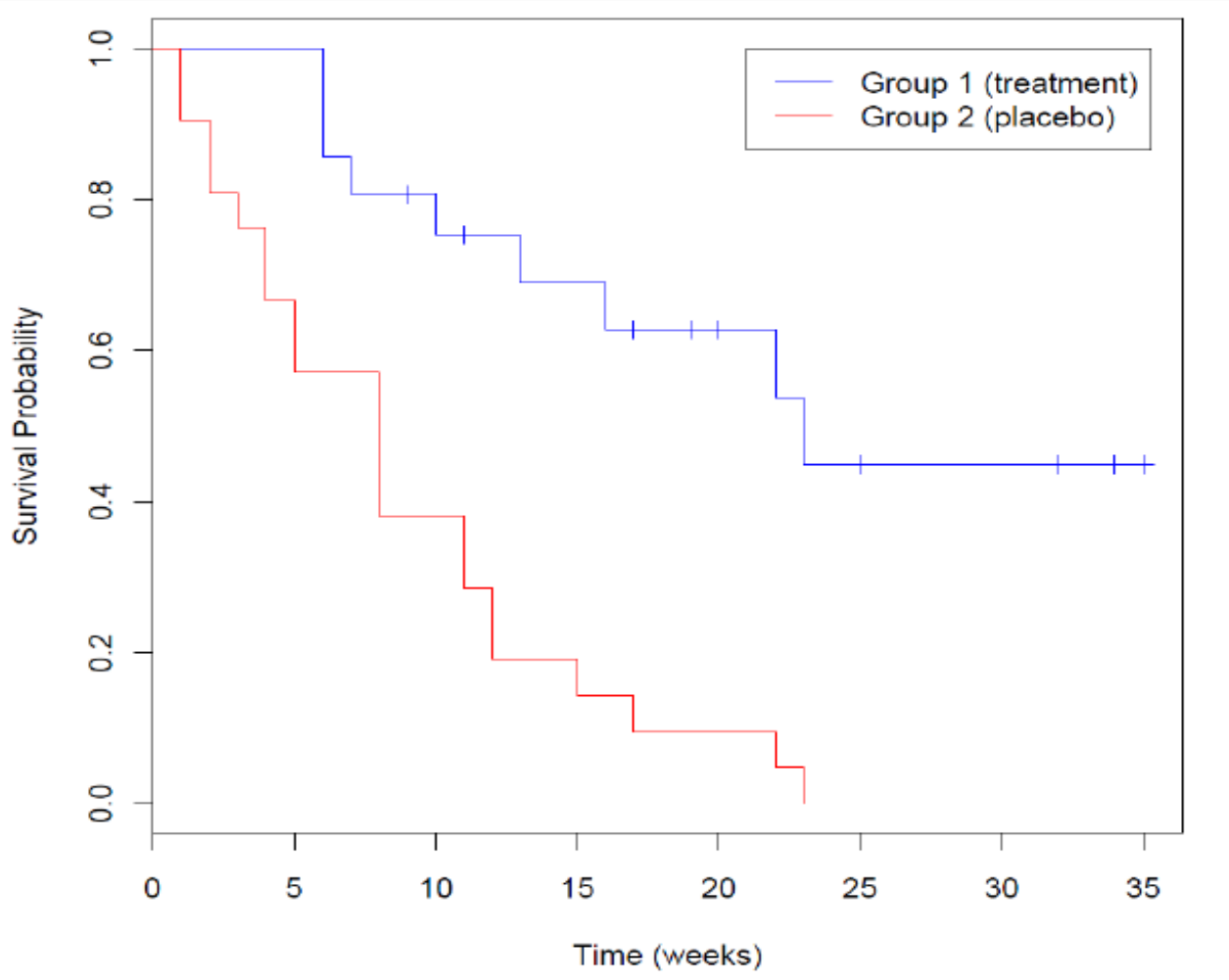
We have to perform an **hypothesis test**

H_0: There is no difference in survival among groups.

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

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Comparing *survival curves*



Do we have any reason to claim that group 1 (treatment) has a ***significant*** better survival prognosis than group 2 (placebo)?

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Log-rank test

We look at **2** groups [→ extension to **several** groups is possible]

When are two KM curves ***statistically*** equivalent?

→ we need a **testing procedure** to compare the two curves
→ when we have evidence that the ***true*** survival curves are different?

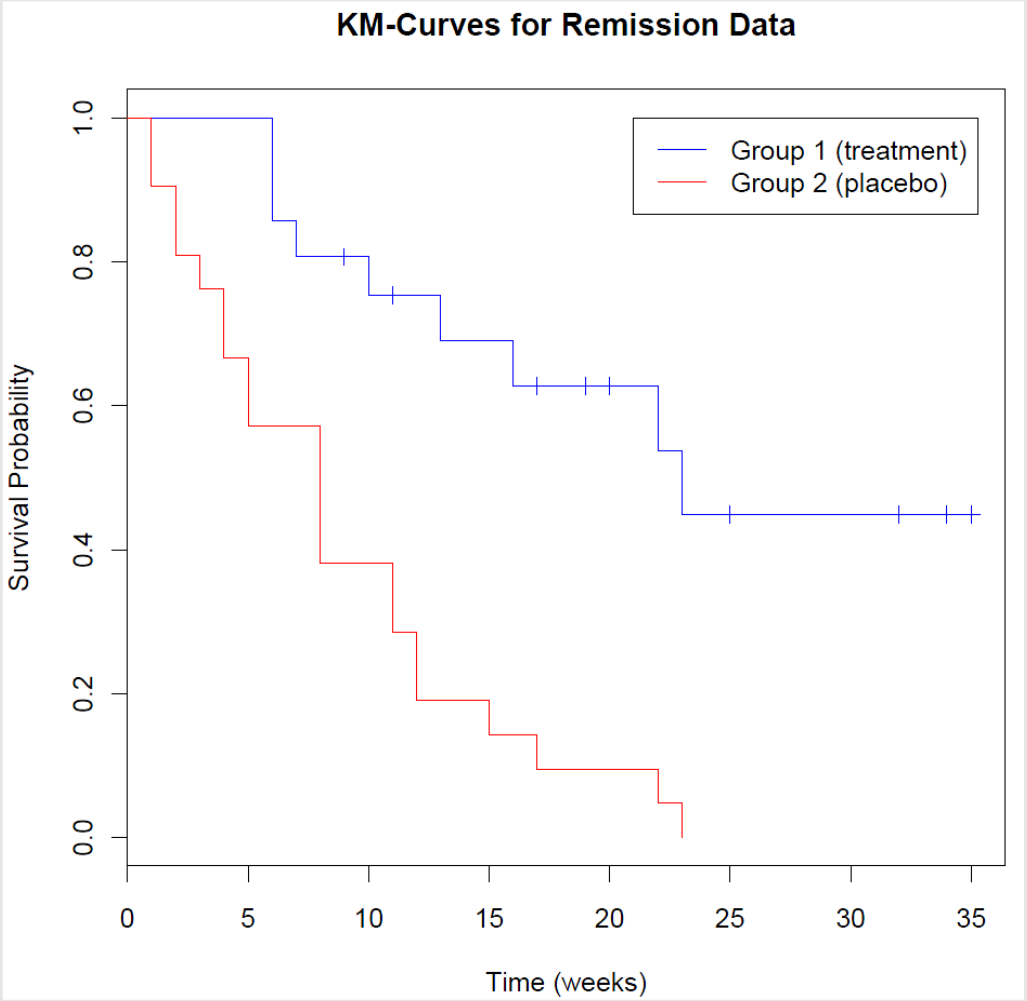
Null hypothesis (H_0): *no difference* between (*true*) survival curves

Goal: To find an expression (depending on the data) from which we know the distribution (or at least approximately) **under the null hypothesis**

Assumption : **Proportional Hazards** over time (see later) !

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Example: remission times (weeks) for two groups of leukemia patients



Remission data: n=42

# failures			# in risk set	
$t_{(j)}$	m_{1j}	m_{2j}	n_{1j}	n_{2j}
1	0	2	21	21
2	0	2	21	19
3	0	1	21	17
4	0	2	21	16
5	0	2	21	14
6	3	0	21	12
7	1	0	17	12
8	0	4	16	12
10	1	0	15	8
11	0	2	13	8
12	0	12	12	6
13	1	0	12	4
15	0	1	11	4
16	1	0	11	3
17	0	1	10	3
22	1	1	7	2
23	1	1	6	1

Remission data: n=42

# failures			# in risk set	
$t_{(j)}$	m_{1j}	m_{2j}	n_{1j}	n_{2j}
1	0	2	21	21
2	0	2	21	19
3	0	1	21	17
4	0	2	21	16
5	0	2	21	14
6	3	0	21	12
7	1	0	17	12
8	0	4	16	12
10	1	0	15	8
11	0	2	13	8
12	0	12	12	6
13	1	0	12	4
15	0	1	11	4
16	1	0	11	3
17	0	1	10	3
22	1	1	7	2
23	1	1	6	1

Expected cell counts:

$$e_{1j} = \left(\frac{n_{1j}}{n_{1j} + n_{2j}}\right) * \left(m_{1j} + m_{2j}\right)$$

$$e_{2j} = \left(\frac{n_{2j}}{n_{1j} + n_{2j}}\right) * \left(m_{1j} + m_{2j}\right)$$

We expect ***no differences between expected and observed events*** under H_0

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Expanded Table (Remission Data)

		# failures		# in risk set		# expected		Observed–expected	
j	$t_{(j)}$	m_{1j}	m_{2j}	n_{1j}	n_{2j}	e_{1j}	e_{2j}	$m_{1j}-e_{1j}$	$m_{2j}-e_{2j}$
1	1	0	2	21	21	(21/42) × 2	(21/42) × 2	–1.00	1.00
2	2	0	2	21	19	(21/40) × 2	(19/40) × 2	–1.05	1.05
3	3	0	1	21	17	(21/38) × 1	(17/38) × 1	–0.55	0.55
4	4	0	2	21	16	(21/37) × 2	(16/37) × 2	–1.14	1.14
5	5	0	2	21	14	(21/35) × 2	(14/35) × 2	–1.20	1.20
6	6	3	0	21	12	(21/33) × 3	(12/33) × 3	1.09	–1.09
7	7	1	0	17	12	(17/29) × 1	(12/29) × 1	0.41	–0.41
8	8	0	4	16	12	(16/28) × 4	(12/28) × 4	–2.29	2.29
9	10	1	0	15	8	(15/23) × 1	(8/23) × 1	0.35	–0.35
10	11	0	2	13	8	(13/21) × 2	(8/21) × 2	–1.24	1.24
11	12	0	2	12	6	(12/18) × 2	(6/18) × 2	–1.33	1.33
12	13	1	0	12	4	(12/16) × 1	(4/16) × 1	0.25	–0.25
13	15	0	1	11	4	(11/15) × 1	(4/15) × 1	–0.73	0.73
14	16	1	0	11	3	(11/14) × 1	(3/14) × 1	0.21	–0.21
15	17	0	1	10	3	(10/13) × 1	(3/13) × 1	–0.77	0.77
16	22	1	1	7	2	(7/9) × 2	(2/9) × 2	–0.56	0.56
17	23	1	1	6	1	(6/7) × 2	(1/7) × 2	–0.71	0.71
Totals		9	21			19.26	10.74	–10.26	–10.26

$$O_i - E_i = \sum_{j=1}^{\#failure\ times} (m_{ij} - e_{ij})$$

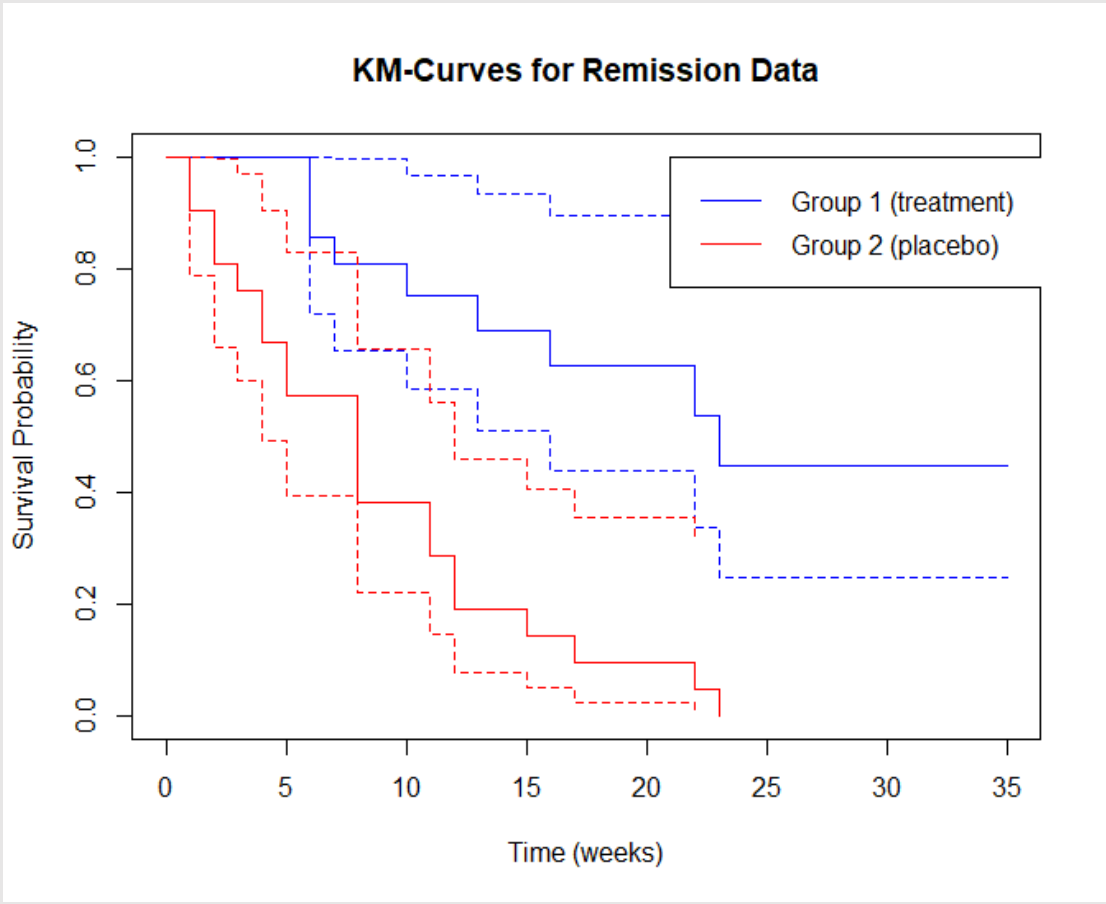
$$Log - rank = \frac{(O_i - E_i)^2}{Var(O_i - E_i)}$$

$$Log - rank \sim \chi_1$$

Remark: Group 1 or 2 are equivalent, we would get the same statistic

The sum is taken over the ordered time intervals defined by the events

```
Call:
survdiff(formula = Surv(time, status) ~ treatment)

                N Observed Expected (O-E)^2/E (O-E)^2/V
treatment=1 21          9     19.3      5.46      16.8
treatment=2 21         21     10.7      9.77      16.8

 Chisq= 16.8  on 1 degrees of freedom, p= 4e-05
```



What does this tell us?

Very low probability of obtaining a value of the test ***at least as extreme*** as the one that was actually observed, under H0.

We can therefore conclude that the treatment and placebo groups have ***significantly*** different survival

Aims of Survival Analysis

- **Estimate time-to-event for a group of individuals**, such as time until hospitalization or death for a group of patients.

- **To compare time-to-event between two or more groups**, such as treated vs. placebo patients in a randomized controlled trial.

- **To assess the relationship of co-variables to time-to-event**, such as: does weight, insulin resistance, or cholesterol influence survival time of CV patients?

The Most-Cited Statistical Papers

Journal of Applied Statistics
Vol. 32, No. 5, 461–474, July 2005

(1) With 25,869 citations (currently cited 1,984 times per year),

Kaplan, E. L. & Meier, P. (1958) Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, 53, pp. 457–481.

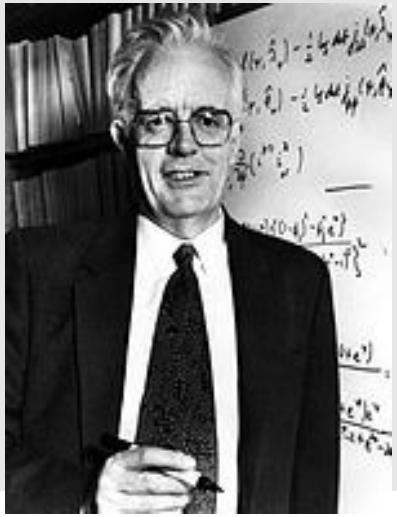
(among the **top five most cited** papers for the entire field of science)

(2) With 18,193 citations (1,342 per year),

Cox, D. R. (1972) Regression models and life tables, *Journal of the Royal Statistical Society, Series B*, 34, pp. 187–220.

Semi-parametric regression approach that estimates the effect of ***covariates*** on the **hazard function**

David Cox



Regression Models and Life-Tables

By D. R. Cox

Imperial College, London

[Read before the Royal Statistical Society, at a meeting organized by the Research Section, on Wednesday, March 8th, 1972, Mr M. J. R. Healy in the Chair]

Summary

The analysis of censored failure times is considered. It is assumed that on each individual are available values of one or more explanatory variables. The hazard function (age-specific failure rate) is taken to be a function of the explanatory variables and unknown regression coefficients multiplied by an arbitrary and unknown function of time. A conditional likelihood is obtained, leading to inferences about the unknown regression coefficients. Some generalizations are outlined.

Why don't we use ***others*** regression methods ?

Logistic regression **[binary outcome]:**

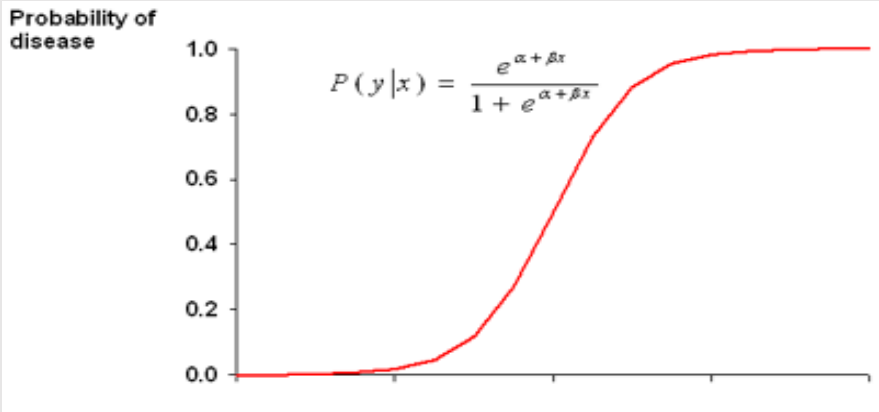
- ignores information about the ***time*** to the event

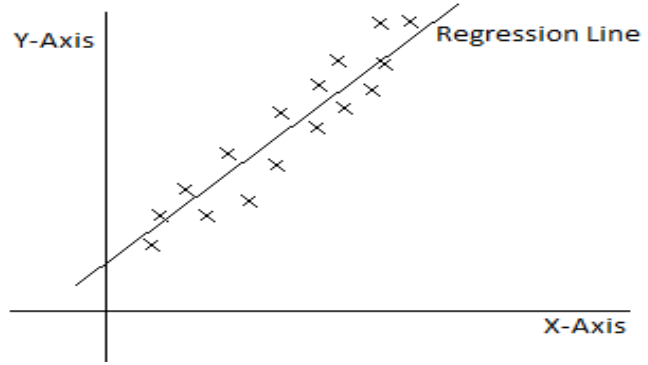
Linear regression **[continuous outcome]:**

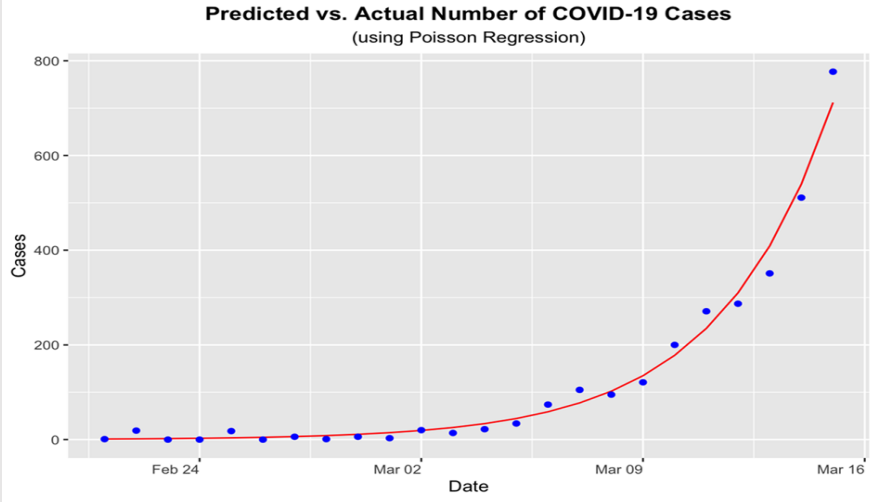
- not suitable for ***non-symmetric*** [>0] distributions [like follow up times]
- does not take into account ***censoring***

Poisson regression **[event counts/rates]:**

- #events/RR ***in a given interval*** (***≠ time to the event***)





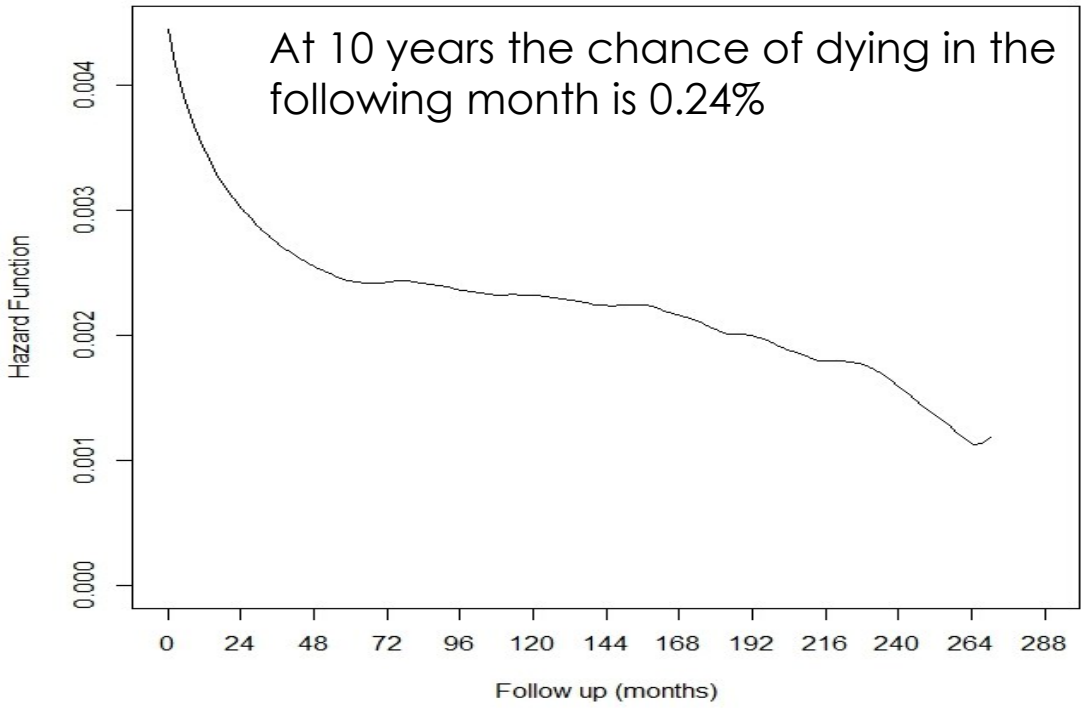


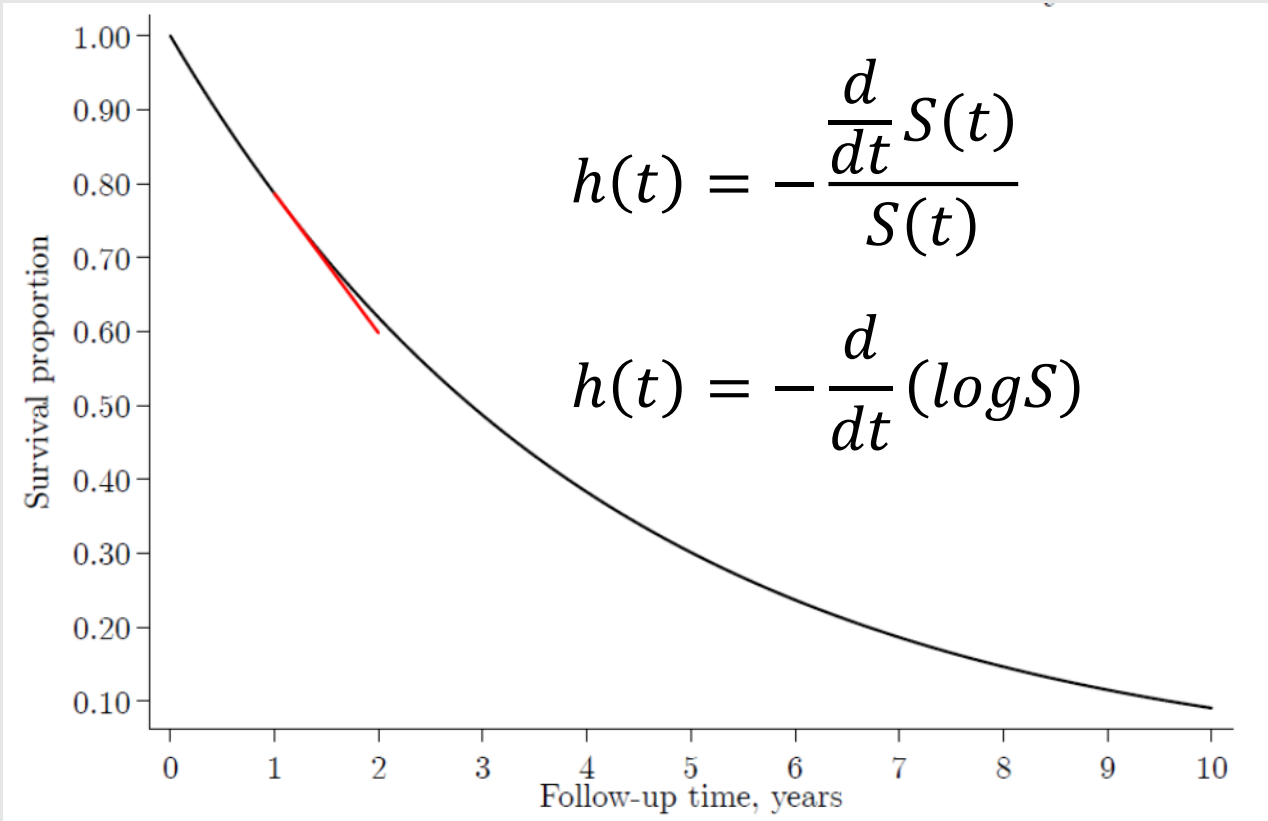
The *dependent* variable of the Cox model

The probability that ***if you survive to t***, you will succumb to the event in the next instant.

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

instantaneous event rate





$$h(t) = -\frac{\frac{d}{dt}S(t)}{S(t)}$$

$$h(t) = -\frac{d}{dt}(logS)$$

Cox Regression Model

The scale on which *linearity* is assumed is the ***log-hazard*** scale:

$$h(t|X) = h_0(t)exp\left(X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \cdots + X_p\beta_p\right)$$

$$log\left(\frac{h(t|X)}{h_0(t)}\right) = X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \cdots + X_p\beta_p$$

- $h_0(t)$ is the **baseline hazard function**

- the exponential function represents the effect of the linear combination of the covariates X on the hazard

The aim is to determine the **joint** effect of the covariates on the hazard or to focus on a **specific** effect.

The dependent variable of the Cox model is the hazard function.

The model assumes that the risk at time *t* for subject *i* is: $h_i(t|X_i) = h_0(t)exp(X_i\beta)$

Remind of **censored** data:
someone who is followed for 18 months is a part of the computations *until the interval that contains the censoring time* (risk set) and not thereafter (***partial likelihood***).

Why *exp*(linear predictor)? To avoid ***negative*** hazard rates.

- Implies that factors are multiplicative, e.g., treatment reduces the hazard by X %.
- Two covariates multiply in effect
- For *biological phenomena* it seems to fit well

The baseline hazard in Cox is estimated **non-parametrically**:
- estimated on the specific dataset
- does not extrapolate….

- To estimate β Cox proposed a **partial** likelihood (PL) procedure based on conditional probability:

$$L(\beta) = \prod_{j=1}^{n} \frac{exp\left(X\left(t_{(j)}\right)\beta\right)}{\sum_{i \in R_j} \exp\left(X_i\left(t_{(j)}\right)\beta\right)}$$

- Maximizing the PL function we obtain:

1. Estimates of β
2. Standard errors for β
3. *p* values for β

$t_{(1)}, \cdots, t_{(n)}$	ordered event times
R_j	**Risk set** at time $t_{(j)}$
$X\left(t_{(j)}\right)$	covariates for the individual who fails at time $t_{(j)}$

(the non-parametric estimate of cumulative baseline hazard could be obtained after β estimation)

In the Cox model the statistical independence between censoring and survival time is assumed *conditional to* the covariates !!

Interpretation of parameter estimates

Let us consider two subjects *i* e *i'*:

$$\eta_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}$$
$$\eta_{i'} = \beta_1 x_{i'1} + \beta_2 x_{i'2} + \ldots + \beta_p x_{i'p}$$

linear part of the Cox model

The **hazard ratio** between them is:

$$\frac{h_i(t|X)}{h_{i'}(t|X)} = \frac{h_0(t)\exp(\eta_i)}{h_0(t)\exp(\eta_{i'})} = \frac{\exp(\eta_i)}{\exp(\eta_{i'})}$$

Suppose to have a single continuous variable X: $h_i(t) = h(t)\exp(\beta x_i)$

The ratio of the hazard for a subject with value x+*1* with respect to one with value x is:

$$\frac{\exp\{\beta(x+1)\}}{\exp(\beta x)} = \exp(\beta)$$

HAZARD RATIO

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Exp(β)=HAZARD RATIO (HR)

If **HR ~1** (95% CI contains 1) : there is **not a significant** impact of the covariate X on the hazard of event

If **HR >1** (95% CI > 1) : presence or increasing values of X **increase** the hazard of event (=decrease survival)

If **HR <1** (95% CI < 1) : presence or increasing values of X **decrease** the hazard of event (=increase survival)

Impact of gender (M=0,F=1) and level of education (school yrs) with respect to time to the first marriage:

Cox model results	β	se(β)	exp(β) HR	lower 95% CI	upper 95% CI
Gender (F vs M)	0,48	0,20	1,61	1,09	2,40
School years	-0,07	0,02	0,93	0,51	0,98

At a given instant in time, the hazard of marriage for women is **1.61 times higher** than men (at the same level of education)

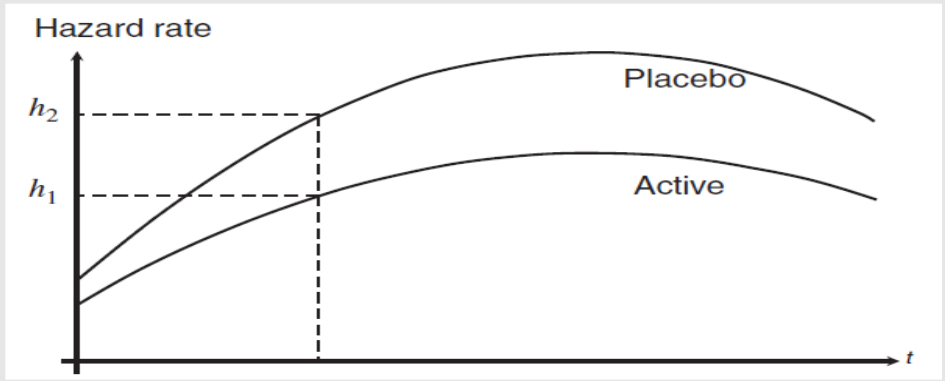
At a given instant in time, the hazard of marriage for women is **61% higher** than for men (at the same level of education)

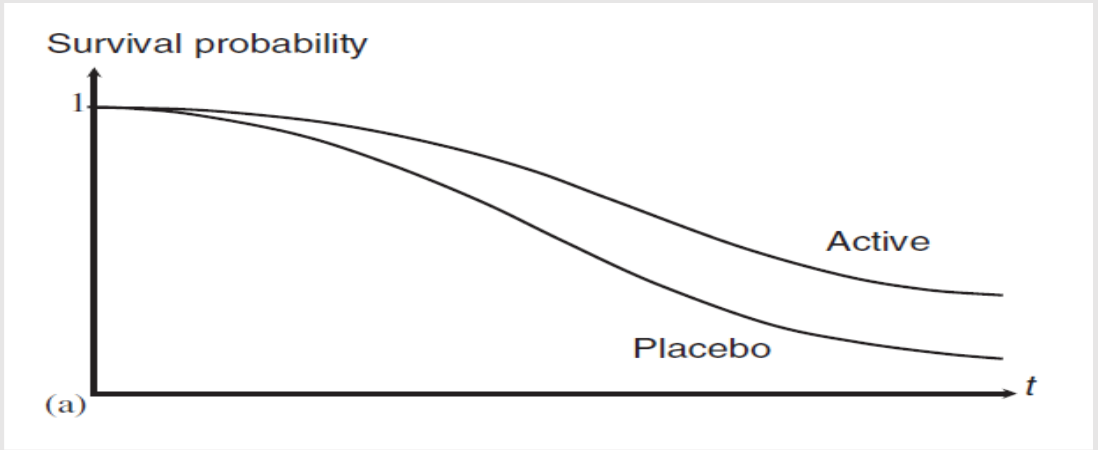
For people (men or women) with an additional +1yr school the hazard of marriage, *at a given instant in time*, is **0.93 times** than for those without…

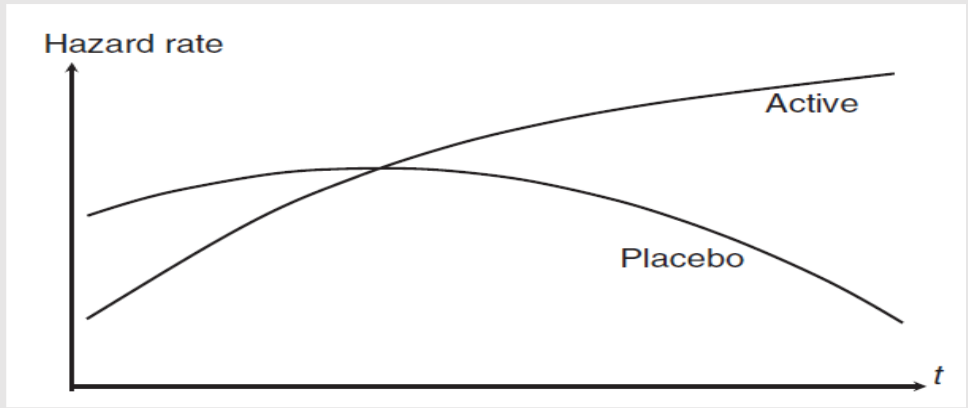
For each extra yr of school the hazard of marriage (men or women) *at a given instant in time* is **7% less**.

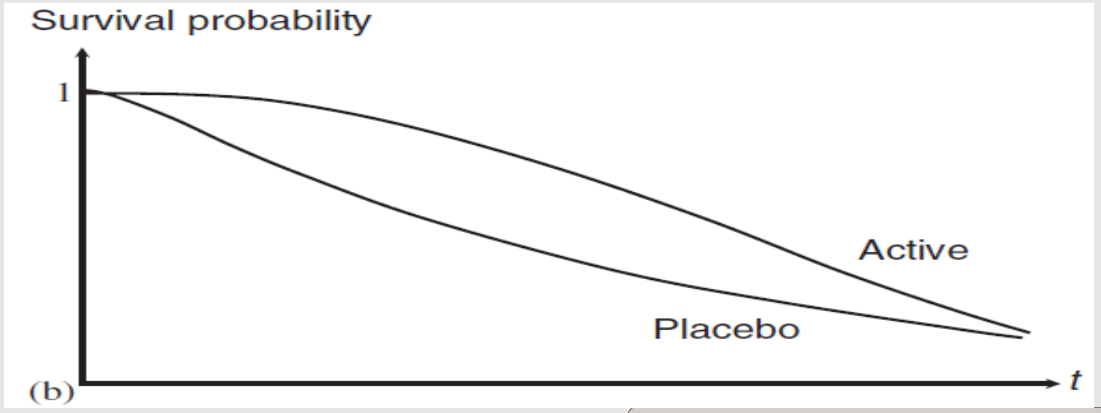
Proportional hazards (PH)

The hazard ***at any given time*** for an individual in one group is proportional to the hazard ***at any given time*** for an individual in the other group. If the hazard functions are proportional **->** survival functions **do not cross** one another…









UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Cox model assumes **proportional hazards** (PH). Covariates X have always *the same relative effect* along time:

$$h(t|X) = h_0(t)\exp(X_1\beta_1 + X_2\beta_2 + \ldots + X_p\beta_p) = h_0(t)\exp(\mathbf{X\beta})$$

The function exp(**Xβ**) does not depend on *t*

Hazard Ratio between two subjects, with covariates X and X* does not depend on *t*:

$$\frac{h_0(t)\exp(\mathbf{X\beta})}{h_0(t)\exp(\mathbf{X}^*\boldsymbol{\beta})} = \exp((\mathbf{X} - \mathbf{X}^*)\boldsymbol{\beta})$$

If PH assumption does not hold, *standard* Cox model could be no longer valid **[we could check for this]** [there are extensions]

$$h_i(t|X_i) = h_0(t)exp(X_i\beta)$$

- β_k is the **difference** in the log-hazard function comparing two subpopulations differing in x_k by "1-unit" and that are similar with respect to all other covariates in the model

- the **effect** expressed by β_k is ***adjusted for*** all other covariates in the model, so it has the interpretation of a log-relative hazard associated with a change in x_k, holding other covariates constant ***at some fixed value***

- is it possible to compare ***hypothetical patients*** with different covariates values and check how their ***estimated survival curves*** appear; [remind: the baseline hazard depends on the study cohort…]

- the Cox PH model is a model for the hazard more than a model for survival time, **although they are related one-to-one if no competing risks exists**

Survival function derived from the Cox regression model

(no competing risks, no time-dependent variables)

Once the β are estimated, we can obtain the corresponding survival function:

$$S(t|x) = S_0(t)\exp(\beta x)$$

$\boldsymbol{S_0(t)}$ is derived from an estimate of the **cumulative baseline hazard**
(a derivation in the **non-parametric** form, similar to the *Nelson-Aalen* formulation)

The estimate of $S_0(t)$ and a fixed set of values for the explanatory variables produce an estimate of the survival function **for a specific person or group**.

The expression for $S(t|x)$ shows that proportional hazard functions dictate that the estimated survival functions ***do not intersect***.

Example of application (I)

Table 3 Univariable and multivariable analysis (primary analysis Cox)

Outcome: Death or CV hosp

Variables	Univariable analysis		Multivariable analysis	
	Hazard ratio (95% CI)	*P*-value	Hazard ratio (95% CI)	*P*-value
Cardiac rehabilitation	0.601 (0.476–0.758)	<0.001	0.578 (0.432–0.773)	<0.001
NSTEMI	1.361 (1.043–1.774)	0.023		
Male	1.168 (0.898–1.517)	0.246		
STEMI	0.908 (0.701–1.176)	0.463		
PCI	1.343 (1.036–1.742)	0.026		
CABG	0.621 (0.465–0.828)	0.001	0.639 (0.466–0.876)	**0.005**
Ejection fraction	0.979 (0.967–0.991)	0.001	0.986 (0.973–0.999)	**0.035**
Diabetes	1.548 (1.219–1.966)	<0.001	1.460 (1.107–1.926)	**0.007**
Hypertension	1.161 (0.887–1.520)	0.276		
Smoking	1.121 (0.860–1.463)	0.398		
Dyslipidaemia	0.897 (0.708–1.138)	0.372		
Beta-blockers	1. 244 (0.910–1.701)	0.171		
ACE-inhibitors/ARBs	1.367 (1.005–1.859)	0.046		
Statins/ezetimibe	0.607 (0.426–0.865)	0.006	0.518 (0.345–0.776)	**0.001**
ASA	0.932 (0.510–1.703)	0.819		
DAPT	1.245 (0.968–1.601)	0.088		
Chronic kidney disease	2.409 (1.823–3.182)	<0.001	2.441 (1.775–3.358)	<0.001
Previous ACS	1.443 (1.111–1.873)	0.006		
Previous PCI	1.718 (1.299–2.272)	<0.001		
Previous CABG	1.884 (1.240–2.861)	0.003		

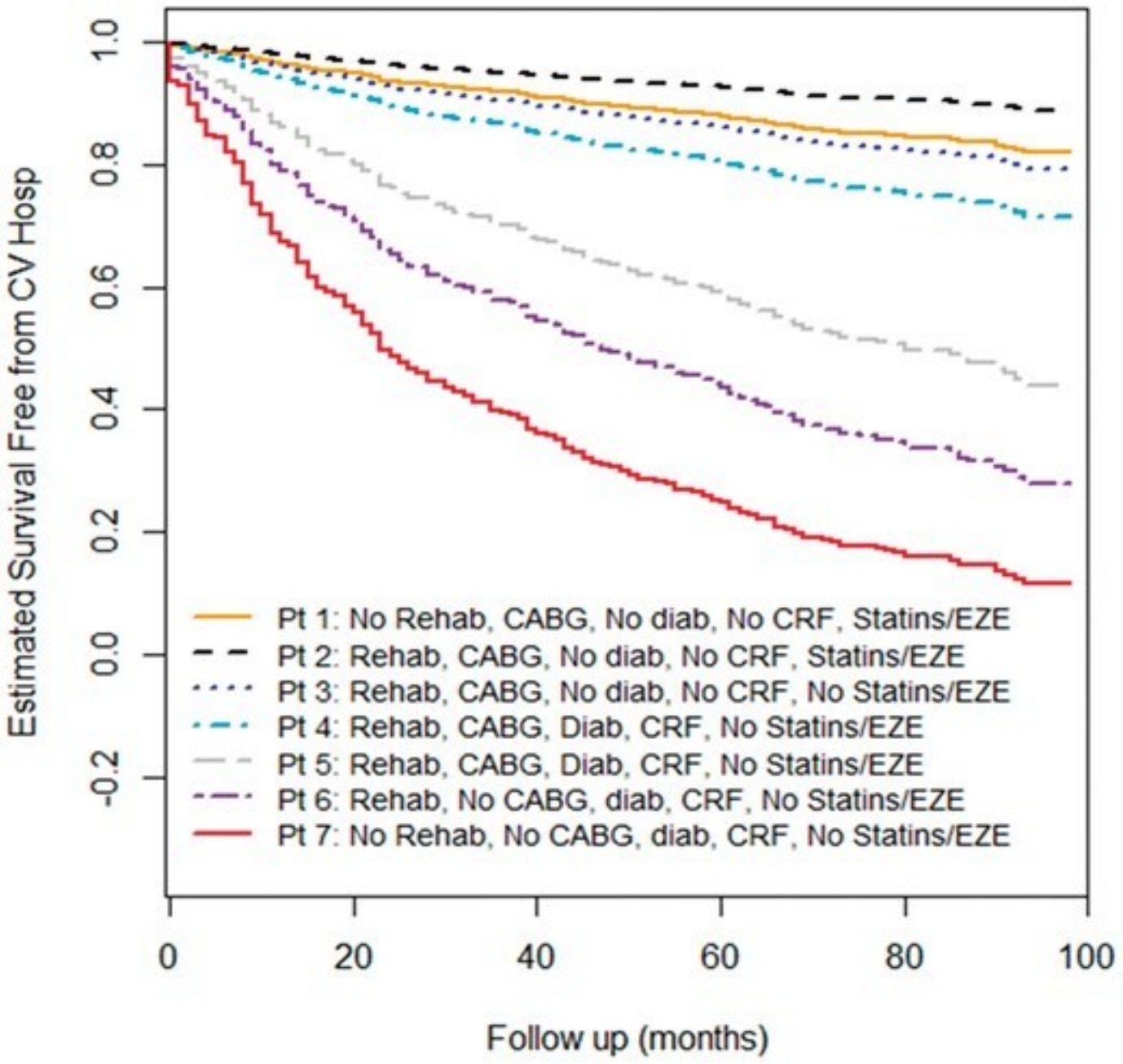
ACE-inhibitors, angiotensin-converting enzyme inhibitor; ACS, acute coronary syndrome; ARBs, angiotensin receptor blockers; ASA, acetylsalicylic acid; CABG, coronary artery bypass graft; CI, confidence interval; DAPT, dual antiplatelet therapy; NSTEMI, non-ST-elevation myocardial infarction; PCI, percutaneous coronary intervention; STEMI, ST-elevation myocardial infarction.

Treatment of interest

Independent covariates

Confounder

Independent covariate



Estimated survival curves from the Cox model.

The curves are estimated for patients having the median ejection fraction (56%) of the population.

CABG : coronary artery bypass graft
CRF: chronic renal failure
Diab: diabetes
EZE: ezetimibe
Rehab: cardiac rehabilitation.

This study demonstrated the positive effects of CR program in the real world showing a decreased risk of CV hospitalizations and mortality during a long-term follow-up.

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Example of application (II)

SCORE2 risk prediction algorithms

1. Model development

Sex-specific, competing risk-adjusted risk models derived in 45 prospective cohorts in 13 countries (~680,000 individuals, and ~30,000 CVD events)

↓

Recalibration to four risk regions in Europe using age-, sex-, and region-specific risk factor values and CVD incidence rates (derived using data on ~10.8 million individuals)

2. Model validation

External validation in 25 prospective cohorts in 15 European countries (~1.1 million individuals, and ~43,000 CVD events)

↓

C-indices ranged from 0.67 (95% confidence interval [CI] 0.65-0.68) to 0.81 (95% CI 0.76-0.86)

SCORE2 risk prediction algorithms key features

- Sex-specific risk prediction models
- Estimate 10-year risk of fatal and non-fatal CVD
- Calibrated to the most contemporary and representative CVD rates
- Available for four distinct European risk regions
- Can be rapidly updated to reflect future CVD incidence and risk factor profiles

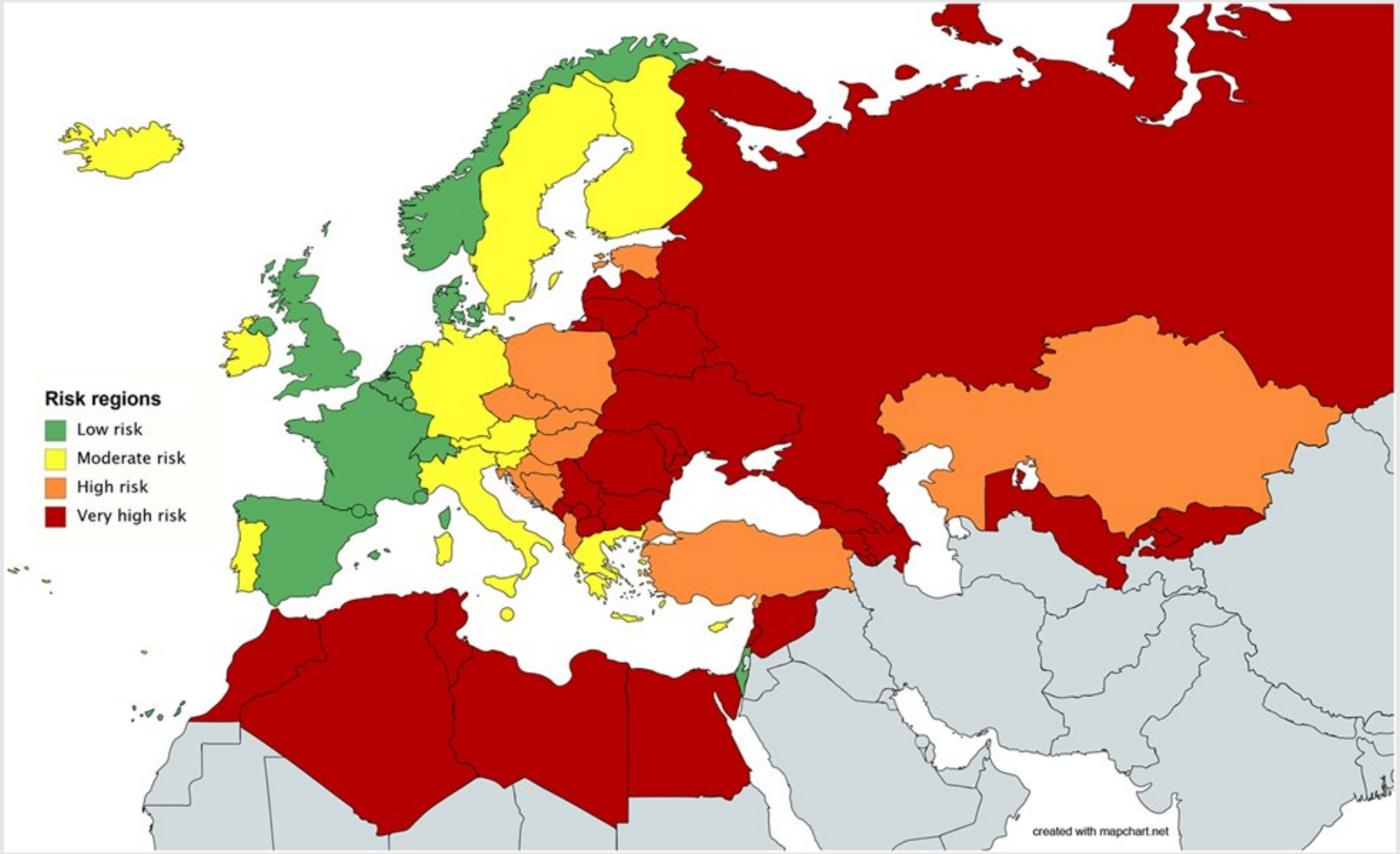
Individual example

Patient risk factors:
50 years old
Smoker
SBP: 140 mmHg
Cholesterol: 5.5 mmol/L
HDL-c: 1.3 mmol/L

10-year risk depending on risk region

Low risk	Moderate risk	High risk	Very high risk	Low risk	Moderate risk	High risk	Very high risk
4.2%	5.1%	6.9%	13.7%	5.9%	7.5%	8.1%	14.0%

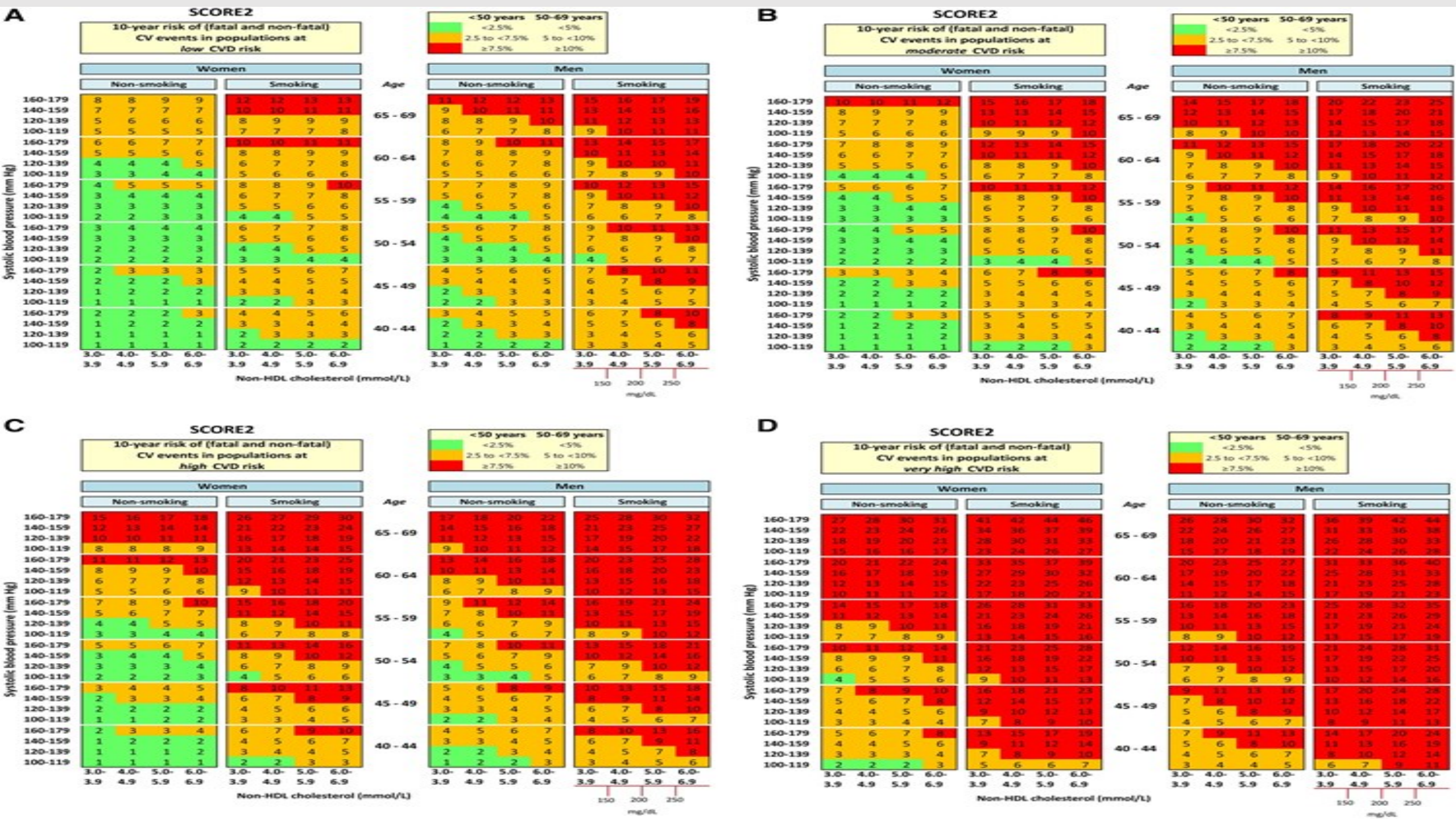
The SCORE2 (OP) algorithms are used to estimate **10-year cardiovascular risk** in individuals aged 40-69 and 70+, respectively. These algorithms, developed by the European Society of Cardiology (ESC), are designed for use across various regions in Europe, including those with low, moderate, high, and very high risk profiles (**different baseline hazard**).



Countries were grouped into **four** risk regions according to their most recently reported WHO **age- and sex-standardized overall CVD mortality rates** per 100,000 population

- **low risk** (<100 CVD deaths per 100,000)
- **moderate risk** (100 to <150 CVD deaths per 100,000)
- **high risk** (150 to <300 CVD deaths per 100,000)
- **very high risk** (≥300 CVD deaths per 100,000)

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute


A
SCORE2
10-year risk of (fatal and non-fatal) CV events in populations at low CVD risk
<50 years 50-69 years
<2.5% <5%
2.5 to <7.5% 5 to <10%
≥7.5% ≥10%
Women
Non-smoking
Smoking
Men
Non-smoking
Smoking
Age
65 - 69
60 - 64
55 - 59
50 - 54
45 - 49
40 - 44
Systolic blood pressure (mm Hg)
160-179
140-159
120-139
100-119
3.0-3.9
4.0-4.9
5.0-5.9
6.0-6.9
Non-HDL cholesterol (mmol/L)
150
200
250
mg/dL
B
SCORE2
10-year risk of (fatal and non-fatal) CV events in populations at moderate CVD risk
<50 years 50-69 years
<2.5% <5%
2.5 to <7.5% 5 to <10%
≥7.5% ≥10%
Women
Non-smoking
Smoking
Men
Non-smoking
Smoking
Age
65 - 69
60 - 64
55 - 59
50 - 54
45 - 49
40 - 44
Systolic blood pressure (mm Hg)
160-179
140-159
120-139
100-119
3.0-3.9
4.0-4.9
5.0-5.9
6.0-6.9
Non-HDL cholesterol (mmol/L)
150
200
250
mg/dL
C
SCORE2
10-year risk of (fatal and non-fatal) CV events in populations at high CVD risk
<50 years 50-69 years
<2.5% <5%
2.5 to <7.5% 5 to <10%
≥7.5% ≥10%
Women
Non-smoking
Smoking
Men
Non-smoking
Smoking
Age
65 - 69
60 - 64
55 - 59
50 - 54
45 - 49
40 - 44
Systolic blood pressure (mm Hg)
160-179
140-159
120-139
100-119
3.0-3.9
4.0-4.9
5.0-5.9
6.0-6.9
Non-HDL cholesterol (mmol/L)
150
200
250
mg/dL
D
SCORE2
10-year risk of (fatal and non-fatal) CV events in populations at very high CVD risk
<50 years 50-69 years
<2.5% <5%
2.5 to <7.5% 5 to <10%
≥7.5% ≥10%
Women
Non-smoking
Smoking
Men
Non-smoking
Smoking
Age
65 - 69
60 - 64
55 - 59
50 - 54
45 - 49
40 - 44
Systolic blood pressure (mm Hg)
160-179
140-159
120-139
100-119
3.0-3.9
4.0-4.9
5.0-5.9
6.0-6.9
Non-HDL cholesterol (mmol/L)
150
200
250
mg/dL

Supplementary materials

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Power/Sample size for the log-rank test

- The logrank test statistics

$$Z = \frac{\sum_{j=1}^{k}(O_j - E_j)}{\sqrt{\sum_{j=1}^{k} V_j}} \sim N(0,1) \text{ under } H_0$$

The power of the logrank test depends on the **number of observed failures** rather than the sample sizes

Under the null hypothesis $H_0 : S_1(t) = S_0(t)$

$$H_1 : S_1(t) = S_0(t)^{\exp(\beta)} \Leftrightarrow h_1(t) = h_0(t)e^{\beta}, \beta \neq 0$$

Hazard functions

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Let:

$$\exp(\beta) = HR = \frac{h_1(t)}{h_2(t)}$$ **Hazard Ratio**

E : *total (expected) number of events*

$\Delta =$ ***HR*** **(effect size)**

α : *significance level*

pw: *power*

p_0, p_1 *proportions of subjects in groups*

$$E = \left(\frac{z_{1-\alpha/2} + z_{pw}}{\ln(\Delta)}\right)^2 * \frac{1}{p_0 * p_1}$$

A new treatment is expected to increase the survival rate at five years **from 0.41**, the value under the standard treatment, **to 0.60**.

$$\ln(\Delta) = \frac{\log(0.60)}{\log(0.41)} = 0.57$$

Assuming α = 0.05 and power =0.90, equal numbers in the two groups, we obtain :

$$\frac{10.51}{0.57^2} * 4 = 133$$

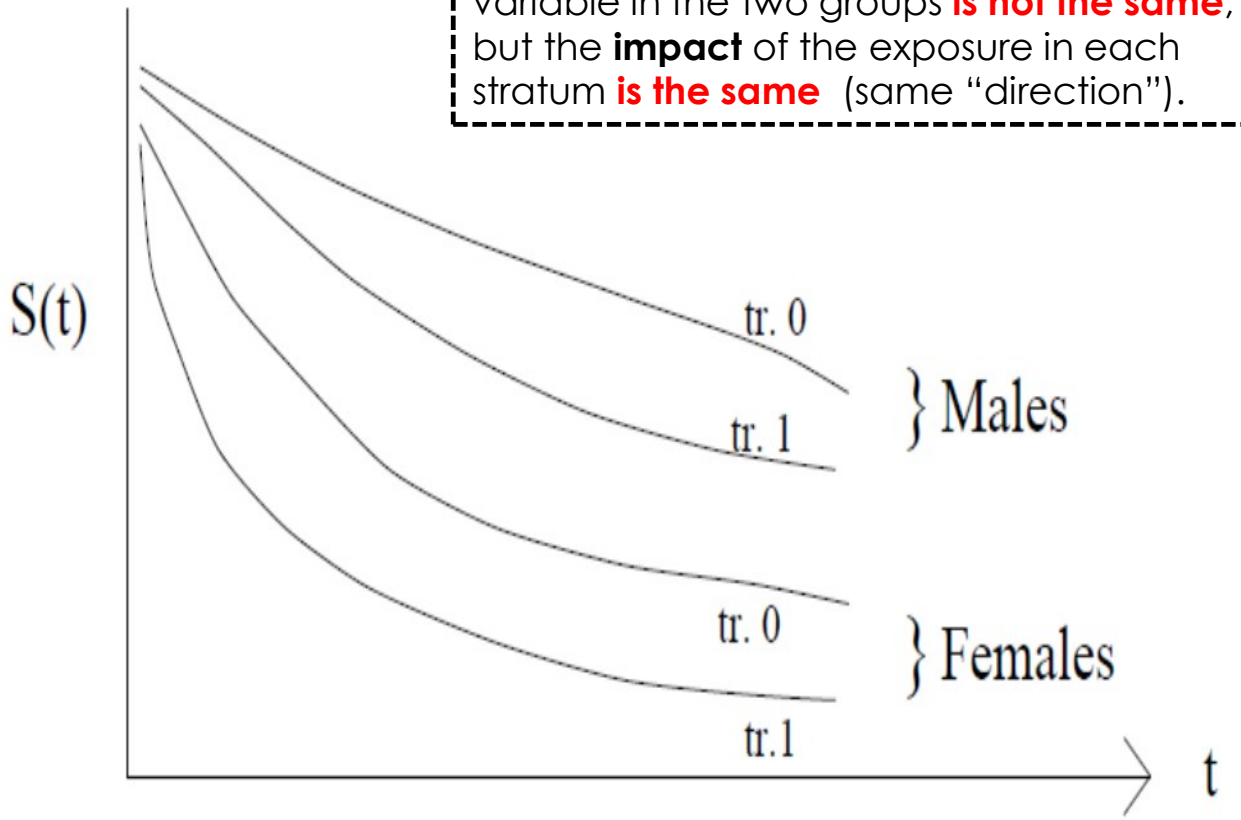
Once the ***expected number of events*** is calculated, then the **total number of subjects** is derived *under some assumptions…*

Stratified Log-rank test

Variation of log rank test:

- Allows controlling for *additional* (*stratified*:categorical) variable [**confounder**]
- Split data into **strata**, based on values of confounder
- Calculate $O-E$ ***within*** strata
- ***Sum*** $O-E$ *across* strata

useful when the distribution of the stratum variable in the two groups **is not the same**, but the **impact** of the exposure in each stratum **is the same** (same "direction").



Stratified log rank test – Example:

- Remission data
- Stratified variable: 3-level variable (LWBC3) indicating low, medium, or high log white blood cell count (coded 1, 2, and 3, respectively)

Treated Group: rx=0 Placebo Group: rx=1

```
->lwbc3 = 1

       |    Events      Events
rx     |  observed    expected
-------+----------------------
0      |       0        2.91
1      |       4        1.09
-------+----------------------
Total  |       4        4.00
```

```
->lwbc3 = 2

       |    Events      Events
rx     |  observed    expected
-------+----------------------
0      |       5        7.36
1      |       5        2.64
-------+----------------------
Total  |      10       10.00
```

```
->lwbc3 = 3

       |    Events      Events
rx     |  observed    expected
-------+----------------------
0      |       4        6.11
1      |      12        9.89
-------+----------------------
Total  |      16       16.00
```

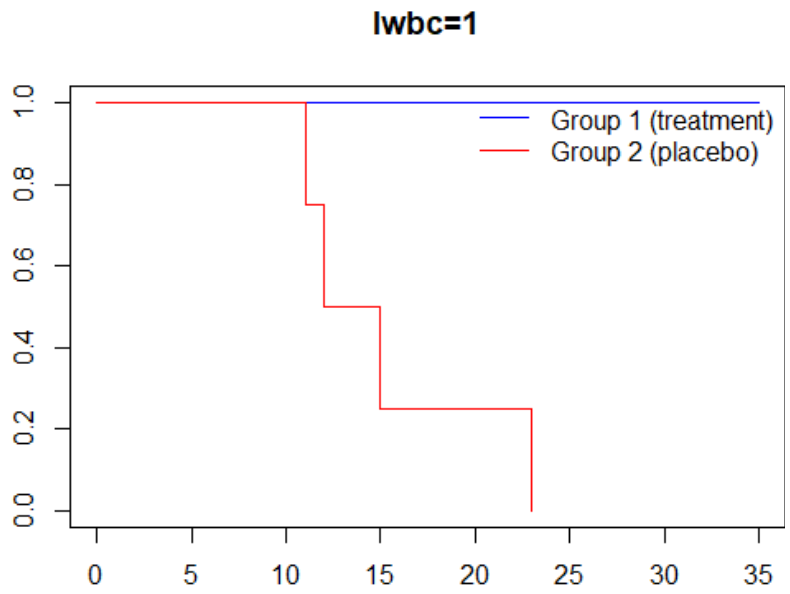
Recap: Non-stratified test : χ2-value of 16.79 and corresponding p-value rounded to 0.0000

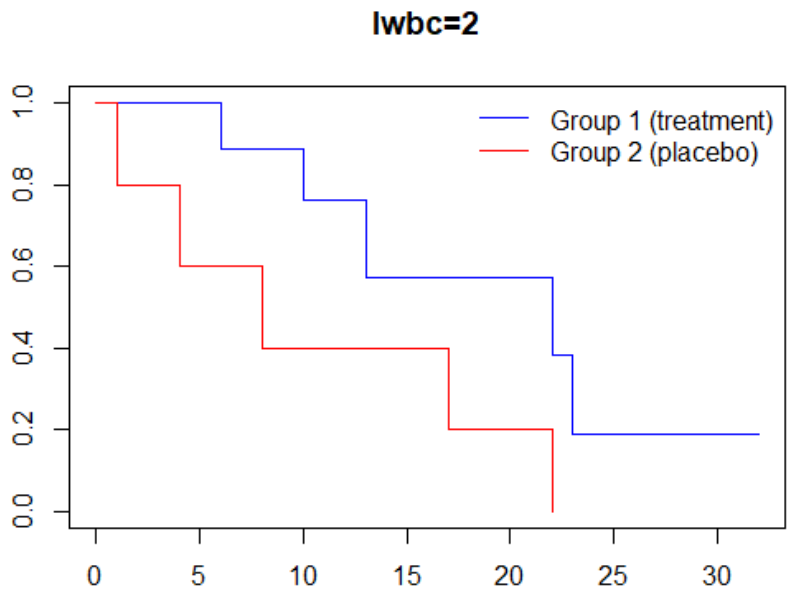
```
Call:
survdiff(formula = Surv(time, status) ~ treatment)

              N Observed Expected (O-E)^2/E (O-E)^2/V
treatment=1 21        9     19.3      5.46      16.8
treatment=2 21       21     10.7      9.77      16.8

 Chisq= 16.8  on 1 degrees of freedom, p= 4e-05
```

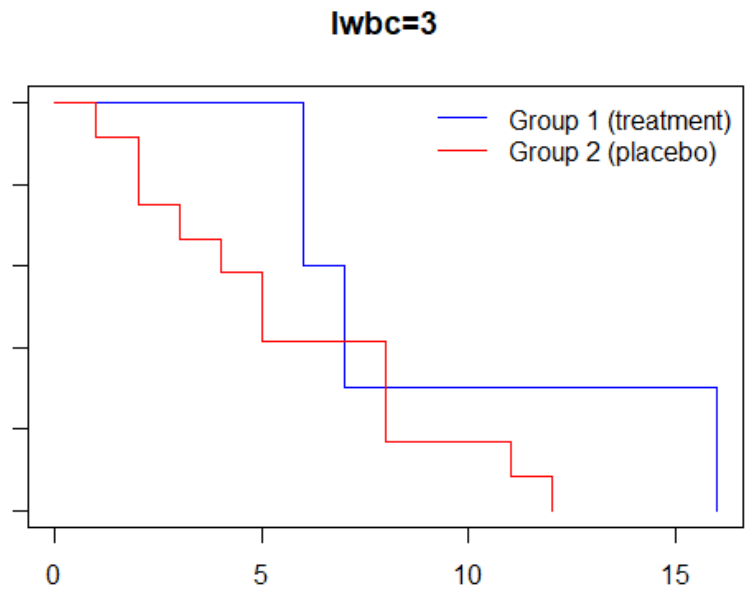
Block 4.2

```
fit <- survdiff(Surv(data$V1,data$V2)~data$V5+strata(lwbc3))
fit
Call:
survdiff(formula = Surv(data$V1, data$V2) ~ data$V5 + strata(lwbc3))

             N Observed Expected (O-E)^2/E (O-E)^2/V
data$V5=0 21         9     16.4      3.33      10.1
data$V5=1 21        21     13.6      4.00      10.1

 Chisq= 10.1  on 1 degrees of freedom, p= 0.001
```

Always significant, same direction of the effect, but **magnitude** of the effect varies across strata (varying sample size..)







Stratified vs. unstratified approach

Log rank unstratified *

$$O_i - E_i = \sum_j (m_{ij} - e_{ij})$$

i = group #, j = jth failure time

Log rank stratified *

$$O_i - E_i = \sum_s \sum_j (m_{ijs} - e_{ijs})$$

i = group #, j = jth failure time,
s = stratum #

Limitations:

- Sample size may be **small** within strata
- **Categorical** stratifying variable and exposure
- **Interactions** ?

*At the denominator there is always an estimate of the variance-covariance matrix

LG test for *Several* Groups

H_0: **All** survival curves are the same

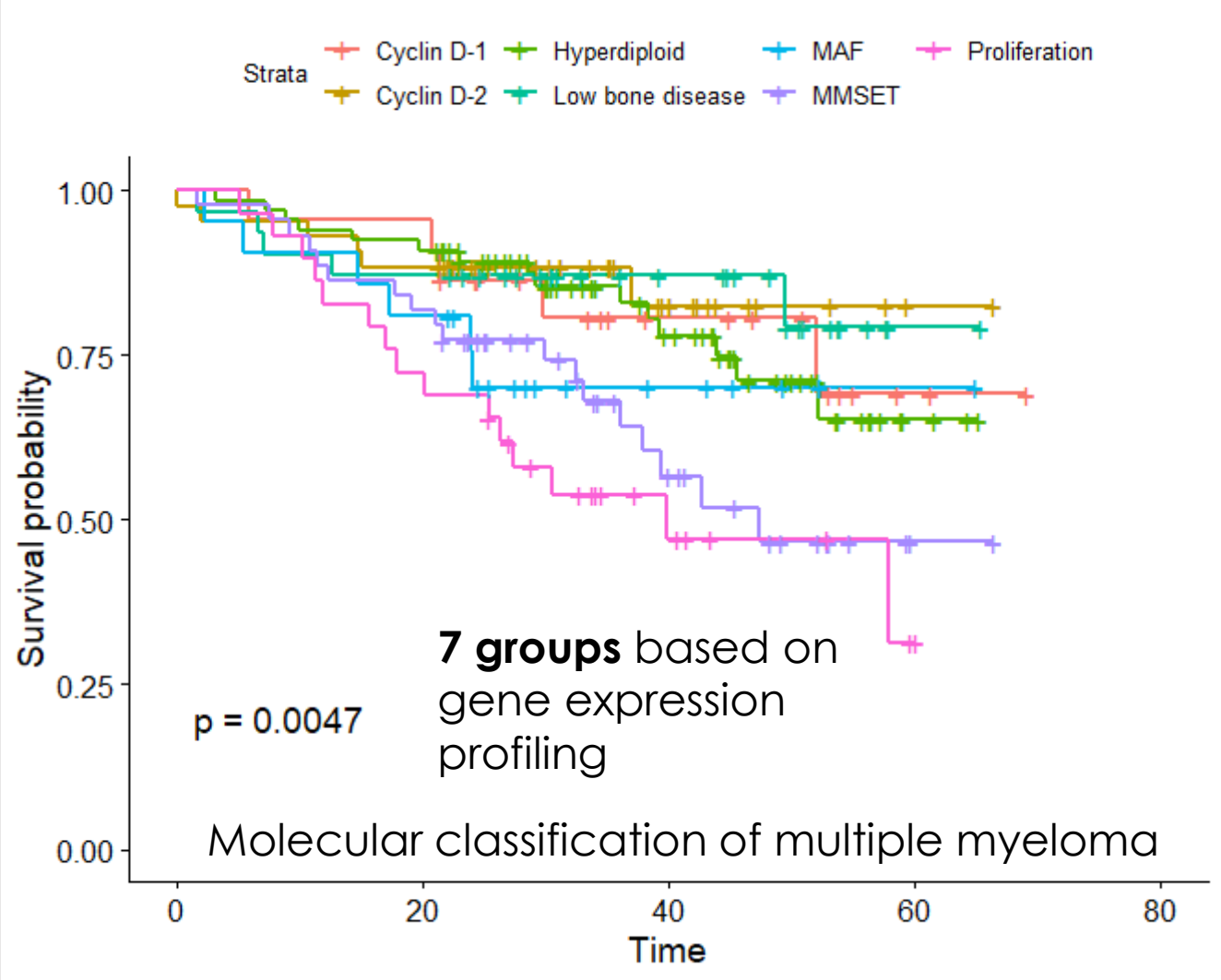
- Suppose we have $K > 2$ groups and we wish to simultaneously compare them with respect to survival time distributions (or equivalently, hazards)

$$H_0: \quad \lambda_1(t) = \lambda_2(t) = \ldots \lambda_K(t), \text{ for all } t > 0$$

(i.e. the survival curves for the all groups are equal everywhere)

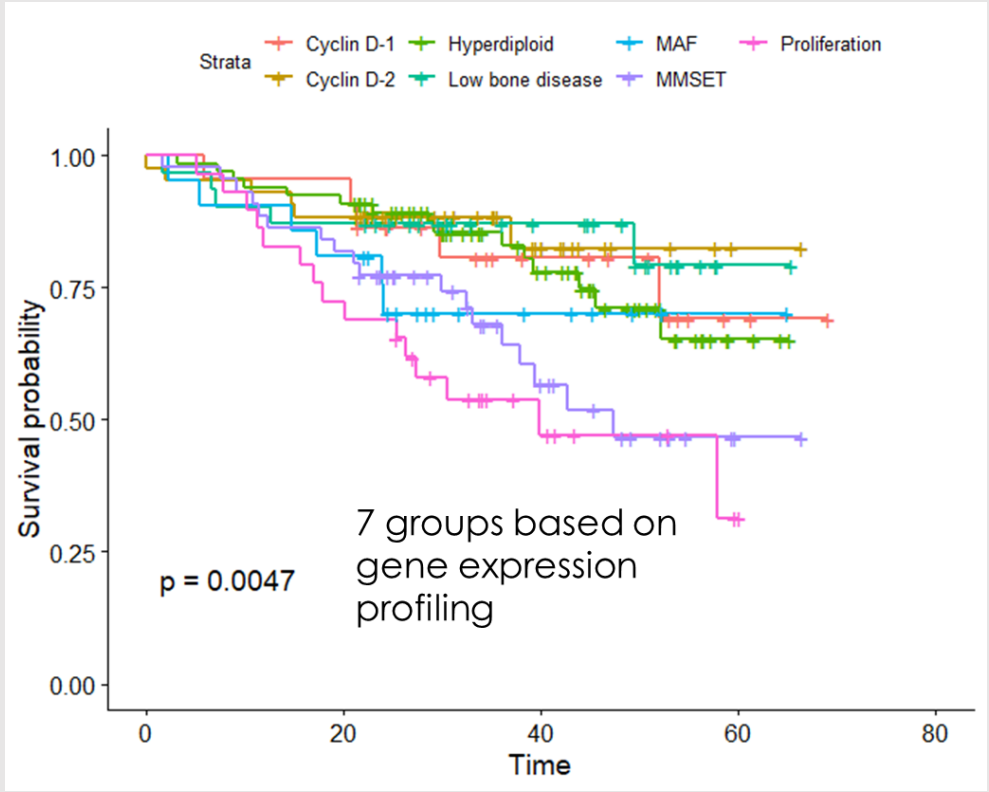
- We are particularly concerned with the alternatives

$$H_A: \quad \lambda_k(t) > \lambda_{k'}(t), \text{ for some } t > 0$$
$$\text{or}$$
$$\lambda_k(t) < \lambda_{k'}(t), \text{ for some } t > 0$$
$$\text{for at least some } k \neq k'$$



- Log-rank statistic for > 2 groups involves computing variances and covariances of $Oi-Ei$
- G (≥2) groups: log-rank statistic ~$\chi 2$ with $G-1$ df

Pairwise comparisons between group levels with corrections for **multiple testing issue [alpha inflation…]**



R function: pairwise_survdiff {survminer}

	Cyclin D-1	Cyclin D-2	Hyperdiploid	Low bone disease	MAF	MMSET
Cyclin D-2	0.723	-	-	-	-	-
Hyperdiploid	0.943	0.723	-	-	-	-
Low bone disease	0.723	0.988	0.644	-	-	-
MAF	0.644	0.447	0.523	0.485	-	-
MMSET	0.328	0.103	0.103	0.103	0.723	-
Proliferation	0.103	**0.038**	**0.038**	***0.062***	0.485	0.527

p value adjustment method: BH

→ Various choices for the adjustment method

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

The Cox model assumes that the **hazards are proportional** (PH), which means that the hazard ratio is **constant over time** with different predictor or covariate levels.

This PH assumption in any covariate is quite a strong assumption. Considering the complexity of biological and physiological responses and associations, this assumption has rarely a solid justification.

If PH doesn't exactly hold for a particular covariate but we fit the PH model anyway, then what we are getting is sort of an ***average HR***, averaged over the event times.

The two most common ways to assess the PH assumption are:

- *Visual assessment* by means of *the* ***log-cumulative hazard plot***
- *Testing* of *scaled* ***Schoenfeld residuals***

Eventually, if the non-PH variable is a categorical one, it could make sense using a **stratified** approach

$$h_i(t|X_i) = h_0(t)exp(X_i\beta)$$

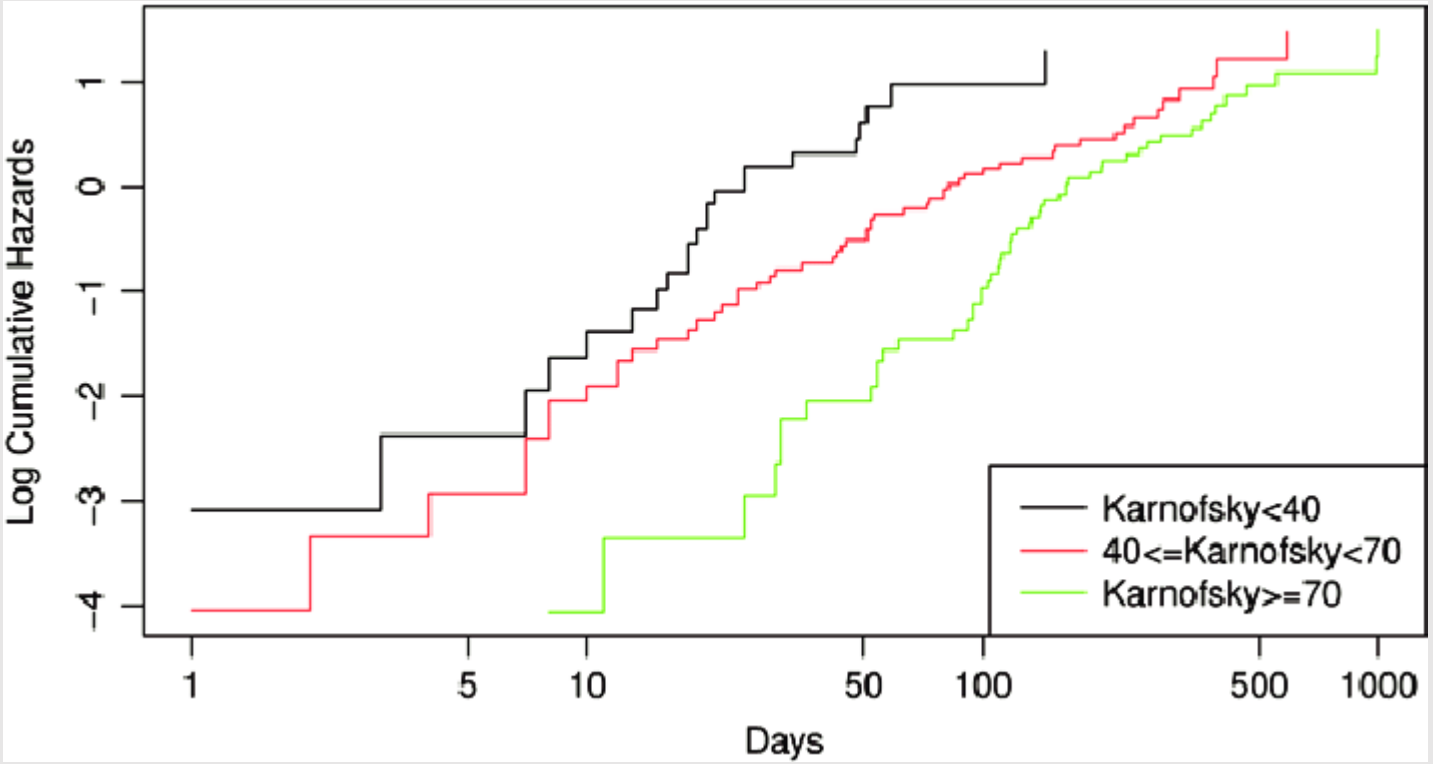
$$\int_0^t h_i(u)du = exp(X_i\beta)\int_0^t h_0(u)du$$

$$H_i(t|X_i) = exp(X_i\beta)H_0(t)$$

Cumulative hazard functions

$$log\big(H_i(t|X_i)\big) = X_i\beta + log\big(H_0(t)\big)$$

If the estimated **log-cumulative hazards** for individuals **with different values** of X (categorical) are plotted against time, the curves will be ***parallel*** if the PH assumption is valid.



- Values of X need to be **categorical/grouped**
- Just a **visual** appreciation

UNITÀ DI BIOSTATISTICA
Dipartimento Universitario Clinico di Scienze Mediche Chirurgiche e della Salute

Just a note about *Schoenfeld* residuals

Time-varying *residuals* from the model are added to the corresponding **time-invariant** coefficient estimate β and smoothed. The result is a **plot** of an estimate of the regression coefficient for the covariate ***over time***. If the plot is ***reasonably flat*** (there is here a formal test), the PH assumption holds.

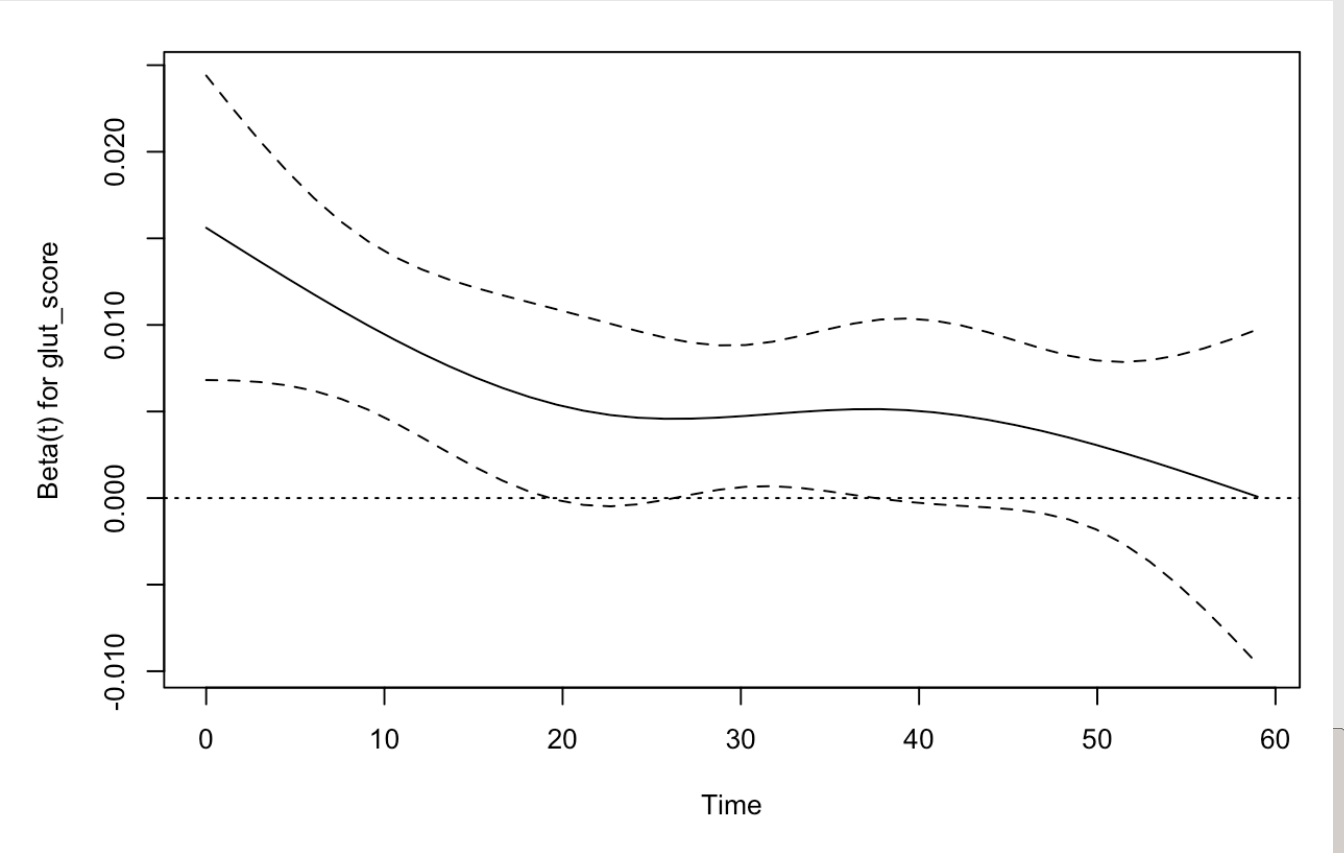
$$s_{k,j}$$

Schoenfeld residual for covariate Xj at time tk

$$E(s_{k,j}) + \hat{\beta}_j \approx \beta_j(t_k)$$

The Schoenfeld residuals are the differences between that individual's covariate values at the event time k and the corresponding risk-weighted average of covariate values among all those at risk at that time.

The word "residual" thus makes sense, as it's the difference between an observed covariate value and what you *might have expected* based on all those at risk at that time.



The Stratified Cox Model

Suppose a confounder C has *k* levels on which we would like **to stratify** when comparing *h(t | E)* and *h(t | not E)* where E is an indicator of "exposure".

$$h_i(t|E) = h_{0i}(t)\exp(E\beta)$$
$$i = 1,..,k$$

1. A [non-parametric] baseline hazard is estimated **within** each stratum (solve ev. non PH hazard)

2. If the confounder is controlled using stratification, there is no way to estimate an **hazard ratio** comparing two levels of the confounder.

3. Stratification generally requires **more data** to obtain the same precision in coefficient estimates