Italian Stata Users Group Meeting

November 12, 2015

Assessment of proportional hazards assumption: restricted mean difference as a potential alternative to the hazard ratio for the analysis of time-to-event endpoint on aggregate data

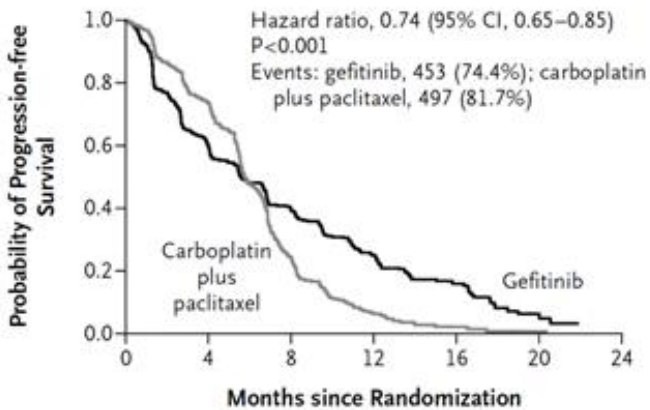
Francesca Ghilotti

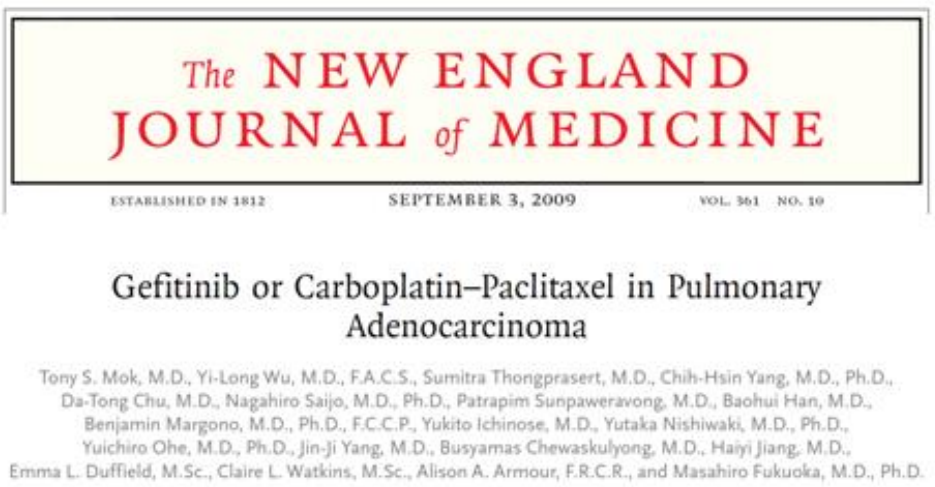
Background

- Survival improvement is an appropriate measure of clinical benefit
- Time-to-event endpoint is the outcome of interest in many oncological clinical studies
- Log-rank and proportional hazards (PH) Cox model are the most common techniques used for analyzing survival time data

Rationale

- The hazards need to be proportional but rarely PH assumption is assessed
- Survival curve convergences and crossings are common in medical research




The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 SEPTEMBER 3, 2009 VOL. 361 NO. 10

Gefitinib or Carboplatin–Paclitaxel in Pulmonary Adenocarcinoma

Tony S. Mok, M.D., Yi-Long Wu, M.D., F.A.C.S., Sumitra Thongprasert, M.D., Chih-Hsin Yang, M.D., Ph.D., Da-Tong Chu, M.D., Nagahiro Saijo, M.D., Ph.D., Patrapim Sunpaweravong, M.D., Baohui Han, M.D., Benjamin Margono, M.D., Ph.D., F.C.C.P., Yukito Ichinose, M.D., Yutaka Nishiwaki, M.D., Ph.D., Yuichiro Ohe, M.D., Ph.D., Jin-Ji Yang, M.D., Busyamas Chewaskulyong, M.D., Haiyi Jiang, M.D., Emma L. Duffield, M.Sc., Claire L. Watkins, M.Sc., Alison A. Armour, F.R.C.R., and Masahiro Fukuoka, M.D., Ph.D.

Objectives

Conduct a **systematic review** to quantify the phenomenon of survival curve convergences and crossings

Propose the use of **meta-regression** as a method to test the **PH assumption** when only **aggregate data** are available

Propose the use of **restricted mean difference** as a potential alternative to the HR in case of non-PH

Systematic review

- Inclusion criteria for the **review**
  - Phase II/III RCTs
  - Advanced non-small-cell lung cancer (NSCLC)
  - Antitumor therapies
  - Time-to-event primary endpoint

Data extraction:

Study design, patient and treatment characteristics, metodological and statistical features

Systematic review

- Inclusion criteria for the **analysis**
  - Number of patients at risk reported at each time-point p
  - At least 3 time-points available

Data extraction:

Survival probabilities from the KM curves at p time-points, number of patients at risk

Estimates of log(HR) and its variance

- Life-table approach
- Censoring uniform within each time interval

$$s^*_{j,i} = s^*_{j,i-1} \cdot \left[1 - \frac{d^*_{j,i}}{n_{j,i-1} - (c^*_{j,i}/2)}\right] \quad (1)$$

$$n_{j,i} = n_{j,i-1} - d^*_{j,i} - c^*_{j,i} \quad (2)$$

Rearranging (1) e (2) gives the number of events $d^*_{j,i}$, the number censored $c^*_{j,i}$ and the number at risk $n^*_{j,i}$ during the interval $[t_{i-1}, t_i)$

more

Williamson, P.R. *Statistics in medicine*, 2002

Estimates of log(HR) and its variance

Logarithm of the Hazard Ratio whitin the i^{th} time interval

$$log(HR)_i = \frac{(d^*_{2,i} - e^*_{2,i})}{v_i}$$

Variance of the log(HR) whitin the i^{th} time interval

$$var(log(HR)_i) = \frac{1}{v_i}$$

more

Assessing the PH assumption

- GRAPHICAL APPROACH
  - **log(-log S)** plot against time

```
twoway (scatter ln_ln1 ln_t, connect(l)) ///
(scatter ln_ln2 ln_t, connect(l))
```

  - **Forest plot** within each study to visualize the relation between the HR and the time of follow-up

```
metan ln_hr se_hr, fixedi eform label(namevar=t)
```

Assessing the PH assumption

- ANALYTICAL APPROACH
  - **Meta-regression** to test for a linear trend with time
    - Outcome: log(HR) at each time-point
    - Explanatory variable: follow-up time
    - Inverse-variance weighting

```
statsby _b e(chi2) e(df_m),by(id):  vwls ln_hr t1,sd(se_hr)

rename _eq2_stat_1 chi2
rename _eq2_stat_2 df_m
gen pvalue=chi2tail(df_m, chi2)
gen z=sqrt(chi2)
list if pvalue<0.1
```

Association between non-PH and study characteristics

- Type of treatment comparison
  - different mechanism of action
  - same mechanism of action

conventional therapy, biologics, tyrosine-kinase inhibitor (TKI), non-conventional target

- Type of endpoint
  - Overall Survival (OS)
  - Progression Free Survival (PFS)

Restricted Mean Survival Time (RMST)

- Select a time-point t^*, up to which we wish to compute the RMST
- For a random time-to-event variable T, we estimate:

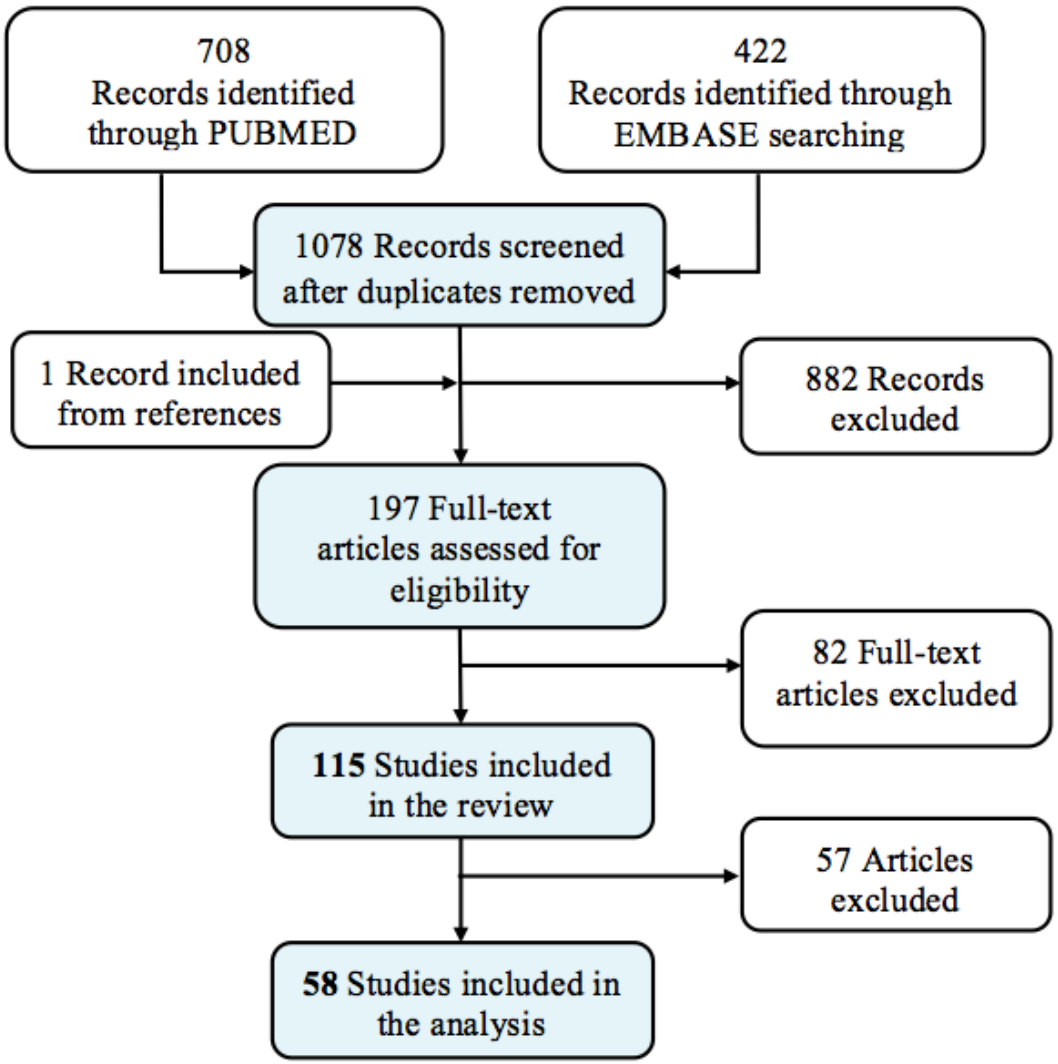
$$\mu(t^*) = E[min(T, t^*)] = \int_0^{t^*} S(t)dt \qquad (3)$$

- Area under the survival curve up to t^*
- Can think of it as the *'t^*-year life expectancy'*
- Difference in RMST between arms could be used as an alternative to the HR

more

Royston, P. and Parmar, M.K. *Statistics in medicine*, 2011

Flow-chart


708
Records identified through PUBMED
422
Records identified through EMBASE searching
1078 Records screened after duplicates removed
1 Record included from references
882 Records excluded
197 Full-text articles assessed for eligibility
82 Full-text articles excluded
115 Studies included in the review
57 Articles excluded
58 Studies included in the analysis

Characteristics of the studies included in the review

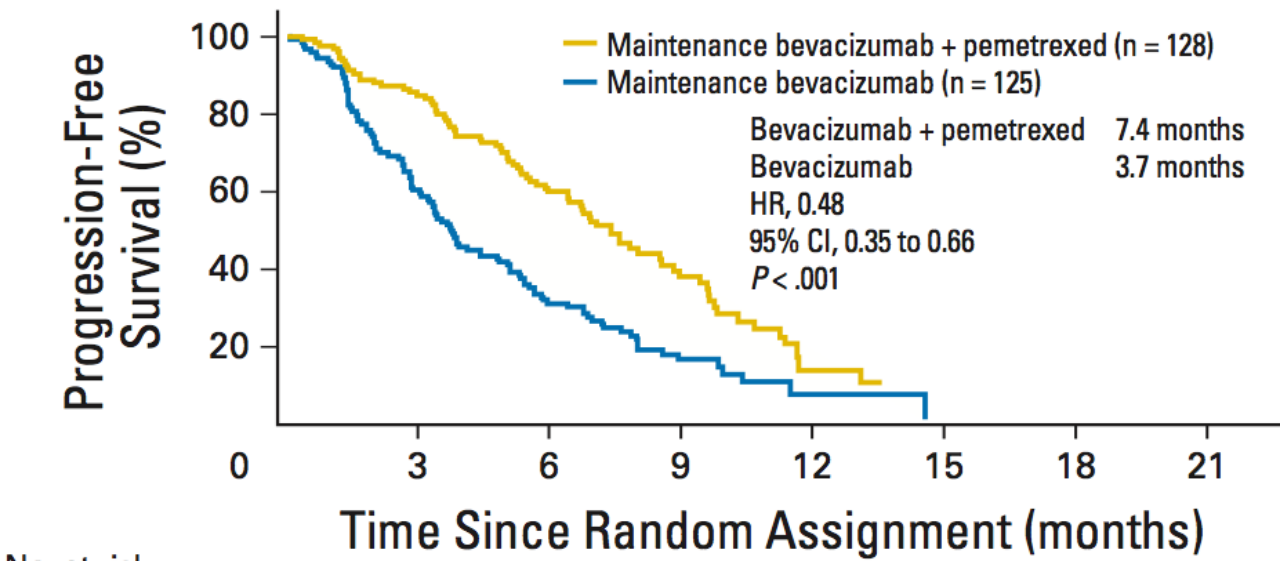
- **Phase:** 33% were phase II studies, 67% were phase III
- **Primary endpoint:** 49% OS, 51% PFS
- **Treatment comparisons:** 41% same mechanism of action, 59% different mechanism
- **Partecipants:** The median number randomized was 332
- **Statistical analysis:** Log-rank test, Cox model

Only 4 (3%) out of 115 studies reported whether PH assumption was satisfied or not

PH assumption assessment

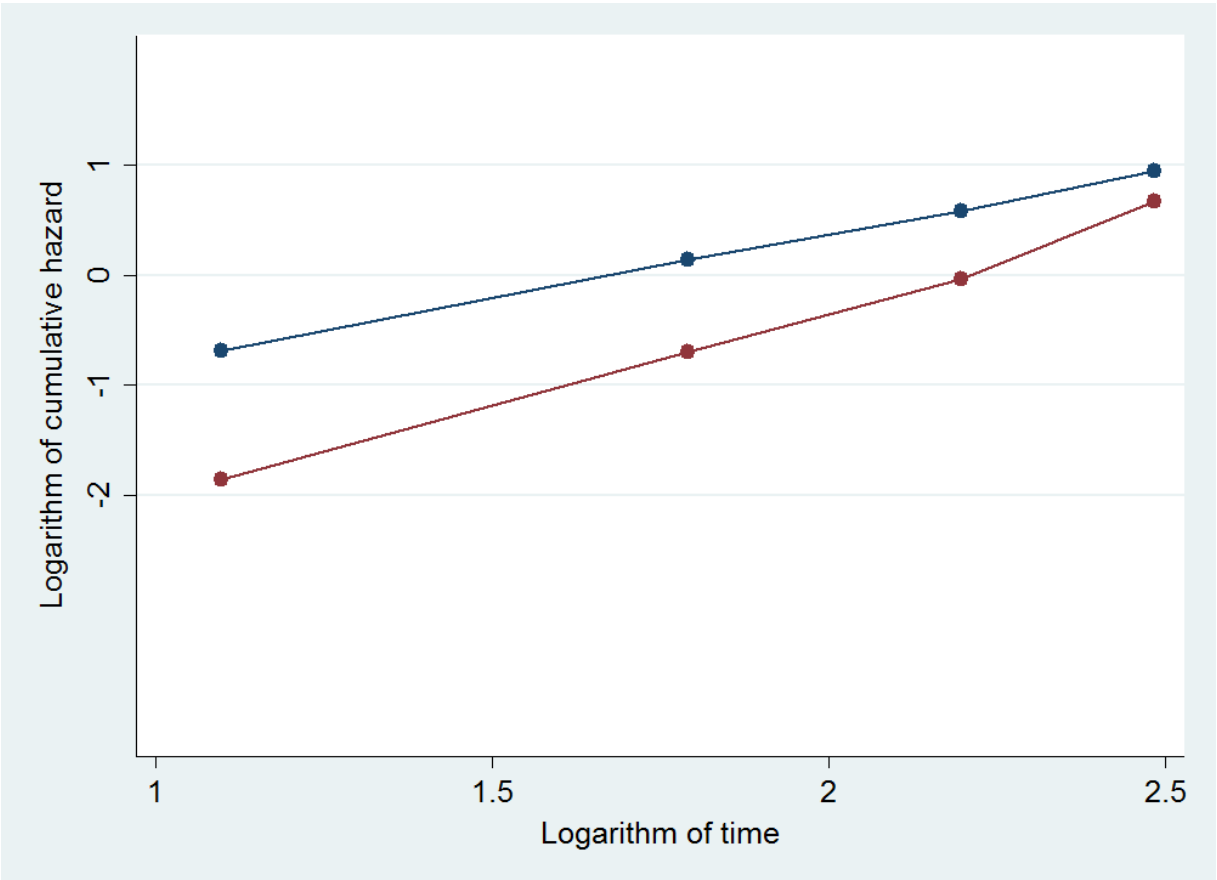
For 12 (19%) out of 62 treatment comparisons non-PH was detected

- Two studies in which PH assumption is violated:



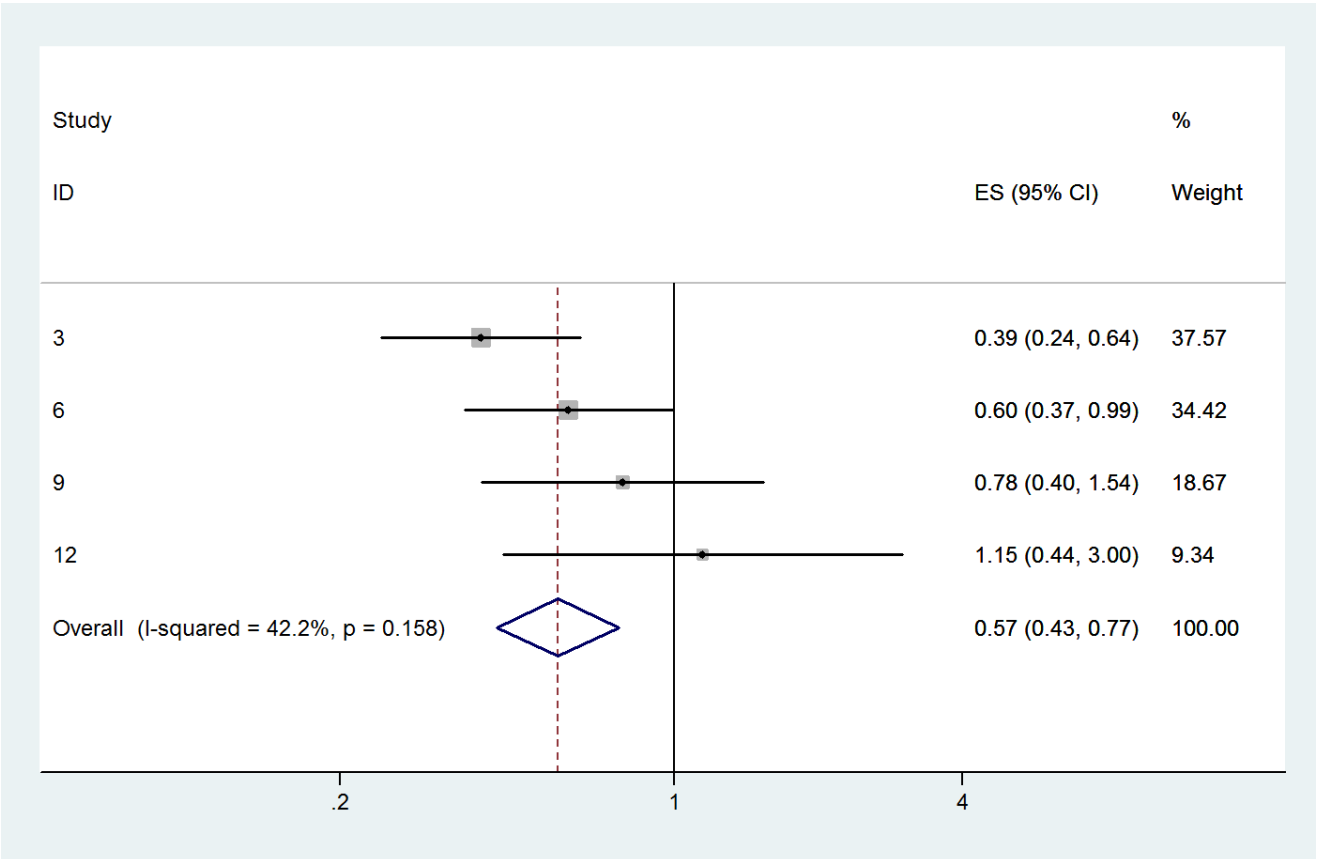
Barlesi, F. *Journal of clinical oncology*, 2013

PH assumption assessment



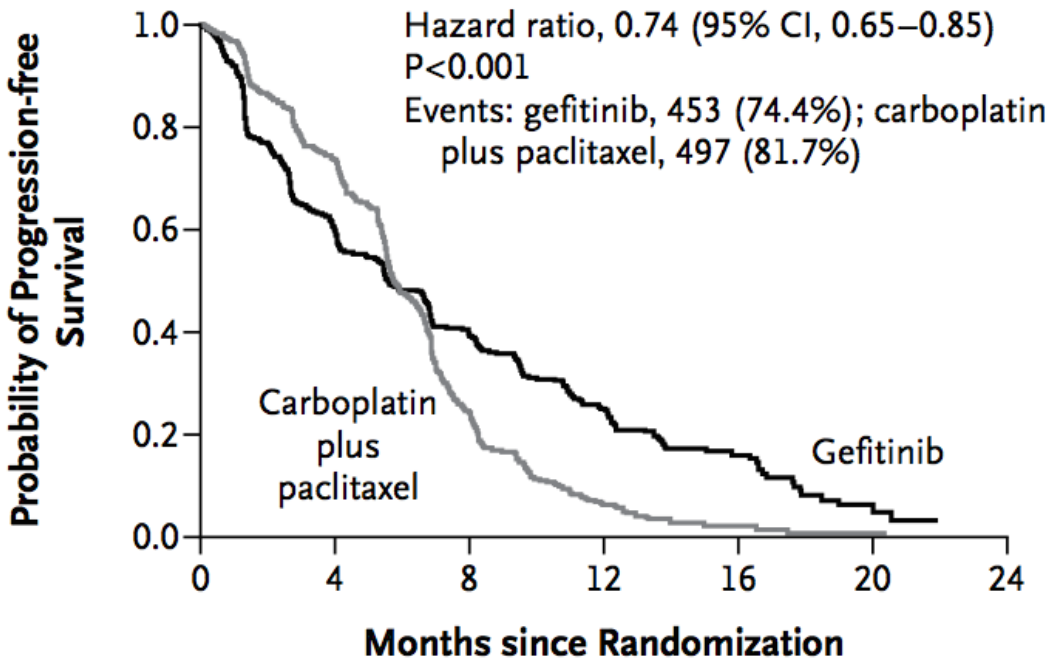
Barlesi, F. *Journal of clinical oncology*, 2013

PH assumption assessment



Barlesi, F. *Journal of clinical oncology*, 2013

PH assumption assessment

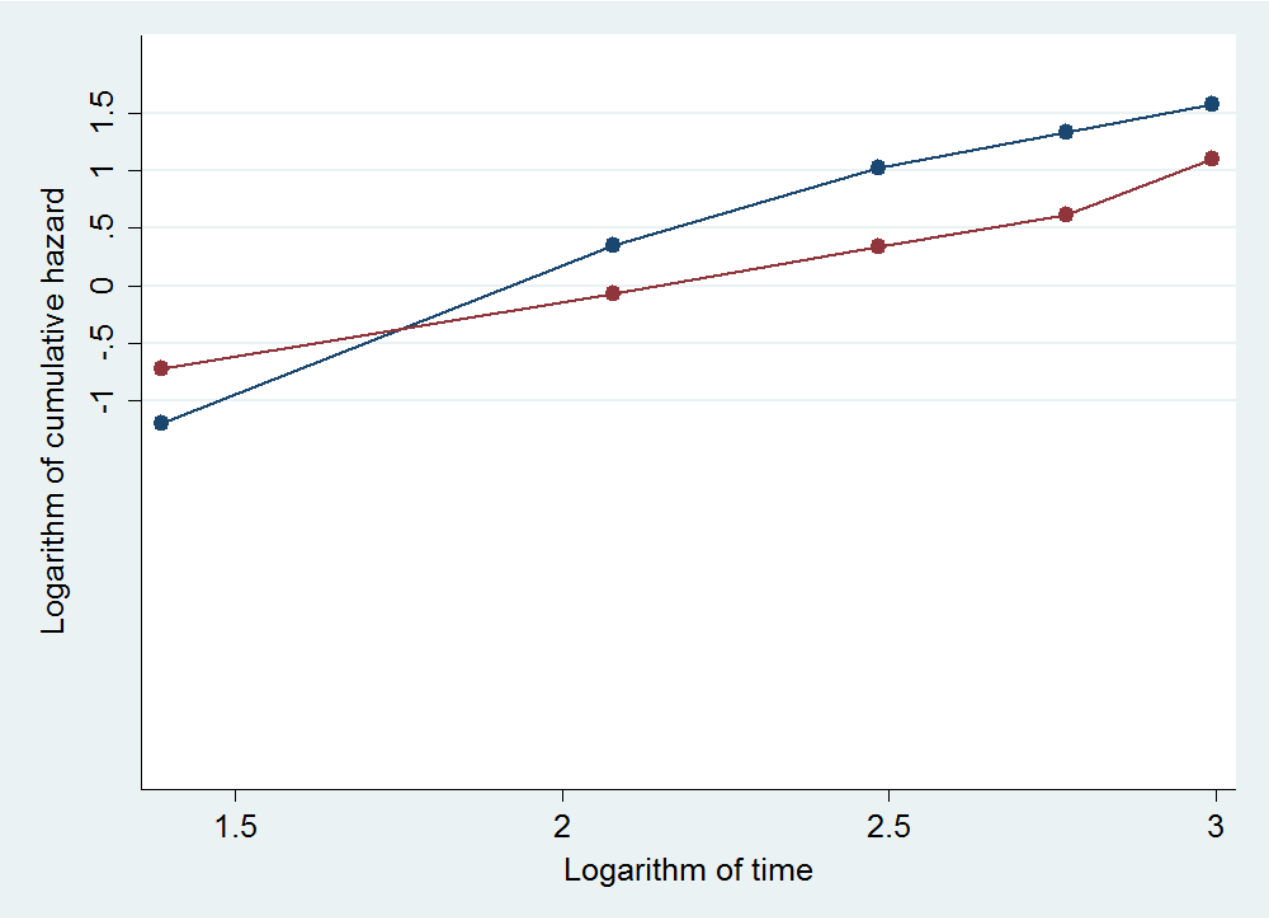


No. at Risk

Gefitinib	609	363	212	76	24	5	0
Carboplatin plus paclitaxel	608	412	118	22	3	1	0

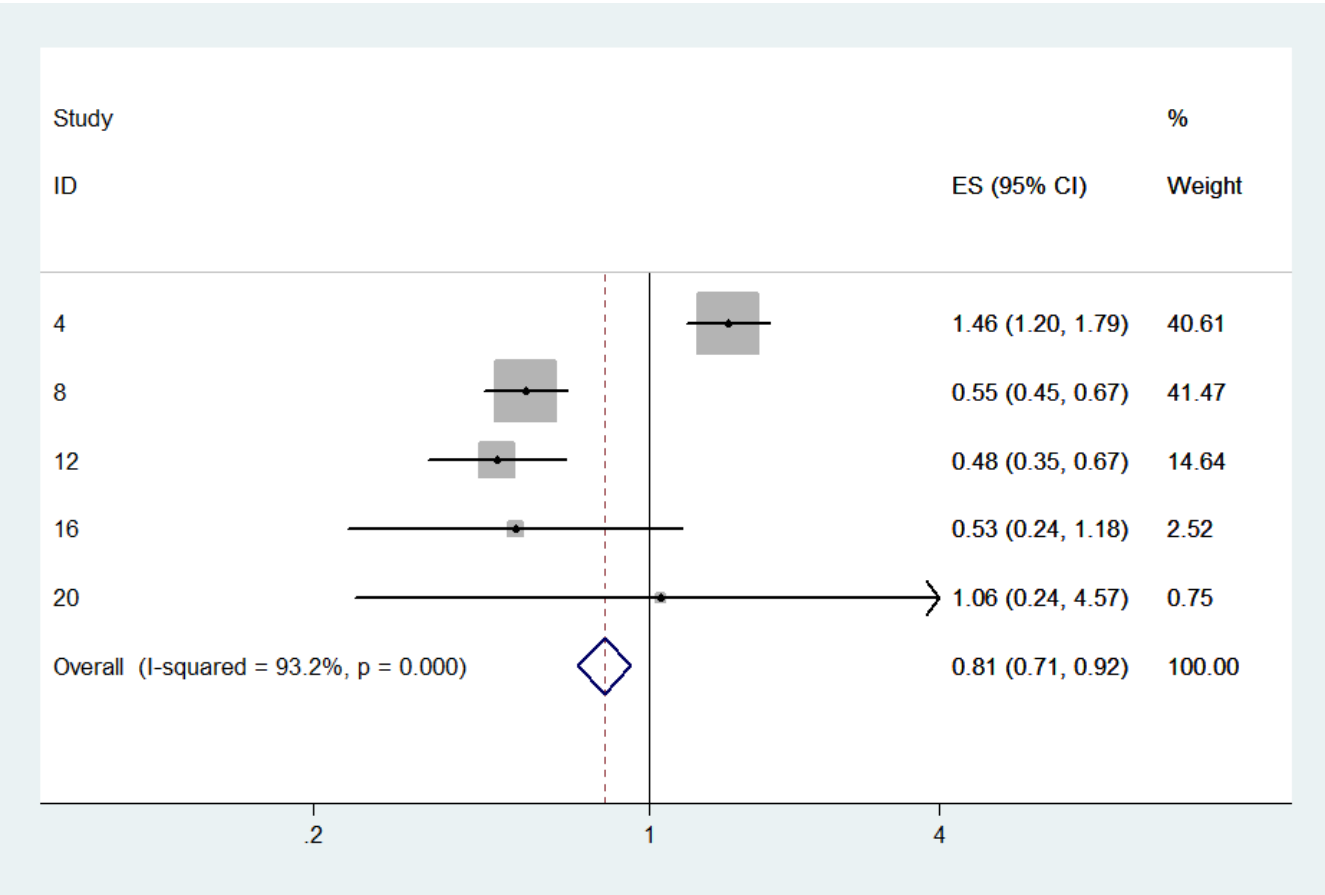
Mok, T.S. *New England Journal of Medicine*, 2009

PH assumption assessment



Mok, T.S. *New England Journal of Medicine*, 2009

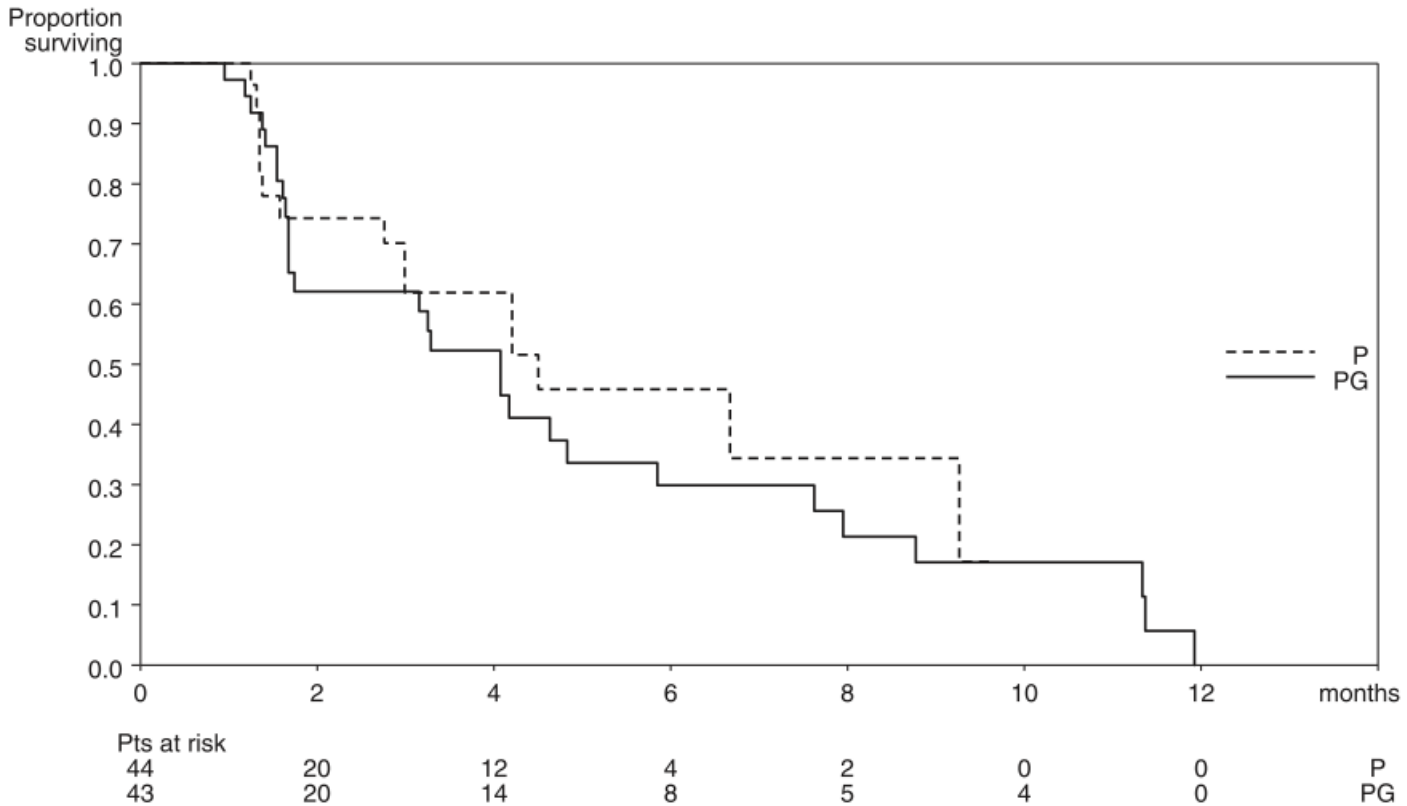
PH assumption assessment



Mok, T.S. *New England Journal of Medicine*, 2009

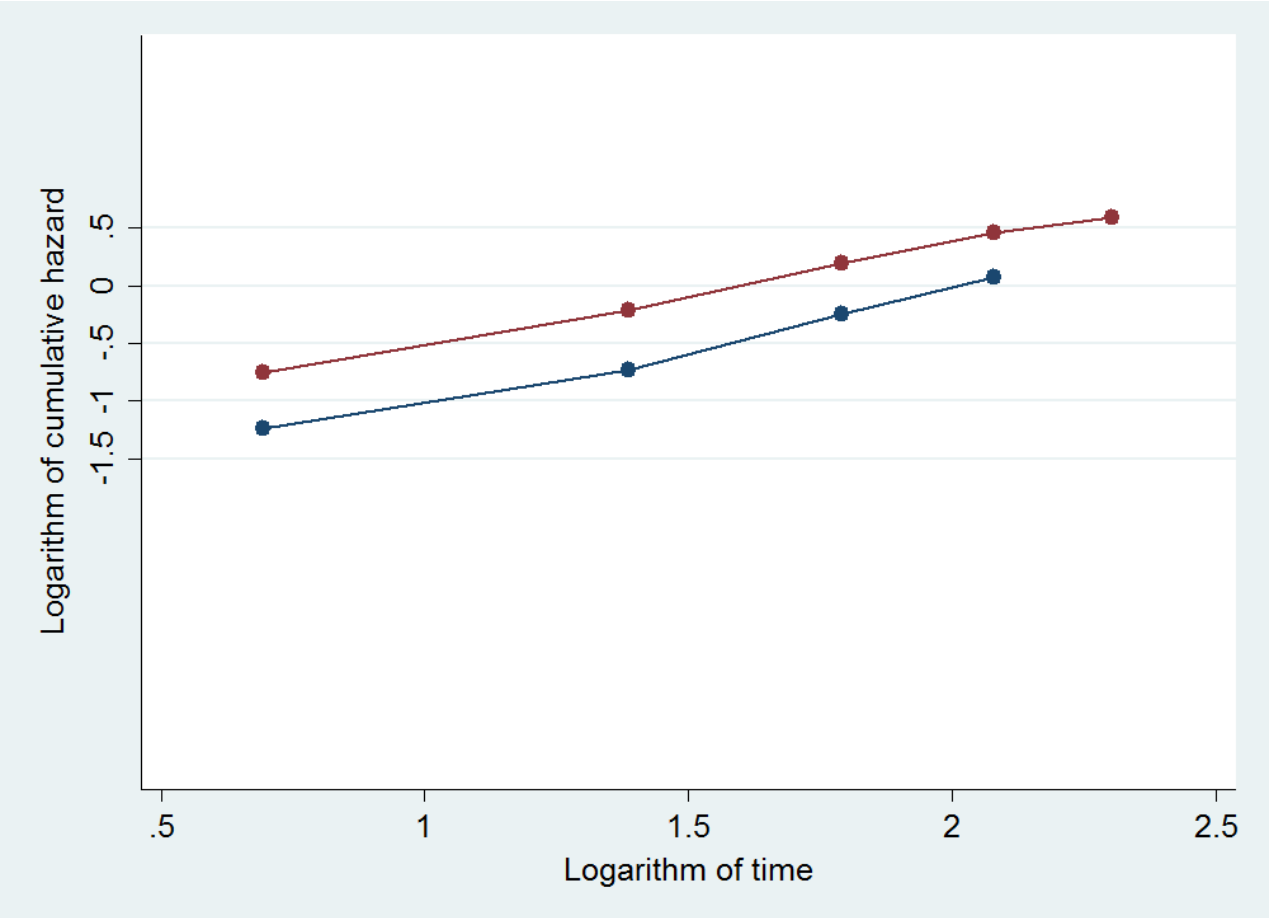
PH assumption assessment

- Two studies in which PH assumption is satisfied:



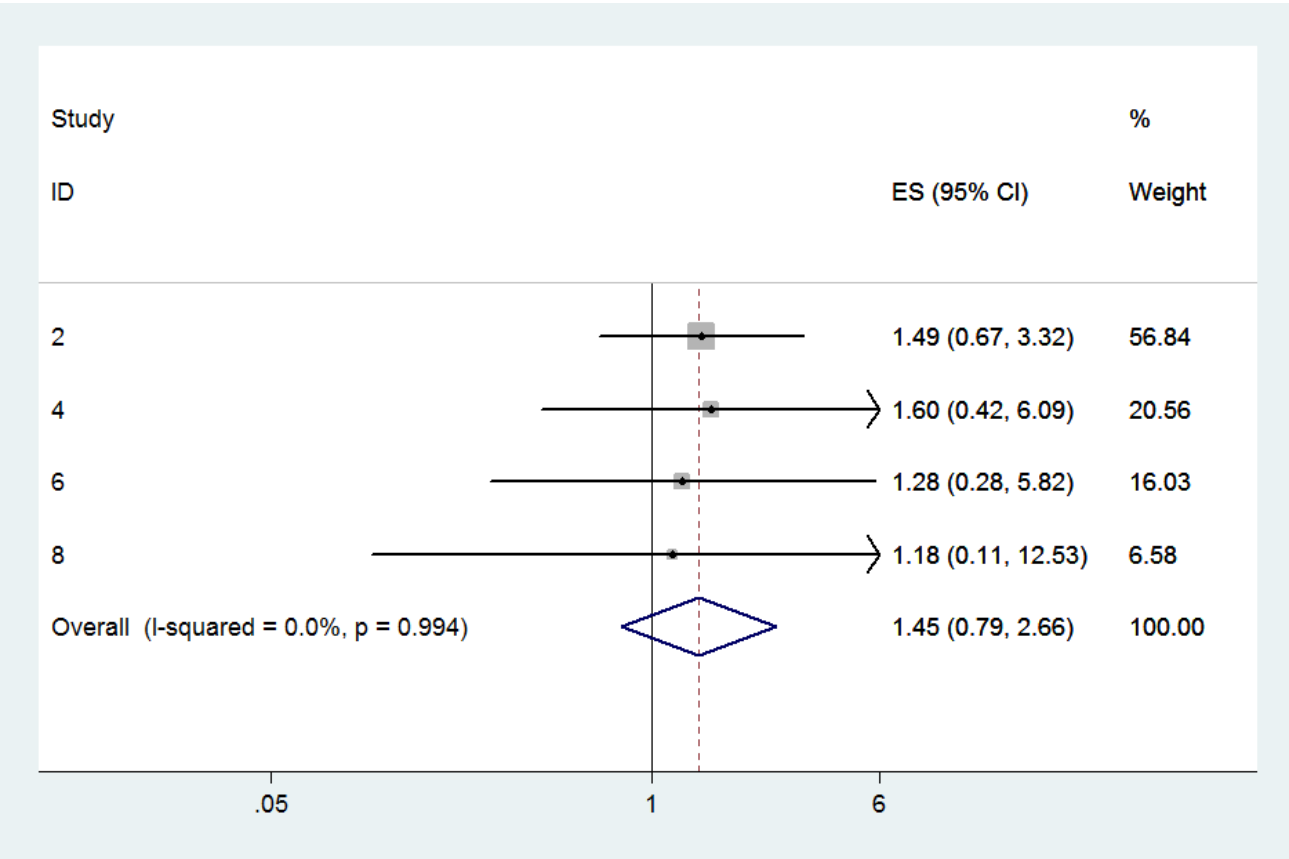
Gridelli, C. *Journal of thoracic oncology*, 2007

PH assumption assessment



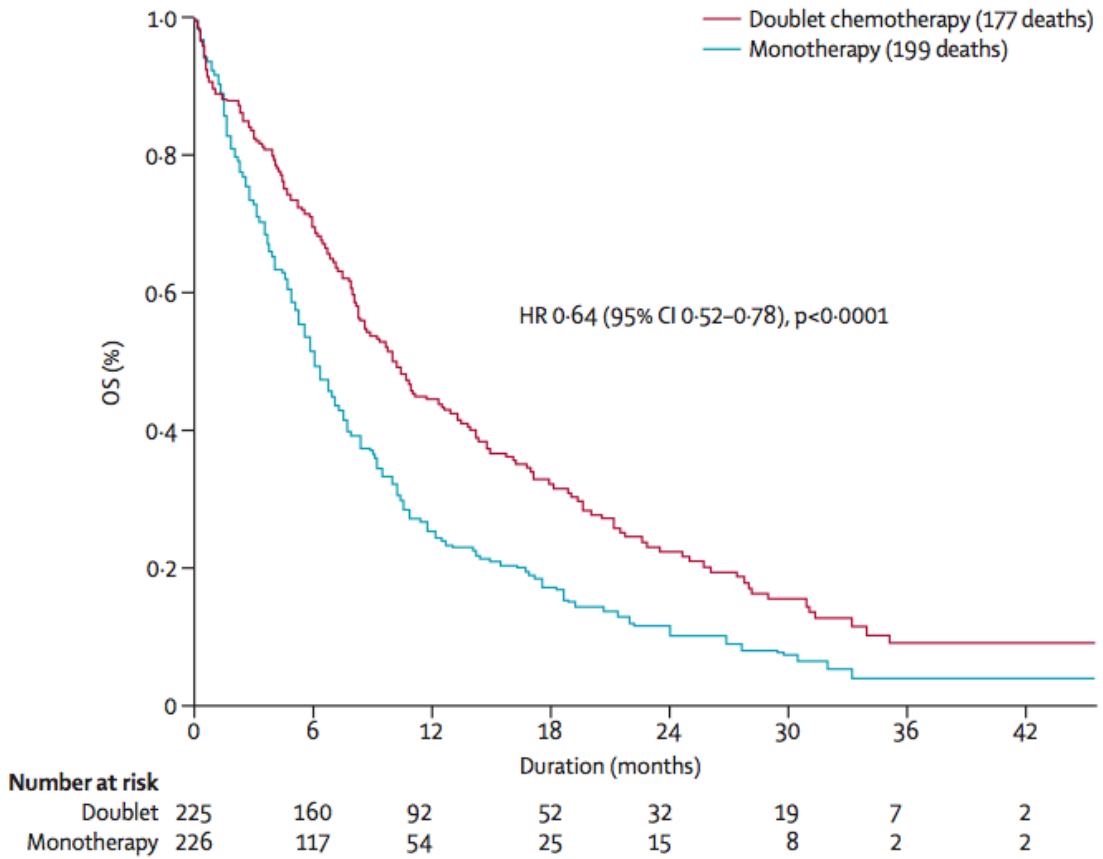
Gridelli, C. *Journal of thoracic oncology*, 2007

PH assumption assessment



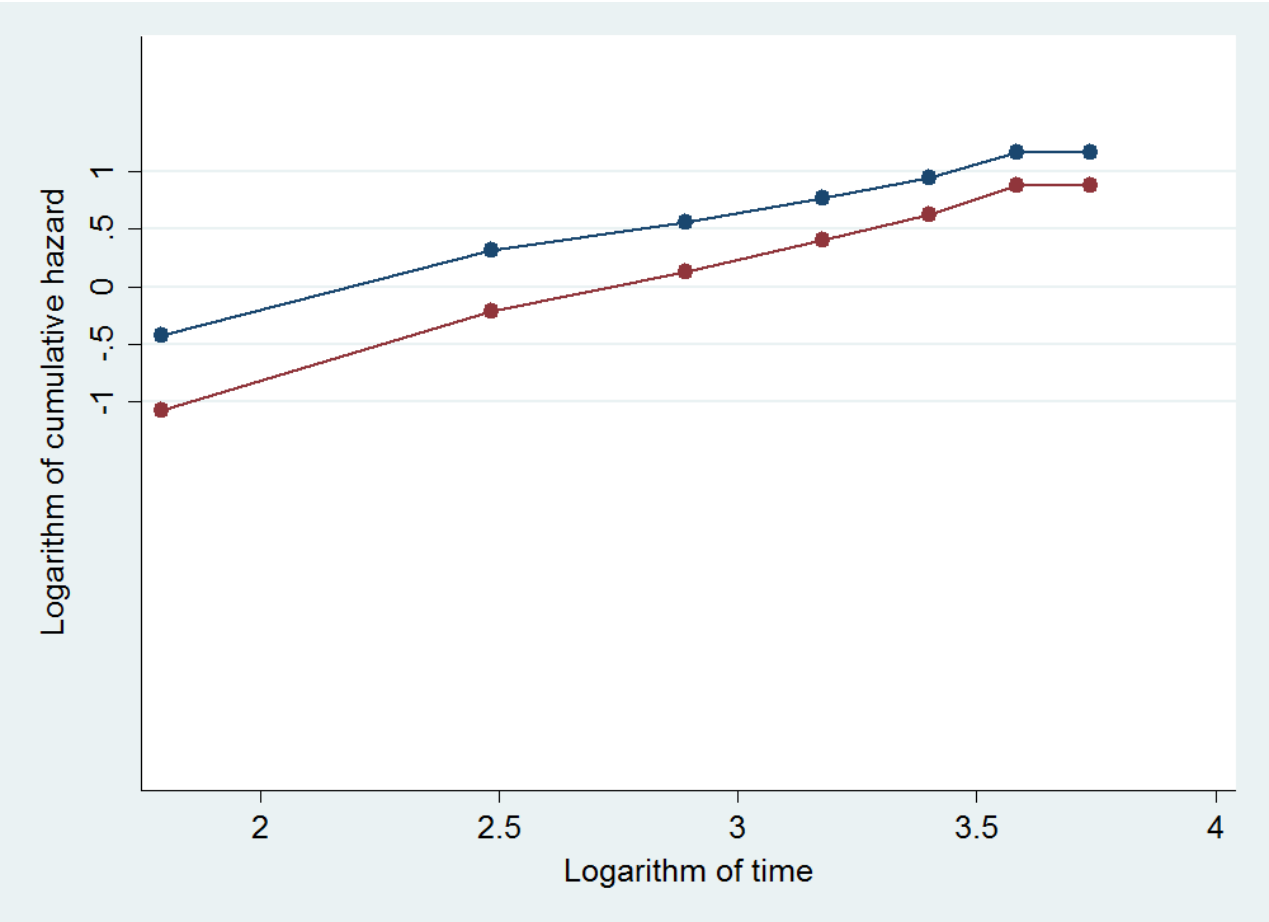
Gridelli, C. *Journal of thoracic oncology*, 2007

PH assumption assessment



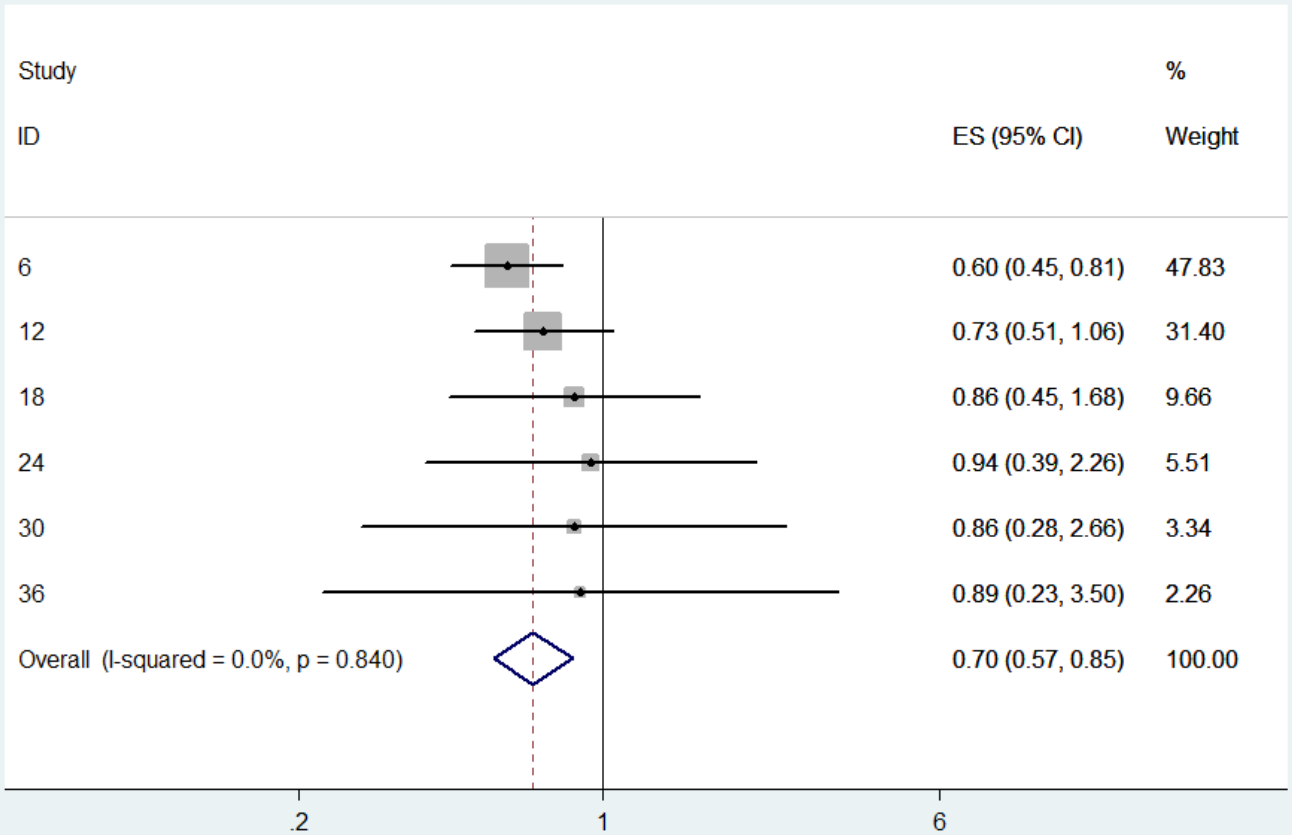
Quoix, E. *The Lancet*, 2011

PH assumption assessment



Quoix, E. *The Lancet*, 2011

PH assumption assessment



Quoix, E. *The Lancet*, 2011

PH assumption results

Table: Association between non-PH and study characteristics

	PH assumption violated		**Fisher's**
	No	Yes	**exact test**
Treatments			
Same treatment comparison	20 (100%)	0 (0%)	
Different treatment comparison	30 (71%)	12 (29%)	0.006
Primary endpoint			
OS	23 (92%)	2 (8%)	
PFS	27 (73%)	10 (27%)	0.101

Table: Comparison between the RMST results and the results reported by authors

Study	RMST results		Median Results			HR Results	
	RMST diff [a]	p-value test Z	Median (control)	Diff (HR)[b]	Diff (KM)[c]	HR	p-value
Wu	6.66	< 0.001	5.6	14.4	8.1	0.28	< 0.001
Solomon	6.13	< 0.001	7.0	8.56	3.9	0.45	< 0.001
Seto	5.13	< 0.001	9.7	8.26	6.3	0.54	0.002
Shaw	3.33	0.004	3.0	3.12	4.7	0.49	< 0.001
Barlesi	2.37	< 0.001	3.7	4.01	3.7	0.48	< 0.001
Lee	1.23	0.15	3.4	1.26	-0.1	0.73	0.04[§]
Jänne	0.82	0.61	5.2	1.30	4.2	0.80	0.21[§]
Reck	0.76	0.018	2.7	0.72	0.8	0.79	0.002
Belani	0.65	0.73	7.1	0.88	0.9	0.89	0.36

[a] Restricted Mean Survival Time difference (months)
[b] Median difference derived from HR (months)
[c] Median difference derived from KM curve (months)
[§] one-sided

Table: Comparison between the RMST results and the results reported by authors

Study	RMST results		Median Results			HR Results	
	RMST diff [a]	p-value test Z	Median (control)	Diff (HR)[b]	Diff (KM)[c]	HR	p-value
Wu	6.66	< 0.001	5.6	14.4	8.1	0.28	< 0.001
Solomon	6.13	< 0.001	7.0	8.56	3.9	0.45	< 0.001
Seto	5.13	< 0.001	9.7	8.26	6.3	0.54	0.002
Shaw	3.33	0.004	3.0	3.12	4.7	0.49	< 0.001
Barlesi	2.37	< 0.001	3.7	4.01	3.7	0.48	< 0.001
Lee	1.23	0.15	3.4	1.26	-0.1	0.73	0.04[§]
Jänne	0.82	0.61	5.2	1.30	4.2	0.80	0.21[§]
Reck	0.76	0.018	2.7	0.72	0.8	0.79	0.002
Belani	0.65	0.73	7.1	0.88	0.9	0.89	0.36

[a] Restricted Mean Survival Time difference (months)
[b] Median difference derived from HR (months)
[c] Median difference derived from KM curve (months)
[§] one-sided

Table: Comparison between the RMST results and the results reported by authors

Study	RMST results		Median Results			HR Results	
	RMST diff [a]	p-value test Z	Median (control)	Diff (HR)[b]	Diff (KM)[c]	HR	p-value
Wu	6.66	< 0.001	5.6	14.4	8.1	0.28	< 0.001
Solomon	6.13	< 0.001	7.0	8.56	3.9	0.45	< 0.001
Seto	5.13	< 0.001	9.7	8.26	6.3	0.54	0.002
Shaw	3.33	0.004	3.0	3.12	4.7	0.49	< 0.001
Barlesi	2.37	< 0.001	3.7	4.01	3.7	0.48	< 0.001
Lee	1.23	0.15	3.4	1.26	-0.1	0.73	0.04[§]
Jänne	0.82	0.61	5.2	1.30	4.2	0.80	0.21[§]
Reck	0.76	0.018	2.7	0.72	0.8	0.79	0.002
Belani	0.65	0.73	7.1	0.88	0.9	0.89	0.36

[a] Restricted Mean Survival Time difference (months)
[b] Median difference derived from HR (months)
[c] Median difference derived from KM curve (months)
[§] one-sided

Pros and Cons

- **New!** Assess the PH assumption using aggregate data
- Conclusions are in line with the log-log plots and with the results reported by authors

- Data constrained by the quality of figures
- Assumption about the mechanism of censoring
- Only studies with patients at risk reported

Future research

- Compare the conclusions obtained with individual patient data (IPD) and with aggregate data
- Investigate how many time-points are needed

Thank you for your attention!

Joined work with:

Rino Bellocco [1]

Eliana Rulli [2]

Valter Torri [2]

1 Karolinska Institutet, University of Milano-Bicocca

2 Mario Negri Institute for Pharmacological Research

Appendix

Numbers at risk during a time interval are:

$$n^*_{j,i} = \frac{(n_{j,i-1} + n_{j,i}) \cdot s^*_{j,i-1}}{(s^*_{j,i-1} + s^*_{j,i})} \tag{4}$$

Number of events during a time interval is:

$$d^*_{j,i} = \frac{(n_{j,i-1} + n_{j,i}) \cdot (s^*_{j,i-1} - s^*_{j,i})}{(s^*_{j,i-1} + s^*_{j,i})} \tag{5}$$

Numbers censored during a time interval are:

$$c^*_{j,i} = \frac{2 \cdot (n_{j,i-1} \cdot s^*_{j,i} - n_{j,i} \cdot s^*_{j,i-1})}{(s^*_{j,i-1} + s^*_{j,i})} \tag{6}$$

Back to 7.

Appendix

$$log(HR)_i = \frac{(d^*_{2,i} - e^*_{2,i})}{v_i} \tag{7}$$

$$var(log(HR)_i) = \frac{1}{v_i} \tag{8}$$

where

$$e^*_{2,i} = (d^*_{2,i} + d^*_{1,i}) \cdot \frac{(n^*_{2,i})}{(n^*_{2,i} + n^*_{1,i})} \tag{9}$$

$$v_i = (d^*_{2,i} + d^*_{1,i}) \cdot \frac{n^*_{2,i} \cdot n^*_{1,i}}{(n^*_{2,i} + n^*_{1,i})^2} \tag{10}$$

Back to 8.

Appendix

The area under the curve for group j was estimated by:

$$\mu_j = \sum_{i=1}^{p} \mu_{j,i} = \sum_{i=1}^{p} \frac{(s^*_{j,i-1} + s^*_{j,i}) \cdot (t_i - t_{i-1})}{2} \qquad (11)$$

To estimate the variability of this quantity the formula reported by Klein was used:

$$V(\mu_j) = \sum_{i=1}^{p} \left[\int_{t_i}^{t^*} S(t)dt \right]^2 \cdot \frac{d_i}{n_i \cdot (n_i - d_i)} \qquad (12)$$

Back to 12.