

Methods to estimate the survival benefit from individual patient data meta-analysis.

Impact on the health economic outcome.

Lueza B^{1,2,4}, Mauguen A², Pignon JP², Le Pécoux C³, Bonastre J¹

¹ Institut Gustave Roussy, Health Economics unit, Villejuif;

² Institut Gustave Roussy, Meta-analyses unit, Villejuif;

³ Institut Gustave Roussy, Oncology-Radiotherapy department, Villejuif;

⁴ ENSAI, Bruz;

Abstract

In economic evaluation, different parametric or non-parametric methods can be used to assess the survival benefit from individual patient data. A challenge, when using individual patient data from a meta-analysis to estimate the survival benefit, is to take into account stratification on clinical trials. Our objective was to compare five different methods to estimate survival benefit and to investigate the impact on the health economic outcome.

We chose to compare four non parametric methods classically used in meta-analyses for survival endpoints (the Kaplan-Meier, the Stewart and the Peto methods) and a parametric method (Weibull) to estimate survival functions. We used individual patient data from the Meta-Analysis of Radiotherapy in Lung Cancer (MAR-LC) in locally advanced non-small cell lung cancer patients. Our data set consisted of 10 phase III trials comparing modified RT to conventional RT in 2,000 patients.

Depending on the chosen method, mean ICERs ranged from 10,320€ to 34,937€ per life year gained. With a willingness to pay for one life year lower than 50,000€, the five acceptability curves led to different conclusions.

The choice of a survival method to estimate health benefit strongly impacts on the health economic results.

Key words: economic evaluation, meta-analysis, cost-effectiveness, survival benefit

Introduction

Meta-analyses aim to estimate an overall treatment effect generally expressed as a pooled hazard ratio combining the results of all randomised controlled trials (RCTs) that studied the same clinical question (Pignon, 2001). The estimation of the pooled hazard ratio is a weighting average of the estimation of treatment effect in each trial. This stratified analysis ensures that only comparable patients, especially in terms of baseline risk, are compared. The principle of the so-called stratification on clinical trials is that the patients in the experimental group of a trial are only compared with the ones in the control group of the same trial (Buyse, 2000).

Some authors insisted on the usefulness of cost-effectiveness analyses alongside RCTs (Gray, 2006). Trial-based evaluations allow dealing with patient heterogeneity and building and validating extrapolation models. The level of scientific evidence seems higher in trials-based evaluations which use real-world data. Therefore, we can expect that economic evaluations performed on individual patient data meta-analyses should have an even higher level of evidence as the data are available for every patients enrolled in different RCTs. However, such economic evaluations raise methodological issues surrounding the estimation of the health benefit. In economic evaluation, the purpose is to combine the risk reduction and the time in order to evaluate the number of life years saved by the treatment. A commonly used effectiveness outcome in economic evaluations of cancer interventions is the survival benefit (Mittmann, 2009; Neymark, 2002). There are many possible methods to estimate this outcome when considering data from a single RCT (Davies, 2011; Barker, 2009; Tappenden, 2006). In case of patient data from a meta-analysis, additional issues must be addressed as for example the stratification on trial.

The results of economic evaluations are likely to be sensitive to the estimation of the survival benefit of an intervention. Several recent papers investigated this issue considering different

parametric methods to extrapolate survival data from RCTs (Latimer, 2011; Connock, 2011; Annemans, 2011). To our knowledge, no previous study compared parametric and non parametric methods to estimate the survival benefit and the implication on the cost-effectiveness ratio. In this paper, we chose to compare five methods of estimation of the survival benefit and to investigate the impact on the health economics outcome. To study this question we used data from the Meta-Analysis of Radiotherapy in Lung Cancer (MAR-LC) which compares conventional radiotherapy (RT) with modified radiotherapy.

Data

Effectiveness

We had access to the patient-level data from the Meta-Analysis of Radiotherapy in Lung Cancer (MAR-LC), performed in the Meta-Analysis Unit at the Gustave Roussy Institute. The database included 2,000 patients with a non-metastatic non-small cell lung cancer treated with radiotherapy and who had been enrolled in 10 phase III RCTs (Le Péchoux, submitted). MAR-LC trials compared conventional RT with modified RT. Modified RT included hyperfractionated RT consisting in increasing the number of fractions per day with a decrease of dose per fraction, and/or accelerated RT, in which the overall treatment time was reduced. MAR-LC primary outcome measure was overall survival. Secondary outcome measures were progression-free survival, late and acute toxicities. The treatment effect was expressed as a pooled hazard ratio based on the Peto's estimate (Yusuf, Peto, 1985) using a fixed-effect model. This method is based on an inverse variance weighted average. Its principle is to give more weight to trials that give more information about the treatment effect. The heterogeneity, i.e. the variation of the treatment effect between the trials, was measured.

Costs

The costs were estimated in the French context from a payer's perspective and expressed in 2010 euro. We assessed the direct medical costs using the healthcare resource use measured

in the MAR-LC for RT, RT-induced toxicities, medical transportations for RT and disease progression (cf. table 1). The unit costs were extracted from literature for medical transportations (Martin, 2003) and disease progression costs (Braud, 2003). Radiotherapy and acute oesophagitis unit costs were computed as the mean lump sums per correspondent diagnosis related groups in the French prospective payment scheme (PPS).

Table 1. Unit cost values and sources

	Healthcare resource available in the MAR-LC	Unit cost	Source
Radiotherapy	Number of fractions received	1 004€ treatment planning + 138€ per fraction 80€ per fraction	PPS 2010
Medical transportations	Number of fractions received	(variable according to use rate)	Martin, 2003
Disease progression	Post-progression survival time	3 073€/month	Braud,2003
Oesophagitis	Acute severe oesophageal toxicity	1 745€ in case of toxicity	PPS 2010

PPS: prospective payment scheme

Costs and effectiveness outcomes were not discounted since the MAR-LC patients' median survival was 14 months.

Methods

Survival benefit

In trial-based economic evaluations, the restricted mean survival (Karrison, 1997; Barker, 2009) is commonly used to assess the survival benefit of a treatment. For our example, it is defined as follows:

$$B = \int_0^T S_{\text{Mod}}(t).dt - \int_0^T S_{\text{Conv}}(t).dt \quad (1)$$

where S_{Conv} represents the survival function in the conventional arm and S_{Mod} in the modified arm. The survival benefit corresponds graphically to the area between the two survival curves (cf Figure 1). In our example, the time horizon was restricted to five years according the primary endpoint of the MAR-LC.

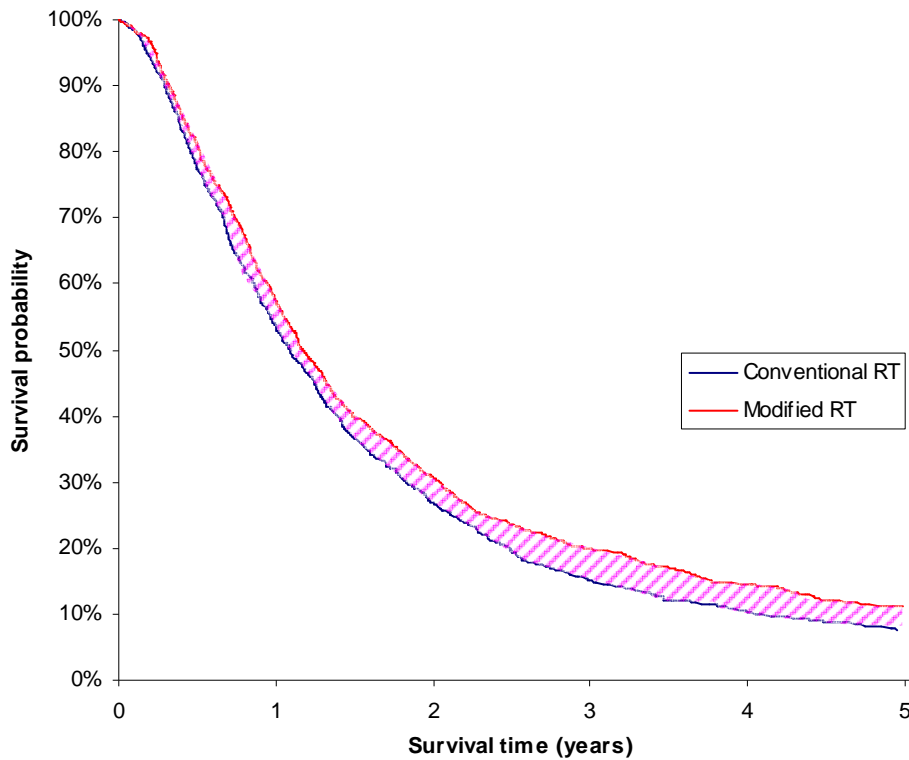


Figure 1. Survival benefit in the MAR-LC (illustration using the Kaplan-Meier method)

As shown in equation (1), the survival benefit depends on the method used to estimate the survival functions S_{Conv} and S_{Mod} . Therefore, it is a relevant issue to measure the impact of the choice of a particular method on the cost-effectiveness ratio. In this paper, five methods of estimation of the survival benefit were compared: four non parametric methods which has been used in meta-analyses for survival endpoints (the Kaplan-Meier, the Stewart, the Peto-year and the Peto-month methods) and a parametric method widely used in economic evaluation (the Weibull method). We focused on how these methods allow taking into account the stratification on trials and whether or not the proportional hazards hypothesis (hazard ratio constant over time) is required.

The first method we used was the standard Kaplan-Meier method (Kaplan, 1958) which estimates the survival probability at each event time. In this method, the proportional hazards assumption is not required but the stratification on trials can not be taken into account to compute the pooled hazard ratio (HR).

The so-called Stewart method is derived from the Kaplan-Meier method (Stewart, 1993). This method requires the proportional hazards assumption, as it utilises an estimation of the HR, considered constant over time.

$$\text{Under proportionality : } S_{\text{Mod}}(t) = S_{\text{Conv}}(t)^{\text{HR}} \quad (2)$$

where S_{Conv} represents the survival function for conventional RT estimated with the KM method and HR the pooled hazard ratio of the treatment effect, stratified on trial. Here, S_{Mod} was estimated after each event time in the control arm.

A third actuarial method developed by R. Peto (Early Breast Cancer Trialists' Collaborative Group, 1992) allows taking into account the stratification on trials and the variation of the HR between periods. The survival probabilities are estimated at predetermined time intervals and not at each event time like in the Kaplan-Meier and the Stewart methods. At each i^{th} period, the survival probability of the whole population p_i and the pooled hazard ratio HR_i are estimated. Stratification on trials is taken into account in the HR_i calculation, based on the inverse variance weighted average, and treatment effect HR_i may vary between periods. Then, based on p_i and HR_i , the probabilities to be alive at i^{th} period in the conventional RT group ($p_{\text{Conv},i}$) and in the modified RT group ($p_{\text{Mod},i}$) are estimated as follows:

$$p_{\text{Conv},i} = p_i - [0,5 \times p_i \times (p_i - 1) \times \log(\text{HR}_i)] \quad (3)$$

$$p_{\text{Mod},i} = p_i + [0,5 \times p_i \times (p_i - 1) \times \log(\text{HR}_i)] \quad (4)$$

$$S_{\text{Conv}}(t) = \prod_{i=1}^t p_{\text{Conv},i} \quad \text{and} \quad S_{\text{Mod}}(t) = \prod_{i=1}^t p_{\text{Mod},i} \quad (5)$$

The Peto method has been computed with both 1-year and 1-month interval length.

Finally, we used the Weibull parametric method using the parameterization advocated by Carroll (Carroll, 2003). This model is widely used in economic evaluation (Dewilde, 2011; Latimer, 2011; Shiroya, 2011; Bagust, 2010). It leads to more robust estimations if Weibull distribution fits the lifetimes' distribution. The shape (α) and scale (λ) parameters were estimated for each trial and for each arm of the trial. The four estimates λ_{Conv} , α_{Conv} , λ_{Mod} and α_{Mod} were computed in combining those obtained from the 10 RCTs of the MAR-LC. We used a fixed effect model in the absence of heterogeneity (Der Simonian-Laird, 1986), according to the generic inverse variance approach (Whitehead, 2002). In the conventional RT arm, the scale parameter $\lambda_{Conv,k}$ and the shape parameter $\alpha_{Conv,k}$ are estimated for each k trial. We pooled the ten scale and the ten shape parameters with an inverse variance weighted sum to obtain pooled λ_{Conv} and α_{Conv} . The same method was used to estimate the survival function for the modified RT arm. Equation (6) gives the survival functions for conventional and modified RT:

$$S_{Conv}(t) = \exp[-\lambda_{Conv} \cdot t^{\alpha_{Conv}}] \quad \text{and} \quad S_{Mod}(t) = \exp[-\lambda_{Mod} \cdot t^{\alpha_{Mod}}] \quad (6)$$

These five methods led to the estimation of the survival functions in both conventional RT and modified RT arms, and thus to the estimation of the mean survival benefit.

Impact on the health economics outcome

We studied the differences between methods comparing Incremental Cost-Effectiveness Ratios (ICER) and their 95% confidence intervals (CI95%). ICER was defined as the difference in costs divided by the mean survival benefit between the two arms. Mean ICERs and 95% confidence intervals were estimated using the non parametric bootstrap method with 1,000 replicates. For each method, the 1,000 replicates were represented in the incremental cost-effectiveness plane to determine the proportion of ICERs in each quadrant. Acceptability curves were also computed for each of the five methods. We also performed a one-way

sensitivity analysis of two parameters which had been extracted from the literature: the use of medical transportations and the cost of treatment of disease progression.

Results

The mean total cost per patient was 25,376€ (CI95%:23,584€ - 27,262€) for conventional RT and 29,186€ (CI95%: 27,350€ - 31,152€) for modified RT.

The forest-plot shown on figure 2 summarizes the MAR-LC effectiveness results. Modified RT was associated with longer overall survival (pooled hazard ratio = 0.88, 95% CI: (0.80-0.97), p=0.009). There was no heterogeneity of treatment effect between trials (p=0.37, I²=8%). No patient subgroups (age, gender, stage, performance status and histology) benefited more from modified RT. However, modified RT was associated with an increased risk of acute severe oesophageal toxicity (odds ratio = 2.44, 95% CI: (1.90-3.14), p<0.001). There was no difference for other acute and late toxicities.

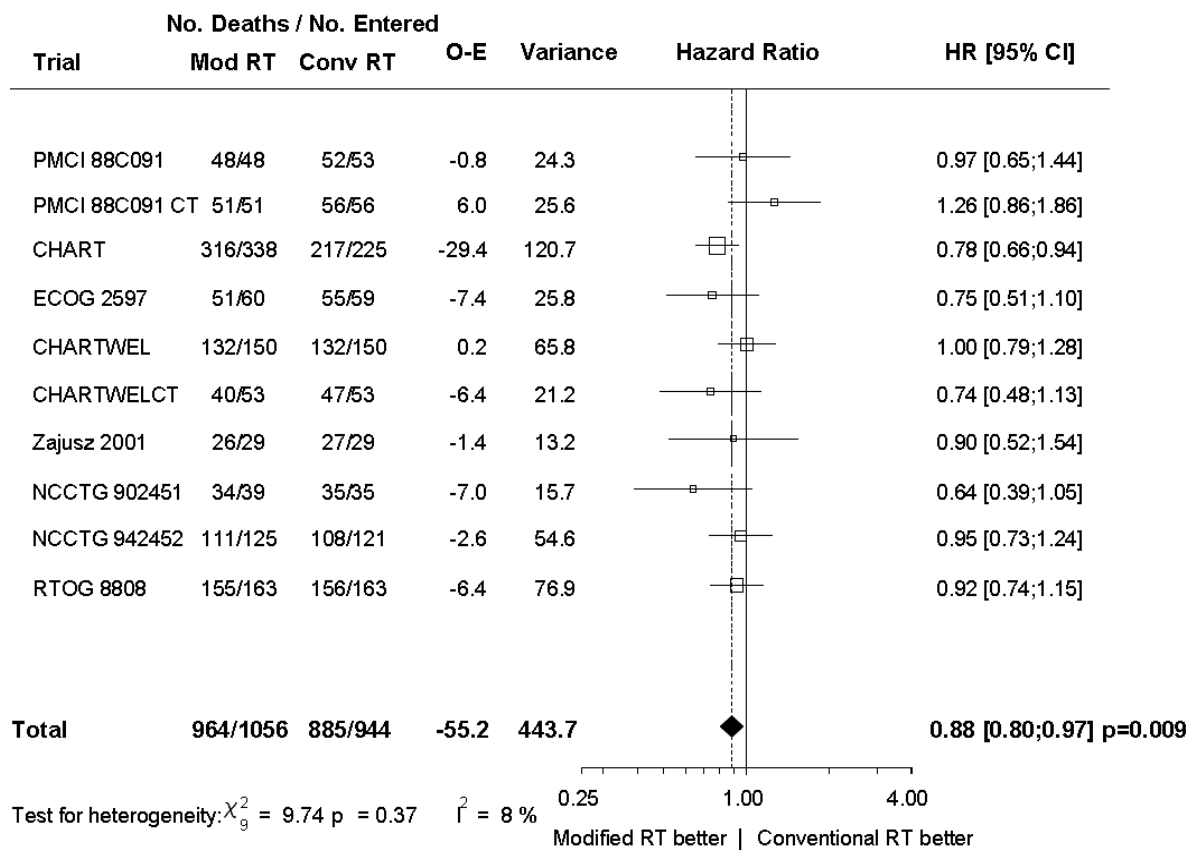


Figure 2. Effect of modified radiotherapy versus conventional radiotherapy on overall survival in the MAR-LC

Figure 3 shows the survival functions for the Peto-year and the Peto-month methods, for conventional RT and modified RT.

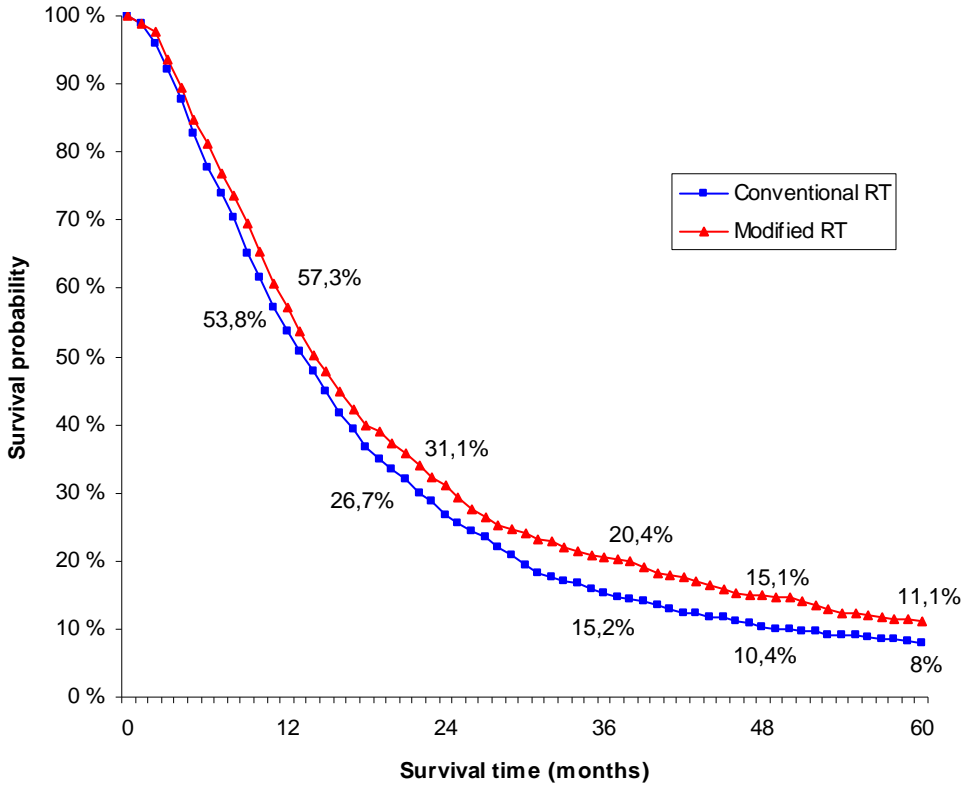
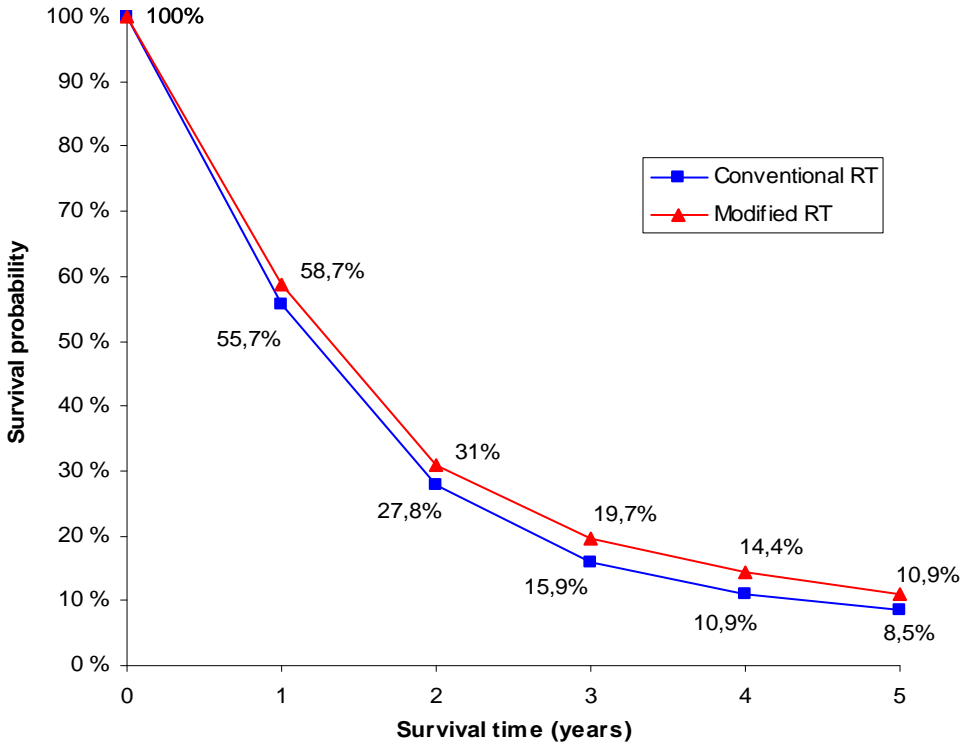


Figure 3. Peto-year and Peto-month survival curves for conventional and modified RT

Survival benefit and ICER estimates are shown in table 2 for each method. The mean survival benefit ranged from 1.6 month with the Peto-year method to 2.4 months with the Weibull method. Small differences between survival benefits led to larger differences between ICERs. Mean ICERs ranged from 10,320€ to 34,937€ per life year gained respectively for the Weibull and the Peto-year methods.

Table 2. Survival benefit and ICER estimates according to survival methods

Survival method	Survival benefit estimate (in month)	ICER (cost per life year gained)
Kaplan-Meier	2.2 [CI95%: 0.7-3.7]	21,356€ [CI95%: 7,265€ - 56,945€]
Stewart	1.9 [CI95%: 0.6-3.4]	26,343€ [CI95%: 8,040€ - 73,061€]
Peto-year	1.6 [CI95%: 0.5-2.8]	34,937€ [CI95%: 10,344€ - 77,277€]
Peto-month	2.0 [CI95%: 0.7-3.5]	29,015€ [CI95%: 7,569€ - 64,997€]
Weibull	2.4 [CI95%: 0.6-4.1]	10,320€ [CI95%: 6,234€ - 59,023€]

For the five methods, all the cost-effectiveness ratios were located in the North-East quadrant of incremental cost-effectiveness plane: modified RT was both more effective and more expensive than conventional RT (figures 4). The dispersion of scatter plots pointed out the variability of ICER estimate amongst methods. Peto-year seemed to be the method with less variability (figure 4-D) whereas Weibull had the largest scattergram (figure 4-C).

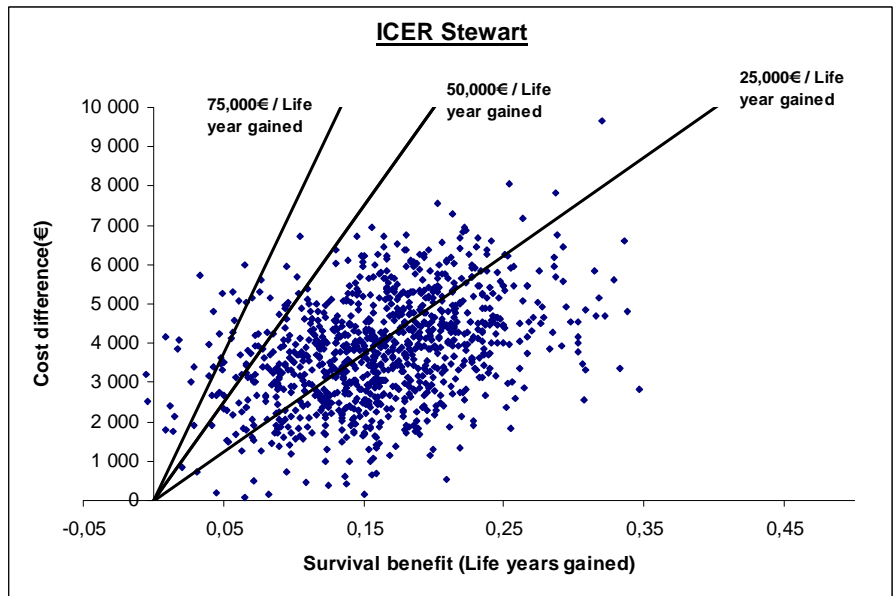
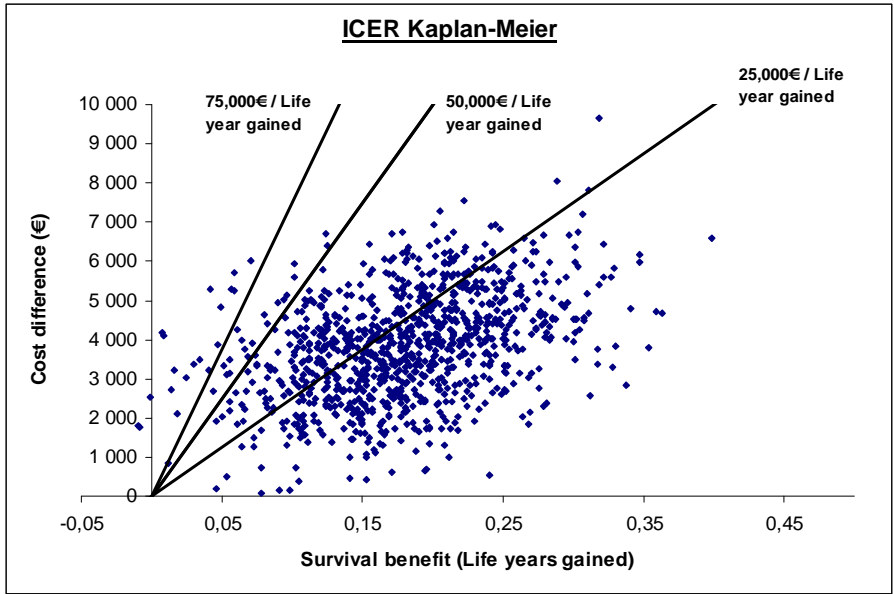


Figure 4-A and 4-B: Kaplan-Meier, Stewart

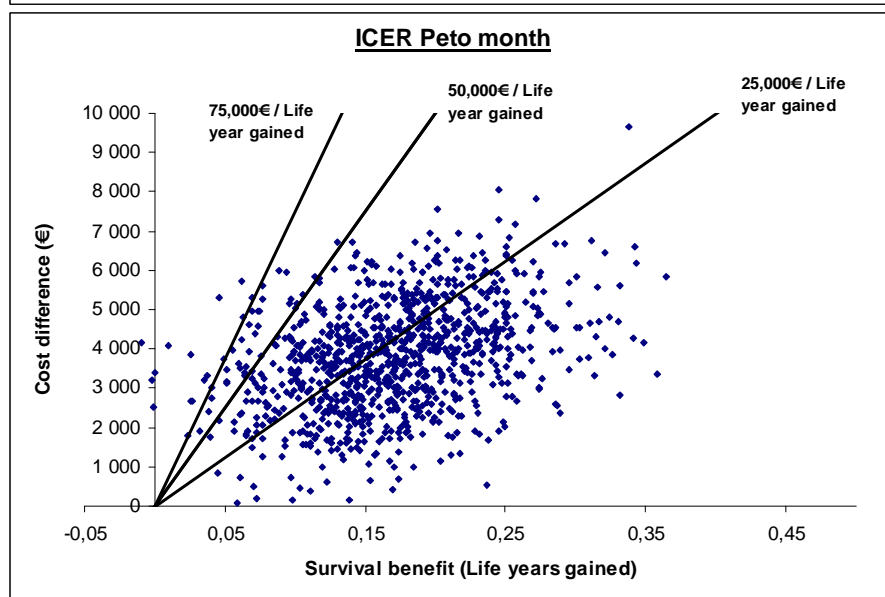
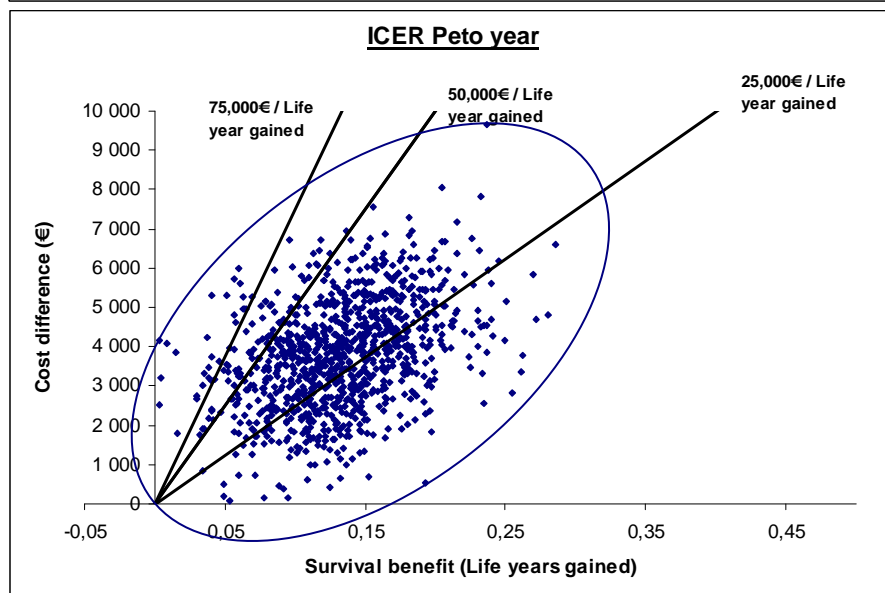
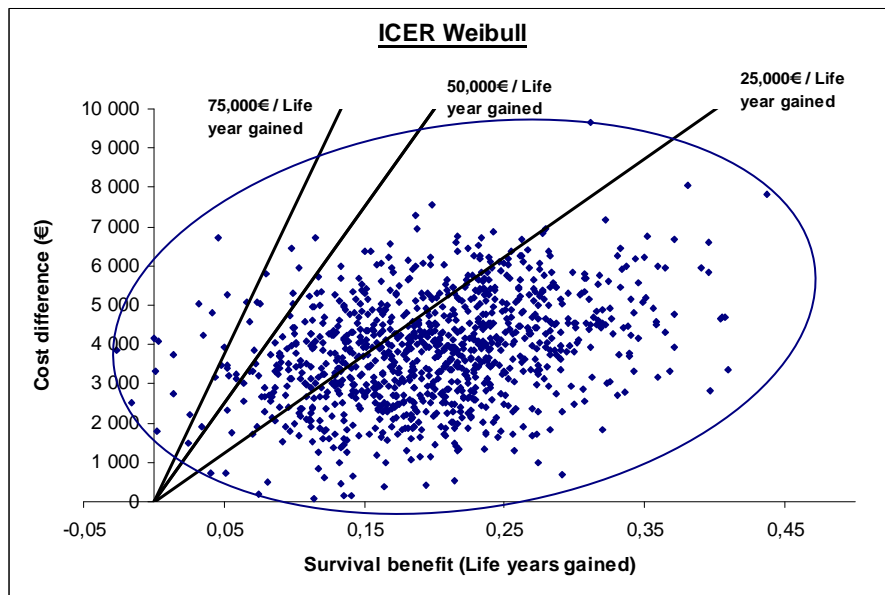


Figure 4-C, 4-D and 4-E. Weibull, Peto-year and Peto-month

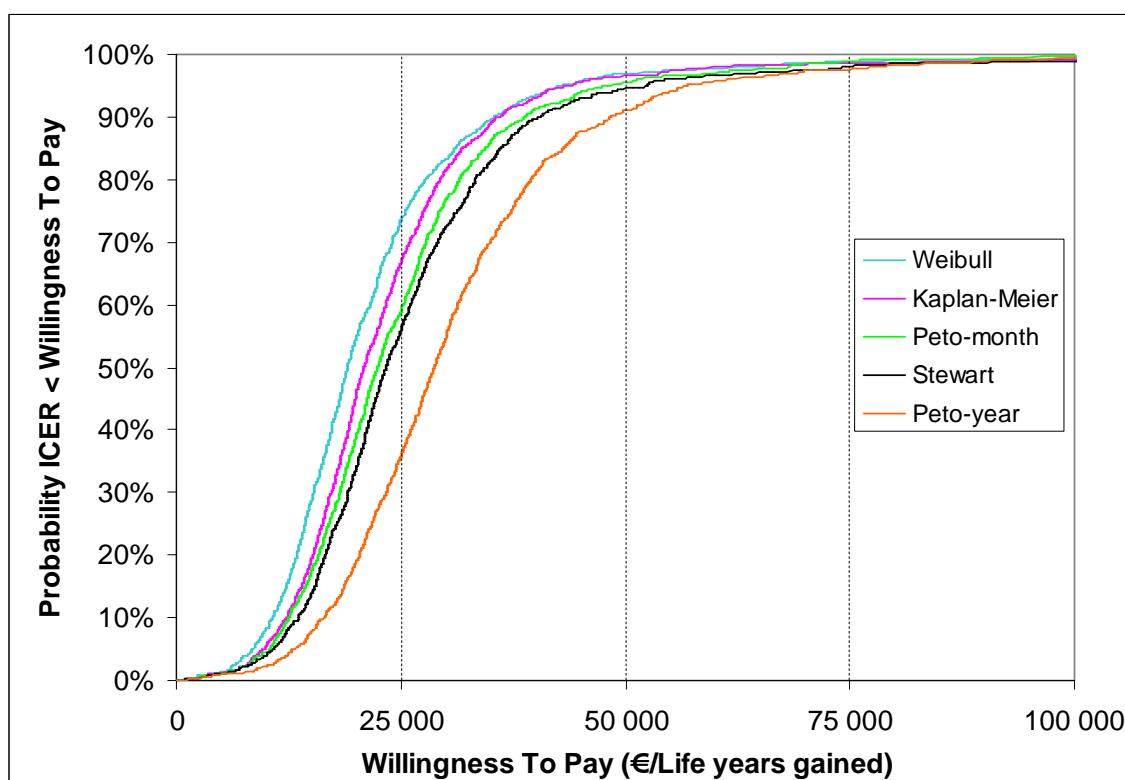


Figure 5. Acceptability curves showing the probability that modified RT is cost-effective at different thresholds of WTP for one life year.

As shown in figure 5, with a willingness to pay (WTP) for one life year higher than 50,000€, all the methods concluded to the efficiency of the modified radiotherapy with a probability higher than 90%. Whereas below the threshold of 50,000€, acceptability curves did not lead to similar conclusions. Indeed, at a 25,000€ threshold, the probability of modified RT being cost-effective ranged from 36% with Peto-year to 74% with the Weibull method (cf. table 3).

Table 3. Probability that modified RT is cost-effective at different threshold of WTP, according the survival method used.

Survival method	WTP for one life year		
	25,000€	50,000€	75,000€
Kaplan-Meier	67%	97%	99%
Stewart	56%	95%	98%
Peto-year	36%	91%	98%
Peto-month	59%	96%	99%
Weibull	74%	97%	99%

In the sensitivity analysis we used the Peto-year method and a threshold of 50,000€ per life year gained. When we varied the rate of use in medical transportation cost estimate (+/- 50%), the probability of modified RT cost-effectiveness ranged from 89% to 92%. Disease progression unit cost was varied from -30% to +30% and led to a probability of efficiency between 85% and 95%.

Discussion

The estimation of health benefit is of high importance in economic evaluation. This is particularly relevant in the field of oncology where new drugs are very expensive and the survival benefit is modest (Ocana, 2010). The estimation of the survival benefit has then to be accurate.

Studying the impact of the choice of a survival analysis method using individual patient data from a meta-analysis of RCTs is innovative. The MAR-LC database allowed us to provide a precise estimate of the survival benefit of modified RT and to compare both non parametric and parametric methods. Our study showed that the ICER estimate and its confidence interval were sensitive to the survival method used. With a threshold value between 20,000€ and 40,000€, the acceptability curves obtained with the five methods had a different shape. Therefore, the choice of a survival method to estimate the health benefit can strongly impact on the health economic outcome. Amongst the five methods that we compared, the Peto-year method appeared more flexible because it took into account stratification on clinical trials and did not require the proportional hazards assumption. Moreover, this method was characterised by less variability in ICER estimation.

Two recent studies have pointed out some methodological issues raised by the estimation of survival benefit for economic evaluation (Latimer, 2011; Connock, 2011). These studies focused on parametric extrapolation of survival beyond the follow-up limits of a clinical trial. Latimer reviewed 45 NICE technology appraisals of advanced or metastatic cancer interventions. He noticed that a wide range of methods have been used and that some of them were not appropriate given the characteristics of the data. He proposed an algorithm to choose an optimal parametric method according to the characteristics of the data under study. Similarly, Connock through two other recent examples of NICE technology appraisals of anticancer drugs (sorafenib for hepatocellular carcinoma and azacitidine for myelodysplastic

syndroms) discussed the issue of parametric extrapolation of survival beyond the data and the impact on the estimation of the survival benefit. Large differences can be observed using different parametric models. The author recommended that multiple parametric models should be explored and compared when estimating the survival benefit for economic evaluation. Unfortunately, these studies could not provide quantitative information on the impact on the health economic outcome.

The most important finding in our study was that the health economic results were discordant using different survival methods even though survival extrapolation was not used. However, some limitations of our study deserve to be mentioned. First, we chose to compare a limited number of parametric and non parametric methods. We selected standard methods used either in economic evaluations or meta-analyses. However other parametric models such as Exponential or Gamma model should also be consider along with the Weibull model. The second limitation relies on the characteristics of our data. There was no treatment effect heterogeneity between MAR-LC trials. This could have attenuated the difference between methods which account for stratification on trials and those which do not (Kaplan-Meier). Similarly a with higher survival benefit and non-proportional hazards, the five methods could have led to even more different health economic results.

Acknowledgements

We thank the MAR-LC trialists who agreed to share their data. We are grateful to Janneke Grutters, Manuela Joore, Bram Ramaekers and Dr Dirk de Ruyscher (Maastricht University, Netherlands) who initiated the cost-effectiveness project based on the MAR-LC database. The MAR-LC project was supported by unrestricted grants from French Programme Hospitalier de Recherche Clinique, Ligue Nationale Contre le Cancer, and Sanofi-Aventis (J-P.P.). The funding sources had no role in study design, data collection, data analysis, data interpretation, or manuscript writing.

References

1. Annemans L, Asukai Y, Barzey V, Kotapati S, Lees M, Van Baardewijk M, Wang Q, Batty AJ, Fisher D. Extrapolation in oncology modelling: novel methods for novel compounds. *ISPOR abstract* 2011;
2. Bagust A, Greenhalgh J, Boland A, Fleeman N, McLeod C, Dickson R, Dundar Y, Proudlove C, Shaw R. Cetuximab for recurrent and/or metastatic squamous cell carcinoma of the head and neck: a NICE single technology appraisal. *Pharmacoeconomics* 2010; 28(6): 439-48.
3. Barker C. The Mean, Median, and Confidence Intervals of the Kaplan-Meier Survival Estimate - Computations and Applications. *The Am. Statistician* 2009; 63 (1) : 78-80
4. Braud A.C. Direct treatment costs for patients with Lung cancer from first recurrence to death in France. *Pharmacoeconomics* 2003; 21(9):671-9.
5. Carroll K. On the use and utility of the Weibull model in the analysis of survival data. *Contr. Cl. Trials*, 2003; 24: 682-701
6. Connock M. Cautions regarding the fitting and interpretations of survival curves. *Pharmacoeconomics* 2011; 29 (10): 827-837.
7. Davies A, Briggs A, Wagner S, Kotapati S, Schneider J, Ebeid O, Levy AR. Mean versus Median Overall Survival (OS) for describing value of new cancer therapies: a case study. *ISPOR abstract* 2011
8. Der Simonian R, Laird N. Meta-analysis in clinical trials. *Contr. Cl. Trials*, 1986; 7: 177-88
9. Dewilde S, Brüggenjürgen B, Nienaber C, Senges J, Welte R, Willich SN. Cost-effectiveness of adjunctive eptifibatide in patients undergoing coronary stenting in Germany. *Eur J Health Econ.* 2011 Apr 12.
10. Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women. *Lancet*, 1992; 339:1-15
11. Gray AM. Cost-effectiveness analyses alongside randomised clinical trials. *Clinical Trials* 2006; 3: 538-542

12. Kaplan EL, Meier P. Non parametric estimation from incomplete observation. *J. Am. Stat. Assoc* 1958
13. Karrison TG. Use of Irwin's Restricted Mean as Index for Comparing Survival in Different Treatment Groups – Interpretations and Power Considerations, *Contr. Cl. Trials* 1997, 18:151-167
14. Latimer N. Survival modelling for economic evaluations: presenting a process for selecting appropriate models to avoid inconsistencies between health technology assessments, *iHEA abstract*, 2011.
15. Latimer N. NICE DSU Technical Support Document 14: Survival Analysis for Economic Evaluations alongside Clinical Trials - Extrapolation with Patient-level Data, June 2011
16. Le Péchoux C, Mauguen A, Baumann M, Schild SE, Parmar M, Turrisi AT, Sause W, Ball D, Belani CP, Behrendt K, Pignon JP on behalf on the MAR-LC Collaborative Group. Evaluation of modified fractionation radiotherapy effect in non metastatic lung cancer: an updated individual patients data meta-analysis on 10 randomized trials and 2685 patients. *J Thorax Oncol* 2011; 6(Suppl. 6):S432-3
17. Martin P. Le coût de la radiothérapie, *Bull Cancer* 2003 ; 90 (11) : 969-75
18. Mittmann N, Au H-J, Tu D et al. Working Group on Economic Analysis of the National Cancer Institute of Canada Clinical Trials Group, Australasian Gastrointestinal Interest Group. Prospective Cost-Effectiveness Analysis of Cetuximab in Metastatic Colorectal Cancer: Evaluation of National Cancer Institute of Canada Clinical Trials Group CO.17 Trial *JNCI* 2009 ; 101 (17) : 1182-92
19. Neymark N, Gorlia T, Adriaenssen I, Baron B, Piccart M. Cost Effectiveness of Paclitaxel/Cisplatin Compared with Cyclophosphamide/Cisplatin in the Treatment of Advanced Ovarian Cancer in Belgium *Pharmacoeconomics* 2002; 20 (7): 485-497
20. Ocana A, Tannock IF. When are “positive” clinical trials in oncology truly positive? *JNCI* 2011, vol.103, 1;16-20
21. Pignon JP, Hill C. Meta-analyses of randomised clinical trials in oncology, *Lancet Oncology* 2001; 2: 475-82

22. Shiroiwa T, Fukuda T, Shimosuma K. Cost-effectiveness analysis of trastuzumab to treat HER2-positive advanced gastric cancer based on the randomised ToGA trial. *Br J Cancer*. 2011 Oct 25;105(9):1273-8.
23. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data : is there a difference? *The Lancet* 1993 Feb 13;341(8842):418-22
24. Tappenden P, Chilcott J, Ward S, Eggington S, Hind D, Hummel S. Methodological issues in the economic analysis of cancer treatments. *Eur J Cancer* 2006; 42: 2867-75
25. Yusuf S, Peto R, Lewis J, Collins R, Sleight T. Beta blockade during and after myocardial infarction : an overview of randomised clinical trials. *Prog Cardiovasc Dis* 1985;27: 335-71
26. Whitehead A. Meta-analysis of Controlled Clinical Trials, Ed John Wiley & Sons, Ltd, 2002