Assessing the Impact of Censoring of Costs and Effects on Health-Care Decision-Making: an Example Using the Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) Study

Elisabeth Fenwick, PhD,[1] Deborah A. Marshall, MHSc, PhD,[2] Gordon Blackhouse, MSc,[2] Humberto Vidaillet, MD[3], April Slee,[4] Lynn Shemanski, PhD,[4] Adrian R. Levy, PhD[5]

[1]University of Glasgow, Glasgow, UK; [2]McMaster University, Center for Evaluation of Medicines, Hamilton, ON, Canada; [3]Marshfield Clinic, Wisconsin, WI, USA; [4]Axio Research, Seattle, WA, USA; [5]University of British Columbia, Vancouver, BC, Canada

ABSTRACT

Objectives: Losses to follow-up and administrative censoring can cloud the interpretation of trial-based economic evaluations. A number of investigators have examined the impact of different levels of adjustment for censoring, including nonadjustment, adjustment of effects only, and adjustment for both costs and effects. Nevertheless, there is a lack of research on the impact of censoring on decision-making. The objective of this study was to estimate the impact of adjustment for censoring on the interpretation of cost-effectiveness results and expected value of perfect information (EVPI), using a trial-based analysis that compared rate- and rhythm-control treatments for persons with atrial fibrillation.

Methods: Three different levels of adjustment for censoring were examined: no censoring of cost and effects, censoring of effects only, and censoring of both costs and effects. In each case, bootstrapping was used to estimate the uncertainty in costs and effects, and the EVPI was calculated to determine the potential worth of further research.

Results: Censoring did not impact the adoption decision. Nevertheless, this was not the case for the decision uncertainty or the EVPI. For a threshold of $50,000 per life-year, the EVPI varied between $626,000 (partial censoring) to $117 million (full censoring) for the eligible US population.

Conclusions: The level of adjustment for censoring in trial-based cost-effectiveness analyses can impact on the decisions to fund a new technology and to devote resources for further research. Only when censoring is taken into account for both costs and effects are these decisions appropriately addressed.

Keywords: censoring, cost-effectiveness acceptability curves, cost-effectiveness analysis, economic evaluation, expected value of perfect information, Kaplan-Meier survival analysis.

Introduction

One of the major challenges faced by decision-makers in all (budget-constrained) health-care systems is the choice between alternative interventions for the same medical indication. Increasingly these decisions are being guided by economic evidence, including results of cost-effectiveness studies. Inevitably, the estimates of the costs and effects involve some uncertainty because of measurement, sampling, and random errors. This leads to a situation in which decision-makers must address two decisions: the first involves identifying the most appropriate method of patient management to fund given the current level of information and uncertainty, and the second involves funding additional research to reduce the uncertainty in the future [1]. A formal framework exists to address these two separate but related decisions [1–3]. Given the objective to maximize health benefit subject to a budget constraint, the appropriate method of patient management is identified, within this framework, according to the expected cost-effectiveness of the interventions (i.e., the point estimate), irrespective of the uncertainty surrounding the estimate [1]. The second decision, that of whether to fund more research, involves an assessment and valuation of the uncertainty surrounding the decision [1].

Prospective collection of patient-level cost data within randomized controlled trials is one approach to obtain the information needed to estimate cost-effectiveness. In the majority of trials, however, these data are incomplete as a result of censoring and this needs to be accounted for in the cost-effectiveness analysis [4]. In this situation, the estimation of the expected costs and effects from the sample becomes more involved with implications for the cost-

effectiveness analysis. Investigators have examined the impact of different levels of adjustment for censoring, ranging from nonadjustment, through adjustment of effects alone or costs alone, to adjustment for both costs and effects (see Young [4] for a comprehensive review and description of approaches for accounting for censored costs). Nevertheless, there is a lack of research on the impact of these adjustments for censoring on the measure of uncertainty and the decision to fund more research.

This article examines the impact of not adjusting for censoring versus adjusting for censoring of effects alone and adjusting for censoring of both effects and costs, on health-care decision-making in terms of the optimal intervention to adopt and whether to fund the collection of additional information. This is demonstrated using patient-level data from a trial-based cost-effectiveness analysis comparing rate-control to rhythm-control treatment for atrial fibrillation (AFFIRM) [5].

Background

Health-Care Decision-Making and Cost-Effectiveness Analysis

Cost-effectiveness is increasingly one of the criteria that are being used to guide adoption choices between alternative health-care interventions. Such cost-effectiveness analyses enumerate the additional resources consumed for an improvement in outcome (e.g., survival or quality-adjusted life-years) associated with one health intervention compared to another. This is expressed as an incremental cost-effectiveness ratio (ICER)—a measure of the additional cost per additional unit of health gain. For example, for a comparison of rate-control versus rhythm-control management for patients with atrial fibrillation, the ICER would be calculated as:

$$ICER = \frac{Mean\ cost_{rate\text{-}control} - Mean\ cost_{rhythm\text{-}control}}{Mean\ survival_{rate\text{-}control} - Mean\ survival_{rhythm\text{-}control}}$$

To decide whether an intervention offers "good" value for money compared to the alternative, the ICER must be compared to a specified monetary threshold. This threshold represents the maximum amount that the decision-maker is willing to pay for one unit of additional health benefit (cost-effectiveness threshold). The intervention is considered cost-effective if the ICER falls below this threshold, and otherwise, it is not considered cost-effective.

By explicitly incorporating this threshold, the decision payoffs (costs and effects) can be combined to form a measure of net monetary benefit (NB) for each intervention (t) [1–3,6–8]:

$$NB = (cost\text{-}effectiveness\ threshold \times health\ benefit) - Cost_t$$

Now the cost-effective intervention is identified as the one associated with the maximum value of NB. The NB approach simplifies the assessment of cost-effectiveness when the decision involves more than two interventions, or when stochastic analyses are undertaken. In addition, the use of net benefits has been suggested to simplify the assessment of uncertainty [9–11].

Inevitably, the estimates of effectiveness and the resources consumed by the interventions are measured with uncertainty. As such, the costs, effects, and any estimate of the cost-effectiveness (ICER or NB) associated with interventions will also be uncertain. When trial data are available, bootstrapping methods are useful for generating the distribution of estimators, especially when the derivation of their distribution is intractable. This provides an estimate of the extent of the uncertainty surrounding the costs and effects individually. The uncertainty surrounding cost-effectiveness is dependent upon the specified amount that a decision-maker is willing to pay for a gain in effect. This uncertainty can be plotted as a function of this threshold on a cost-effectiveness acceptability curve as the, essentially Bayesian, probability that the intervention is cost-effective [12–15].

In a health-care system with the objective to maximize health benefit subject to the budget constraint, interventions should be selected on the basis of the expected costs, outcomes, and cost-effectiveness, rather than measures of uncertainty [1–3]. It is when considering the second decision faced by the decision-maker, that of whether to fund the collection of additional information through research, that the uncertainty surrounding the costs, effects, and cost-effectiveness is important [1–3]. At this point, the finite probability that the decision is incorrect (or error probability, given as the complement of the cost-effectiveness acceptability curve) and the consequences of an incorrect decision, in terms of opportunity costs, play a crucial role [3]. Bayesian value-of-information analysis provides a method to assess the expected opportunity losses associated with the existing (uncertain) evidence base, to determine whether further research should be conducted and how this should be designed [2,3]. The techniques involve establishing the difference between the expected value of a decision made on the basis of the existing evidence and the expected value of a decision made on the basis of further information [16]. This difference is then compared to the cost of collecting the additional evidence; where the value of further information exceeds the costs of collecting it, the research is deemed worthwhile [1].

Perfect information surrounding all elements of the decision would, by definition, eliminate all uncertainty.

The expected value of perfect information (EVPI) is therefore equivalent to the expected cost of the current uncertainty surrounding the decision and provides a measure of the maximum return to further research; providing a necessary condition for determining whether further research is potentially worthwhile. See Claxton and colleagues [1–3] and Ades et al. [16] for more details on value-of-information analysis.

Censoring

In most trials, some proportion of the patient sample will drop out of the study or be lost to follow-up, leaving the data incomplete at the time of study termination. In trials with mortality as the end point, data from patients who are alive at the end of follow-up or who are lost during the study are considered "right censored" [17]. All that is known for censored patients is that their survival is longer than the follow-up time in the trial. If no adjustment is made for censoring, survival times will be underestimated and the estimate of survival will be biased. Kaplan-Meier survivor curve [18] and Cox regression [19] methods are accepted as the standard approaches for estimation adjusting from right-censored survival data. These traditional nonparametric approaches assume that censoring is completely random and independent of the risk of event under observation (i.e., death) at any time [17,20].

Similarly, censoring can lead to biased estimates of trial-based costs unless appropriately accounted for in the analysis. If data from all patients (both censored and uncensored) are used to estimate mean costs, these will be underestimated because they will exclude all potential costs that might have been incurred by subjects after they are censored. If data from only uncensored patients (i.e., those observed to die during the study) are used, the estimates will be biased toward cost estimates of patients with shorter survival times, because patients with longer survival times are more likely to be censored [21].

Previous approaches to adjust cost estimates for censoring have been based on the application of survival analysis techniques to cost data [22]. These approaches treated patient costs as survival times but failed to recognize that cost data typically violate the assumption of independent censoring. This violation occurs because patients accrue costs at different rates, with patients in better health accruing costs at lower rates than those in worse health [23]. To address the issue of dependent censoring with cost data, Lin et al. [21] and Etzioni et al. [24] presented the Kaplan-Meier sample average (KMSA) estimator. In the KMSA estimator, the duration of the study is partitioned into a number of time intervals. For each time interval, the mean cost for all patients who are noncensored at the start of the interval is multiplied by the probability of survival during the period as calculated at the beginning of the period. These products are summed over all time intervals to produce an estimate of total censoring-adjusted costs. This approach has been shown to provide consistent estimators of average costs that are asymptotically normal under the assumption that censoring occurred at the boundaries of the defined time intervals [21,25–29].

Methods

Data Source

Data for this study were from the clinical and cost-effectiveness analyses of the Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) trial, and salient features are summarized below [5,30–32]. The objective of AFFIRM was to compare the effectiveness of rate-control versus rhythm-control in patients with atrial fibrillation and having one or more risk factors for stroke or death [5,30]. The study sample of 4060 patients was similar to the target population (61% men with a mean age of 69.7 years) with associated cardiovascular comorbid conditions. The primary outcome was survival time, with follow-up reported for 97.6% of patients (71 withdrew consent and 26 had unknown vital status at termination). Only 16% died during follow-up, which averaged 3.5 years; therefore, 84% of the sample was censored.

Health-care resource use considered in the cost-effectiveness analysis included hospitalizations, cardiac procedures, electric or chemical cardioversion, short-stay and emergency department visits, and medications used to treat atrial fibrillation. The perspective adopted was that of a third-party payer [5]. The patient cost was calculated by multiplying each health-care resource use component by the unit cost (obtained from multiple publicly available sources) and summing the results for each patient. The mean cost was calculated across all patients. Future costs and effects were discounted to their present value at a rate of 3% per annum [33,34]. The cost-effectiveness analysis reported a nonsignificant mean survival gain (0.08-year survival gain, 95% confidence interval [CI] −0.02 to 0.17 years; $P = 0.10$) and lower mean cost (−$5077, 95% CI −$7423 to −$2801) for rate-control subjects. Rhythm-control was both more costly and less effective than (i.e., dominated by) rate-control and, as such, no ICER was calculated [35].

Data analysis

The incremental costs, survival, and ICER (measured as the cost per life-year gained) of rate- versus rhythm-control strategy were estimated for three scenarios that differed in the level of adjustment for censoring.

No adjustment for censoring in either costs or survival. The number of life-years for each subject was estimated to be equal to the time from randomization

to last contact (death or lost to follow-up). The mean number of life-years per patient for each treatment strategy was estimated as the sum of the number of life-years for each patient divided by the total number of subjects enrolled in the study. The total cost per patient was calculated by multiplying each component of health-care resource use by the unit cost and summing the results for each patient. The mean cost per patient was estimated as the sum of the total cost for each patient divided by the total number of patients at the start of the study. Mean costs were estimated from observed costs for all patients (both censored and uncensored) with no adjustment for censoring.

Partial adjustment—survival adjusted for censoring, and costs not adjusted for censoring. The mean number of life-years for each treatment strategy was adjusted for censoring using the Kaplan-Meier product limit estimator to calculate survival. The time horizon for both treatment and control groups was standardized at 5.65 years, the longest follow-up observed in AFFIRM. Mean survival time (in years) was estimated as the area below the Kaplan-Meier survival curve for each treatment group. Costs were not adjusted for censoring. This was the approach applied in the published cost-effectiveness analysis of the AFFIRM trial, as a result of a lack of access to the individual cost data at the time that the analysis was undertaken [5].

Full adjustment—both survival and costs adjusted for censoring. Life-years were adjusted for censoring as described above. Costs were adjusted for censoring using the KMSA method [21,24].

For each scenario, a sampling distribution of costs and effects was estimated for each intervention by nonparametric bootstrapping with 10,000 replicates. The results for rate- compared to rhythm-control are presented as a scatter plot on the incremental cost-effectiveness plane and as cost-effectiveness acceptability curves [12,36]. The mean values of these sampling distributions are used to calculate the ICER and, in comparison to the cost-effectiveness threshold, identify the appropriate intervention to fund given current levels of information. To address the second decision, concerning funding of further research, the population EVPI is calculated for each scenario. For simplicity, the calculation of EVPI was undertaken using NB determined for a range of values of the ceiling ratio ($20,000, $50,000, and $100,000 per life-year).

Each individual value in the distribution (i.e., each bootstrap replicate) represents a possible future resolution of the current uncertainty (i.e., a possible future realization of perfect information) for which the appropriate intervention can be determined, on the basis of maximum NB. Nevertheless, it is not known at which particular realization the uncertainty will resolve. As such, the expected value of a decision with perfect information is calculated by averaging these maximum NB over the distribution. The EVPI is simply the difference between the expected value of the decision taken with perfect information and that taken with current information, which is based on the expected NB [16]:

$$E_\theta \max_t NB(t, \theta) - \max_t E_\theta NB(t, \theta)$$

where t = interventions, and θ = uncertainty in parameters.

Because information provided by research is a public good (once generated it can be used to inform the decision for all patients), the societal value of research should be calculated across the population of potential programmed participants [1,2]. Here, the estimate of this potential population is based on an estimate of the population from the United States who would be similar to the trial subjects (in terms of age and diagnosis) and would face the same treatment decision regarding management with either rate- or rhythm-control. This population was estimated to be 4.16 million patients over 5 years, based on a prevalence estimate of 2.3 million people [37] and an incidence of 500,000 cases per year [38]. The population calculation involves discounting at 3% per annum (after the first year) to account for time preference. The population EVPI is presented for three values for the cost-effectiveness threshold: $20,000, $50,000, and $100,000 per life-year.

Results

Figure 1 presents the Kaplan–Meier curves illustrating the survival over time for the scenario when survival is not adjusted (no adjustment) and the scenarios when survival is adjusted (partial adjustment and full adjustment).

Table 1 reports the mean survival, total cost, and ICER associated with rate- versus rhythm-control treatment for the three scenarios that differed in how adjustments were made for censoring. These are all computed as sample averages.

In each case, rate-control is less costly on average than rhythm-control. In the two scenarios where censoring of survival is taken into account, rate-control is also more effective on average and dominates rhythm-control. When no account is made for censoring, rate-control is associated with a nonsignificant shorter mean survival (–0.0009 years, 95% CI –0.08 to 0.07) compared to rhythm-control. Nonetheless, after considering the joint distribution of costs and effects, rate-control remains the favored approach because of the lower cost compared to rhythm-control (–$4800, 95% CI –$6624 to –$2923), despite the absence of a statis-
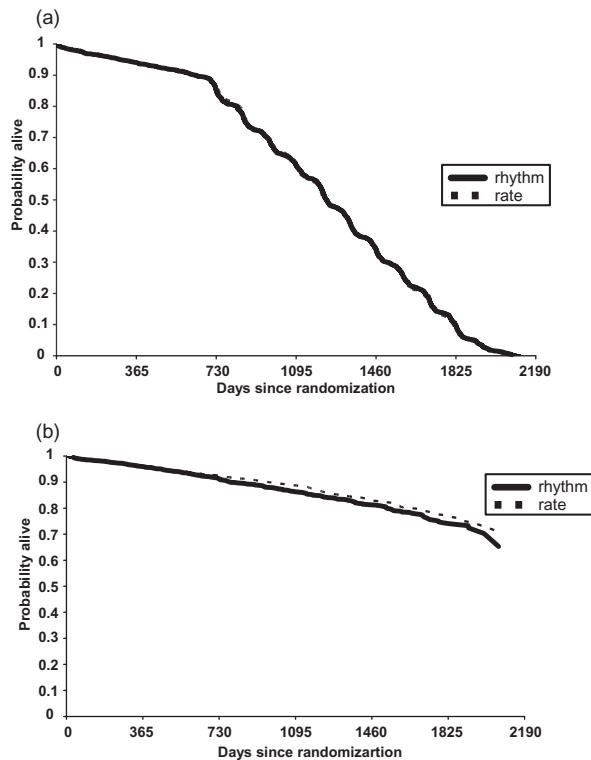
Figure 1 (a) Kaplan–Meier curves nonadjusted survival time. (b) Kaplan–Meier curves adjusted survival time.

tically significant difference in survival between the two strategies. Therefore, in each case, the decision made on the basis of expected ICER estimates is to adopt rate-control, either because it dominates (where censoring is applied) or because the associated ICER (saving almost $5 million for every life-year given up in the case of no censoring) exceeds any reasonable threshold. When both the incremental cost and incremental survival are negative, the decision rule becomes select if ICER greater than threshold.

Figure 2 illustrates the incremental cost-effectiveness plane, for the comparison between rate-control and rhythm-control, for each scenario. Each point represents one replicate (incremental cost and

incremental survival) from the bootstrap. Figure 2a represents the case where censoring is not taken into account. The location and spread of the incremental cost-effect pairs in the vertical direction indicates that there is no uncertainty regarding the existence of cost-savings with the rate-control strategy compared to the rhythm-control strategy (all points fall below the horizontal axis), although there is some uncertainty about the magnitude of the cost-savings (incremental savings vary from $2923 to $6624). With regard to effectiveness, there is uncertainty regarding whether and the extent to which rate-control confers a survival benefit compared to rhythm-control (from −0.08 to 0.07 years). This is consistent with the finding of a nonsignificant difference in survival gain between the two treatment groups. Approximately one-half of replicates (51.1%) were located to the left of the vertical axis (negative incremental survival), indicating that there was considerable uncertainty surrounding the effectiveness of rate-control.

Figure 2b represents the case where censoring is taken into account in terms of survival only. The location of the replicates in the vertical (cost) plane is the same as the scenario with no account for censoring. Within the horizontal plane, the location of the replicates indicates that there is much less uncertainty about whether rate-control is effective compared to rhythm-control (now only 4.6% of replicates involved negative incremental survival). Nevertheless, the spread of the replicates indicates that there is slightly more uncertainty surrounding the extent of the survival difference between the two treatments (−0.01 to 0.17 years).

Figure 2c represents the case where censoring is taken into account in terms of both costs and survival. The location and spread of the incremental cost-effect pairs within the vertical plane indicates that there is uncertainty regarding the existence and extent of cost-savings with the rate-control strategy in comparison to the rhythm-control strategy (+$1810 to −$8438). The majority (90.4%) of the replicates were located below the horizontal axis (negative incremental cost), indicating that rate-control was most often cost-saving compared to rhythm-control. Within the horizontal plane,

Table 1 Mean expected survival, total cost, and incremental cost-effectiveness ratio (ICER) for each censoring scenario

Censoring	Total cost ($)		Mean survival (years)		Incremental (rate—rhythm)		ICER of rate vs. rhythm ($ per life-year)	Decision
	Rate	Rhythm	Rate	Rhythm	Cost ($)	Survival (years)		
None	20,595	25,375	3.1869	3.1878	−4,800 (−6,624 to −2,923)	−0.0009 (−0.08 to 0.07)	4,983,477	Chose rate-control
Partial*†	20,595	25,375	4.6749	4.5983	−4,800 (−6,624 to −2,923)	0.08 (−0.01 to 0.17)	Rate-control dominates	Chose rate-control
Full‡	32,048	35,509	4.6749	4.5983	−3,461 (−8,438 to 1,810)	0.08 (−0.01 to 0.17)	Rate-control dominates	Chose rate-control

*Survival only.
†These figures vary slightly from those published in the original cost-effectiveness article [5]. This is due to the use of a different bootstrapping sample for the two analyses.
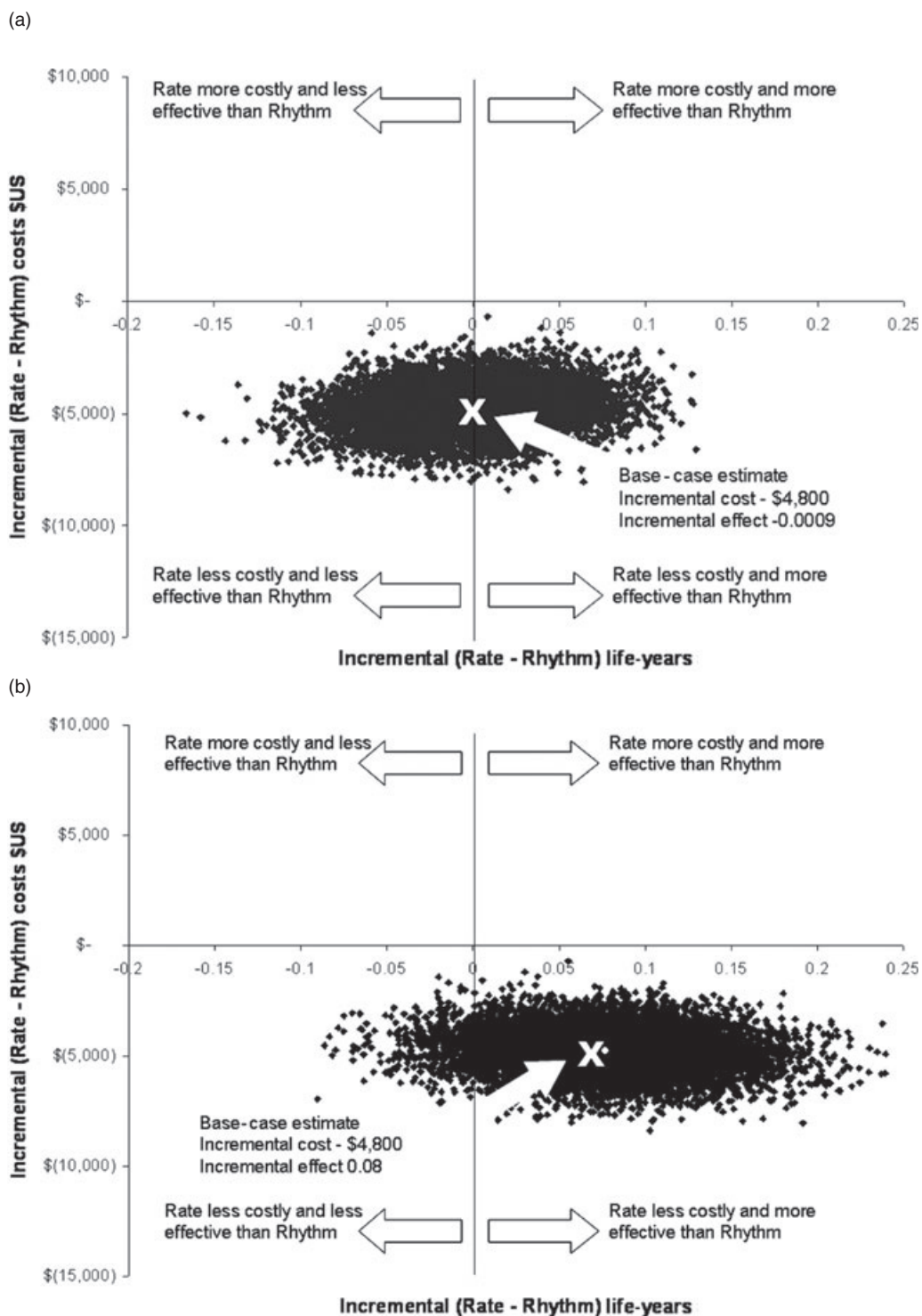‡Both survival and costs.

(a)

(b)

Figure 2 (a) Incremental cost-effectiveness plane of rate-control versus rhythm-control: no censoring. (b) Incremental cost-effectiveness plane of rate-control versus rhythm-control: partial—censoring of survival only. (c) Incremental cost-effectiveness plane of rate-control versus rhythm-control: full—censoring of costs and survival.

the replicates were identical to censoring of survival only (Fig. 2b). In addition, there was a small proportion of replicates (0.48%) that were located both above the horizontal axis (positive incremental cost) and to the left of the vertical axis (negative incremental survival), indicating the potential for rate-control to be both more costly and less effective than (dominated by) rhythm-control.
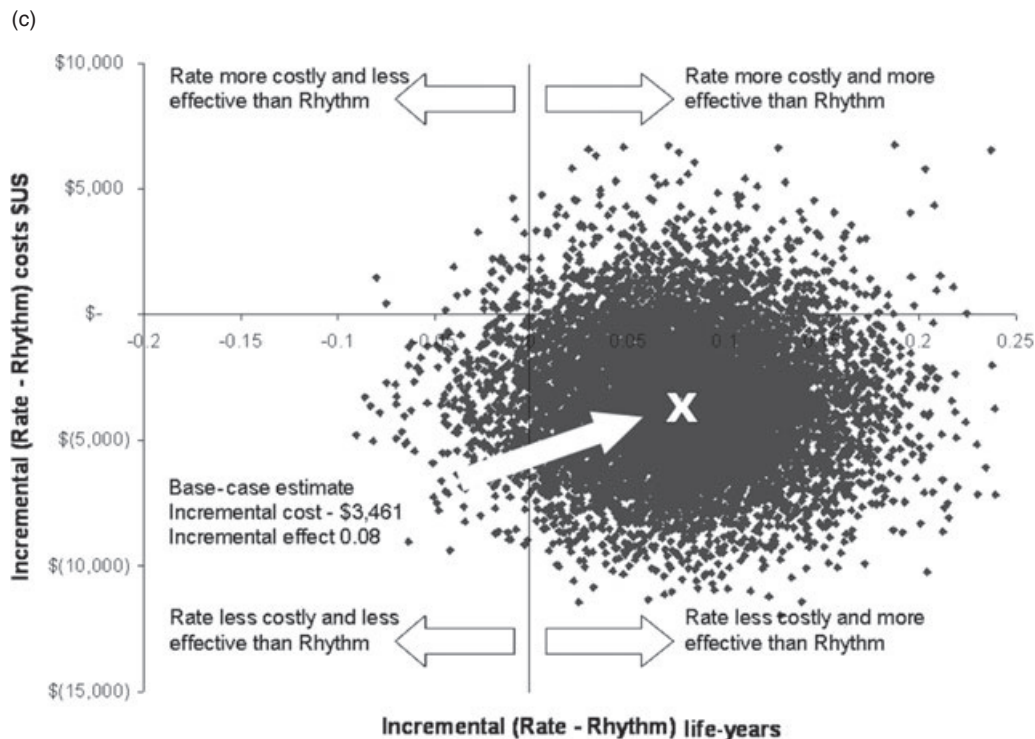
(c)

Figure 2 Continued.

Figure 3 illustrates the cost-effectiveness acceptability curves for rate-control compared to rhythm-control for each of the levels of censoring adjustment, calculated from the bootstrap replicates as the proportion where rate-control is associated with the maximum NB. Given the data, over the range specified for the threshold (λ) ($0 to $100,000), the probability that rate-control is cost-effective compared to rhythm-control is high (>89%) regardless of the level of adjust-

ment for censoring (note the discontinuation of the axis). Nevertheless, the level of adjustment for censoring has an impact upon the extent of the decision uncertainty. When censoring is taken into account for survival only (partial censoring), the probability that rate-control is cost-effective remains above 95% irrespective of the threshold. This reflects the minimal uncertainty regarding the effectiveness of rate-control compared to rhythm-control (only 4.6% of the cost-
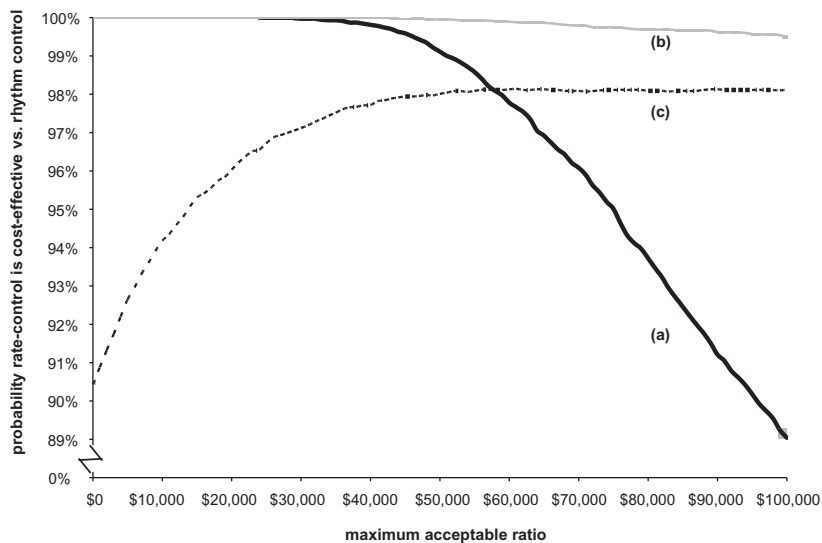

Figure 3 Cost-effectiveness acceptability curves for rate-control versus rhythm-control: (a) no censoring, (b) partial—censoring of survival only, and (c) full—censoring of costs and survival.

Table 2 Expected value of perfect information (EVPI) for the eligible population for each censoring scenario at different thresholds (λ)

Censoring	EVPI ($) at		
	$\lambda = \$25,000$	$\lambda = \$50,000$	$\lambda = \$100,000$
No	11,000	23 million	831 million
Partial*	0	626,000	33 million
Full†	179 million	117 million	159 million

*Survival only.
†Both survival and costs.

effectiveness pairs involved negative incremental survival). In each of the other cases, the probability falls below 95% over some range of the threshold ($\lambda > \$73,000$ for the scenario with no account for censoring, $\lambda < \$17,000$ for the scenario with censoring taken into account for both costs and survival), indicating the presence of some decision uncertainty.

In addition, the cost-effectiveness acceptability curve for the full-censoring scenario has a different shape to those for the other scenarios, an increasing rather than decreasing probability that rate-control is cost-effective as the threshold increases. This reflects the uncertainty concerning the existence of cost savings associated with rate-control compared to rhythm-control. In this scenario, 9.6% of the cost-effectiveness pairs involved positive incremental cost; as such, the probability that rate-control is cost-effective given a threshold of zero (the decision-maker is only interested in cost-savings) is below 100% (90.4%). Nevertheless, as the value of the threshold increases, these cost-effect pairs begin to look cost-effective and the probability that rate-control is cost-effective increases. See Fenwick et al. for a full discussion of the relationship between cost-effectiveness pairs in the cost-effectiveness plane and the shape of cost-effectiveness acceptability curves [15].

Table 2 details the EVPI for the population associated with the decision, given a threshold (λ) of $25,000, $50,000, and $100,000 per life-year gained. Using the example of $50,000 per life-year gained as the threshold (λ), when no account is made of censoring, the EVPI surrounding the decision is $23 million. When censoring is taken into account in both costs and survival, the EVPI surrounding the decision is $117 million. Nevertheless, when censoring is only taken into account for survival, the EVPI surrounding the decision is only $626,000. Figure 4 illustrates the population EVPI associated with the different censoring scenarios over a range of values for the threshold. In the cases where censoring for cost is not taken into account (no censoring and partial censoring), the EVPI rises. This is because both the uncertainty surrounding the decision (error probability) and the value of the threshold (value of the consequences of an error) are rising. In the case where censoring is taken into account for both costs and survival, the EVPI falls initially as the threshold rises, because the reduction in uncertainty outweighs the increased valuation of the consequences associated with an incorrect decision. As the threshold increases beyond $53,000 per life-year, the EVPI rises with the threshold. This corresponds to the point where the cost-effectiveness acceptability curve levels off (the reduction in the decision uncertainty slows), and thus reflects the fact that the value of the consequences of an error (measured by the threshold) outweighs the reduction in the error probability over this range.

Discussion

In this article we examine the impact of the level of censoring adjustment undertaken on the two decisions faced by a rational decision-maker using data from a
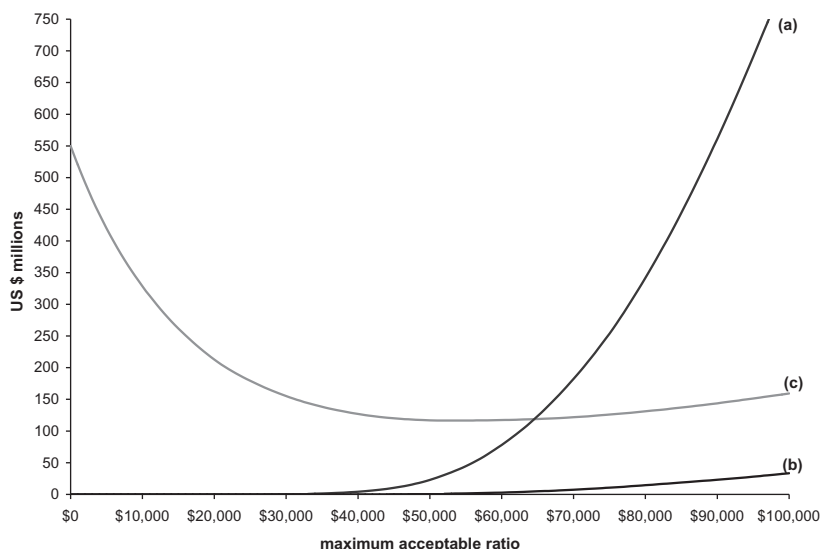

Figure 4 Population EVPI: (a) no censoring, (b) censoring of survival only, and (c) censoring of costs and survival. EVPI, expected value of perfect information.

large trial of two strategies for treating atrial fibrillation [5]. The trial showed a small, nonsignificant difference in survival between the two strategies. Only 16% of the subjects died; thus 84% were censored. This high amount of censoring is a common situation in the literature.

We found that the level of adjustment for censoring did not affect the decision about which intervention to adopt, despite the converse effectiveness result associated with the uncensored analysis (that rhythm-control was more effective than rate-control). Nevertheless, the level of censoring may impact the decision to fund additional research, because the estimated population EVPI varies by several orders of magnitude depending on the censoring scenario. For the partial-censoring scenario (survival only), EVPI is estimated as $626,000 for a threshold of $50,000 per life-year gained, suggesting that further research, the cost of which is likely to exceed this value, may not be worthwhile. Nevertheless, for the scenarios involving either no censoring or full censoring, the estimates of EVPI are $23 and $117 million dollars respectively, and as such, further research is likely to be worthwhile. In this case, the variations in the estimates of the EVPI across the three censoring scenarios are driven mainly by differences in the error probabilities surrounding the decisions (as measured by the inverse of the cost-effectiveness acceptability curve), although differences in the consequences of making an error (as measured by the net benefits associated with the decisions) have some impact on the results. Thus, the EVPI estimates are similar for no censoring and partial censoring, over the range of thresholds for which the uncertainty associated with each is similar (up to a value of approximately $30,000 per life-year), and the EVPI estimate is greatest for the no-censoring scenario for values of the threshold above $65,000 per QALY (approximately), where the uncertainty surrounding the decision exceeds that associated with the other scenarios. The partial-censoring scenario is associated with the lowest level of decision uncertainty (see Fig. 3), and this is reflected in the lower EVPI estimates generated for the partial-censoring scenario.

These results should be interpreted within the context that EVPI alone is not sufficient for determining the worth of further research. In the situation where the EVPI suggests that further research would be potentially worthwhile, additional analysis could be performed to determine the EVPI for a particular parameter or group of parameters (e.g., economic parameters, clinical parameters) to assess the (potential) worth of research focused on different facets of the decision [1,16]. The process involves determining the increase in the expected value of the decision associated with resolving the uncertainty concerning a parameter or group of parameters. Nevertheless, perfect information is not achievable with a finite

sample size, and the expected value of partial perfect information still provides only a maximum value for further research which can be compared to the cost to determine whether the research is potentially worthwhile (necessary condition). Determining whether specific research, with a finite sample size, is worthwhile requires an analysis of the expected value of sample information. This involves valuing the reduction in uncertainty, and hence the increase in the expected value of the decision, actually achievable through research, and depends upon the extent to which uncertainty and the associated consequences are actually reduced by the information provided from research (the informativeness of the research) [16].

There are other methods of addressing censoring in cost-effectiveness analysis that we did not examine in this article. Bang and Tsiatis [27] suggested a number of nonparametric-solutions estimators of costs in the presence of censoring based upon inverse weighting techniques. Unlike those of Lin et al. [21], their estimators are shown to be consistent regardless of censoring pattern. Bang and Tsiatis [27] provide both an estimator based on total costs only (simple weighted case estimator), and an estimator that partitions the study into intervals similar to Lin et al. [21]. Willan et al. [39] provide methods to estimate the mean and variance of the incremental net benefit statistic under conditions of censoring. These methods can be applied when the health outcome is either mean survival time or mean quality-adjusted survival. More recently, the application of regression techniques to estimate costs and effects in the presence of censoring has been discussed in the literature [28,40,41]. One of the advantages of using regression techniques is the ability to include covariates in the estimation of costs and effects in the presence of censoring.

We selected the KMSA approach described by Lin et al. [21] for censoring, because this approach has been shown to provide consistent estimators of average costs if it is assumed that censoring occurs at the boundaries of the intervals. Furthermore, it is the most commonly applied approach in the literature. Nevertheless, the method does have limitations because there is no reason to expect censoring to occur at the boundaries of the selection intervals. Therefore, the assumption about the consistency of the estimators will likely be violated to some degree in most cases [21].

Conclusion

This analysis illustrates that it is only when censoring is taken into account for both costs and effects (the full-censoring scenario) that the decision uncertainty and the value of information are appropriately identified. In particular, the partial-censoring approach used in the analysis of the AFFIRM trial [5] underestimated

the decision uncertainty and hence the EVPI, although in this particular example, there was no impact on the decision regarding cost-effectiveness. This partial approach is often employed in cost-effectiveness analyses. The results shown here suggest that there may be implications for the results of these studies.

Supplementary Material

Supplementary material for this article can be found at: http://www.ispor.org/valueinhealth_index.asp.

References

1 Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence (NICE). Lancet 2002;360:711–5.

2 Claxton K, Posnett J. An economic approach to clinical trial design and research-priority setting. Health Econ 1996;5:513–24.

3 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ 1999;18:341–64.

4 Young TA. Estimating mean total costs in the presence of censoring. A comparative assessment of methods. Pharmacoeconomics 2005;23:1229–42.

5 Marshall DA, Levy AR, Vidaillet H, et al. Cost-effectiveness of rhythm versus rate control in atrial fibrillation. Ann Intern Med 2004;141:653–61.

6 Phelps CE, Mushlin A. On the (near) equivalence of cost-effectiveness and cost-benefit analyses. Int J Technol Assess Health Care 1991;7:12–21.

7 Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making 1998;18(Suppl. 2):S68–80.

8 Tambour M, Zethraus N, Johannesson M. A note on confidence intervals in cost-effectiveness analysis. Int J Technol Assess Health Care 1998;14:467–71.

9 Briggs AH, Gray AM. Handling uncertainty in economic evaluations of healthcare interventions. Br Med J 1999;319:635–8.

10 Willan AR, O'Brien BJ. Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. Health Econ 1996;5:297–305.

11 Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Econ 1997;6:327–40.

12 Briggs AH, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. Health Econ 1998;7:723–40.

13 van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ 1994;3:309–19.

14 Fenwick E. Marshall DA, Levy AR, Nichol G. Using and interpreting cost-effectiveness acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. BMC Health Serv Res 2006;6:52.

15 Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves—facts, fallacies and frequently asked questions. Health Econ 2004;13:405–15.

16 Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Making 2004;24:207–27.

17 Leung KM, Elashoff RM, Afifi AA. Censoring issues in survival analysis. Annu Rev Public Health 1997;18:83–104.

18 Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457–81.

19 Cox DR. Regression models and life tables. J R Stat Soc Ser B (Methods) 1972;34:187–220.

20 Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data. New York: John Wiley & Sons, 1980.

21 Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. Biometrics 1997;53:419–34.

22 Fenn P, McGuire A, Phillips V, et al. The analysis of censored treatment cost data in economic evaluation. Med Care 1995;33:851–69.

23 Etzioni R, Feuer EJ, Sullivan S, et al. On the use of survival analysis to estimate medical care costs. J Health Econ 1999;18:365–80.

24 Etzioni R, Urban N, Baker M. Estimating the costs attributable to a disease with application to ovarian cancer. J Clin Epidemiol 1996;32:95–103.

25 O'Hagan A, Stevens JW. On estimators of medical costs with censored data. J Health Econ 2004;23:615–25.

26 Baser O, Gardiner JC, Bradley CJ, et al. Longitudinal analysis of censored medical cost data. Health Econ 2006;15:513–25.

27 Bang H, Tsiatis AA. Estimating medical costs with censored data. Biometrika 2000;87:329–43.

28 Willan AR, Lin DY, Manca A. Regression methods for cost-effectiveness analysis with censored data. Stat Med 2005;24:131–45.

29 Raikou M, McGuire A. Estimating medical care costs under conditions of censoring. J Health Econ 2004;23:443–70.

30 The Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) Investigators. A comparison of rate control and rhythm control in patients with atrial fibrillation. N Engl J Med 2002;347:1825–33.

31 The Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) Investigators. Base-

line characteristics of patients with atrial fibrillation: the AFFIRM Study. Am Heart J 2002;143:991–1001.

32 The Planning and Steering Committees of the AFFIRM Study for the NHLBI AFFIRM Investigators. Atrial fibrillation follow-up investigation of rhythm management—the AFFIRM study design. Am J Cardiol 1997;79:1198–202.

33 Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-Effectiveness in Health and Medicine. Oxford: Oxford University Press, 1996.

34 Drummond MF, Sculpher MJ, Torrance GW, et al. Methods for the Economic Evaluation of Health Care Programmes, (3rd ed). Oxford: Oxford University Press, 2005.

35 Karlsson G, Johannesson M. The decision rules of cost-effectiveness analysis. Pharmacoeconomics 1996; 9:113–20.

36 Black W. The CE plane: a graphic representation of cost-effectiveness. Med Decis Making 1990;10:212–4.

37 Go AS, Hylek EM, Phillips KA, et al. Prevalence of diagnosed atrial fibrillation in adults. JAMA 2001; 285:2370–5.

38 Greenlee RT, Vidaillet H. Recent progress in the epidemiology of atrial fibrillation. Curr Opin Cardiol 2005;20:7–14.

39 Willan AR, Chen EB, Cook RJ, Lin DY. Incremental net benefit in randomized clinical trials with quality-adjusted survival. Stat Med 2003;22:353–62.

40 Carides GW, Heyse JF, Iglewicz B. A regression-based method for estimating mean treatment cost in the presence of right-censoring. Biostatistics 2000;1:299–313.

41 Lin DY. Regression analysis of incomplete medical cost data. Stat Med 2003;22:1181–200.