Approaches to subgroup economic analyses in clinical trials- the case of cardiovascular disease prevention

by

Borislava Mihaylova, Andrew Briggs and Alastair Gray

Health Economics Research Centre, University of Oxford, UK

Paper presented at the Health Economics' Study Group Meeting,
Oxford 5th -7th January, 2005

Abstract

Clinical trials often report similar relative treatment effect across different patient populations in terms of relative risks reduction or hazard ratios. However, the absolute benefits of these interventions could differ widely in populations at different initial disease risk. It is the absolute additional benefits and costs that are relevant when undertaking economic analyses to estimate the cost-effectiveness of interventions. Such economic analyses frequently report cost-effectiveness for different population groups, e.g. by age and sex, but it is not always clear which methods are employed or whether they are appropriate. Therefore, there is a need to evaluate different approaches to subgroup analysis in economic evaluation and to suggest which might be most appropriate.

In this study we suggest that it will often be better to define subgroups by multivariate risk than by single risk factors. Three main approaches to subgroup analyses are then identified - within subgroup analysis, use of trial wide relative effects, and empirical Bayes (shrinkage) subgroup estimates. Each is then applied to a data set for a cardiovascular disease preventive intervention, and the results are reported and compared in terms of central estimates, impact on the uncertainty of the results, and ability to reliably inform decision-makers on the cost-effectiveness of the intervention in patient subgroups. We argue that the use of trial-wide estimates to inform the subgroup analyses, where appropriate, avoids the pitfalls of underpowered within subgroup analyses, and that these results are more useful for decision-making compared to the empirical Bayes subgroup estimates.

Introduction

Well designed and conducted randomized controlled trials are the gold standard when evaluating the effects of health interventions and traditionally much effort has been put into improving their design and analysis. Therefore, such trials are considered to be the most rigorous way of determining whether the treatment is causally related to the observed outcomes, and their results are well accepted by the clinical audience. At the same time, given limited healthcare resources, cost-effectiveness analyses are increasingly being used to inform on the value-for money of health interventions and to aid the decision on the appropriate treatment thresholds. As both

effectiveness and cost-effectiveness analyses provide evidence that is used to decide for or against the provision of an intervention, they should be of a similarly high quality.

In the field of cardiovascular disease, clinical trials often report similar relative treatment effects across different patient populations in terms of relative risks reduction or hazard ratios (1-3). Relative effects are often considered to more closely describe the underlying nature of the treatment effect and to provide a more stable measure of treatment effect across a heterogeneous population. However, in the presence of different baseline disease risk, these similar relative treatment effects will result in different absolute benefit of interventions(4;5). Economic analyses estimate the incremental costs per unit of health benefit, and so their interest is in estimating the absolute impact of the interventions on the health outcomes and costs.

In this paper we aim to explore the appropriate methods to estimate the cost-effectiveness of cholesterol-lowering with statin across a heterogeneous patient population with widely differing baseline disease risk, using data from a large clinical trial.

The methods employed to evaluate the heterogeneity of the cost-effectiveness of interventions based on trial data have not received much attention. Few approaches have been suggested in the literature. Tests for qualitative and/or quantitative interactions have been suggested as a useful initial step in the investigations. Unfortunately the power of these tests is limited and they can not accommodate evidence beyond the one provided in the study of interest. Independent analyses in well specified patient subgroups that report within subgroup central estimates and confidence intervals have also been published. These analyses are generally criticized for their potential to propagate spurious results (6). This within-subgroup approach has also been used in economic analysis, mainly with reference to important and easily identifiable patient populations (e.g. diabetic patients (7) (8)). Random effect models have received significant attention in the health economics literature with respect to appropriate analysis of multinational (or multi-centre) trials. Some authors show that this method provides a more correct way of estimating the trial-wide treatment effect given the possible within-country (or within-centre) clustering (9) (10), as well as providing results that are potentially more generalisable outside the trial setting. Other authors add that this approach also allows the evaluation of country-specific absolute benefits, costs and

cost-effectiveness (see Manca et al (11), Pinto et al (12)). These group-specific estimates (often referred to as shrinkage or empirical Bayes estimates) have been known in the literature to provide prediction with smaller overall mean squared error (13-15). And finally, decision models constructed upon trial risk equations have been widely used to estimate cost-effectiveness for different sets of patient and/or treatment characteristics. All of these approaches but the last could be regarded as trial-based, but all have the deficiency of using only part of the data or introducing additional assumptions; the last approach uses all trial data, but is somewhat separated from the original trial design, employs a number of structural assumptions and is often referred to as a modeling study.

Another problem when addressing the heterogeneity of cost-effectiveness is the specification of the underlying population heterogeneity. Population sub-groups are frequently specified on the basis of a single factor, such as diagnosis of diabetes, presence or absence of previous disease, gender, or smoking status. However, as risk is generally multifactorial, no single risk factor is likely to define the full range of risk in the population of interest; simultaneously, any sub-group defined using a single risk factor may well span a wide range of risk, due to the presence or absence of other risk factors. Insufficient attention has been paid to these issues, or more generally to methods that might better define the heterogeneity across the population.

In this paper we investigate approaches to describing the heterogeneity of baseline disease risk and evaluating cost-effectiveness in the trial population with respect to this heterogeneity.

Methods

We use the 20,536-large Heart Protection Study, which investigated the effect of 40mg simvastatin daily on major vascular events (MVEs, defined as heart attacks, strokes or revascularizations)(1). The aim is to estimate the incremental cost per MVE avoided during the mean follow-up of five years for patients at different cardiovascular disease risk. A number of simplifying assumptions have been made. First, administrative censoring is ignored as it is minimal (only in the last year) and is identical across treatment and risk strata. Second, we have found that the incremental statin treatment cost is similar across risk groups, and so this has not

been explicitly modeled; instead the mean trial-wide incremental treatment cost is added to the vascular related hospitalization cost savings for the purpose of subgroup economic analyses. Third, as no differences in the non-vascular hospitalization costs were observed between trial arms, we have decided to exclude these costs from the economic analysis on the grounds that they would have increased the random error and would not have provided additional information.

The risk of first within trial MVE was modeled using Cox proportional hazards models in order to define the hetereogeneity of the baseline disease risk in the trial population. The explanatory variables employed were baseline prior vascular disease or diabetes, sex, age, plasma LDL and HDL cholesterol, midpoint of systolic and diastolic blood pressure, smoking status, plasma creatinine, statin allocation and other possible risk factors. The 5-year risk of MVE was then predicted for each trial participant, and all participants were ordered in terms of increasing risk and split into five equal risk groups (with about 4,100 trial participants within each group).

After specifying the patient subgroups that provide a good description of heterogeneity in the trial, three approaches to subgroup analysis are studied (a) within subgroup estimates, (b) empirical Bayes subgroup estimates, and (c) subgroup estimates based on trial-wide treatment effects and the placebo rates of costs and outcomes in the respective subgroups.

In the first approach (Model 1), the within subgroup MVEs avoided and vascular hospitalization costs saved are estimated based on linear regression models for MVEs and vascular hospitalization costs (CVE) with the constant term and the treatment effect fully interacted with the subgroups investigated (e. g. subgroup specific constant terms and treatment effects). The baseline and treatment subgroup effects are fixed effects in this model. The correlation between costs and effects is accounted for by concurrent estimation of both linear models on the same patient sample for the purpose of the evaluation of the mean cost-effectiveness and its uncertainty.

Model 1 Linear fixed effects model

$MVE_i = \mu o + so_k + to + tso_k + eo_i$	$CVE_i = \mu c + sc_k + tc + tsc_k + ec_i$
MVE_i, number MVEs for patient i	CVE_i vascular related hospitalisation costs for patient i
μo, intercept	μc, intercept
so_k, the kth subgroup effect	sc_k, the kth subgroup effect
to, statin allocation	tc, statin allocation
tso_k, the treatment effect for the kth subgroup	tsc_k, the treatment effect for the kth subgroup
eo_i, error term for patient $i, eo_i \ : \ N(0,\sigma_o^2)$	ec_i, error term for patient i, $\quad ec_i : \ N(0,\sigma_c^2)$

$ICER_k = {tsc_k}/{tso_k}$, incremental cost-effectiveness ratio in k-th risk group, k=1..5, risk group

In the second approach (Model 2) a linear mixed effects model is structured in which the subgroup treatment effects are modeled as random effects. In this model it is assumed that the subgroup effects (on the intercept and treatment) are drawn from normal distributions with means 0 and variances σ^2_{so} and σ^2_{tso}; and σ^2_{sc} and σ^2_{tsc} respectively for outcomes and costs. In addition, correlation between the random effects is allowed for as it is expected that risk groups with higher baseline rate of events could gain larger absolute benefit from intervention. The empirical Bayes (or shrunken) estimates for the treatment effects on MVEs and hospitalization costs of the risk subgroups are subsequently evaluated.

Model 2 Model with subgroup effects defined as random effects

$MVE_i = \mu o + so_k + to + tso_k + eo_k$	$CVE_i = \mu o + sc_k + tc + tsc_k + ec_k$
MVE_i, number MVEs for patient i	CVE_i vascular related hospitalisation costs for patient i
μo, intercept	μc, intercept
so_k, the kth subgroup effect,	sc_k, the kth subgroup effect,
$so_k : \ N(0,\sigma^2_{so})$	$sc_k : \ N(0,\sigma^2_{sc})$
to, statin allocation	tc, statin allocation
tso_k, the treatment effect for the kth subgroup	tsc_k, the treatment effect for the kth subgroup
$tso_k : \ N(0,\sigma^2_{tso})$	$tsc_k : \ N(0,\sigma^2_{tsc})$
eo_i, error term for patient $i, ec_i : \ N(0,\sigma^2_o)$	ec_i, error term for patient i, $\quad ec_i : \ N(0,\sigma^2_c)$

$ICER_k = {\dot{tsc}_k}/{\dot{tso}_k}$, where $\dot{tsc}_k$ and $\dot{tso}_k$ are the shrinkage subgroup estimates

$ICER_k$, Incremental cost-effectiveness ratio in k-th risk group, $k = 1..5$, risk group

While the first two approaches employ different statistical models to estimate the subgroup specific incremental outcomes and costs, the third approach (Model 3) employs structural assumptions that the intervention has a similar relative impact on MVEs and a similar relative impact on the cost of vascular related hospitalizations. Therefore the ratio of MVEs (or vascular related hospitalization costs) in the treatment and in the placebo groups of the risk subgroups are assumed to be equal to the respective ratios between the trial arms from the whole trial. Concerning impact on MVEs, this assumption has been supported by a number of other large RCTs that have investigated the effect of statins in primary or secondary CVD prevention populations and have shown that the relative effect is similar in different population groups (1) (2). As the MVEs are the main determinant of the vascular related hospitalization costs it is reasonable to expect that the proportional reductions in hospitalisation costs associated with vascular events would also be similar across different subgroups. Hence, the absolute reduction in the MVEs (or the costs of vascular event hospitalisations) in any particular subgroup is derived by applying the overall proportional reduction in MVEs (or vascular event costs) observed among all participants to the MVEs (or vascular event costs) observed in the placebo group for that particular subgroup.

Model 3 Model based on constant relative treatment effects on MVEs and vascular hospitalization costs

$MVE_i = \mu o + to + eo_i$	$CVE_i = \mu c + tc + ec_i$
MVE_i, MVEs for patient i	CVE_i, vascular hospitalisation costs of patient i
μo, intercept with estimated coefficient co_μ	μc, intercept with estimated coefficient cc_μ
to, statin allocation with estimated coefficient co_t	tc, statin allocation with estimated coefficient cc_t
eo_i, error term for patient i, $eo_i : N(0,\sigma_o^2)$	ec_i, error term for patient i, $ec_i : N(0,\sigma_c^2)$

The estimated relative treatment effects are $\frac{co_\mu + co_t}{co_\mu}$ (for outcome) and $\frac{cc_\mu + cc_t}{cc_\mu}$ (for vascular costs).

Therefore, the MVE avoided (vascular costs saved) are estimated as $(1-\frac{co_\mu + co_t}{co_\mu})*MVE_{pk}$ and $(1-\frac{cc_\mu + cc_t}{cc_\mu})*CVE_{pk}$, where MVE_{pk} and CVE_{pk} are the mean MVEs and mean vascular hospitalisation costs of placebo patients in the kth subgroup.

As it is expected that health outcomes (MVEs) and vascular related hospitalization costs are strongly correlated and that the distributions of outcomes and costs are skewed, a non-parametric

method is employed to evaluate the confidence intervals of our estimates. We use the non-parametric bootstrap method and draw 10,000 re-samples (with replacement) from the trial population. For each of these re-samples the Cox-proportional hazards model is re-estimated, participants are split into 5 equal-frequency groups based on their predicted 5-year risk of MVE and the three methods (both for MVEs and vascular related hospitalization costs) are applied. The results are summarized across the bootstraps after an appropriate ordering to provide 95% confidence intervals for the estimates of the incremental cost per MVE avoided. The widths of the confidence intervals for different models are compared.

The analyses were performed with the SAS 9.1 statistical software and the linear mixed effects models were fitted with proc MIXED (ridge-stabilized Newton-Raphson algorithm is used to optimize the likelihood function).

Results

Model selection techniques and expert opinion concerning the possible interdependence of the available explanatory variables guided the choice of the chosen multivariate risk model. The predicted 5-year risks of a MVE based on the final multivariate risk model were 12% (risk group 1), 18%, 22%, 28%, and 42% (risk group 5) respectively. Similar Cox proportional hazards models with statin allocation but using only single risk factors produced a much narrower spread of the baseline disease risk. For example, the predicted 5-year risk of a MVE for patients with LDL-cholesterol less than3.0 mmol/l; between 3.0 and 3.5 mmol/l and more than 3.5mmol/l ranged from 22% to 27%.

The cost-effectiveness results and the estimated uncertainty for the 5 multivariate risk groups under the three estimation approaches are presented in Table 1. We have chosen to model and evaluate separately the incremental outcomes (number of MVEs) and vascular related hospitalization cost savings (as opposed to modeling directly the net benefit). The incremental statin treatment cost per a trial participant randomized to simvastatin was £1,497 (SE 8) as compared to a trial participant allocated to placebo, and was added to the estimates of MVEs saved and vascular hospitalization costs avoided. An allocation to simvastatin produced a highly significant 25% (20–29; $p<0{\cdot}0001$) proportional reduction in MVEs. As a consequence of the

highly significant effect of statin allocation on vascular events, a highly significant 22% (95 CI 16-27) reduction in the costs of hospitalisations for vascular events was also observed. These are the trial wide treatment effects that were used in Model 3 together with the rate of MVEs and the vascular related hospitalisation costs in the placebo subgroups of these risk groups to evaluate the MVEs avoided, hospital cost savings and overall cost-effectiveness.

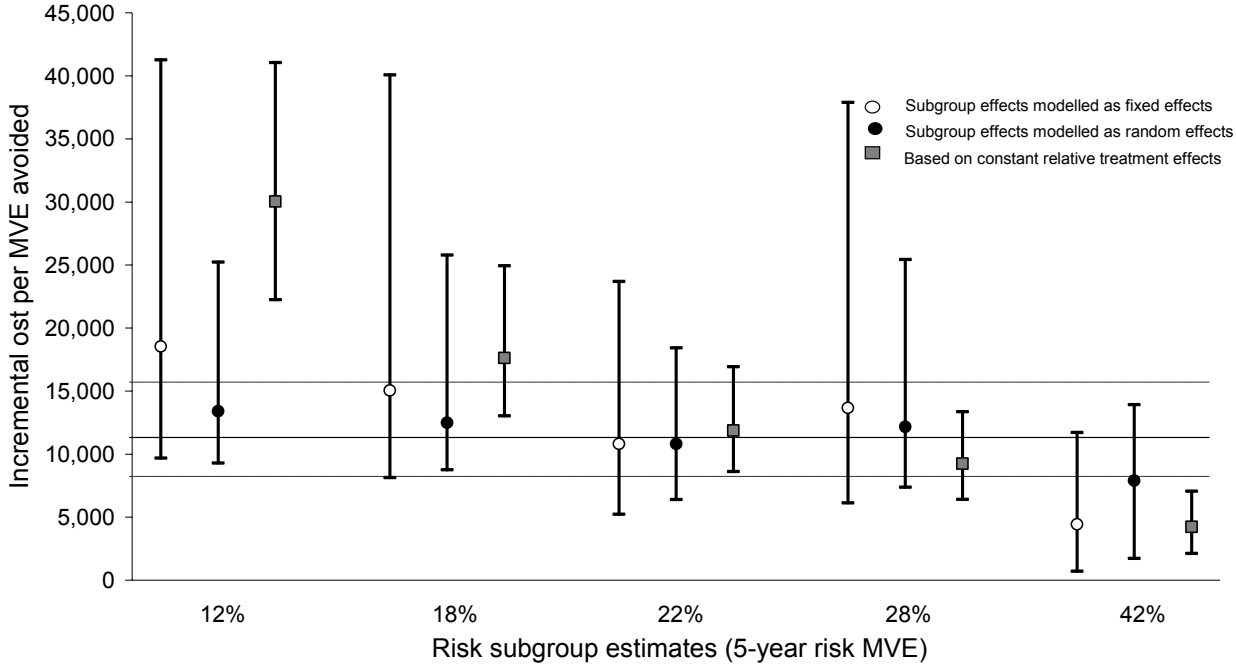
Table 1 Cost effectiveness results

Risk group	MVEs Avoided (SE)	Savings from vascular related hospitalizations (SE)	ICER (£ per MVE avoided) (95% CI)	Width of CI
Overall	**0.082 (0.009)**	**£501 (78)**	**11,153 (8,162, 15,648)**	**7,486**
Model 1 Linear regression model with fixed subgroup effects				
1 (low risk)	0.057 (0.015)	£432 (130)	18,540 (9,687, 41,282)	31,594
2	0.075 (0.020)	£374 (179)	15,050 (8,133, 40,075)	31,942
3	0.107 (0.023)	£344 (194)	10,811 (5,231, 23,694)	18,462
4	0.076 (0.024)	£453 (193)	13,668 (6,130, 37,894)	31,763
5 (high risk)	0.132 (0.029)	£911 (227)	4,438 (708, 11,721)	11,013
Model 2 Linear regression model with random subgroup effects				
1 (low risk)	0.076 (0.014)	£483 (94)	13,401 (9,294, 25,227)	15,933
2	0.083 (0.014)	£459 (139)	12,493 (8,755, 25,789)	17,035
3	0.097 (0.017)	£445 (137)	10,815 (6,387, 18,429)	12,042
4	0.083 (0.016)	£482 (137)	12,169 (7,370, 25,427)	18,058
5 (high risk)	0.108 (0.026)	£645 (232)	7,906 (1,728, 13,921)	12,193
Model 3 Use of constant treatment effects on MVEs and vascular hospitalisation costs				
1 (low risk)	0.041 (0.005)	£264 (50)	30,040 (22,245, 41,063)	18,818
2	0.064 (0.008)	£376 (67)	17,625 (13,037, 24,933)	11,896
3	0.088 (0.010)	£457 (78)	11,873 (8,607, 16,926)	8,319
4	0.101 (0.012)	£563 (91)	9,244 (6,407, 13,355)	6,949
5 (high risk)	0.153 (0.017)	£847 (137)	4,239 (2,131, 7,055)	4,923

It is important to note that the variances estimated for the random effects were small ($\sigma^2_{so} = 0.0267\ (0.0190)$, $\sigma^2_{tso} = 0.0004(0.0006)$, and $\sigma^2_{o} = 0.457(0.005)$ for MVE; $\sigma^2_{sc} = 906{,}936\ (653{,}773)$, $\sigma^2_{tsc} = 906{,}936\ (653{,}773)$, and $\sigma^2_{c} = 30{,}667{,}902(302{,}732)$ for vascular related hospitalization costs). Nevertheless they have an important impact on the subgroup estimates in the mixed effects

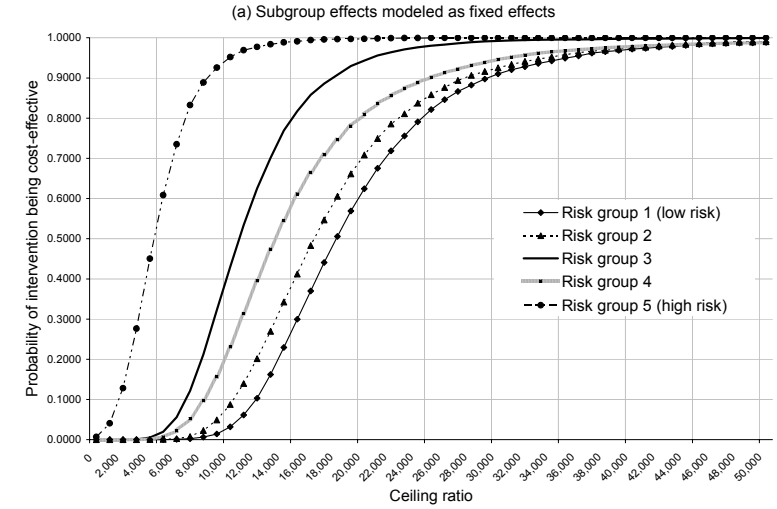
model. The estimated subgroup cost-effectiveness and their confidence intervals are presented in Figure 1 together with the overall trial-wide cost-effectiveness and its 95% confidence bands.

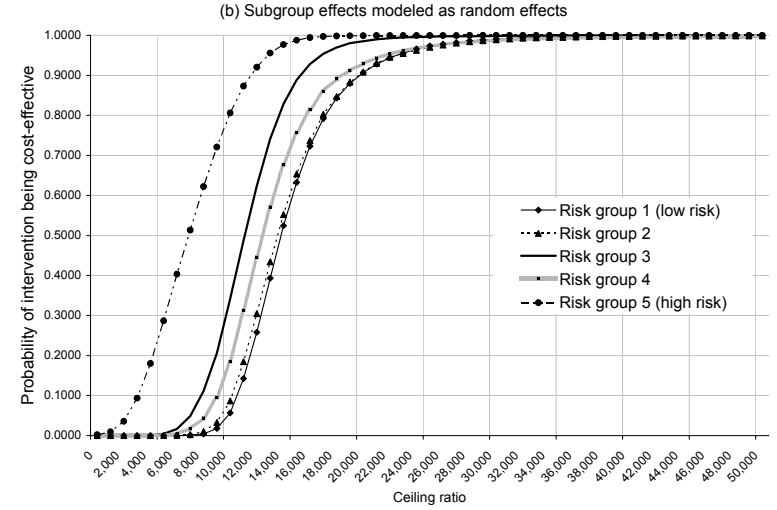
Figure 1 Subgroup incremental cost-effectiveness ratios and their confidence intervals

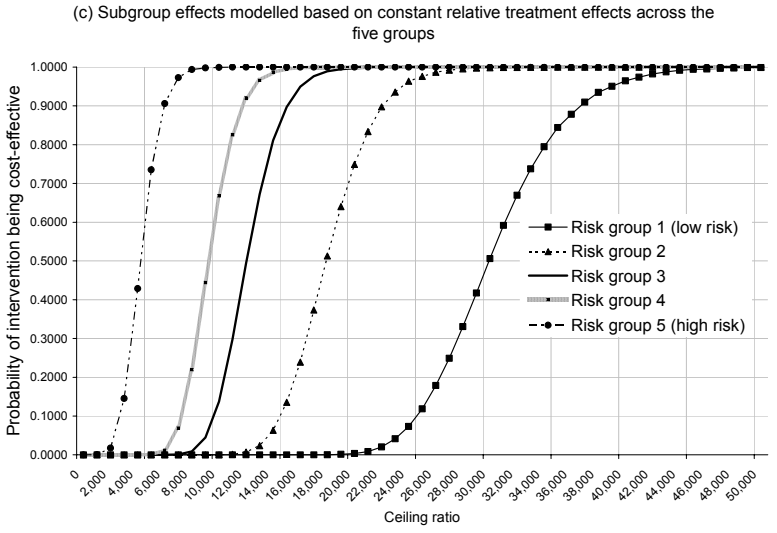


The cost-effectiveness acceptability curves are presented in Figure 2. The shape of the cost-effectiveness acceptability curves shows that the approach in which subgroup effects are modeled as random effects makes the cost-effectiveness estimates in the five risk groups seem more similar as compared to the pure within subgroup analysis. The approach that employs the constant treatment effects on the outcomes and cost-savings, on other hand, clearly differentiates the five risk groups in terms of their cost-effectiveness.

Fig. 2 Subgroup cost-effectiveness acceptability curves







The results show that even in this very large trial the within subgroup estimates have largely overlapping confidence intervals and are not statistically distinguishable. The use of a trial-wide treatment effect across the five risk subgroups resolves a large part of the uncertainty in the estimated cost-effectiveness (decrease of confidence intervals by 40-78% compared to the within subgroup analysis) and clearly outlines a trend of increasing cost-effectiveness with an increased baseline disease risk. The impact of the assumption of constant treatment effect across different patient groups on the confidence interval of the summary measure confirms that a large part of uncertainty in the subgroup estimates is due to the uncertainty in the treatment effect. The approach that uses random subgroup effects generally decreases the uncertainty of the subgroups' estimates (confidence intervals decreased by up to 50%) but the mean estimates are also shrunken towards the overall trial cost-effectiveness and therefore no sub-group effects are distinguishable.

Discussion

In the field of cardiovascular disease prevention there is a continuum of cardiovascular risk across the population and therefore interventions with similar relative risk will have different absolute benefits in population groups with varying disease risk. It is a recognized issue that the methods for estimating the absolute benefit of interventions to individuals, given results from clinical trials, are not well investigated (16). Trial-wide economic analysis, therefore, could mask serious differences in cost-effectiveness across patient groups at different baseline disease risk. In these cases cost-effectiveness of interventions needs to be evaluated in relation to the baseline disease risk in order to inform the decision on the appropriate treatment threshold given the available resources.

We have used the 20,536-large randomized controlled trial to explore the cost-effectiveness of allocation to 40mg simvastatin on subsequent MVEs. The heterogeneity in the baseline cardiovascular disease risk (as modeled by the Cox proportional hazards model) and the stable relative treatment impact across the risk groups is expected to lead to marked differences in cost-effectiveness measured as the incremental cost of avoiding a MVE. An important element of the analysis was the way we specified the heterogeneity of the baseline disease risk in the population. Importantly, although it is well known that cardiovascular risk is best described by multiple factors(17-19) analyses are still being performed based on single risk factors.

In this paper we compare three methods for trial based subgroup economic analysis that have the potential to estimate the cost-effectiveness of the intervention in patient groups at different disease risk. We have concentrated on the estimation methods (and have ignored the potential of the tests for interaction) as we believe that they potentially could inform decisions in the area we are interested in. We have also chosen to concentrate on methods applicable to patient-level trial data that more naturally capture the multifactorial heterogeneity in the baseline risk represented as compared to modeling studies that explicitly model this hererogeneity (20).

The main issues related to within subgroup economic analyses correspond to those of the similar effectiveness analysis – they are generally underpowered and often not pre-specified in the design and analysis of the study. As a result only very large subgroup differences could be detected, while the risk of spurious results is increased. Our results support these claims. Despite the large size of the trial, the confidence intervals of the within subgroup cost-effectiveness estimates are largely overlapping and no subgroup differences are statistically detectable. Furthermore, the absolute benefits of risk group four are somewhat worse than those of the lower risk group three, which is against expectation and is likely to be a play of chance.

When modeling subgroup effects as random, the main assumption is that the subgroup mean effects are a random sample from a distribution (e.g. Gaussian). Therefore, the shrinkage estimates are regarded as a compromise between ignoring the subgroups (e.g. report only a trial wide effect) and estimating the within subgroup effects that have been achieved based on the data available (under the additional distributional assumption and therefore the estimates are referred to as empirical Bayes estimates) (13). The way we structure our risk groups (by splitting the trial sample into equal frequency groups) could be regarded as a violation of the assumption of normality because we do not draw these risk groups at random from the overall population. On the other hand if the trial population is representative in any way from a larger population the equal frequency groups would postulate a reasonable spread of risk in the population of interest that should facilitate the estimation. In addition, it has been shown that the methods are ‘robust to the assumption of normal distribution, because they depend primarily on the means and variances of the sampling distributions and of the unknown parameters’ (13).

An additional limitation of our approach could be the use of only 5 groups to characterize heterogeneity. In all our simulations convergence has been reached, although the random effects were evaluated at 0 in about 15% of the simulations of health outcomes and in about 13% of simulations for hospitalization costs. It is important to admit that we have simulated based on more risk groups (17 groups) but the variances of the random effects were estimated at 0 for the trial data (although not for all of the bootstrap simulations), which would mean that the overall trial result is recommended for all subgroups (potentially due to the large increase in the random errors). Further work will investigate whether this is a result of the way we have modeled the highly skewed outcomes and costs or it is due to the highly diminished power to estimate subgroup effects.

Compared to the within subgroup estimates, the shrinkage estimates based on mixed effects models are usually drawn towards the overall mean (that is why they are called 'shrinkage'), and therefore make the mean estimates more convergent. At the same time, because of the use of the whole trial in the estimation, these estimates usually have tighter confidence intervals (by about 35-50% in our case). Our results (see Table 1) clearly show that the variability of the mean estimates between the groups has been substantially reduced, with cost-effectiveness ratios of between £8,000 and £13,400 based on shrinkage estimation as compared to £4,400 to £18,500 based on within subgroup estimates. Unfortunately, although as a whole the uncertainty diminishes, in the highest-risk group (risk group 5) the confidence interval of the estimated cost-effectiveness ratio increases by about 10%. It is important to note that the within subgroup estimates of the absolute treatment effects on the MVEs and vascular hospitalization costs are largest in this risk group; they are also furthest from the overall treatment effect and are more drawn to the overall treatment effects within the shrinkage estimation.

The use of constant treatment effect across different population groups have been widely used in epidemiology (for example to power trials or to estimate absolute treatment effects as NNT, etc.) but in the field of economic evaluation it has not been directly used for the purpose of trial based subgroup analysis. A common treatment effect (on some scale), though, is often used in decision models constructed to aid the estimation of cost-effectiveness of interventions (21;22). In this paper we are not explicitly exploring subgroup approaches based on modeling: these clearly

have a number of advantages, but also employ stronger assumptions and may require modeling of more parameters (e.g. baseline rates, interrelation between health events, etc.). We have chosen to investigate the performance of an approach that uses a single structural assumption (namely constant relative treatment effect across different populations) to supplement the subgroup data for the purpose of economic analyses. Our results, based on constant relative treatment impacts, clearly demonstrate a trend of increasing cost-effectiveness with increasing baseline disease risk, and also significantly decrease the uncertainty in the cost-effectiveness estimates. The large reduction in uncertainty indicates that the within-subgroup uncertainty is largely driven by the treatment effect.

It is important to note that when estimating uncertainty through the bootstrap the treatment effects on MVEs and vascular hospitalization costs (as well as the incremental treatment effect) are re-estimated and therefore the underlying uncertainty related to these ratios is a part of the analysis.

Clearly we have simplified our analysis by assuming that the incremental statin cost is constant across different patient populations, thus avoiding modeling this component of cost-effectiveness. Two main reasons justify this assumption. First, the within subgroup estimates of the incremental statin cost between trial arms were very similar and do not alter materially the cost-effectiveness results, and second, by using a single estimate (based on the whole trial) we have improved the precision of the cost-effectiveness estimates. In any case we believe that this assumption could be separated from the issues discussed in this paper, namely how best to estimate the subgroup specific cost-effectiveness when there is some evidence that it varies in the population.

Figure 1 graphically presents the estimates of cost-effectiveness in the five risk groups together with their confidence limits. Clearly, both the mixed effects model and constant relative effects model have had a pronounced impact on the uncertainty as compared to the within subgroup estimates (based on fixed effect models). While the constant relative effects model has decreased the uncertainty in all subgroups, the random effect shrinkage estimates have failed to decrease it in the highest risk group. Table 1 shows that the standard error for the shrinkage estimate of the incremental MVEs has slightly decreased, but for the vascular hospitalization cost has slightly

increased. Therefore, the increase of the uncertainty can be due to the joint effect of impacts on costs and effects or could be simply due to simulation error (e.g. the 95% CLs were not stabilized based on 10,000 bootstraps). Figure 2 presents the cost-effectiveness acceptability curves for the five risk groups based on the three estimation models. Fig. 2 part (a) clearly shows that the estimates based on fixed effects (or the within subgroup estimates) clearly differentiate the highest risk group (90% cost effective at about £8,000 per a MVE avoided) but is unable to differentiate well the rest of the risk groups (and the order of risk groups 3 and 4 is somewhat exchanged). Fig. 2 part (b) shows that the shrinkage estimates based on random effects models are drawn very much closer and the ranges of cost effectiveness at 90% probability is within £10,000, while in part (c) the subgroup estimates based on constant treatment effects provide a much clearer differentiation between the five risk groups.

In all approaches we have relied heavily on the central limit theorem and the large trial sample when we modeled the number of MVEs and vascular related hospitalization costs as normally distributed variables while both are highly skewed (and the distribution of MVEs is more likely to be Poisson than Normal). This simplification aims to more clearly present the impact of different methods. Although all methods employed have proven to be relatively stable under distributional misspecifications, further work will explore the impact of employing more complex models that could better fit the data.

By comparing how these three methods for estimating trial-based subgroup cost-effectiveness perform in a large RCT we aimed to shed some light on their appropriateness. We have concentrated on a clinical area with relatively strong evidence for similar relative treatment impact of the intervention on the outcome (MVEs) and related savings in hospitalization costs combined with similar incremental treatment cost across these subgroups. We have shown that the use of this evidence has achieved the largest and clearest differentiation between the groups, compared to analyses based on within subgroup data or on shrinkage estimation. Clearly, this differentiation could be largely driven by the assumptions and therefore good quality evidence is needed to support them. In comparison, the analysis based on the subgroup data has proven to be largely underpowered to inform any decision that aims to differentiate between groups, while the shrinkage estimates were acting in the direction of decreasing the variability (i.e. biasing the

mean estimates) between the groups in order to improve their precision. By doing so, it provided subgroups estimates that were essentially indistinguishable in spite of the good background knowledge that there should be marked differences. Therefore, although the shrinkage estimates could provide a useful subgroup inference when we want to avoid spuriously significant subgroup results, we find that the method failed to provide an acceptable estimate of the cost-effectiveness of statins for patients at different cardiovascular disease risk.

Conclusions

Exploring the heterogeneity in cost-effectiveness estimates across the population is important if these analyses are to inform decisions for optimal provision of interventions. The methods to support the analyses, though, have not been widely addressed. We have compared three methods that estimate cost-effectiveness in heterogeneous risk populations. The within subgroup analysis makes no assumptions, but suffers from the loss of power associated with splitting the data and is often unable to differentiate moderate but important subgroup effects, such as those observed in the area of cardiovascular disease prevention. The random subgroup effects analysis and the analysis based on constant treatment effects across subgroups employ all the data to estimate subgroups, but impose assumptions on the estimation process. The random subgroup effects analysis is based on an initial belief of a lack of subgroup effect and therefore, similar to within subgroup approach, acts in the direction of diminishing observed differences. As a consequence, although differences in the absolute impact of statin therapy on patients at different baseline risk are expected, these were not clearly observed in our study based on the mixed effects model. The use of trial-wide treatment effect in subgroups at different baseline disease risk assumes lack of variation in the relative treatment effect across the subgroups which is possibly a strong assumption and needs support from external sources.

We conclude that in cases where no clear evidence for subgroup effect exists the shrinkage subgroup estimates could provide acceptable initial estimates of cost-effectiveness in different patient groups. However, in the presence of strong evidence for similar treatment impact across different patient groups this evidence should be used to inform the subgroup specific cost-effectiveness estimates. We argue that the use of trial-wide treatment effect to inform the

subgroup analyses, when applicable, is the most appropriate approach when estimating cost-effectiveness from heterogeneous trial populations – this approach avoids the pitfalls of underpowered subgroup analyses, and provides results that are more useful for decision-making than the shrinkage subgroup estimates.

(1) Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. Lancet 2002; 360(9326):7-22.

(2) Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. N Engl J Med 1998; 339(19):1349-1357.

(3) Yusuf S, Sleight P, Pogue J, Bosch J, Davies R, Dagenais G. Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. The Heart Outcomes Prevention Evaluation Study Investigators. N Engl J Med 2000; 342(3):145-153.

(4) Wilhelmsen L, Pyorala K, Wedel H, Cook T, Pedersen T, Kjekshus J. Risk factors for a major coronary event after myocardial infarction in the Scandinavian Simvastatin Survival Study (4S). Impact of predicted risk on the benefit of cholesterol-lowering treatment. Eur Heart J 2001; 22(13):1119-1127.

(5) West of Scotland Coronary Prevention Study: identification of high-risk groups and comparison with other cardiovascular intervention trials. Lancet 1996; 348(9038):1339-1342.

(6) Assmann SF, Pocock S, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. Lancet 2000; 355:1064-1069.

(7) Jonsson B, Cook JR, Pedersen TR. The cost-effectiveness of lipid lowering in patients with diabetes: results from the 4S trial. Diabetologia 1999; 42(11):1293-1301.

(8) Herman WH, Alexander CM, Cook JR, Boccuzzi SJ, Musliner TA, Pedersen TR et al. Effect of simvastatin treatment on cardiovascular resource utilization in impaired fasting glucose and diabetes. Findings from the Scandinavian Simvastatin Survival Study. Diabetes Care 1999; 22(11):1771-1778.

(9) Senn S. Some controversies in planning and analysing multi-centre trials. Stat Med 1998; 17:1753-1765.

(10) Localio AR, Berlin JA, Ten Have TR, Kimmel SE. Adjustments for center in multicenter studies: an overview. Ann Intern Med 2001; 135(2):112-123.

(11) Manca A, Rice N, Sculpher MJ, Briggs A. Assessing generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models. Health Economics. In press.

(12) Pinto EM, Willan AR, O'Briaen BJ. Cost-effectiveness analysis for multinational clinical trials. 2004. (in submission)

(13) Efron B, Morris C. Data Analysis Using Stein's Estimator and Its Generalizations. Journal of the American Statistical Association 1975; 70(350):311-319.

(14) Morris CN. Parametric Empirical Bayes Inference: Theory and Applications. Journal of the American Statistical Association 1983; 78(381):47-55.

(15) Efron B, Morris C. Limiting the risk of Bayes and empirical Bayes estimators. II. The empirical Bayes case. 1972.

(16) O'Connell D, Glasziou P, Hill S, Sarunak J, Lowe J, Henry D. Results of clinical trials and systematic reviews: to whom do they apply? In: Stevens A, Abrams K, Brazier J, Fitzpatrick R, Lilford R, editors. The Advanced Handbook of Methods in Evidence Based Healthcare. SAGE Publications Ltd, 2001.

(17) Anderson KM, Wilson PWF, Odell PM, Kannel WB. An Updated Coronary Risk Profile. A statement for Health Professionals. Circulation 1991; 83(1):356-362.

(18) Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. Am Heart J 1991; 121(1 Pt 2):293-298.

(19) Odell PM, Anderson KM, Kannel WB. New models for predicting cardiovascular events. J Clin Epidemiol 1994; 47:583-592.

(20) Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. Stat Med 2003; 22(23):3687-3709.

(21) Pharoah PD, Hollingworth W. Cost effectiveness of lowering cholesterol concentration with statins in patients with and without pre-existing coronary heart disease: life table method applied to health authority population. BMJ 1996; 312(7044):1443-1448.

(22) Johannesson M, Jonsson B, Kjekshus J, Olsson AG, Pedersen TR, Wedel H. Cost effectiveness of simvastatin treatment to lower cholesterol levels in patients with coronary heart disease. Scandinavian Simvastatin Survival Study Group. N Engl J Med 1997; 336(5):332-336.