No 37 May 1998

OHE briefing

FROM EFFICACY TO COST-EFFECTIVENESS

1. INTRODUCTION

In introducing the session Professor Michael Drummond pointed out that traditionally we obtain assessments of efficacy of pharmaceuticals from Phase 3 clinical trials. This is because evidence of efficacy is required for drug licensing. The advantage of such trials is high internal validity, i.e. relative freedom from dangers of confounding bias, as compared to other forms of research. Economic evaluations, however, require evidence of *effectiveness*, that is outcomes as realised in regular clinical practice.

Most official guidelines for cost-effectiveness studies ask for evidence of effectiveness, which is rarely available at the time the product is first launched. Moreover, effectiveness trials are often expensive and difficult to conduct. Economists therefore often use modelling approaches to obtain estimates of cost-effectiveness outcomes from efficacy data. There has been criticism of modelling, notably from the Food and Drug Administration (DDMAC, 1995) and the New England Journal of Medicine (Kassirer and Angel, 1994). However, the report of the Public Health Service Panel (Gold et al, 1996) in the USA gave strong support to modelling approaches. Therefore the session will explore the issues involved in moving from the assessment of efficacy to cost-effectiveness.

This OHE Briefing summarises the presentations and discussion at the session on 'Efficacy to cost-effectiveness' at the Conference of the International Society for Technology Assessment in Health Care (ISTAHC) in Barcelona on 26 May 1997.

The participants were:

Professor Michael Drummond, Director of the Centre for Health Economics, University of York, UK

Dr Kevin Schulman, Director of the Clinical Economics Research Unit, Georgetown University Medical Center, USA

Professor Milton Weinstein, Henry J. Kaiser Professor of Health Policy and Management, Harvard School of Public Health, USA

Professor Bengt Jonsson, Stockholm School of Economics, Sweden.

2. WITHIN-TRIAL ASSESSMENTS OF COST-EFFECTIVENESS

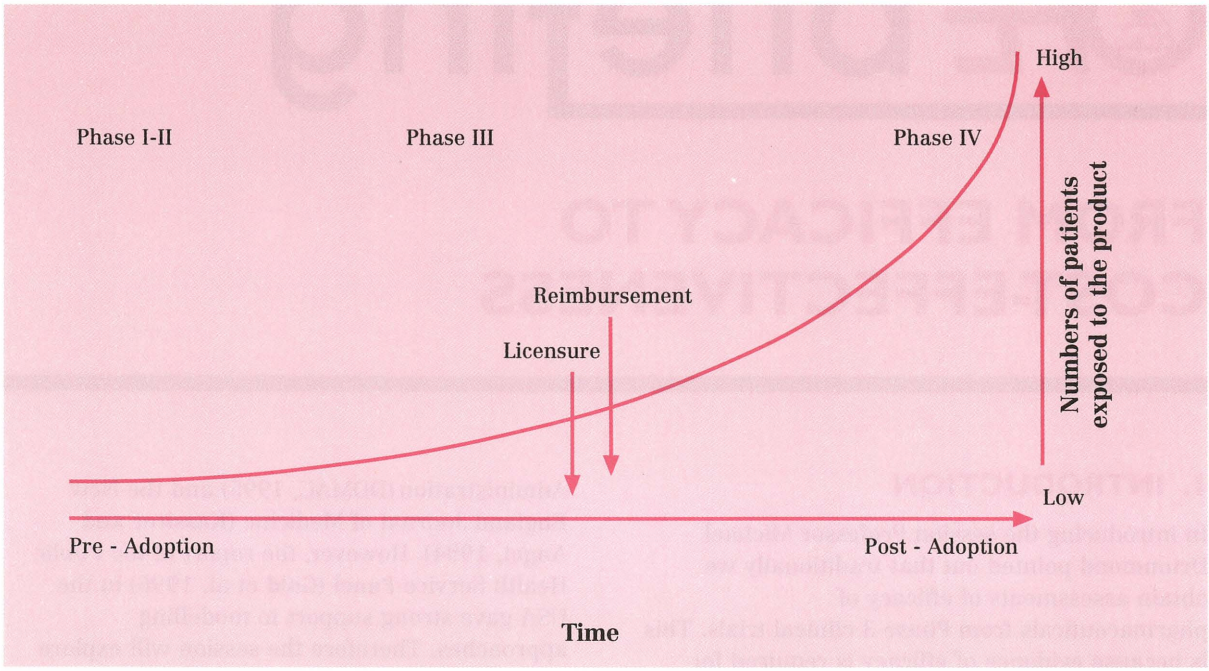
Dr Kevin Schulman addressed the design of economic evaluation in clinical trials, setting out the rationale for conducting economic evaluation within Phase 3 clinical trials and then discussing why the trials that economists had worked on were different from 'traditional' Phase 3 trials.

Use of economic evaluation in different phases of development

As a first principle economic analysis should be undertaken prior to initiating a clinical development programme to assess the economic potential of new clinical initiatives. To prepare appropriate protocols for Phase 3 trials we need to understand how the disease affects the population and how it is treated. This can be done during Phase 2 trials. In Phase 2 studies the concentration is on safety, but there is an opportunity to learn a lot about patients with the specific clinical diagnosis.

Phase 3 studies are the pivotal trials for the generation of information for registration, reimbursement and marketing. Phase 4 studies are comparative validation studies where we can reassess our preliminary economic analysis from Phase 3 studies to study the drug in a real world clinical setting, or in another health care system, to see if the data is similar to that obtained in the Phase 3 trial setting.

Figure 1 **Generalisability of clinical economic data**



The model curve illustrated in Figure 1 is an information time line. As we move along the horizontal axis from pre-adoption to post-adoption the number of patients introduced to a product grows. Through the time of licensing, which usually occurs after Phase 3, only a maximum of a few thousand people have been exposed to a clinical product. Once that product has been marketed, however, millions of people might be exposed to it. When using economic analysis from a Phase 3 trial we are trying to take data from a very limited population of patients and extrapolate to a very much larger population. The ideal economic study would take place after Phase 4. This is because only when everyone has learned how to use the product appropriately are we best able to define the appropriate population of patients that can benefit clinically from the product and the appropriate indication for the product. The result would be that the external validity of the analysis will be extremely high. However, the initial decision as to whether or not a national Health Authority or other third party payer is going to reimburse or use the therapy is often made soon after the conclusion of the Phase 3 studies. Economic evaluation in Phase 3 clinical trials has to be designed to help decision makers address this reimbursement question.

Trial design issues

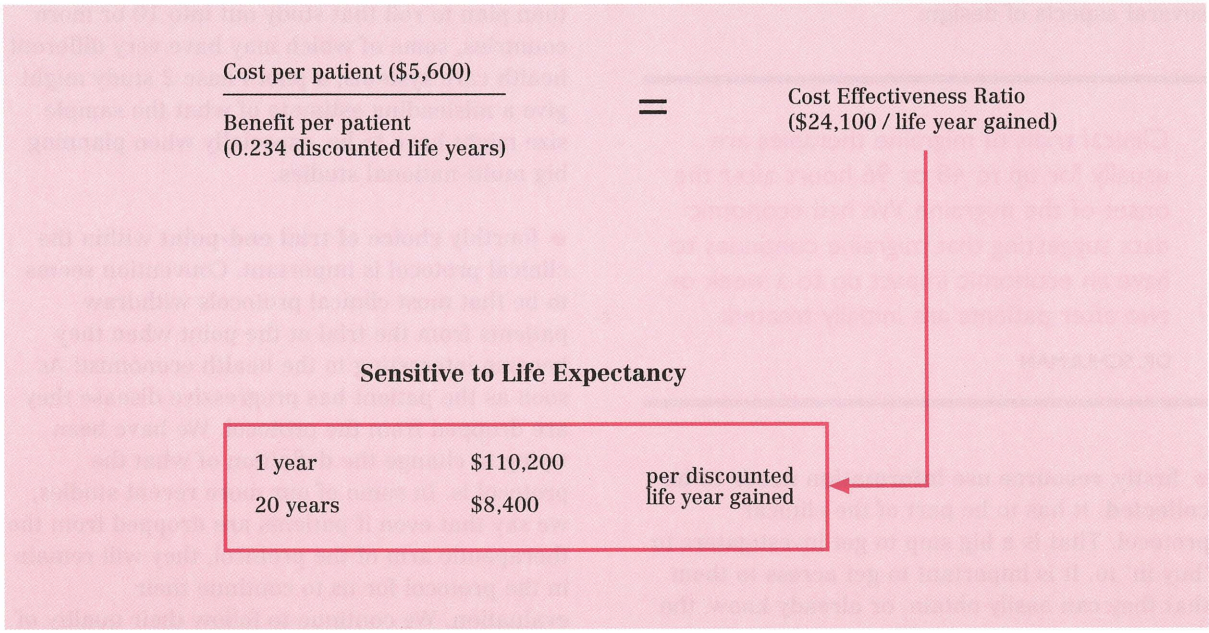
The design of trials to accommodate economic analysis is very challenging. We can't just append economic evaluation to a pre-arranged clinical trial if we are going to have data that is as reliable as possible. We begin from an economic perspective by trying to understand what the therapy means, how it is used in a trial setting, and how we might recommend the therapy be used in practice. We have to understand the economic profile of the therapy. For example, what is the expected duration of treatment and the nature of treatment benefit? Is the new product an acute treatment that will have long term benefits or a chronic therapy that patients will receive for the rest of their lives? What are the resource requirements and what are the resource offsets? Is this an in-patient or out-patient therapy? If I use this new technology, what resources will it substitute for, and can these be captured within the clinical protocol. Different answers to these questions will lead to different designs for the trial.

> From a clinical perspective we could design a trial around whatever end-point we can capture, but from an economic evaluation perspective we want to design a trail that makes sense to decision makers about what happens to patients. It is hard to do this for new therapeutic compounds.
>
> DR. SCHULMAN

In particular two issues are important:

- **the primary end-point for the trial.** What is the hypothesis? It could be a clinical hypothesis, relating to some change in a clinical outcome

Figure 2 **HA-1A results**



Source: Data taken from Schulman et al (1991).

measure, such as survival. It could be a economic hypothesis, relating to a change in resource utilisation, such as length of stay. Many trials of supportive therapies for cancer treatment have impact on resource use as their primary outcome measure.

- **the time horizon for the trial.** This is a crucial aspect of an economic evaluation. From a clinical perspective we could design a trial around whatever end-point we can capture, but from an economic evaluation perspective we want to design a trial that makes sense to decision makers in terms of describing what happens to patients. To do this we have to follow patients for a long enough period of time to capture the broader effects of therapy. It is easier to do this when developing a third, fourth or fifth product in a class, when there is a good picture of what happens to patients who receive this type of treatment. It is harder to do this for new therapeutic compounds. To give two examples:

  - we undertook an economic evaluation several years ago around a new medicine HA-1A that never got marketed (Schulman et al, 1991). In the economic analysis, the patients treated with the product were followed up for 28 days. As shown in Figure 2, our analysis based on the clinical trial resulted in a cost-effectiveness ratio of $24,100 per year life gained. However, in sensitivity analysis we found that the important question arising from this analysis was how long did patients survive after receiving the therapy? This end-point was not included in the clinical trial. If the patients had survived for only a year, instead of a cost effectiveness ratio of $24,100 it would have been up over $100,000. On the other hand, if patients had survived for 20 years the cost-effectiveness ratio would have been $8,400. At this latter figure it would have been regarded as economically advantageous by almost every reimbursement authority in the world. So the most critical factor in the economic evaluation was not the number of days in the hospital, but how long the patient survived. The investigators developing this study did not consider this issue in advance, and hence survival beyond hospital discharge was not captured in the time horizon for the clinical trial;

  - we used data from the Medicaid programme in the United States to consider patients with a migraine diagnosis. We examined their records for 60 days prior to, and after, their initial treatment for migraine to see if they had incurred resource use, whether related to the migraine or not. We had 2,400 subjects with about 2,700 migraine events. Resource use continued for 7 to 21 days after the migraines occurred. Yet clinical trials of migraine therapies are usually for up to 48 or 96 hours after the onset of the migraine. Here we had economic data suggesting that migraine continues to have an economic impact up to a week or two after patients are initially treated.

How do we go about designing economic evaluations as part of Phase 3 studies? There are several aspects of design:

> Clinical trials of migraine therapies are usually for up to 48 or 96 hours after the onset of the migraine. We had economic data suggesting that migraine continues to have an economic impact up to a week or two after patients are initially treated.
>
> DR. SCHULMAN

- firstly, **resource use information needs to be collected**. It has to be part of the clinical protocol. That is a big step to get investigaters to 'buy in' to. It is important to get across to them that they can easily obtain, or already know, the data about resource use. They may know how many days a patient is in hospital if they are following them. What they don't know about are how much things cost, but we normally assess the cost of these resources outside the clinical trial through a separate data collection exercise. Even in the United States, where detailed cost information is normally available for billing purposes, we often develop a separate parallel mechanism to collect the cost of those resources and therefore do not burden the clinical investigators with collecting cost data.

- the second aspect of design is **how to capture quality of life and utility information**. There is a choice of two options in a Phase 3 clinical trial. One option is to collect periodic quality of life or utility information from patients in the study. The alternative is to collect some information on the health states of patients during the clinical trial, but to value those health states outside the trial by using a population sample. This could involve either developing a health state measurement that we use within the trial, or using a measure that has external weights, like the Health Utilities Index. In most of our trials we have tried to assess preference weights directly from patients in the study at a minimum, with options for other strategies for collecting societal weights.

- thirdly, **sample size has to be calculated.** This usually involves a lengthy discussion. We can calculate a sample size for an economic end-point in the same way that clinical investigators can. Ideally we would have some pilot data available from Phase 2 trials that would help inform us. However, one limitation, especially when thinking about a multi-national Phase 3 study, is that pilot data might be from only one or two centres, in one or two countries. If we then plan to roll that study out into 10 or more countries, some of which may have very different health care systems, a pilot Phase 2 study might give a misleading estimate of what the sample size might have to be, especially when planning big multi-national studies.

- fourthly **choice of trial end-point** within the clinical protocol is important. Convention seems to be that most clinical protocols withdraw patients from the trial at the point when they become interesting to the health economist! As soon as the patient has progressive disease they are dropped from the protocol. We have been trying to change the definition of what the protocol is. In some of our more recent studies, we say that even if patients are dropped from the therapeutic arm of the protocol, they will remain in the protocol for us to continue their evaluation. We continue to follow their quality of life and resource use after they have met the primary clinical end-point for the study. We have also used 'salvage protocols' in trials where we weren't able to extend the protocol. Here we offered the patient the opportunity to continue to participate in the daily information collection exercise, so we could try to estimate the cost of progressive disease to these patients.

To summarise on data collection strategy. Routinely plan from the outset to collect relevant clinical, resource use and quality of life information concurrently. Consider modifying the clinical protocol in order to:

- determine relevant clinical end-points for the trial, as discussed above;

- remove any economic biases in the protocol, such as that patients in trial arm A have to stay in the hospital for five days, but patients in trial arm B go home as soon the doctor thinks they are ready to. Wording like that was very common in protocols 5 or 6 years ago. It is less common now, but we still need to look for it;

- increase generalisability. Clinical trial populations are pretty selective to increase the internal validity of the trial database. The question is how much can we loosen those constraints in designing Phase 3 trials, given that we need economic data with external validity. We have been successful in loosening some bounds but this is an ongoing challenge. There is a lot of scope within Phase 3 studies however to make the populations more representative of the people who are likely to be given the therapy once the drug is licensed.

Issues in data analysis

What do we do once we obtain all the data? This leads us from economics to statistics. The issues include the following:

- no matter how well we do our clinical trials, we are going to have **missing data**. Therefore we have to develop some method of accounting for missing data. Remember, however, that we are interested in following patients longitudely over time. The epidemiologists already do this. If you are only interested in survival then you can get this missing data from National Data sets;
- there are **censored data**. Most trials stop long before the end of the patient's lifetime. There are now a whole variety of methods, both parametric and non-parametric, to try to understand how to analyse resource data over a longitudinal period;
- **statistical testing** and the derivation of confidence intervals is a complex topic. There are now a variety of different methods out there in order to look at statistical bounds on a cost-effectiveness ratio;
- there may be a need to project effects from a short term clinical trial to a patient's lifetime.

Limitations of Phase 3 analyses

There is a major problem of the external validity of any Phase 3 analysis. It will involve at best thousands of patients for a product that may be used by many millions. To help overcome this economists want to make trials bigger, and to follow the patients longer. This is expensive and is going against the trend in the industry at the current time, which is to make the trials as small as possible and as simple as possible. This is an ongoing struggle.

Our analytic methods are evolving rapidly, and we are increasing the generalisability of Phase 3 results. As we do this, people want to go running to policy makers to get them to use this information. Remember however that there are still only one or two Phase 3 trials. This only gives data to support a limited number of hypothesis. The data can't answer every question that people might be interested in.

To conclude, Phase 3 trials are often the only opportunity to collect economic data before adoption and reimbursement decisions are made. This is the unique case for a Phase 3 data collection exercise. We need to re-design these Phase 3 trials to meet the needs and expectations of policy makers. Just because we do obtain data in Phase 3 doesn't, however, mean we don't need to collect them again in Phase 4. There are still many audiences to address and a need for validation of the data collected in Phase 3.

> Clinical trial populations are pretty selective to increase the internal validity of the trial database. The question is how much can we loosen those constraints in designing Phase 3 trials, given that we need economic data with external validity.
>
> DR. SCHULMAN

3. MODELLING IN COST-EFFECTIVENESS ANALYSIS

Professor Milton Weinstein discussed the use of modelling in cost-effectiveness analysis and expanded on the limitations of clinical trials as sources of information for cost-effectiveness analysis. He agreed with Kevin Schulman's analysis of how to improve clinical trials but went on to say, 'what do you do to extend the results?' He also tackled the prior question as to 'why extend the results?'

The recommendations of the cost-effectiveness panel

The recommendation from the Cost-effectiveness Panel of the US Public Health Service (Gold et al, 1996) specifically related to the use of modelling is as follows:

'When direct primary or secondary empirical evaluation of effectiveness is not possible, the use of modelling to estimate effectiveness is a valid mode of scientific enquiry for cost-effectiveness analyses' p168.

Indeed direct primary or secondary empirical evaluation of all the information that you need to know about cost, effectiveness and quality of life for a cost-effectiveness analysis is *never* possible.

The limitations of RCTs

These include:

- **a restrictive target population**. In order to increase the statistical power of the trial and achieve internal validity it is usually necessary to restrict greatly the target population in a clinical trial, by one or more of clinical indication, age, gender, or use of prior treatments. As a result we can question the implications of the trial results for populations which differ from those included within the target population;
- **a limited time horizon**. We heard earlier about the 28 day trial. We have to decide what to say about what happens after the trial ends. There are various modelling approaches. The point is that one has to model;

- **a limited set of intervention options.** It is rare to find a clinical trial with more than two, three, or maybe four arms. Yet very often the therapeutic or diagnostic alternatives are not viewed by clinicians in isolation of other possible courses of action. They may be combined with other treatment options into various strategies. These may involve particular sequences of treatments or combinations of treatments. Diagnostic procedures may be used in conjunction with therapeutic manouevres. In order to capture these permutations one has to go beyond the trial. I am involved in a modelling study that has around 48 different strategies. You could never do a clinical trial with 48 arms. So modelling is inevitable;

- clinical trials have strange **protocol-driven resource use** characteristics. This is partly because the protocol demands certain uses of resources in order to measure the variables being recorded, but also because these trials are conducted in academic medical centres, where for one reason or another, the clinicians who happen to be investigators do different things than would be done by a clinician in the general community. One has to address this.

Extrapolation of target population

Extrapolations can be done for a number of variables, including age, sex, and the risk factors for the disease. Very often a clinical trial evaluates treatment in a very high risk population in order to increase the number of events. This is to get a better statistical power. Sometimes a lower risk population is used if it might be regarded as unethical to withhold treatment from a high risk population. For whatever reason, the full spectrum of risk factors for disease for the end-point that you are trying to avoid is usually not represented in the clinical trial.

Therefore we have to adjust for the fact that subjects in clinical trials are usually healthier than the average member of the population. In the Veterans Administration clinical trial in the United States of coronary by-pass surgery versus medical management (Weinstein and Stason, 1982), it was hoped to use the data to estimate the first part of a survival curve which could then be used to calculate life expectancy. The patients in the control arm of that trial had coronary artery disease, mostly three vessel disease or left main coronary artery disease. They were not healthy people. Yet the 3 year survival of the subgroup with two-vessel disease was better than that of the average US adult of the same age! Here were patients with coronary disease sufficiently advanced to warrant coronary by-pass surgery in the early days of this technology when people were very conservative about the application of by-pass with longer survival than the average member of the population. This opened my eyes to the 'healthy subject' effect. It relates not only to survival but also to their health related quality of life, which may be better than the average member of the population, and also to their resource utilisation for related and perhaps unrelated medical conditions, which maybe less in the trial than it would be for less healthy people with the disease. This may offset some of the protocol induced effects on resource use mentioned above, but we don't know whether they cancel each other out or even in which direction the overall impact is.

> The rule of thumb is that most clinical protocols withdraw patients from the trial at the point when they become interesting to the health economist.
>
> DR. SCHULMAN

Extension of the time horizon

As pointed out above, we are not interested in knowing what happens to patient survival for the next 28 days only. We want to know if any survival advantage persists. Usually we don't know that from the clinical trial. In some cost-effectiveness studies we have introduced what we regard as the most conservative possible assumption. This we call the stop and drop assumption. It assumes that, if there is a survival advantage at the end of trial, all the excess survivors drop dead the day after the trial ends. That is the most conservative assumption you can make about the results of a clinical trial in which survival or death is an end point. That is almost obviously too conservative in most circumstances. It doesn't make sense simply to look at the gain in life expectancy during the time of a trial, so what do you do? Do you assume the survival curves are parallel? Do you assume the survival curves come together gradually? The point is that you need to make some assumptions. These will preferably appeal to other sources of data on related interventions.

Life expectancy is not the only factor we need to estimate beyond the period of the trial. We also need projections of health-related quality of life and resource utilization in future years. These estimates are usually based on projections of the

patient's transition from one health state to another, or on projections of the incidence of important events, such as myocardial infarctions, strokes, cancer recurrences, or opportunistic infections in AIDS.

Modelling within RCTs

Whereas we normally think of modelling being used to project beyond the period of a clinical trial, or to link intermediate to final endpoints, it is worth noting that models are often used in statistical analysis of RCT results.

For example, assumptions are often made about the functional form of the relationship between variables and the nature of the treatment effect in order to decide on the most appropriate form of the statistical analysis. Therefore, it is wrong to make a dichotomy between observed data (as in the trial) and modelled data (which are based on extrapolation). Even the analysis and interpretation of observed data requires some modelling.

Validation of models

My next topic is how models can be validated. This is a question that people who do models are always asked. Yet clinical trialists don't get asked this question. The clinical trial is obviously valid, it was blinded, double blinded and you did all the 'right' things. In a model how do you know that what was done was 'right'? You don't always know, and that is why we do a lot of *sensitivity analyses*. Ultimately that is what you have to fall back on. We don't know for sure that the assumptions that we are making are right so we examine what happens, if anything, when you make different assumptions.

Face validity is also very important. Models have bugs in them sometimes and even the best of modellers and their programmers make mistakes. How do you know if there is a mistake in a model? You don't worry about it so much in a clinical trial, because you can go back and look at the patient records. There are only a relatively small finite number of them to go back and check. In a model if some equation has a mistake in it, or there is a logical error in the way the equations are put together, how do we know? One way to check is face validation, do the results make sense? Do they respond to sensitivity analyses the way they ought to?

Predictive validity is another type of validation. Does the model predict or postdict what actually happened in the past? For example, the Coronary Heart Disease Policy Model began in 1980 (Weinstein et al, 1987). After 1990 came and went, we looked back to see how well the 1980 model predicted what happened in 1990 (Hunink et al, 1997). Of course it didn't predict what happened in 1990 very well. This made us concerned that there was a mistake in the model, but it turned out there wasn't – at least not a mistake that was accounting for the discrepancy. The discrepancy arose because things changed in the real world. Risk factors for coronary heart disease changed. Treatments got better. There were new treatments that didn't exist in 1980 or weren't widely applied in 1980. Beta-blocker drugs, for example, were not widely used in 1980, but were much more widely used in 1990.

> 'When direct primary or secondary empirical evaluation of effectiveness is not possible, the use of modelling to estimate effectiveness is a valid mode of scientific enquiry for cost-effectiveness analyses'.
>
> GOLD ET AL (1996) P168

We went back and we tried to see whether changing risk factors and treatment assumptions could explain what happened between 1980 and 1990. In fact we were able to account for more than 90 per cent of the discrepancy between what the model predicted and what actually happened. The model had over predicted the number of coronary heart disease deaths by a fairly substantial amount. By plugging in the risk and treatment trends that were observed between 1980 and 1990 the discrepancy all but disappeared. We now have more confidence that our model was valid. There was nothing technically wrong with it.

Peer review and disclosure

To what extent is the modeller obligated to release the model into the public domain? I believe that there is an obligation to reveal the assumptions, data sources and actual data that are in the model. Publishers of journals don't allow you to put all those in the journals. As a result the US Cost-effectiveness Panel recommended that there be a technical appendix for each cost-effectiveness analysis using a model. You maybe able to put some appendices within the journal article, but it is more than likely that you are going to need some additional material. The Panel recommended that the technical report be made available on request from authors, giving all the detail that anyone might conceivably want.

More controversial is whether the actual model ought to be released in the public domain for

review. The Panel had mixed minds on this and didn't reach a conclusion. We did suggest that it should be available for the purpose of peer review. This doesn't mean, however, that it ought to be generally disseminated. It makes me very nervous to think that somebody might have on their hard drive a copy of the Coronary Heart Disease Policy Model. They could do an analysis using it and say, 'we did this study using the Coronary Heart Disease Policy Model'. I'm not sure they really would know how to use the Coronary Heart Disease Policy Model properly. On the other hand if another academic was reviewing a paper that used our model, we should be willing to give him or her a copy. In principle it ought to be available for peer review.

In conclusion, I think modelling is an unavoidable fact of life if you are trying to make resource allocation decisions involving medical and healthcare technologies. You have to make assumptions. Modelling is a way of structuring the assumptions and being explicit about these assumptions, to extrapolate conclusions that might be useful to real world decision makers.

4. EFFICACY TO COST-EFFECTIVENESS

Professor Bengt Jonsson considered the issue of efficacy to cost-effectiveness from the perspective of the economist, who is concerned with the issue of resource allocation. He took a practical example to illustrate the points which were made earlier and then made general recommendations on the choice of outcome measures.

The example was from the field of cholesterol-lowering. The initial problem was posed thus:

'We have a clinical study and a published paper in the Lancet which shows that there is a 30 per cent reduction in the relative risk of mortality from any cause in patients with angina pectoris or prior myocardinal infarction through drug treatment with simvastatin in a trial over a 5 year period. The clinical benefit is clear but is it cost effective?'

That is without the drug you have about 87 per cent survival probability after 5 years. If you take the drug you have about 91 per cent probability, an increase of about 4 percentage points.

What would be the data requirements for an economic evaluation? We need some cost data, (cost of intervention data, cost of disease or events data, and indirect costs when they are relevant, i.e. when you are dealing with a population of working age), and we need outcome data, in terms of survival and of quality of life.

> 'Whereas we normally think of modelling being used to project beyond the period of a clinical trial, or to link intermediate to final end points, it is worth noting that models are often used in statistical analysis of RCT results'.
>
> PROFESSOR WEINSTEIN

Major shortcomings using cost data from clinical trials

There are often major short comings from using costing data from clinical trials. These have been elaborated on already:

- the trial situation does not correspond to clinical practice. We obtain efficacy data corresponding to the management of patients in the clinical trial. Economists would like to know what happens when patients are managed in regular medical practice;
- intervention costs are usually very poorly documented. The resources, including medicines, being consumed by the patients as part of the intervention are often not well recorded in clinical trials, which is a little surprising;
- cost savings or cost increases, (disease costs) are often difficult to assess due to limited sample sizes. Very seldom do you have a large enough sample size to look at them;
- indirect costs (i.e. productivity losses) in general are often omitted from the analysis altogether.

Intervention costs

This clinical trial (the Scandinavian Simvastatin Survival Study (4S) trial) (Jonsson et al, 1996, Johannesson et al, 1996) was quite good because, unusually, it kept track of how much medication the patients took. In other studies when we ask how much of the drugs did the patient take, we are usually told that it is in the protocol. However we cannot assume full compliance. Effectiveness is determined by the amount of drug that actually got into the patient. From the costing point of view, we have to know, for example, how many prescriptions the patient had. One of problems here is that very often drugs in the clinical trial are given free to the patients but in actual medical practice, patients have to pay for them.

Compliance is a difficult issue. For example when looking at the cost-effectiveness of HRT the economist is often asked what will happen if the patient doesn't fully comply with the treatment

regimen. The response is that it depends on how compliance affects effectiveness and how it affects cost. Very seldom are clinical trials large enough to allow a study of the relationship between compliance and efficacy. Related to compliance is the question of how to take account of side effects. If side effects are significant, their consequences for costs and quality of life should be assessed.

In the 4S trial the patients were already being managed by physicians, and having regular tests and physician visits. It was quite reasonable to assume that the only additional cost here was the drug cost. But in most situations it is much more complicated. In the planning of the 4S trial it had been decided to collect prospective data on hospitalisations. They were thus able to show, within the five to six year follow-up of the trial, that there was a reduction in hospital admissions for cardiovascular diseases of 26 per cent, most of it coronary heart disease. The hospital days were reduced by 34 per cent, as there was some reduction in average length of stay, giving a reduction in overall hospital costs of 32 per cent.

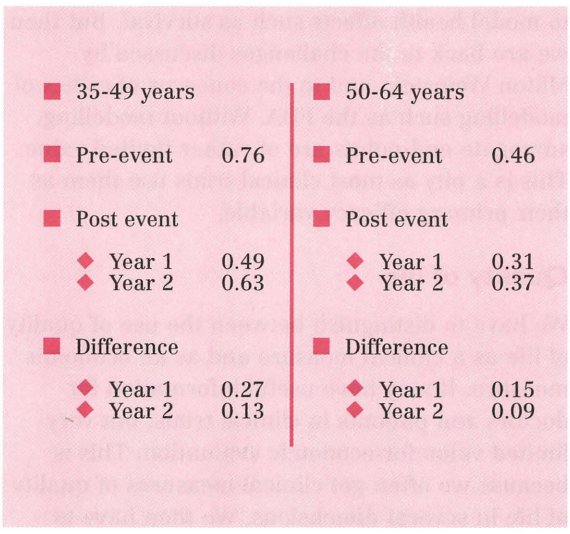
Disease costing

How do we get from data on resources to costs? There are problems of:

- the definition of resources;
- the valuation of resources;
- the aggregation of resources over countries.

The 4S trial was conducted in all the Scandanavian countries. Very often we have an even more complicated situation with more countries and very different healthcare systems.

Figure 3 **Proportion of full time workers**



So how do we solve this? If we have data on the number of hospitalisations, and the number of hospital days, we can choose to calculate cost using either a cost per bed-day or a cost per discharge. The classification of hospitalisations according to DRG groups facilitates costings, particularly within international trials. There have been a lot of developments in the grouping of hospitalisations in DRGs and this is a big step forward. It means that we don't have to undertake detailed costing in all of the hospitals. If we can just group discharges according to DRGs then we have a definition of resource use for costing which can be transported consistently across different countries. The problem is that we lack a similar type of classification for patients and resource use in ambulatory care. This means that there is a lot of what economists call 'bean counting' going on. We count drugs, we count visits, we count tests, we count everything! We have to try to progress because it is not very efficient to define resources at this very specific level, particularly when we know that there are a lot of joint costs for patients. One of the best things we could develop is a good DRG type classification system for patients in ambulatory care, which would provide reasonable cost and resource use data.

Indirect costs

There is major controversy as to whether indirect costs should be included in a cost-effectiveness study or not. However, if we don't include them, the first question we get asked is 'why didn't you include them?' If we do include them, then the first question we usually get is 'why did you include them?' So we often include them.

It is relevant to look for differences in labour force participation between the intervention and control groups. If we are studying people of working age, the indirect costs for most diseases may well be higher than the direct costs. The important thing here is to collect data on the actual differences in time spent on different activities, such as employment and different types of leisure activities. What we try to measure is the difference between flows of activity in the two groups. A problem is that clinical trials usually give you point estimates.

A typical example of this was with the 4S study. Every six months patients were asked about their work status. How do you use these point estimates to calculate indirect costs? We started from coronary heart disease events, and compared work status at the 6 month point estimate before the event, and at work status at the 6 month point estimates after the event. As

patients were being asked questions as to whether they were in full-time work, part-time work or half-time work, we could construct pre-event and post-event figures by type of employment. For example, 76 per cent of the subjects between the ages of 35 and 49 were working full-time. For patients aged 50-64, pre-event, 46 per cent worked full-time, which is surprisingly high. Figure 3 shows that after an event their work participation rate goes down, and then it goes up again a little. This is because even if they have their second myocardial infarction they usually recover. We can calculate the difference in proportion of full-time workers pre and post event and measure indirect costs that way.

This is one example of how to get from a clinical trial which has some data about resources to cost estimates that can be used in a cost-effectiveness analysis. It is quite a lot of work, using a lot of assumptions. Thus economic evaluation is partly a science and partly a question of informed judgement.

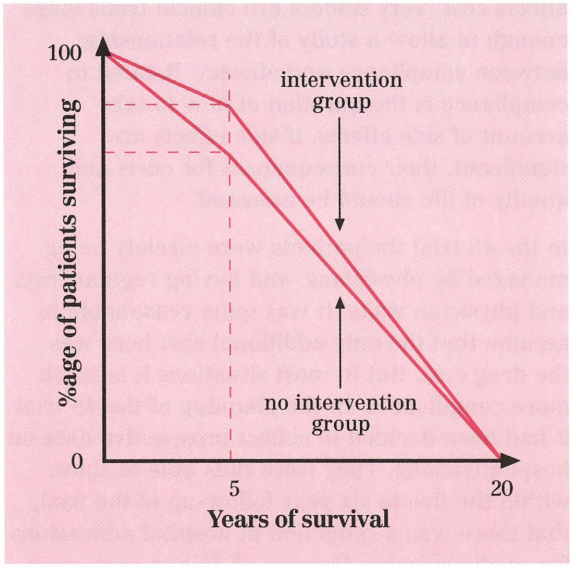
> Modelling is a fact of life if you are trying to make resource allocation decision involving medical and healthcare technologies. You have to make assumptions. Modelling is a way of structuring the assumptions and being explicit about these assumptions, to extrapolate conclusions that might be useful to real world decision makers.
>
> PROFESSOR WEINSTEIN

Outcome measurement

The outcome measures collected in a trial need to be relevant to the person who is going to make decisions based on the results of the trial (Johannesson et al, 1996). This has a number of components:

- a cost-effectiveness ratio has to have a relevant comparator. If we have a cost per event avoided, (for example cost per myocardial infarction avoided) what are going to compare it to?
- it may be difficult to determine the willingness to pay for a surrogate outcome measure? What is the decision maker willing to pay for a certain reduction in blood pressure, for example?
- the outcome measures could be complementary. It may not be enough just to have one outcome measure. Different decision makers could be interested in different kinds of outcome measures.

Figure 4 **Gain in life expectancy**



Surrogate end-points

Clinical trials are full of surrogate end-points. It is often difficult for an economist to know what they mean and sometimes difficult even to pronounce them. Examples include cost per millimol in reduction in total cholesterol, or cost per millimetre reduction in blood pressure, or cost per milligram per square centimetre increase in bone mineral density. The problem with all this is that we can't compare the results of different studies unless you have specific budgets to increase bone mineral density, reduce blood pressure and lower cholesterol, respectively.

If you really have a decision maker with a budget to maximise bone mineral density then such a surrogate end-point is a good outcome measure. Otherwise it could be very difficult to use in resource allocation decisions. However, it could be useful if combined with epidemiological data to model health effects such as survival. But then we are back to the challenges discussed by Milton Weinstein and to the concerns of critics of modelling such as the FDA. Without modelling, surrogate end-points are of rather limited value. This is a pity as most clinical trials use them as their primary efficacy variable.

Quality of life

We have to distinguish between the use of quality of life as a clinical measure and as an economic measure. It may have useful information for doctors and patients in clinical trials, but very limited value for economic evaluation. This is because we often get clinical measures of quality of life in several dimensions. We then have to

choose one for the cost-effectiveness ratio. If we combine them into one measure, it is difficult to interpret them and to validate their relevance. We could, of course, use them as a description of healthcare states and have them translated into utilities or willingness to pay. However, that is tricky and presents many of the same problems as other modelling exercises.

Clinical events

Instead of using surrogate end-points or clinical quality of life measures, we can use the clinical events avoided (such as myocardial infarctions and hip fractures) as outcome measures.

It is easier for a decision maker to think of their willingness to pay to reduce myocardial infarctions, or hip fractures, or to achieve symptom-free weeks for asthma sufferers. The problems are twofold – we have to define clinical events with real meaning, and we have to define a time perspective. Mortality as a clinical event cannot be avoided, only postponed. Clinical events as outcome measures are easier to understand than surrogate end-points or clinical quality of life measures but still offer limited comparisons for resource allocation. We cannot easily compare the cost per hip fracture avoided with the cost per MI avoided.

The need for modelling

We can model from clinical outcome to impact on survival, to increase the value of clinical information. We can model to compensate for limitations in data collection. Let us take, as examples, death and survival. Death is an event which will only occur once, but all events are not equal. If all deaths were equal we could do studies of the cost per death avoided, but it is never avoided. We look for gain in life expectancy. The problem is that most clinical trials end before all patients are dead. Within trial analysis can never be complete. If trials did go on until all the patients were dead, then we would probably find that most of the investigators would be dead as well – along with the health economists who were hoping to analyse the data. So we have to get from survival to life expectancy by using modelling. This was the situation with the 4S study. The 4S trial provided survival data for the first five years. We then had to model to estimate the overall gain in life expectancy. We did this by assuming an average life expectancy of 10 years at the end of the study. The model is illustrated in Figure 4.

> It is relevant to look for differences in labour force participation between the intervention and control groups. If we are studying people of working age, the indirect costs for most diseases may well be higher than the direct costs.
>
> PROFESSOR JONSSON

The results of the calculation of life expectancy were as follows:

- the gain in life expectancy from the use of simvastatin was 0.065 years (undiscounted) in the period of the trial;
- modelling from this gave a total gain in life expectancy of 0.3777 life-years undiscounted;
- this gave a discounted gain in life expectancy from simvastatin treatment of 0.240 life years;
- cost per life year gained including direct costs only was £5500;
- cost per life year gained including direct and indirect costs was £1600 per life year gained.

In conclusion:

- clinical efficacy trials are an important, often essential, source of information for cost-effectiveness studies;
- clinical trials give 'point estimates' but economic evaluations require data on the 'area under the curve';
- trial data must be supported with other data to answer questions about the efficient allocation of resources for health.

Concluding comments on the issues raised

In closing the session Professor Michael Drummond noted that the discussion demonstrates that, if we really need to move from efficacy to cost-effectiveness, a number of adjustments are required to traditional Phase 3 clinical trials. Some of these adjustments may be made to the design of the trial itself and the nature of data collection. Other adjustments may be made through modelling after the trial has been completed.

In particular:

- we need to recognize that, because of the questions they are designed to answer, Phase 3 trials only include a restrictive patient population, with a limited range of measured end-points, studied over a limited time horizon;
- where possible, the time horizon of the trial should be set so that important changes affecting cost-effectiveness can be observed;

- the range of data collected needs, where possible, to encompass resource use and quality of life, although the costs of this need to be considered carefully;

- because of the cost of studying all the relevant groups of patients, extending the time horizon, or including the appropriate range of measurements, modelling is often required;

- because of the assumptions inherent in modelling, particular attention should be placed on transparency and peer-review. However, there are still unresolved issues, such as the best way of validating models.

References:

DDMAC (1995) Principles for the Review of Pharmaceutical Promotion (DRAFT), March 20, 1995, FDA: Washington DC.

Gold M R, Siegel J E, Russell L B, Weinstein M C (1996) (eds). Cost-effectiveness in health and medicine. New York: Oxford University Press.

Hunink M G, Goldman L, Tosteson A N et al. (1997). The recent decline in mortality from coronary heart disease, 1980-1990. Journal of the American Medical Association; 277:535-542.

Johannesson M, Jonsson B, Karlsson G (1996). Outcome measurement in economic evaluation. Health Economics; 5: 279-296.

Johannesson M. Jonsson B, Kjekshus J, Olsson A G, Pedersen T R, Wedel H (1997). Cost-effectiveness of simvastatin treatment to lower cholesterol levels in patients with coronary heart disease. New England Journal of Medicine; 336(5): 332-336.

Jonsson B, Johannesson M, Kjekshus J, Olsson A G, Pedersen T R, Wedel H (1996). Cost-effectiveness of cholesterol lowering: results from the Scandinavian Simvastatin Survival Study (4S). European Heart Journal; 17: 1001-1007.

Kassirer J P and Angel M (1994). The journal's policy on cost-effectiveness analyses (editorial). New England Journal of Medicine; 331: 669-670.

Schulman K A, Glick M A, Rubin M, Eisenberg J M (1991). Cost-effectiveness of HA-1A monoclonal antibody for gram-negative effects. Journal of the American Medical Association; 226:3466-3471.

Weinstein M C, Coxson P G, Williams T M, Pass T M, Stason W B, Goldman L (1987). Forecasting coronary heart disease incidence, mortality, and cost: The coronary heart disease model. American Journal of Public Health; 77: 1417-26.

Weinstein M C, Stason W B (1982). Cost-effectiveness of coronary artery bypass surgery. Circulation; 66 (supp): 36-66.

SELECTED OHE PUBLICATIONS

The Pros and Cons of Modelling in Economic Evaluation
edited by Adrian Towse and Mike Drummond, 1997 (price £5.00).

Uses of Models in Economic Evaluations of Medicines and other health technologies
by Brian Rittenhouse, 1996 (price £7.50).

Handling Uncertainty in the Results of Economic Evaluation
by Andrew Briggs, 1995 (price £5.00).

Managed Care – a Model for the UK?
by Dr Peter West, 1998 (price £10.00).

Guidelines for the Economic Evaluation of Pharmaceuticals: Can the UK learn from Australia and Canada?
Edited by Adrian Towse, 1997 (price £10.00).

Health Economics: An Introduction to Economic Evaluation
by Gisela Kobelt, 1996 (price £2.00).

Competition and Contestability between Acute Hospitals
by Jon Sussex, 1998 (price £10.00).

Compendium of Health Statistics (10th Edition)
by Sean McGuigan, 1997 [price £195.00 (£60.00 to public sector organisations) plus P&P for non-UK].

Theory and Evidence on Cost Sharing in Health Care: an Economic Perspective
by Martin Chalkley and Ray Robinson, 1997 (price £10.00).

Prescribing, Budgets and Fundholding in General Practice
by Darrin Baines, Keith Tolley and David Whynes, 1997 (price £7.50).

A full list of OHE publications is available from the address below.

OHE publications are available from the Office of Health Economics, 12 Whitehall, London SW1A 2DY. To order, send a cheque or telephone (0171 930 9203) or fax (0171 747 1419) with credit card details.

Acknowledgement of funding

The ISTAHC session was sponsored by the IFPMA Health Economics Advisory Group, with financial support from Glaxo Wellcome, Pharmacia & Upjohn and Rhône Poulenc Rorer.

This briefing was edited by Adrian Towse, Director of the Office of Health Economics, London, UK, and Michael Drummond, Professor of Health Economics and Director of the Centre for Health Economics, University of York, UK.

Office of Health Economics,
12 Whitehall, London SW1A 2DY.
Telephone: 0171 930 9203
Fax: 0171 747 1419