Prevalent and multiple future incident cohorts in cost-effectiveness analysis

Martin Hoyle & Rob Anderson

Peninsular Technology Assessment Group (PenTAG),
Universities of Plymouth and Exeter

Contact details

Peninsula Technology Assessment Group
Peninsula Medical School
Noy Scott House
Barrack Road
Exeter
EX2 5DW
UK

Tel: UK 01392 406969
Fax: UK 01392 406936
Email: martin.hoyle@pms.ac.uk

Abstract

Aim For cost-effectiveness analysis, we aim to account for all future costs and benefits for all patients who are currently eligible for a new technology and who will become eligible in the future.

Methods We adapt the fundamental concept from epidemiology of the incidence and prevalence of a disease to cost-effectiveness analysis. We define the prevalent cohort as those patients eligible to switch from the comparator to the new technology at the time the new technology is introduced. Next, we introduce the concept of multiple future incident cohorts. The incident cohort starting *t* years in the future consists of those patients who first become eligible for the new technology *t* years in the future. Currently cost-effectiveness analyses worldwide consider only either the prevalent cohort, the incident cohort in only the first year, or a mixture of the two. We suggest that analyses capture the costs and benefits arising from the prevalent cohort and all future incident cohorts.

Results On average, patients in the prevalent cohort are older and at a more advanced stage of illness than patients in the incident cohort. If the cost and benefit discount rates differ, the cost-effectiveness of all technologies will be substantially affected by our methodology. Otherwise, the incremental cost-effectiveness rate will not change for acute conditions, but may change substantially for chronic conditions, particularly for chronic progressive conditions.

Conclusions If our methodology had been used in the past, some technologies would have appeared substantially more cost-effective, others substantially less cost-effective. We acknowledge the extra work required to implement our methodology. Accordingly, we suggest scenarios when our methodology can be safely ignored, and when it should be implemented. If possible, parameter values (e.g. average age) for both the incident and prevalent cohorts should be obtained from the literature. Otherwise, we describe how such parameters can be estimated. We suggest how uncertainty in the additional parameters can be incorporated in the probabilistic sensitivity analysis.

Key words: chronic disease, incremental cost-effectiveness ratio, probabilistic sensitivity analysis, progressive disease, technology assessment, technology diffusion

Introduction

The estimation of an incremental cost effectiveness ratio (ICER), defined as the difference in the costs of two technologies divided by the difference in their health effects, is the primary goal of the most commonly conducted form of health economic evaluation. It has also become a major determinant of whether a health technology is publicly funded in a number of countries (NICE 2004, Pharmaceutical Benefits Advisory Committee, Australia (PBAC), Canadian Agency for Drugs and Technologies in Health (CADTH), George et al. 2001, Devlin & Parkin 2004). The lower the ICER, the more cost-effective a given technology relative to its comparator, and the more likely it will be deemed 'good value for money' in a given health system. While ICERs can be generated purely on the basis of patient-level data from a single economic study (usually alongside a clinical trial), it is increasingly recognised that to inform decision-making at a regional or national level ICERs need to be based on rigorously informed decision model-based analyses which compare the incremental costs and effects of all relevant comparators, and typically for the remainder of patients' lifetimes (Sculpher et al 2006).

The prevalence and incidence of a disease are fundamental concepts in epidemiology. The prevalence is the number of cases in a population at a specified point in time, and the incidence is the number of new cases arising in a given period in a population (Beaglehole et al 1993). By analogy, we introduce the fundamental concepts of the prevalent cohort and future incident cohorts to cost-effectiveness analysis. We define the prevalent cohort as those patients eligible for the new technology when the technology is first introduced, i.e. at the date of technology assessment. Any given patient will be eligible from the time when the technology is first clinically appropriate (e.g. just diagnosed with multiple sclerosis and eligible for drug treatment, or when first eligible for a hip replacement) until the time when the new technology is no longer appropriate (e.g. patient dies, or the disease has reached such a severe state that the drug is no longer effective, or the patient is too old to receive a hip replacement). Next, we define the incident cohort starting t years in the future (i.e. t years after the date of technology assessment) as comprising those patients who first become eligible for the new technology (e.g. diagnosed) t years in the future.

Of course, the absolute sizes of the incident and prevalent cohorts increase with the rate of incidence of the disease. However, since cost-effectiveness analysis focuses on the ratio of costs to benefits, here we need consider only the relative number of patients in the prevalent and incident cohorts. There will always be future incident cohorts of patients who may benefit from a new treatment. However, the prevalent cohort may be large, small or even negligible relative to the size of the incident cohort. The prevalent cohort is small relative to the incident cohort *when a technology is appropriate only at a fixed moment in time*, for example;

- Drugs or non-drug technology for acute conditions or episodes, e.g. for acute infection such as the common cold, for treatment in accident and emergency.
- Screening people at certain fixed ages e.g. screening for breast cancer, neonatal screening for cystic fibrosis, health checks on retirement.
- Drugs administered during or immediately after an operation, e.g. calcineurin inhibitor e.g. ciclosporin.
- For rapidly progressing, high-mortality conditions, e.g. chemotherapy drugs for multiforme glioblastoma (type of brain tumour).

By contrast, the prevalent cohort is large relative to the incident cohort for technologies that are used to treat *chronic* conditions, for example;

- "Maintenance therapies" aimed at arresting decline and managing symptoms in conditions such as asthma, multiple sclerosis, diabetes, HIV, rheumatoid arthritis.
- Medical devices such as hip replacements, cochlear implants, pacemakers.
- Periodic screening policies that apply to a population of broad age range (e.g. cervical screening, breast screening).

The current ISPOR Guidance on good practice in decision analytic modelling focuses dominantly on the structure of the model, the validation of the model estimates/inputs, and the choice between alternative simulation models (e.g. Monte Carlo vs. cohort) (Weinstein et al 2003). However, there is no specific advice on what populations should go into a decision model. Additionally, the National Institute for Health and Clinical Excellence (NICE) reports no specific guidance on the nature of the cohorts (NICE 2004). Currently, cost-effectiveness analyses that compare a new versus a comparator technology tend to model a single cohort separately for patients that receive the current, comparator technology, and for patients that receive the new technology. A literature search of papers of cost-effectiveness analyses for chronic conditions revealed no studies that modelled both incident and prevalent cohorts. All studies modelled just a single cohort, presumably either the first future incident cohort, or the prevalent cohort. Furthermore, we found no papers that discussed the nature of cohorts in cost-effectiveness analysis. It is essential that our estimates of costs and benefits accurately reflect those incurred by the patients who will be affected by the new or comparator technologies in the future. Therefore, in this paper, we suggest that both the prevalent and all future incident cohorts should be modelled. Note that we do not suggest that our methods are appropriate for trial-based economic evaluations, because the cost-effectiveness of the technologies are based on patients in the trial alone.

In this paper, we describe the mathematics for estimating the ICER that includes the costs and benefits for both a prevalent and future incident cohorts. We consider a new technology versus a comparator technology, but the 'comparator technology' could represent no treatment. The technologies can be either a drug or not a drug (e.g. cochlear implant, hip replacement). We describe the additional data that would need to be collected. If some of this data is unavailable, we suggest how this data can be estimated. Probabilistic sensitivity analysis, which quantifies the uncertainty in the ICER, is widely regarded as an essential component of cost-effectiveness analysis, and is a NICE requirement in the UK (NICE 2004; Claxton et al. 2005). We describe parameters related to the structure of the patient cohorts that should be included in the probabilistic sensitivity analysis.

ICER for incident cohorts

First future incident cohort

Suppose the incremental costs per patient between the new and comparator technology (where the comparator technology could be no technology, i.e. best supportive care), at cycles 0, 1, 2,…., H are ΔK_0, ΔK_1, ΔK_2, …., ΔK_H and incremental benefits ΔB_0, ΔB_1, ΔB_2,…., ΔB_H . The time horizon is H cycles. Then the ICER as currently calculated for health technology assessments for the first future incident cohort, given discount rate for costs of r^*_C and benefits r^*_B over a cycle;

ICER (first future incident cohort) = $$\frac{\sum_{j=0}^{H} {v^*}_C{}^{j+1/2} \Delta K_j}{\sum_{j=0}^{H} {v^*}_B{}^{j+1/2} \Delta B_j}$$

where we define ${v^*}_C = \frac{1}{1+{r^*}_C}$, ${v^*}_B = \frac{1}{1+{r^*}_B}$, and the $1/2$ term in the exponents reflect the fact that, on average, costs and benefits are incurred half way through each cycle.

First two future incident cohorts

Now assume a new cohort of patients presents for treatment in the second year after the introduction of the new technology. This is the second future incident cohort. Then assuming that the number of patients in the two cohorts is equal;

ICER (two incident cohorts) = $$\frac{\sum_{j=0}^{H} {v^*}_C{}^{j+1/2} \Delta K_j + v'_C \sum_{j=0}^{H} {v^*}_C{}^{j+1/2} \Delta K^*_j}{\sum_{j=0}^{H} {v^*}_B{}^{j+1/2} \Delta B_j + v'_B \sum_{j=0}^{H} {v^*}_B{}^{j+1/2} \Delta B^*_j}$$

where incremental costs and benefits in the second cohort at cycle *j* are ΔK^*_j and ΔB^*_j. Here, $v'_C = \frac{1}{1+r'_C}$ and $v'_B = \frac{1}{1+r'_B}$, where the discount rates for costs and benefits over a year between the current time and the time of the start of the second future incident cohort are r'_C and benefits r'_B respectively. Henceforth, we assume that incremental costs and benefits are independent of the cohort, i.e. $\Delta K^*_j = \Delta K_j$ and $\Delta B^*_j = \Delta B_j$. But note that this assumption would be violated, for example, if a medical condition is diagnosed at an earlier age over time. Then;

ICER (first two future incident cohorts) = $\frac{(1+v'_C)}{(1+v'_B)}$ ICER (first future incident cohort)

Henceforth, we assume $r'_C = r_C$ and $r'_B = r_B$, then;

ICER (first two future incident cohorts) = $\frac{(1+v_C)}{(1+v_B)}$ ICER (first future incident cohort)

(Equation 1)

where $v_C = \frac{1}{1+r_C}$ and $v_B = \frac{1}{1+r_B}$.

All future incident cohorts

Now assume, more realistically, that a new cohort of patients will become *eligible* for treatment with the new or comparator technologies at the start of each of *T* years in the future. The new and comparator technologies are assumed to become obsolete

after T years, and are then replaced by another technology. In the general case, assume that the number of eligible patients at the start of each cohort, relative to the number of eligible patients at the start of the first year, is given by n_t, at year t, so that $n_0 = 1$. Assume further that the probability that a patient is given the new technology in the t^{th} year in the future is p_t. The n_t could increase with year t, e.g. to model the increasing numbers of Type 2 diabetes patients in the future as obesity becomes more widespread. The graph of the volume of sales of a drug, i.e. the product $n_t p_t$, against year t is generally $\cap$ -shaped (Danzon & Kim 2002). The annual volume of a drug sold typically increases for the first decade after drug launch, reflecting the diffusion of new drug after launch. Diffusion is particularly rapid in France, Canada and the US, but slower in Germany and the UK. The annual volume of a drug sold in the second decade after launch reflects post-patent experience and declines as patients switch to newer drugs (Danzon & Kim 2002; Salomon et al. 2004). Then, the relative number of patients in the incident cohort starting t years in the future affected by the new technology is $n_t p_t$. By analogy with Equation 1;

$$\text{ICER (all future incident cohorts)} = \frac{\sum_{t=0}^{T} n_t p_t v_C^{\,t}}{\sum_{t=0}^{T} n_t p_t v_B^{\,t}} \text{ ICER (first future incident cohort)}$$

(Equation 2a)

Therefore, the ICERs for all future incident cohorts combined and for the first future incident cohort are equal if the cost and benefit discount rates, r_c and r_b, are equal. Alternatively, if $r_C > r_B$;

$$\left(\frac{v_C}{v_B}\right)^T \leq \frac{\text{ICER (multiple incidence cohorts)}}{\text{ICER (current incidence cohort)}} \leq 1 \qquad \text{(Equation 2b)}$$

where the lower bound occurs in the unlikely event that $n_t p_t$ for all i is negligible compared to $n_T p_T$, i.e. when the number of patients treated with the new technology is negligible at any time before T compared to the number treated at time T. Conversely, the upper bound occurs when $n_t p_t$ for all i is negligible compared to $n_0 p_0$. If $r_C < r_B$;

$$1 \leq \frac{\text{ICER (multiple incidence cohorts)}}{\text{ICER (current incidence cohort)}} \leq \left(\frac{v_C}{v_B}\right)^T$$

When the product $n_t p_t$ is equal for all t, Equation 2a simplifies to;

$$\text{ICER (all future incident cohorts)} = \left(\frac{1-v_B}{1-v_C}\right)\left(\frac{1-v_C^{\,T+1}}{1-v_B^{\,T+1}}\right) \text{ ICER (first future incident cohort)} \qquad \text{(Equation 2c)}$$

Now, if we assume that $n_t p_t$ follows a $\cap$ -shaped quadratic curve then Equation 2c again is applicable (see Appendix). If independent estimates of $n_t p_t$ are available then they should be used in Equation 2a, otherwise Equation 2c is appropriate. Equation 2c is convenient since we need estimate just the single parameter T, not the $n_t p_t$ for all t.

Example of ICER given different discount rates for costs and benefits
Suppose we are assessing cost-effectiveness of a technology in a jurisdiction where costs are discounted at 6% per year and health benefits at 1.5% per year. Suppose the technology is for an acute condition. Then the prevalent cohort is negligible, and we calculate the ICER based on all future incident cohorts from Equation 2. In all examples, we assume the technology is used over T = 30 years. Assuming either that $n_t p_t$ are equal for all t or that $n_t p_t$ follow a quadratic curve over time, then from Equation 2c;

$$\frac{\text{ICER (multiple incidence cohorts)}}{\text{ICER (current incidence cohort)}} \approx 0.57$$

The change in ICER is substantial. In general, for any n_t and p_t, (Equation 2b);

$$0.27 \leq \frac{\text{ICER (multiple incidence cohorts)}}{\text{ICER (current incidence cohort)}} \leq 1$$

ICER for prevalent cohort

In addition to the patients who will become eligible for the new technology in the future, there may be patients who are eligible at the time the technology is introduced. Such patients would switch from the current to the new technology. Denoting the incremental *per patient* costs and benefits of the prevalent cohort at cycle j = 1….H by ΔC_j, and ΔQ_j, the ICER for the prevalent cohort is;

$$\frac{\sum_{j=0}^{H} v^{*}{}_{C}{}^{j+1/2} \Delta C_j}{\sum_{j=0}^{H} v^{*}{}_{B}{}^{j+1/2} \Delta Q_j} \qquad \text{(Equation 3)}$$

ICER for incident and prevalent cohorts combined

Given that the ICER equals total incremental costs divided by total incremental benefits;

ICER (prevalent and all future incident cohorts) =

$$= \frac{\bar{p}N\sum_{j=0}^{H} v^{*}{}_{C}{}^{j+1/2} \Delta C_j + \left(\sum_{t=0}^{T} n_t p_t v_C{}^{t}\right)\left(\sum_{j=0}^{H} v^{*}{}_{C}{}^{j+1/2} \Delta K_j\right)}{\bar{p}N\sum_{j=0}^{H} v^{*}{}_{B}{}^{j+1/2} \Delta Q_j + \left(\sum_{t=0}^{T} n_t p_t v_B{}^{t}\right)\left(\sum_{j=0}^{H} v^{*}{}_{B}{}^{j+1/2} \Delta B_j\right)} \qquad \text{(Equation 4)}$$

where N is the number of patients in the prevalent cohort that are eligible for treatment, relative to the number of patients in the first future incident cohort, and $\bar{p}$ is the probability that a patient in the eligible prevalent cohort is given the new technology. Note that if the cost and benefit discount rates are equal, then Equation 4 implies that the ICER for the prevalent and incident cohorts combined will lie

between the ICER for the prevalent cohort alone and the ICER for the first future incident cohort alone.

We now introduce parameters to allow us to estimate *N* and $\overline{p}$. Denote the average age of patients at the start of any incident cohort as *A* (and assume that this has changed little over time). Suppose that a patient is eligible for treatment with the new technology over an average period of *M* years, from age *A* to age *A+M*. When *M* is small, e.g. treatments for acute infection, the costs and benefits of the incident and prevalent cohorts are similar, because the patients' initial parameters, such as the average age and average severity of condition are similar between the incident and prevalent cohorts (see below). Conversely, when *M* is large, for example, for long-term therapies for chronic conditions, the costs and benefits of the incident and prevalent cohorts can be substantially different for a variety of reasons. Hence the ICER for the prevalent cohort is similar to the ICER for the incident cohort for acute conditions, but can be very different for chronic conditions. On average, we expect that patients in the prevalent cohort are approximately half way through their treatment with the comparator technology. Correspondingly, we can expect that patients at the start of an incident cohort (i.e. at the start of their treatment), to be treated for approximately twice the length of time as patients in the prevalent cohort.

If the number of patients in the prevalent cohort that are eligible for treatment, relative to the number of patients in the first future incident cohort, *N*, is known from the literature, then this value should be used. Otherwise, we now describe how to estimate *N*. Denote the probability that a patient who is treated with the *comparator* technology survives from age A, at the start of an incident cohort, to age *A + t* as $s(A, A+t)$. Such data are often available from cost-effectiveness models. Then;

$$N = n_{-1}s(A, A+1) + n_{-2}s(A, A+2) + n_{-3}s(A, A+3) + ... + n_{-M+1}s(A, A+M-1)$$

(Equation 5)

This shows that when *M* is large, for conditions that require a long period of treatment, e.g. multiple sclerosis, *N* is large, and when *M* is small, for conditions that require short-term treatment, e.g. acute infection, *N* is small.

In Equation 4 we make the simplifying assumption that the proportion of patients in a given incident cohort that are given the new technology, p_t, does not change over cycle *j*. This assumption becomes less accurate as the eligible treatment period, *M* increases. We also assume that the probability that a patient in the eligible prevalent cohort is given the new technology, $\overline{p}$, does not change over cycle *j*. We estimate $\overline{p}$ as the weighted average of the p_t, with the weights equal to the number of patients in the prevalent cohort *t* years in the future;

$$\overline{p} = \frac{\sum_{t=0}^{M-1}\sum_{i=0}^{M-1} p_t n_{-i} s(A, A+t+i)}{\sum_{t=0}^{M-1}\sum_{i=0}^{M-1} n_{-i} s(A, A+t+i)} \qquad \text{(Equation 6)}$$

where subscript *i* refers to the incident cohort that started *i* years in the past. Now suppose that the cost and benefit discount rates are equal, i.e. $v_C = v_B = v$. Then Equation 4 becomes;

ICER (prevalent and multiple incident cohorts) =

$$\frac{\dfrac{\overline{p}N}{\left(\sum_{t=0}^{T} n_t p_t v^t\right)}\sum_{j=0}^{H} v^{*j+1/2}\Delta C_j + \left(\sum_{j=0}^{H} v^{*j+1/2}\Delta K_j\right)}{\dfrac{\overline{p}N}{\left(\sum_{t=0}^{T} n_t p_t v^t\right)}\sum_{j=0}^{H} v^{*j+1/2}\Delta Q_j + \left(\sum_{j=0}^{H} v^{*j+1/2}\Delta B_j\right)}$$

From which it is clear that the prevalent cohort is negligible when $\dfrac{\overline{p}N}{\left(\sum_{t=0}^{T} n_t p_t v^t\right)}$ is small. This is true when *T* is very large, or *M* is very small. We now consider three cases;

1: Parameters for incident and prevalent cohorts are known
2: Parameters for incident cohort only are known
3: Parameters for prevalent cohort only are known

Case 1: Parameters for incident and prevalent cohorts known

Suppose we know the model parameters for both the incident and prevalent cohorts from a literature search. Then we can calculate ΔC_j, ΔQ_j, ΔK_j, and ΔB_j. We then calculate the ICER for the incident and prevalent cohorts combined from Equation 4, using an estimate of the p_t and hence $\overline{p}$, as explained in the Discussion.

Case 2: Parameters for incident cohort only known

Suppose we know the parameter values for the incident cohort only. We now describe two methods for estimating the incremental costs and benefits for the prevalent cohort, ΔC_j, ΔQ_j. As above, we then calculate the ICER for the incident and prevalent cohorts combined from Equation 4.

Method A – estimate costs and benefits for each incident cohort in prevalent cohort

Suppose the costs per patient in the incident cohort are K_j and K'_j at cycle j =1…H for the new and comparator technologies respectively (Fig. 1). As above, we assume that these costs are the same across all incident cohorts. We cannot simply assume that the new technology future costs from incident cohorts that started *t* years in the past, $K_{t,j}$, (where *j* is the number of cycles since the start of the incident cohort that started *t* years in the past, i.e. $j \geq t$) are given by K_j, because this would assume that patients had been treated with the new technology in the past, which is incorrect. Instead, we estimate the $K_{t,j}$ as described in the following algorithm, which can be coded as a macro. For the incident cohort that started *t* years in the past, and that contributes to the prevalent cohort $t = 1 \ldots M - 1$;

1. Set the average age at the start of treatment on the new technology equal to the average age of the comparator incident cohort at the time of assessment.
2. Repeat for the number of patients across the Markov states.

3. Record the total resulting new technology costs and benefits going into the future, i.e. $K_{t,j}$ and $Q_{t,j}$.

Next, the prevalent costs and benefits for the new technology at cycle *j* are calculated as $C_j = \sum_{t=1}^{M-1} K_{t,j} n_t$ and $Q_j = \sum_{t=1}^{M-1} Q_{t,j} n_t$ and for the comparator technology as $C'_j = \sum_{t=1}^{M-1} K'_{t,j} n_t$ and $Q'_j = \sum_{t=1}^{M-1} Q'_{t,j} n_t$ (Fig. 1). More precisely, the costs and benefits for the new technology, C_j and Q_j, should be adjusted to reflect the fact that patients in the prevalent cohort have just started the new technology. For example these patients may incur upfront costs for treating adverse events from the new technology. Conversely, patients in the prevalent cohort of the comparator drug will not incur such start-up costs. The incremental prevalent costs and benefits are calculated as $\Delta C_j = C_j - C'_j$ and $\Delta Q_j = Q_j - Q'_j$.

Method B – estimation of initial parameter values for prevalent cohort
Here, we estimate the parameter values that specify the characteristics of patients at the start of the prevalent cohort. This method is slightly less accurate than Method A, but simpler to implement in cost-effectiveness analyses.

Estimation of average age of prevalent cohort
We denote the average age of patients at the start of an incident cohort as *A*, and the probability that a patient, aged *A*, who starts on the comparator technology survives to the end of the t^{th} year as $s(A, A+t)$. The average age of patients at the start of the prevalent cohort is then;

$$A_p = \frac{\left[(A+1)n_{-1}s(A,A+1) + (A+2)n_{-2}s(A,A+2) + \ldots + (A+M-1)n_{-M+1}s(A,A+M-1)\right]}{\left[n_{-1}s(A,A+1) + n_{-2}s(A,A+2) + \ldots + n_{-M+1}s(A,A+M-1)\right]}$$

$$= \frac{1}{N}\left[(A+1)n_{-1}s(A,A+1) + (A+2)n_{-2}s(A,A+2) + \ldots + (A+M-1)n_{-M+1}s(A,A+M-1)\right]$$

(Equation 7)

If the n_t are known, then these values should be used. Otherwise it is reasonable to assume that they are all equal.

Estimation of distribution of patients across Markov states in prevalent cohort
Suppose we know the distribution of patients at the start of the incident cohort across Markov health states, I_x for state x=1….I, where $\sum_{x=1}^{I} I_x = 1$. The Markov states could, for example, represent the disability levels of a chronic disease, e.g. the Expanded Disability Status Scale for multiple sclerosis. It is easy to calculate the probability that a patient moves from state *x* to state *y* in *t* years, given that the patient is still alive after *t* years, *q(x,y,t)*, where $\sum_{y} q(x,y,t) = 1$ for all *t* and *x*. Then the proportion of patients at the start of the prevalent cohort in state *x* equals the number of patients in the prevalent cohort in state *x* divided by the total number of patients in the prevalent cohort;

$$I^{P}{}_{x} = \frac{\left[s(A,A+1)n_{-1}\sum_{y=1..I} I_y p(y,x,1) + s(A,A+2)n_{-2}\sum_{y=1..I} I_y p(y,x,2) + ... + s(A,A+M-1)n_{-M+1}\sum_{y=1..I} I_y p(y,x,M-1) \right]}{\left[s(A,A+1)n_{-1} + s(A,A+2)n_{-2} + ... + s(A,A+M-1)n_{-M+1} \right]}$$

$$= \frac{1}{N}\left[s(A,A+1)n_{-1}\sum_{y=1..I} I_y p(y,x,1) + s(A,A+2)n_{-2}\sum_{y=1..I} I_y p(y,x,2) + ... + s(A,A+M-1)n_{-M+1}\sum_{y=1..I} I_y p(y,x,M-1) \right]$$

(Equation 8)

Other parameters

Other parameters that could differ between the prevalent and incident cohorts include risk factors such as blood pressure, mix of ethnicity, proportion of patients who are smokers. Estimation of the probability density function of ages of patients in the prevalent cohort given the probability density function of ages in the incident cohort is given in the Appendix. Also estimation of the sex ratio in the prevalent cohort given the ratio in the incident cohort is given in the Appendix.

As described in Method A, to be more accurate, the costs and benefits for the new technology, C_j and Q_j, should be adjusted to reflect the fact that patients in the prevalent cohort have just started the new technology.

Case 3: Parameters for prevalent cohort only known

Suppose we know the parameter values for the prevalent cohort only, e.g. A_p and $I^P{}_x$. We recommend the method described in this section to estimate the incremental costs and benefits for the incident cohort, ΔK_j, ΔB_j. We also outline a more complex method to estimate these quantities in the Appendix. As above, once we estimate ΔK_j, ΔB_j, we calculate the ICER for the incident and prevalent cohorts combined from Equation 4. In this case, it is not possible to calculate the ΔK_j, ΔB_j directly, in the manner described in Method A above for estimating ΔC_j, ΔQ_j. Instead, we must estimate the initial parameters for the incident cohort, as follows.

We estimate the parameters for the incident cohort, I_x and A, from $I^P{}_x$ and A_p as follows. We define a single aggregate objective function as the sum of the squares of the differences between the calculated (using the unknown I_x and A) and known values of $I^P{}_x$ and A_P. A_P and A are related by Equation 7, and $I^P{}_x$ and I_x are related by Equation 8. The objective function is then minimised using 'Solver' in Excel® (Microsoft Corporation, Redmond, WA, USA) or similar software by varying I_x and A, subject to the constraints $\sum_{x=1}^{I} I_x = 1$, $0 \le I_x \le 1$, and $A > 0$.

Application to example cost-effectiveness model

Here, we apply the methods described above on an example cost-effectiveness model of a new drug versus an existing comparator drug to treat a chronic progressive condition, e.g. multiple sclerosis. We consider first that the parameters for the incident cohorts only are known (Case 2), then that the parameters for the prevalent cohort only are known (Case 3).

Structure of example model

The structure of the Markov model is given as follows;

New drug period of use: assume that new drug is used for the next T = 30 years, and initially assume that all eligible patients are switched immediately from the comparator to the new drug, i.e. p_t = 1 for all t.
Death rates: taken from UK life tables. For simplicity, disease does not affect mortality.
Sex ratio: all male for simplicity,
Model cycle length: annual,
Time horizon (H): until death of all patients.
Period over which drug taken: from diagnosis until death.
Health states (x): labelled 1 to 5, with 1 denoting mild illness, 5 denoting serious illness.

Transition probabilities between Markov health states;

Comparator drug

$$\begin{pmatrix} 1-q & q & 0 & 0 & 0 \\ 0 & 1-q & q & 0 & 0 \\ 0 & 0 & 1-q & q & 0 \\ 0 & 0 & 0 & 1-q & q \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

New drug

$$\begin{pmatrix} 1-(RR)q & (RR)q & 0 & 0 & 0 \\ 0 & 1-(RR)q & (RR)q & 0 & 0 \\ 0 & 0 & 1-(RR)q & (RR)q & 0 \\ 0 & 0 & 0 & 1-(RR)q & (RR)q \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

where q is the probability of disease progression a year, and RR is the relative risk of disease progression of the new drug versus the comparator. Set p = 0.1, RR = 0.5.

Costs: state cost increases with disease severity, and is £10,000 * x per person per year for state x = 1…5. In addition, drug costs are £0 and £10,000 per person per year for the comparator and new drug respectively.

Utilities: utility decreases with disease severity, and is 1 – 0.2 * x per person per year for states x = 1…5.

Discounting: $r_C = r_B$ = 3.5% as required in the UK (NICE 2004).

Relative number of patients in incident cohort in year t, n_t: For simplicity, all set equal.

Case 2: Parameters for incident cohort only known

Suppose from the literature, we find that;

- the average age at diagnosis, i.e. at the start of an incident cohort, A = 30 years.
- 75% of patients at the start of the incident cohort are in disability state x = 1, 25% are in disability state x = 2, and no patients are in the most severe disability states x = 3 to 5.

Therefore, we modelled patients from age 30 to death or age 100. This gives M = 70 years over which patients are eligible to be treated with the new drug. Given the initial parameter values above, we find that the ICER for the first future incident cohort alone is approx. £25,000 per quality-adjusted life year (Table 2). Assuming all patient are given the new drug in future, i.e. p_t = 1 for all t, then all patients in the

prevalent cohort are also given the new drug in the future, i.e. $\overline{p}$ = 1. The terms $\sum_{t=0}^{T} n_t p_t v_C{}^t$ and $\sum_{t=0}^{T} n_t p_t v_B{}^t$ in Equation 4 are then both approx. 19.

Method A – estimate costs and benefits for each incident cohort in prevalent cohort

The total discounted costs and benefits for the prevalent cohort were calculated using the algorithm described above, implemented in Visual Basic for Applications in Microsoft Excel (Figs. 2). The ICER for the prevalent cohort alone was calculated as £94,000 per quality-adjusted life year (Table 2), which is substantially higher than the £25,000 for the incident cohorts alone. This demonstrates that the ICERs for the incident and prevalent cohorts can differ radically. The ICER for both the prevalent and all incident cohorts combined was £57,000 (Table 2).

Method B – estimation of initial parameter values for prevalent cohort

The survival probabilities from age *A* to *A*+*t* under the comparator drug, $s(A, A+t)$, were calculated from the example cost-effectiveness model. Hence, we estimate that the prevalent cohort is *N* = 47 times the size of a single incident cohort (Equation 5). Next, the average age of patients in the prevalent cohort, A_P, was calculated as approx. 56 years (Equation 7). The initial distribution of patients across the disability states in the prevalent cohort, I^P_x, was calculated from Equation 8. As expected patients in the prevalent cohort are at a more advanced stage of illness (Fig. 3). The proportion of patients at the start of the prevalent cohort in disability states 1 to 5 was 0.14, 0.19, 0.18, 0.15, 0.33 respectively. The ICER for the prevalent cohort alone was then calculated from these parameter values as approx. £145,000 per quality-adjusted life year (Table 2), which is substantially higher than the ICER for the incident cohorts alone (£25,000) and the ICER for the prevalent cohort alone calculated by Method A (£94,000). In Fig. 4, we explain graphically why the ICER for the prevalent cohort is greater than for the incident cohort. As recommended, the ICER for the prevalent and incident cohorts combined is £73,000 per quality-adjusted life year (Equation 4, Table 2). Note that the agreement between the prevalent methods A and B would probably have been closer if we had modelled a distribution of ages for cohorts as described in the Appendix, rather than just a mean age.

Case 3: Parameters for prevalent cohort only known

Now assume we know the parameters for the prevalent cohort but not for the incident cohort. For example, suppose we know that A_P = 56 and that the proportion of patients initially in disability states 1 to 5, i.e. the I^P_i are 0.14, 0.19, 0.18, 0.15, 0.33 respectively. As described above, we used the 'Solver' algorithm from Microsoft Excel to find that the objective function is minimised by the correct values *A* = 56 and the I_x are 0.75, 0.25, 0.00, 0.00, 0.00.

***p_t* quadratic**

Now suppose the probability that a patient eligible for treatment takes the new drug at time *t* follows a $\cap$-shaped quadratic curve, with equation $n_t p_t = \frac{t(T-t)}{(T/2)^2}$, as previously. As before, set n_t = 1 for all *t*, and *T* = 30 years, then

$\sum_{t=0}^{T} n_t p_t v_C{}^t = \sum_{t=0}^{T} n_t p_t v_B{}^t$ = 12.2, and hence $\bar{p}$ = 0.64 (Equation 6). Hence the ICER for the prevalent cohort calculated by Method A and all incident cohorts combined is £57,000 per quality-adjusted life year (Equation 4) and calculated by Method B £73,000. These values are almost identical to the values given p_t = 1 for all t (Table 2).

Discussion

In this paper we have argued that the cost-effectiveness of a treatment should be assessed in relation to all patients whose costs and benefits will be affected; both those currently eligible and those who will become eligible for the new treatment in the future. On average, patients in the prevalent cohort are older and at a more advanced stage of disease than patients in the incident cohort. Furthermore, the longer the period, M, over which the technology is applicable, i.e. for more chronic conditions, the greater these differences. In summary, the suggestions in this paper are particularly important to implement in cost-effectiveness analysis in any of the following circumstances;

- for chronic conditions, particularly for chronic progressive conditions, e.g. Alzheimer's disease, multiple sclerosis, Parkinson's disease, chronic lymphocytic leukaemia, cystic fibrosis, idiopathic pulmonary arterial hypertension, eczema, cardiovascular disease, Crohn's disease, diabetes, epilepsy, chronic Hepatitis B, schizophrenia, rheumatoid arthritis.
- when the discount rates for costs and benefits differ.

In these cases, the ICER as calculated in this paper may differ substantially from the ICER as traditionally calculated. In particular, we have described a simplified but realistic example cost-effectiveness analysis of a chronic progressive condition, assuming equal cost and benefit discount rates. In this example, the ICER as calculated by our method is approximately 2.3 times the ICER as traditionally calculated by assuming just a single incident cohort, and 0.6 times the other traditional method of assuming a single prevalent cohort (Table 2).

There has been much debate about appropriate values for the cost and benefit discount rates, r_C and r_B. Most economists agree that $r_C = r_B$ (Gravelle & Smith 2001). However, some suggest $r_C > r_B$ (Gravelle & Smith 2001; Brouwer et al. 2005). In particular, Brouwer recommend r_C = 3.5% and r_B = 1.5%, and Gravelle & Smith (2001) suggest that r_B should be 2-5% less than r_C. r_C and r_B, are equal in most countries in the base case analysis (ISPOR 2007). For example, in the UK, NICE prescribe $r_C = r_B$ = 3.5%, based on guidance from the UK Treasury (NICE 2004), and this has been defended by Claxton et al. (2006). Notable exceptions are France, Norway, Portugal and Russia, where r_C and r_B differ. Furthermore, r_C and r_B often differ in sensitivity analyses.

An obvious question is: when the prevalent cohort is not negligible, when is the ICER for the prevalent cohort greater than the ICER for the incident cohort, and vice versa? Below, we suggest an answer to this question for three types of treatments/conditions. First, we have shown that for the example cost-effectiveness model of a continuous treatment for a progressive chronic disease, the prevalent cohort ICER is substantially greater than the incident cohort ICER, because at each cycle, the ratio of incremental costs to incremental benefits is greater for the

prevalent cohort (Fig. 4). We believe that this is probably a general result for a progressive chronic condition, although this question warrants further analysis. Second, we consider a continuous treatment for a non-progressive chronic condition, such as asthma. Suppose there are two health states A and B, and patients are in the worse state A under the comparator drug and the better state B under the new drug. Suppose further that life expectancy is independent of the drug and that costs are a function just of the drug (higher for the new drug) and whether the patient is in state A (higher) or state B (lower). Further, suppose that patient utility is a function of just the state, and is higher in state B than in state A. In this case, the ratios of incremental costs and benefits $\frac{\Delta K_j}{\Delta B_j}$ and $\frac{\Delta C_j}{\Delta Q_j}$ are constant over cycle *j* and are the same for the incident and prevalent cohorts. Hence the prevalent cohort ICER equals the incident cohort ICER. Third, consider the scenario where the majority of costs are incurred up front for chronic conditions. This is particularly appropriate for non-drug technologies, such as cardiac pacemakers for heart conditions and cochlear implants for deafness. Again, suppose there are two health states A and B, and suppose that patients are in the worse state A under the comparator technology and in the better state B under the new technology. Again, suppose that life expectancy is independent of the technology. Suppose the cost of the technology, e.g. cost of cochlear implant itself plus cost of implantation surgery, is incurred in the first cycle, and is greater for the new than the old technology. Health state costs can be higher or lower in state A than in state B. Patient utility is as described above. In this case, for the incident and prevalent cohorts, the ratios of incremental costs and benefits $\frac{\Delta K_j}{\Delta B_j}$ and $\frac{\Delta C_j}{\Delta Q_j}$ are high in the first cycle, and far smaller in all future cycles. The ratios for the two cohorts are equal by cycle. However, given that patients are older in the prevalent than in the incident cohort, in the prevalent cohort, there will be fewer cycles with low incremental cost/benefit ratios. Hence, the prevalent cohort ICER is greater than the incident cohort ICER.

Another question is whether the ICER calculated here will be greater or smaller than the ICER as traditionally calculated. In general, it is not possible to say. Some technologies will appear more cost-effective, and others less cost-effective. Consider first the case when the prevalent cohort is negligible compared to the incident cohort, e.g. treatments for acute conditions. Then, so long as the discount rate for costs equals the rate for benefits, the ICER will be unchanged. Alternatively, if the discount rate for costs is greater than the rate for benefits, the ICER will be less than traditionally calculated. Now, assume that the prevalent cohort is not negligible. In previous model-based cost-effectiveness analyses, either;

1: all patient-related parameters (e.g. average age, average disability level) refer to the prevalent cohort, or
2: all patient-related parameters refer to the incident cohort, or
3: some parameters refer to the prevalent cohort and the rest to the incident cohort.

Again, assuming equal discount rates, in the common scenario that the prevalent cohort ICER is greater than the incident cohort ICER, the ICER as calculated here is lower than the ICER calculated in case 1, greater than in case 2, and uncertain in case 3. Conversely, in the less likely event that the prevalent cohort ICER is lower than the incident cohort ICER, then these conclusions are reversed. However, in a review of the simulated populations in model-based cost-effectiveness analyses, we found very few cost-effectiveness studies that explicitly state whether model parameters were derived from incident or prevalent cohorts (see companion paper).

Therefore, our analysis suggests that the ICER as calculated in previous cost-effectiveness analyses may be substantially different from the ICER as calculated according to the methods of this paper. As a side issue, note that we have assumed that the costs in all future incident cohorts are equal, and similarly for the benefits. This assumption would be violated if, for example, one component of the costs is predicted to increase in the future at a different rate to the other components of the costs. Then we must adjust Equations 2a and 4 appropriately.

One disadvantage of our suggested methods is that they require estimation of additional parameters. We suggest the following algorithm to allow the analyst to decide when it is necessary to implement our methodology in full. First, if the cost and benefit discount rates differ, our suggested method should be followed. Specifically, we must estimate the *relative* sizes of the affected patient populations ($n_t p_t$) for each year in the future up to year $t = T$ (Equation 2a). This could be estimated by analysing trends in the volumes of sales of similar technologies in the past. On the other hand, if such data is not available, we suggest above that $n_t p_t$ can be assumed a quadratic function of year t. We then require only an estimate of the lifetime of the new technology, T (Equation 2c). This could also be estimated from the experiences of similar past technologies. Variability in $n_t p_t$ and/or T could be incorporated in the probabilistic sensitivity analysis. Next, assume the cost and benefit discount rates are equal. When the size of the prevalent cohort is negligible compared to the size of the incident cohort, then the ICER for the prevalent cohort and all future incident cohorts combined can be approximated by the ICER for the first future incident cohort alone. When the prevalent cohort is not small, the analyst should first compare the ICERs for the prevalent and incident cohorts. Given that the ICER for both types of cohort combined lies between the ICER for the prevalent cohort and the ICER for the first future incident cohort when the cost and benefit discount rates are equal (see above), if the two ICERs are similar, then the ICER for the prevalent cohort and all future incident cohorts combined can be approximated by the ICER for either the prevalent cohort or the ICER for the first future incident cohort. If the ICERs for the prevalent cohort and first future incident cohort are not similar, then our methodology should be adopted in full.

We require estimates of n_t and p_t separately for each year in the future up to year $t = T$ in order to estimate $\bar{p}$ (Equation 6). We suggest that, without further data, it is reasonable to assume that the n_t are equal for all t. The p_t are then estimated as described in the estimation of $n_t p_t$ above. Next, we must estimate the size of the prevalent cohort relative to the size of the first future incident cohort, N, and patient-related parameters, such as the average age and average disability status for both the incident and prevalent cohort. Uncertainty in N should also be reflected in the probabilistic sensitivity analysis. Given that the ICER can be strongly influenced by our methodology, we suggest that the extra effort in estimating these parameters and in adjusting the cost-effectiveness analysis is well worth while. Nonetheless, we are mindful of the extra analytical effort and data requirements that are implied by our methods. We have therefore also provided some practical tools for estimating the costs and effects for incident or prevalent patient cohorts when full data on the other type of cohort is unavailable. Ideally, however, cost-effectiveness analyses in these situations should be grounded in rigorous empirical studies which yield separate effectiveness estimates and other data from both incident, newly eligible, patients and those prevalent patients who are switching to the new treatment.

An Excel spreadsheet implementing the example cost-effectiveness model is available from the authors on request.

References

Beaglehole, R., Bonita, R & Kjellström, T. (1993) *Basic epidemiology*. World Health Organisation, Geneva.

Brouwer WBF, Niessen LW, Postma MJ, Rutten FFH (2005) Need for differential discounting of costs and health effects in cost effectiveness analyses. *British Medical Journal*, 331:446-448.

Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, et al. (2005) Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Economics* 14:339-347.

Claxton K, Sculpher M, Culyer A, McCabe C, Briggs A, Akehurst R, Buxton M, Brazier J (2006) Discounting and cost-effectiveness in NICE - stepping back to sort out a confusion. *Health Economics*, 15:1-4.

Danzon, PM, Kim, JD. (2002) The life cycle of pharmaceuticals: a cross-national perspective. Office of Health Economics, Whitehall, London.

Devlin N, Parkin D. (2004) Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Economics 13:437-52.

George B, Harris A, Mitchell A. (2001) Cost-effectiveness analysis and the consistency of decision making. Evidence from pharmaceutical reimbursement in Australia (1991–1996). Pharmacoeconomics 19:1103–9.

Gravelle H & Smith D (2001) Discounting for health effects in cost-benefit and cost-effectiveness analysis. *Health Economics* ,10:587-599.

ISPOR (2007) International Society for Pharmacoeconomics and Outcomes Research http://www.ispor.org/PEguidelines/COMP1.asp (accessed 5 May 2007).

NICE. National Institute for Health and Clinical Excellence. (2004) *Guide to the Methods of Technology Appraisal*. London, www.nice.org.uk/page.aspx?o=201974 (accessed 7 Jun 2007).

Salomon, J.A., Weinstein, M.C., Goldie, S.J. (2004) Taking account of future technology in cost effectiveness analysis. *British Medical Journal*, 329, 733-736.

Sculpher MJ, Claxton K, Drummond M, McCabe C. (2006) Whither trial-based economic evaluation for health care decision making. Health Economics 15:677-687.

Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, Luce BR (2003) Principles of good practice for decision analytic modeling in health-care evaluation: Report of the ISPOR task force on good research practices-modeling studies. Value in Health, 6:9-17.

Appendix

Proof that if $n_t p_t$ is given as a quadratic curve, Equation 2c is applicable

Suppose $n_t p_t = at(T-t)$, with some constant *a*. Using the identity $\int_0^T xb^x dx = \frac{1}{(\ln b)^2}\left(Tb^T \ln b - b^T + 1\right)$ for some constant *b*, Equation 2a becomes;

ICER (all future incident cohorts) = $\left(\frac{\ln v_B}{\ln v_c}\right)^3 \frac{\left(Tv_c{}^T \ln v_c + T\ln v_c - 2v_c{}^T + 2\right)}{\left(Tv_B{}^T \ln v_B + T\ln v_B - 2v_B{}^T + 2\right)}$ ICER (first future incident cohort)

For all realistic values of v_c, v_B and *T*;

ICER (all future incident cohorts) $\approx \frac{r_B}{r_C}\left(\frac{1-v_C{}^{T+1}}{1-v_B{}^{T+1}}\right)$ ICER (first future incident cohort)

as Equation 2c.

Estimation of probability density function of age of patients in prevalent cohort

Suppose we know the probability density function of patient ages at the start of the incident cohort as $f(\alpha)$. Then the probability density function of patient ages at the start of the prevalent cohort is;

$$g(\alpha) = \frac{1}{N}\left[f(\alpha-1)s(\alpha-1,\alpha)n_{-1} + f(\alpha-2)s(\alpha-2,\alpha)n_{-2} + ... + f(\alpha-M+1)s(\alpha-M+1,\alpha)n_{-M+1}\right]$$

Where the relative number of patients in the prevalent cohort, *N*, is now generalised to;

$$N = n_{-1}\sum_\beta f(\beta)s(\beta,\beta+1) + n_{-2}\sum_\beta f(\beta)s(\beta,\beta+2) + n_{-3}\sum_\beta f(\beta)s(\beta,\beta+3) + ... + n_{-M+1}\sum_\beta f(\beta)s(\beta,\beta+M-1)$$

with summations over all ages β.

Equation 4 is adjusted to give;

ICER (prevalent and incident cohorts)

$$= \frac{\sum_\alpha g(\alpha)\left[PN\sum_{j=0}^{H} v^*{}_C{}^{j+1/2}\Delta C_j\right]_\alpha + \sum_\alpha f(\alpha)\left[\sum_{t=0}^{T} n_t p_t v_C{}^t \sum_{j=0}^{H} v^*{}_C{}^{j+1/2}\Delta K_j\right]_\alpha}{\sum_\alpha g(\alpha)\left[PN\sum_{j=0}^{H} v^*{}_B{}^{j+1/2}\Delta Q_j\right]_\alpha + \sum_\alpha f(\alpha)\left[\sum_{t=0}^{T} n_t p_t v_B{}^t \sum_{j=0}^{H} v^*{}_C{}^{j+1/2}\Delta B_j\right]_\alpha}$$

Estimation of sex ratio of prevalent cohort

Define the proportion of patients at the start of an incident cohort that are male and female as p_M. and p_F, and define the probability that a patient, aged *A*, who is under the comparator technology survives to the end of the i^{th} year as $s^m(A, A+i)$ and $s^f(A, A+i)$ for males and females respectively. Then by definition, $s(A, A+t) = p_m s^m(A, A+t) + p_f s^f(A, A+t)$, and the total number of male patients in the prevalent cohort is;

$$p_m\left[s^m(A, A+1)n_{-1} + s^m(A, A+2)n_{-2} + ... + s^m(A, A+M-1)n_{-(M-1)}\right]$$

Hence the proportion of patients at the start of the prevalent cohort that are male is;

$$P_m = \frac{p_m}{N}\left[s^m(A, A+1)n_{-1} + s^m(A, A+2)n_{-2} + ... + s^m(A, A+M-1)n_{-(M-1)}\right] \text{ (Equation A1)}$$

Estimation of average age at the start of incident cohort, *A*

Denote the probability that a patient, aged A_P, who takes the comparator technology survives to the end of the t^{th} year as $S(A_p, A_p + t)$. These values can be read directly from the cost-effectiveness model. First we estimate the $s(A, A + t)$;

$$S(A_p, A_p + M - 1) = \frac{n_{-1}s(A, A+M)}{\left[\left[n_{-1}s(A, A+1) + n_{-2}s(A, A+2) + ... + n_{-M+1}s(A, A+M-1)\right]\right]} = \frac{n_{-1}}{N}s(A, A+M)$$

$$S(A_p, A_p + M - 2) = \frac{1}{N}\left[n_{-2}s[A, A+M] + n_{-1}s[A, A+M-1]\right]$$

$$S(A_p, A_p + M - 3) = \frac{1}{N}\left[n_{-3}s(A, A+M) + n_{-2}s(A, A+M-1) + n_{-1}s(A, A+M-2)\right]$$

……

$$S(A_p, A_p + 1) = \frac{1}{N}\left[\frac{s(A, A+(M+1))}{(1+g)^M} + \frac{s(A, A+M)}{(1+g)^{M-1}} + \frac{s(A, A+(M-1))}{(1+g)^{M-2}} ....\right]$$

These expressions comprises *M-1* equations in *M-1* unknown variables. Hence, in principle we can solve for the $s(A, A+t)$, for *t* = 1…..*M*. Then we can solve Equation 7 for the average age of the incident cohort, *A*, by trying different values of *A*.

Estimation of sex ratio of incident cohort

Suppose we know the proportion of patients in the prevalent cohort that are male, P_M, and we want to calculate the proportion of patients in the incident cohort that are male, p_M. First, we calculate the $s^m(A, A+t)$ exactly as for the $s(A, A+t)$, but using the $S^m(A_p, A_p+t)$. We then solve Equation A1 for p_M.

Estimation of distribution of patients across Markov states in incident cohort

NB: This is work in progress

Table 1. Key parameters.

Parameter	Definition
ΔK_j, ΔB_j	incremental incident cohort cost and benefit per patient between the new and comparator technology at cycle j = 0, 1, 2,…., H
ΔC_j, ΔQ_j	incremental prevalent cohort cost and benefit per patient at cycle j = 0, 1, 2,…., H
K_j, K'_j	incident cohort cost per patient for the new and comparator technology at cycle j = 0, 1, 2,…., H
C_j , Q_j	prevalent cohort cost and benefit per patient for the new technology at cycle j = 0, 1, 2,…., H
C'_j , Q'_j	prevalent cohort cost and benefit per patient for the comparator technology at cycle j = 0, 1, 2,…., H
H	time horizon in cycles
r_C , r_B	discount rate for costs and benefits over a year
v_C , v_B	$= \frac{1}{1+r_C}$, $= \frac{1}{1+r_B}$
r^*_C , r^*_B	discount rate for costs and benefits over a cycle
v^*_C , v^*_B	$= \frac{1}{1+r^*_C}$, $= \frac{1}{1+r^*_B}$
T	expected lifetime of new technology in years
n_t	number of eligible patients at the start of the incident cohort starting t = 1….T years in the future, relative to the number of eligible patients at the start of the first year
p_t	probability a patient is given the new technology t = 1….T years in the future
$\overline{p}$	probability that a patient in the eligible prevalent cohort is given the new technology
N	number of patients in the prevalent cohort that are eligible for treatment, relative to the number of patients in the first future incident cohort
A, A_P	average age of patients at the start of incident and prevalent cohorts respectively
$s(A, A+t)$	probability a patient who is treated with the comparator technology survives from age A to $A + t$
M	number of years over which patients are eligible to be treated with the new technology
I_x , I_x^P	distribution of patients at the start of the incident and prevalent cohorts respectively across Markov health states x=1….I

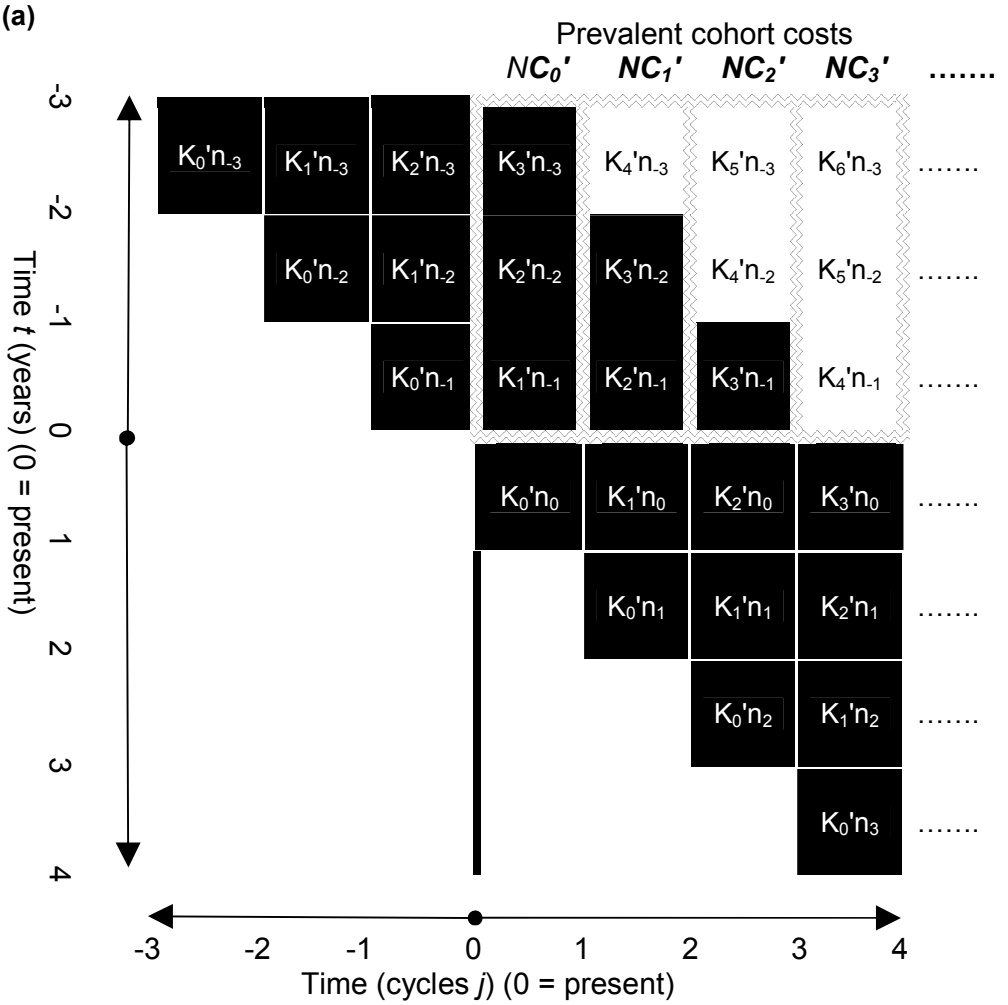
$q(x,y,t)$	probability a patient moves from health state x to state y in t years, given patient still alive after t years

Table 2. Example model cost-effectiveness of a new drug versus a comparator drug according to the cohort modelled. Currently, cost effectiveness analyses consider only either the prevalent cohort or the first future incident cohort. We suggest that both the prevalent and all future incident cohorts should be modelled.

Cohort	**Incremental cost (£)**	**Incremental benefit (QALYs)**	**ICER (£ / QALY) (nearest 1,000)**
first future incident	76,913	3.1	25,000*
all future incident combined	1,464,093	59	25,000*
prevalent method A	4,819,811	51	94,000
prevalent method B	5,734,304	40	145,000
prevalent method A + all incident	6,283,904	112	57,000
prevalent method B + all incident	7,198,398	98	73,000

* ICERs for the first future incident cohort and all future incident cohorts are equal when the discount rates for costs and benefits are equal.

Figure 1. Prevalent and incident cohort costs for (a) the comparator and (b) the new technology. Incident cohorts are shown in separate rows. The average age of patients at the start of each incident cohort is *A*. For simplicity, one cycle equals one year in this example. Here, the technology is applicable over $M = 4$ years (black cells), and the prevalent cohort comprises $M - 1 = 3$ incident cohorts. The future prevalent cohort comparator and new technology total costs at cycle j, NC_j' and NC_j equal the sum of the costs in the respective highlighted boxes. In (b), all costs before the assessment time refer to the comparator technology, because the new technology was not used then.



(b)

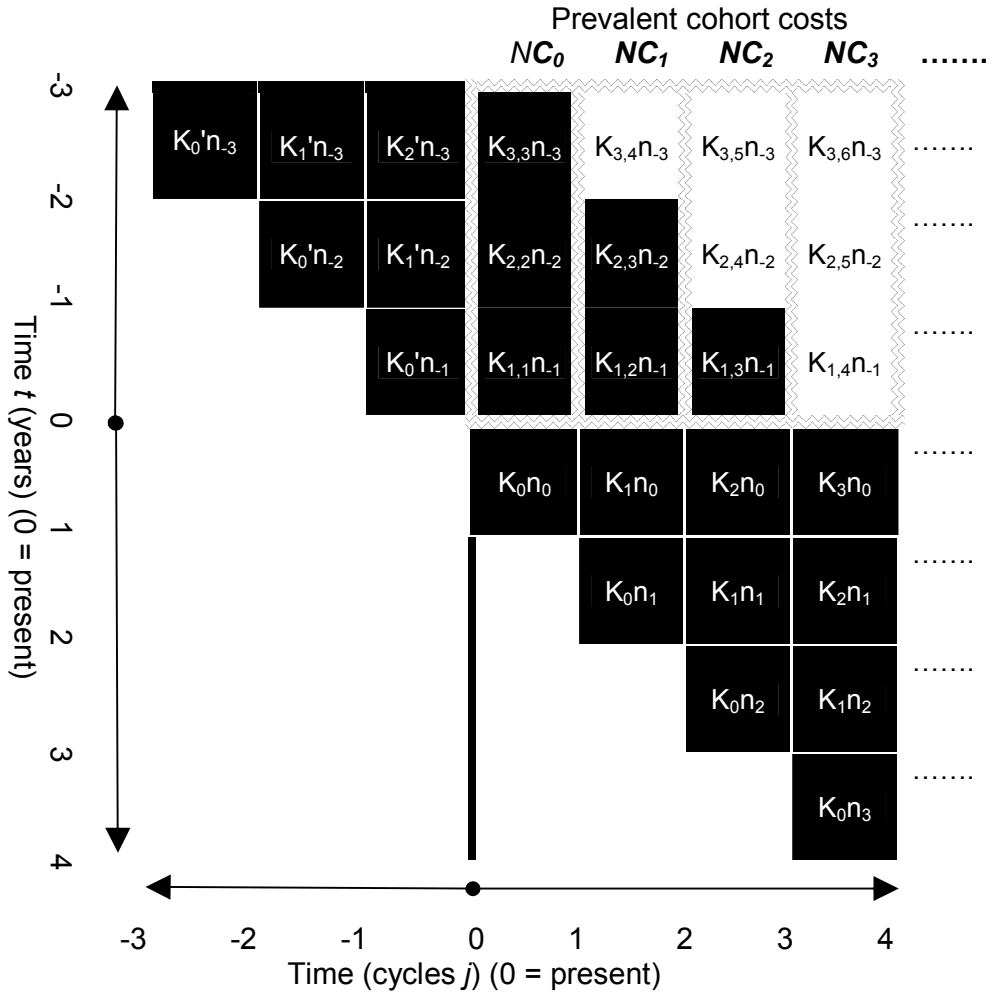
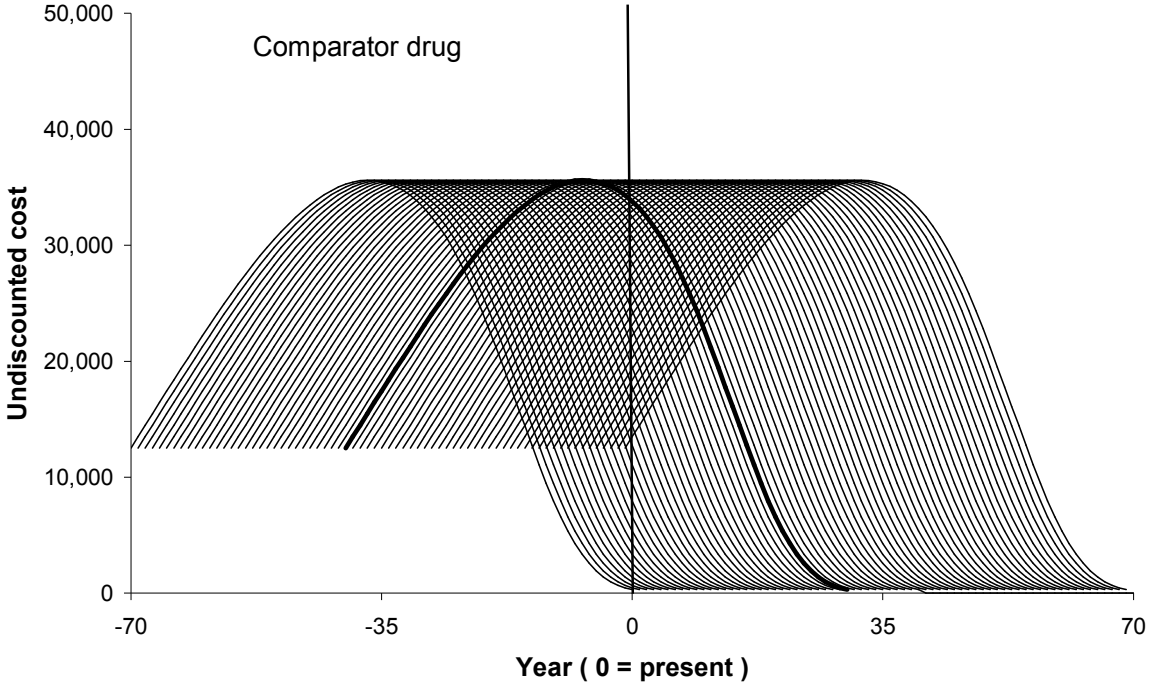
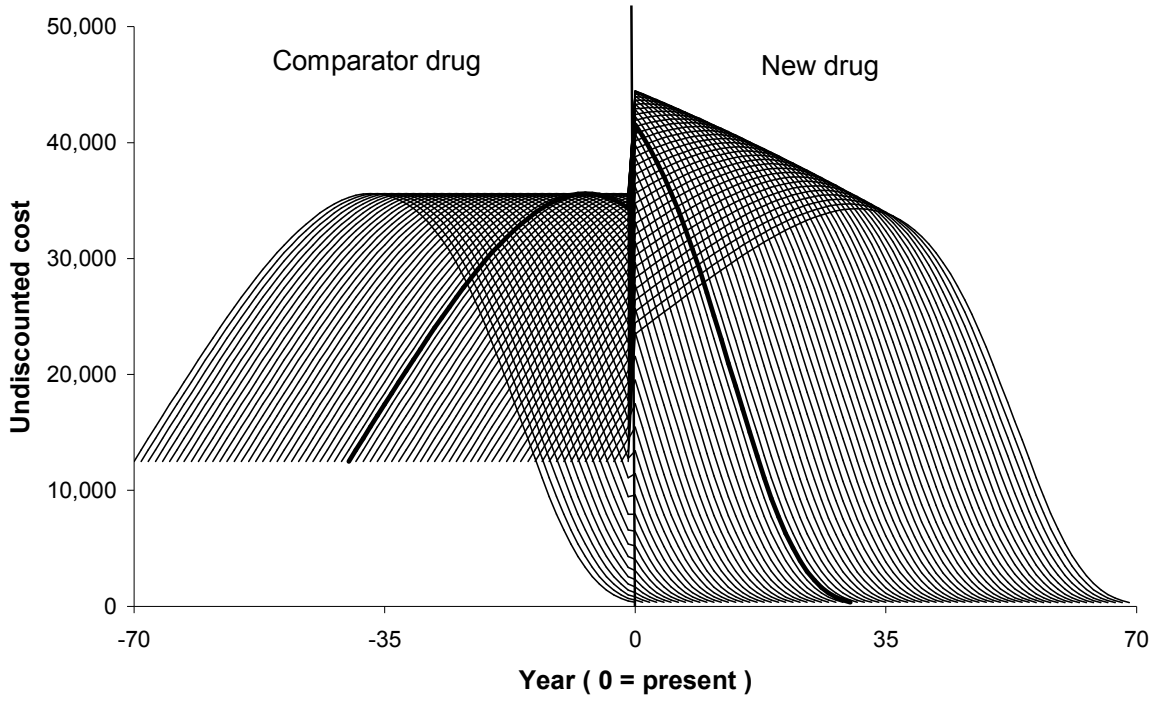

Prevalent cohort costs
NC0 NC1 NC2 NC3 …….
K0'n-3 K1'n-3 K2'n-3 K3,3n-3 K3,4n-3 K3,5n-3 K3,6n-3 …….
K0'n-2 K1'n-2 K2,2n-2 K2,3n-2 K2,4n-2 K2,5n-2 …….
K0'n-1 K1,1n-1 K1,2n-1 K1,3n-1 K1,4n-1 …….
K0n0 K1n0 K2n0 K3n0 …….
K0n1 K1n1 K2n1 …….
K0n2 K1n2 …….
K0n3 …….
Time t (years) (0 = present)
-3 -2 -1 0 1 2 3 4
Time (cycles j) (0 = present)

Figure 2. Undiscounted costs over time in the example cost-effectiveness model. (a) displays the per patient comparator drug costs showing separately all incident cohorts that started in the past. The costs in the future, i.e. to the right of the vertical line, comprise the costs of the prevalent cohort. For clarity, a single example incident cohort is displayed in bold. Costs initially rise as disease becomes more severe, thus incurring higher health state-related costs. Costs eventually fall to zero as patients die. (b) displays the same data for times in the past, but costs for the new drug in the future, i.e. for the new drug costs in the prevalent cohort. (c) displays comparator drug costs. In (c), the downward sloping line represents total costs in the prevalent cohort (summing over costs in all incident cohorts that started in the past), and the upward sloping line represents total costs in all future incident cohorts. To demonstrate scale, the incident cohorts that make up these quantities, some of which are shown in (a), are just visible at the bottom of the graph. We assume that there are the same number of patients in all incident cohorts.

(a)



(b)


Comparator drug
New drug
Undiscounted cost
50,000
40,000
30,000
20,000
10,000
0
-70
-35
0
35
70
Year (0 = present)

(c)

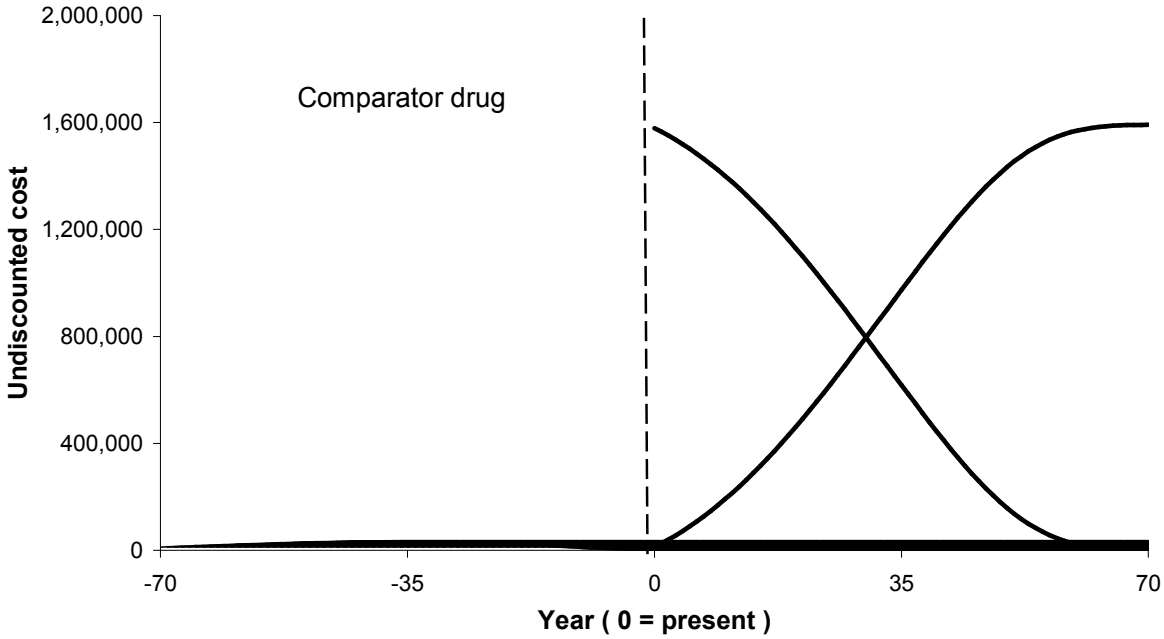

Comparator drug
Undiscounted cost
2,000,000
1,600,000
1,200,000
800,000
400,000
0
-70
-35
0
35
70
Year (0 = present)

Figure 3. Average disability state of patients by year in future for the example cost-effectiveness model. Broken lines represent patients in the prevalent cohort, continuous lines represent patients in the incident cohort. Bold lines represent the new drug, which slows the worsening of the illness, and thin lines represent the comparator drug.

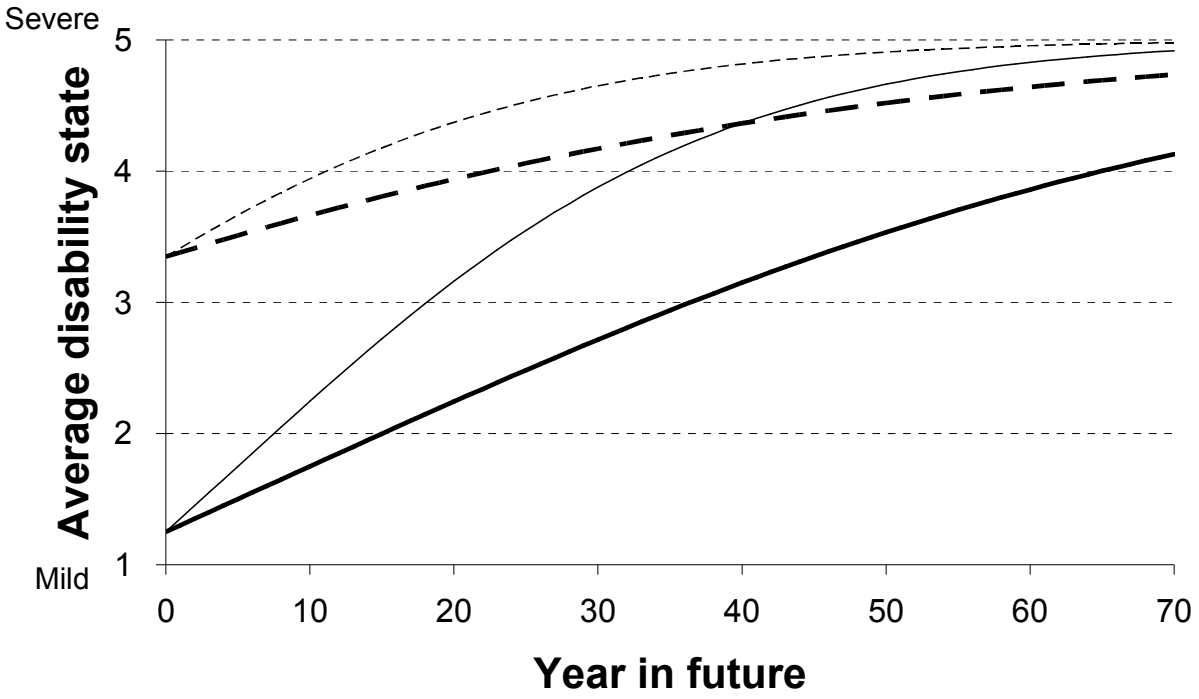
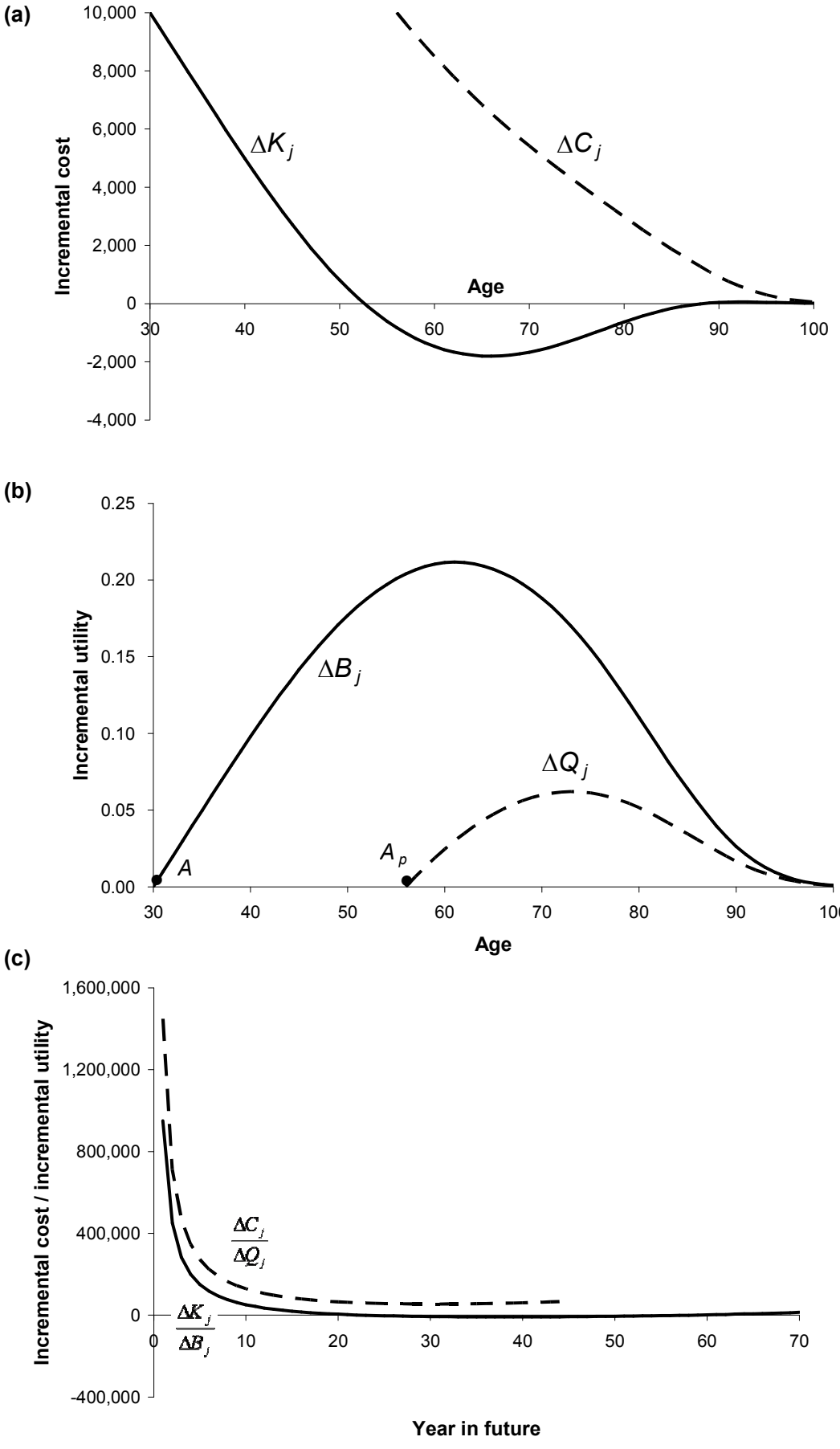


Figure 4. Explanation of why the ICER for the prevalent cohort is greater than for the incident cohort in the example cost-effectiveness model under Method B. (a) shows the per patient incremental undiscounted costs and (b) utilities and (c) ratio of costs to utilities. The continuous and broken lines represent the incident and prevalent cohorts respectively. In (a) both curves start at 10,000, which is the different in costs between the drugs per patient. Both curves fall as patients taking the comparator drug progress in disability more quickly than patients taking the new drug, thereby incurring higher state costs. Both curves tend to zero as all patients die. In (b) both curves start at zero because patients taking the two drugs start at the same level of disability. Both curves then increase as patients taking the comparator drug progress in disability more quickly than patients taking the new drug, i.e. their quality of life deteriorates more quickly. Again, both curves tend to zero as all patients die. (c) demonstrates that the ratio of incremental costs to utility at each cycle is always greater for the prevalent than the incident cohort. Together with the fact that the prevalent cohort ICER is more heavily weighted to the costs and benefits at earlier cycles compared to the incident cohort ICER (because of greater mortality at greater ages), the ICER for the prevalent cohort is greater than for the incident cohort.


(a)
Incremental cost
10,000
8,000
6,000
4,000
2,000
0
-2,000
-4,000
30
40
50
60
70
80
90
100
Age
ΔK_j
ΔC_j
(b)
Incremental utility
0.25
0.20
0.15
0.10
0.05
0.00
30
40
50
60
70
80
90
100
Age
ΔB_j
ΔQ_j
A
A_p
(c)
Incremental cost / incremental utility
1,600,000
1,200,000
800,000
400,000
0
-400,000
0
10
20
30
40
50
60
70
Year in future
$\frac{\Delta C_j}{\Delta Q_j}$
$\frac{\Delta K_j}{\Delta B_j}$