Unrelated Medical Costs in Life-Years Gained
Should They be Included in Economic Evaluations of Healthcare Interventions?

David R. Rappange,1 Pieter H.M. van Baal,2 N. Job A. van Exel,1 Talitha L. Feenstra,2 Frans F.H. Rutten1 and Werner B.F. Brouwer1

1 Department of Health Policy & Management and Institute for Medical Technology Assessment, Erasmus University Medical Center, Rotterdam, the Netherlands
2 Centre for Prevention and Health, Services Research, National Institute for Public Health and the Environment, Bilthoven, the Netherlands

Abstract

Which costs and benefits to consider in economic evaluations of healthcare interventions remains an area of much controversy. Unrelated medical costs in life-years gained is an important cost category that is normally ignored in economic evaluations, irrespective of the perspective chosen for the analysis. National guidelines for pharmacoeconomic research largely endorse this practice, either by explicitly requiring researchers to exclude these costs from the analysis or by leaving inclusion or exclusion up to the discretion of the analyst. However, the inclusion of unrelated medical costs in life-years gained appears to be gaining support in the literature.

This article provides an overview of the discussions to date. The inclusion of unrelated medical costs in life-years gained seems warranted, in terms of both optimality and internal and external consistency. We use an example of a smoking-cessation intervention to highlight the consequences of different practices of accounting for costs and effects in economic evaluations. Only inclusion of all costs and effects of unrelated medical care in life-years gained can be considered both internally and externally consistent. Including or excluding unrelated future medical costs may have important distributional consequences, especially for interventions that substantially increase length of life. Regarding practical objections against inclusion of future costs, it is important to note that it is becoming increasingly possible to accurately estimate unrelated medical costs in life-years gained. We therefore conclude that the inclusion of unrelated medical costs should become the new standard.
Economic evaluations are usually used to inform policy makers about the costs and benefits of a given change in resource allocation (e.g. introducing a particular healthcare intervention). The general idea behind such evaluations is that a particular intervention should only be introduced if the associated benefits are found to outweigh the associated costs. Therefore, it is obviously important to determine which costs and benefits should be included in an economic evaluation. The theoretical welfare economic answer is simple: if we really want to know whether the benefits outweigh the costs, we need to include all the costs and benefits, and neither exclude anything relevant nor count anything twice. However, in practice, the issue of what to consider in economic evaluations of health technologies has been and remains an area of much controversy. Some[1-3] promote a broad societal perspective that theoretically should comprise all relevant costs and effects. Others[4] suggest a narrower perspective, most notably a healthcare perspective, as being most relevant in the context of healthcare decision making. If the latter option is chosen, costs (and effects) falling outside of this perspective are systematically ignored and deemed irrelevant for the healthcare decision maker. For example, the inclusion or exclusion of productivity costs depends on the perspective chosen,[1,5,6] as they do not fall under the healthcare budget, but do represent real societal costs.

Unrelated future medical costs are a potentially important yet often ignored cost that is relevant in both these perspectives. These costs, also referred to as indirect medical costs or survivor medical costs, are an indirect result of an intervention that has successfully prolonged the life of an individual. During these added years of life, this individual, just like any other person, may fall ill and consume healthcare. This healthcare consumption may be termed either related or unrelated to the life-saving intervention. For example, Bob is a 60-year-old male with acute heart failure. Immediate treatment, involving bypass surgery, has prevented Bob from dying. He recovers completely, although he does require lifelong medication for his heart condition. Bob will now live on to the age of 80 years; however, at age 75 years, Bob trips over and breaks his hip. After a total hip replacement and much therapy, Bob is mobile again. When calculating the cost effectiveness of the bypass surgery that prolongs Bob’s life by 20 years, we normally take into account the related costs during these added life-years (i.e. the costs of the required medication). The unrelated medical costs are those related to the hip replacement. Should they be included in the analysis of the heart surgery as well?

The answer may be considered straightforward. Indeed, since all medical consumption during the gained life-years would not have occurred if the initial intervention (i.e. the bypass surgery) had not taken place, and since it represents actual medical resource use, the inclusion of these costs seems required in order to reach optimal and informed funding decisions. But while this inclusion of related costs in gained life-years is uncontroversial,[7] the inclusion of unrelated costs is still very contentious. Using the example above, some would argue that the bypass surgery should be judged ‘in isolation’ and that it can not be blamed for subsequent decisions or interventions such as hip replacement.[8] Others have argued that inclusion is irrelevant so long as the practice of inclusion or exclusion is consistently executed, since it would only entail adding a constant to cost-effectiveness ratios, which does not alter the relative efficiency of different interventions and, therefore, does not affect subsequent prioritization of interventions (or decision making).[9] However, arguments in favour of inclusion of these costs as being real healthcare costs are increasing, and practical ways to do so have been explored.[10-13] This controversy is persistent and many guidelines for economic evaluations of health technologies either instruct analysts to exclude these costs or
leave inclusion up to the discretion of the analyst. For instance, the Dutch guidelines for pharmacoeconomic evaluations explicitly state that unrelated medical costs in life-years gained should be excluded from the analysis. Indeed, excluding these unrelated medical costs is common practice in most economic evaluation studies.

However, recent arguments in favour of including such survivor costs of unrelated medical care appear to be gaining support. One of the main arguments in favour of including unrelated future medical costs has been labelled the internal consistency argument. This argues that what is being projected as gains (benefits) in an economic evaluation needs to be consistent with what is being counted as costs in that same evaluation. Since most projections of QALY gains (e.g. based on average healthy life expectancy in the population) implicitly assume normal medical care consumption during added life-years (without which the healthy life expectancy cannot be attained), it would be inconsistent to exclude the associated costs. For example, if we calculate the QALY gain for Bob due to the bypass surgery by using the actual predicted health level until the age of 80 years, achieving this health level requires the hip replacement. Projecting the gain due to the hip replacement but not the associated costs would thus be inconsistent. Another crucial argument, most forcefully put forward by Meltzer, is that ignoring future medical costs is not consistent with lifetime utility maximization. Thus, ignoring these costs does not result in optimal decision making.

Therefore, it appears timely to reconsider the current practice of ignoring unrelated medical costs. This will especially change the outcomes of economic evaluations of interventions that substantially prolong life, for example, in the curative sector (i.e. neonatal surgery) and in the preventive sector (i.e. reducing risk factors such as obesity or smoking).

We provide an overview of the literature so far and highlight the consequences of different practices of accounting for costs and effects in cost-effectiveness analyses, using a preventive intervention example based on a recent study by Van Baal et al. The structure of the paper is as follows. Section 1 provides a brief description of different cost categories that can be distinguished in economic evaluations; section 2 highlights the debate about unrelated future medical costs as well as some current recommendations regarding their inclusion in US, Dutch, UK and Swedish guidelines. Section 3 presents some recent developments in the literature and in section 4, we use an example of a smoking-cessation intervention to present four different cost-utility ratios to demonstrate the consequences of different practices of accounting for unrelated future medical costs. Finally, section 5 concludes this paper by drawing some lessons for the inclusion of these controversial costs in economic evaluations.

1. Economic Evaluations and Cost Categories

Economic evaluations in healthcare compare the costs and effects of a given medical intervention with the costs and effects of a relevant alternative (comparator). Health effects are usually valued in some common denominator (e.g. QALYs). The ratio of additional monetary costs to QALYs gained (i.e. the incremental cost-effectiveness ratio [ICER]) can then be calculated. Ideally, the ICER enables the investigator to judge whether the incremental health effects of the intervention justify its incremental monetary costs. In calculating costs, it is important to distinguish different types of costs. A common categorization of costs is shown in figure 1, which presents the different types that are directly required as input in the intervention (direct medical and non-
1.1 Direct Medical Costs

Direct medical costs are usually defined as those that directly relate to the intervention or condition under study and fall within the formal healthcare sector. They occur in normal years of life (i.e. the years that a patient would have lived with or without the treatment). All costs falling within the formal healthcare sector that are related to the intervention or condition (e.g. formal caregivers’ time, costs of diagnostic tests, drugs and other hospital materials) belong to this category independent of who is financing these costs. This cost category is a central part of any cost-effectiveness analysis, whether performed from a societal or a healthcare perspective.

1.2 Direct Non-Medical Costs

As figure 1 shows, the direct costs relate to resources that directly contribute to the intervention. Indirect costs (which can also be savings, e.g. when productivity changes) can best be viewed as consequences of the intervention and the related health improvement. Whether or not these different costs are deemed relevant in an intervention is partly determined by the given intervention (e.g. small health changes may not be associated with changes in productivity) and, importantly, influenced by the perspective chosen for the analysis.

A societal perspective implies the inclusion of all relevant costs and effects of a medical intervention regardless of whether they fall within or outside the healthcare sector, and, in principle, all these costs are relevant. When a healthcare perspective is used, the non-medical costs are normally deemed irrelevant. We discuss these different cost categories further in the following sections.

1.3 Indirect Non-Medical Costs

Indirect non-medical costs are societal costs or savings that occur as a result of changes in a patient’s living environment. From a societal perspective, these costs should be included in economic evaluations since they are a direct input into the total treatment. These types of costs can be considered irrelevant when a narrower perspective is adopted (e.g. the healthcare perspective, which is largely the reference case in the UK).

While we do not wish to fully address here which perspective is most appropriate, it is clear that some costs may be more relevant for healthcare decision makers than others. On the other hand, simply and completely leaving certain costs out of an economic evaluation may still be considered a short-sighted strategy. Those costs in normal life-years unrelated to the intervention are normally excluded from the analysis, since these costs are independent of changes in length and quality of life (QOL) due to the intervention under study. Therefore, these costs are generally the same for all treatment strategies and cancel out.
tient’s productivity level. These so-called productivity costs occur when an illness leads to absenteeism, presenteeism (reduced productivity at work), disability or premature death in a productive person, whether paid or unpaid. Other costs may include, for example, those that occur outside the healthcare sector, such as costs for special education or, in cases of addiction, police or legal costs. There is, at least from a societal perspective, general theoretical consensus that all costs belonging to this cost category should be included. However, there continues to be much discussion on how to estimate these costs.[1,6,23-30]

Another type of cost that may be considered as indirect non-medical costs is currently receiving considerable attention in the literature: costs of non-medical consumption (net of production) in life-years gained (e.g. costs of food, housing and clothing). From a full societal perspective, there appears no valid reason to exclude these future non-medical costs. However, inclusion is uncommon and exclusion has been advocated for feasibility reasons[7] and for consistency between the cost and effect sides of the cost-effectiveness ratio.[11] Consensus on how to handle these costs has yet to be achieved.[9-11,20]

From the healthcare perspective, indirect non-medical costs can be readily ignored (although the line between these costs and future medical costs is somewhat vague – i.e. the costs of food and housing of people who are institutionalized can be seen as ‘medical costs’).

1.4 Indirect Medical Costs

Indirect medical costs usually describe medical costs that result from care consumption in life-years gained. These costs are relevant when an intervention prolongs the life of a patient, and include costs that would not have occurred without the life-prolonging intervention. Moreover, these costs are relevant from both a societal and a healthcare perspective, as they involve real societal costs that fall under the healthcare budget. These costs are often further described as ‘related’ and ‘unrelated’ indirect medical costs.

Related indirect medical costs are related to the life-prolonging intervention. In our earlier example, these would be the costs of Bob’s heart condition medication. Another example would be the continued use of anti-rejection medication after a life-saving liver transplantation. This care consumption is directly related to the intervention that was required for the condition in the first place. The inclusion of these related indirect medical costs is normally advocated in an economic evaluation.[7]

Unrelated indirect medical costs are a result of consumption of medical care in added life-years, but are unrelated to the intervention that was required to treat the initial condition (e.g. the costs of Bob’s hip replacement). Obviously, these costs would not have occurred if the patient’s life had not been extended, but other than that, these medical consumption costs are unrelated to the intervention under study. The inclusion of these costs remains a matter of much debate,[13] and in the remainder of this paper we focus on the inclusion of these costs.4

2. Unrelated Indirect Medical Costs

The debate on the inclusion of unrelated medical costs in life-years gained in economic evaluations is not new. Weinstein et al.[31] argued that this cost category is often unjustly excluded, “… if treatment results in prolonged life because a condition has been cured or early disease has been avoided, then the cost of treating later disease that would not

---

4 It should be noted that the cost categorization outlined here is common but certainly not perfect or complete. For example, during life-years gained, people may also require informal care, which would be something like ‘indirect direct non-medical costs’, etc. It appears to becoming more common to use less aggregate cost categories (e.g. productivity costs, informal care, travel costs, etc.) with more meaningful labels.
otherwise have arisen must be considered.” In contrast, Russell[8] argues that an intervention should be judged on its own ‘merits’ if the aim of the evaluation is to establish whether an intervention produces good value for money. She claims that it would be incorrect to attribute additional costs in life-years gained to an intervention just because it is successful in prolonging a patient’s life. Russell[8] maintains that “When the question is … simply whether the proposed program is a good use of society’s resources, its indirect effects on medical expenditures are no more relevant than its indirect effects on expenditures for food, clothes, or housing.” Indeed, the latter type of costs (i.e. survivor consumption costs) is not usually included in economic evaluations; however, this is another matter that is also being fiercely debated.[18,32-36] In practice, the viewpoint of Russell appears to have received quite some support. For instance, Mushlin and Fintor[37] evaluated nine cost-effectiveness studies of breast cancer screening and found that none of the studies had included unrelated medical care in life-years gained. The authors concluded that the studies “… all avoided the nonsensical conclusion that it is almost always more cost effective to do nothing than to screen and attempt to cure.” It must be noted that while statements such as this are sometimes tempting, especially in the pursuit of getting effective screening programmes established, it is rather nonsensical to perform a cost-effectiveness analysis in which real costs are deliberately ignored, in an attempt to attain the desired results. Obviously, excluding certain costs while keeping the effects constant will result in more favourable cost-effective ratios, but then the ultimate strategy would be to ignore all costs. Rather than such nonsensical strategies, studies should be performed in a way that is methodologically sound and that results in relevant and complete information for the decision maker. But what should be the role of unrelated medical costs in gained life-years in that context? Garber and Phelps[9] tried to shed some light on this issue. They showed that, under certain strict assumptions, the inclusion or exclusion of survivor medical costs will not affect the ranking of cost-effectiveness ratios and that, therefore, they can be safely omitted from the analysis. They claimed that if the unrelated future healthcare costs are truly conditionally independent of prior expenditures (i.e. independent of the expenditures of the intervention under study) and the practice is consistently executed, including these costs will only add a constant figure to all cost-effectiveness ratios. Thus, although the cost-effectiveness ratio will increase, the ranking of the ratios remains unaltered.

However, some have questioned these results. First, the US Panel[2] argued that this would only hold if interventions were compared that were intended to treat people with similar personal characteristics, such as age. Interventions aimed at different age groups and thus adding life-years in different life phases would most likely entail different costs per additional life-year. In that case, the inclusion or exclusion of these age-specific future costs would indeed matter and, therefore, the ranking of the cost-effectiveness ratios could be influenced.[2] Second, the important assumption in the model of Garber and Phelps,[9] that future unrelated medical costs are costs that are conditionally independent of prior expenditures, would probably not hold in practice.[2,18] In this respect, the Panel[2] argued that “It is fair to ask whether the pattern of future expenditures is ever truly unaffected by an intervention that has a large impact on longevity.” Finally, related to this point, the Panel[2] claimed that if standard practice would be to exclude unrelated medical costs in life-years gained, it would seem almost impossible to achieve exclusion in a consistent and comprehensive manner since there “… are practical and conceptual problems in disentangling the ‘related’ and ‘unrelated’ components of costs for ‘related’ diseases …” Likewise, Weinstein and Manning[19] il-
illustrated the difficulty in separating unrelated costs from related costs: “In the analysis of a heart disease program, for example, one would need to include future induced costs that are affected (conditionally upon survival) by the intervention, but exclude all heart disease costs that are conditionally independent.”

Meltzer\textsuperscript{[18]} directly addressed the questions and doubts raised by the Garber and Phelps\textsuperscript{[9]} model. Using a more general model than Garber and Phelps, and relaxing some of their assumptions, he reached a completely different conclusion: that both the absolute and relative outcomes of cost-utility ratios of interventions are significantly altered when future unrelated medical costs are included and that their inclusion is required in order to reach optimal decisions.\textsuperscript{5}

In areas of so much debate, guidelines can play an important role in advocating uniform studies and harmonizing the applied methodology. However, guidelines in different jurisdictions differ remarkably in important aspects, including how they handle unrelated medical costs in gained life-years. We briefly describe how these costs are handled in guidelines from the US, Holland, UK and Sweden.\textsuperscript{[2,4,14,38]}

2.1 Guidelines for Pharmacoeconomic Research

The influential US guidelines\textsuperscript{[2]} most comprehensively addressed the issue of unrelated medical costs in gained life-years, but ultimately left the decision of whether to include them or not up to the discretion of the analyst. On one hand, the US Panel\textsuperscript{[2]} wrote that including these costs seems self-evident, since medical care consumption in added life-years is only possible because of the treatment under study; ignoring these future costs would not be an adequate reflection of an intervention’s true cost to society. However, on the other hand, the Panel pointed out the large amount of support for excluding these costs, reasoning that it seems politically inappropriate to consider unrelated medical consumption in life-years gained. They claimed that living longer entails additional consumption costs, including medical consumption. If these costs would be accounted for, then it would make sense to consider all consumption costs, which is uncommon. This reflects the argument of Russell.\textsuperscript{[8]}

Alongside these theoretical considerations, the US Panel\textsuperscript{[2]} also highlighted several practical issues surrounding the inclusion of unrelated medical costs in life-years gained. The first is the aforementioned difficulty of distinguishing between related and unrelated medical care, which makes complete exclusion of unrelated medical costs extremely difficult. Another practical issue that might hamper the inclusion of unrelated medical costs is the lack of comprehensive data of future unrelated medical care and the uncertainty of how to estimate the associated costs. Finally, the Panel argued that, if unrelated medical care costs are to be included, then so should all costs in added life-years, in line with Meltzer’s argument.\textsuperscript{[18]}

In short, the US Panel\textsuperscript{[2]} recommended omitting this cost category if it is small and the influence on the cost-effectiveness ratio can be considered negligible. However, if inclusion is expected to have a significant effect on the cost-effectiveness ratio, the Panel recommended that a sensitivity analysis should be conducted. The difficulty of reaching firm recommendations in this controversial area is illustrated by this recommendation within the US guidelines, which also prescribe the use of a broad societal perspective and argue for a clear reference case – a standard set of methodological practices. Leaving the decision of whether or not to include these costs

\textsuperscript{5} Note that Meltzer\textsuperscript{[18]} extends the discussion by also arguing in favour of including all future non-medical costs, such as survivor consumption and survivor earnings.
up to individual analysts can interfere with this societal perspective and lead to incomparability between studies if analysts make different decisions from each other.

Like the US guidelines, the Dutch pharmacoeconomic guidelines advocate use of a societal perspective. However, while this implies that the Dutch guidelines are inclusive regarding most types of costs, the guidelines explicitly state that unrelated indirect medical costs should be excluded from the analysis. It appears that the controversy in the literature as well as the aforementioned practical difficulties in recommending otherwise have probably resulted in this guidance, which may, nonetheless, be considered rather restrictive and disappointing in the advocated societal context.

The guidelines of the National Institute for Health and Clinical Excellence (NICE) adhere to a narrower perspective than that of the US and Dutch guidelines. NICE recommends the use of a healthcare budget allocation perspective, where only the costs that fall within the UK NHS and Personal Social Services are taken into consideration. In the Guide to Methods of Technology Appraisal that was published recently, it is explicitly stated that “Costs related to the condition of interest and incurred in additional years of life gained as a result of treatment should be included in the reference case analysis. Costs that are considered to be unrelated to the condition or technology of interest should be excluded,” while these costs clearly fall under the NHS budget.

In contrast, the Pharmaceutical Benefits Board (LFN) in Sweden has published guidelines in which the societal perspective and the inclusion of all relevant costs are advocated. In these guidelines, “All relevant costs associated with treatment and illness should be identified, quantified and evaluated. … If treatment affects survival, then the costs for increased survival – total consumption less total production during gained life years – should be included.”

Current guidelines thus largely encourage ignoring unrelated medical costs in gained life-years, either by requiring researchers to exclude these costs from the analysis or by leaving it up to the analyst to decide whether to include them. Both the theoretical controversy and the practical problems associated with including these costs seem to have contributed to this current situation. In terms of the latter, the practical issues may not be easy to solve. For example, Meltzer and Johannesson agree with the US Panel that the lack of adequate data to estimate future unrelated medical costs is still troublesome. However, they also claim that “In any case, it seems difficult to argue that including an implicit estimate of zero by omitting these costs would be preferable to an imprecise estimate, especially with appropriate sensitivity analysis.” Moreover, it is important to note that there is progress in this area. The recent literature on healthcare costs of aging, demonstrating that healthcare costs are not merely dependent on age and sex but importantly on time to death, can also be of use here (van Baal et al., unpublished data). In terms of the more theoretical debate on inclusion of unrelated future medical costs, new arguments in favour of including these costs appear to be gaining support, as will be highlighted in section 3.

3. Recent Developments in this Debate

The influential US guidelines have spurred debate in the literature regarding the inclusion of unrelated medical costs in gained life-years. For instance, Meltzer and Johannesson address the different arguments put forward by the US Panel for not firmly advocating inclusion of these costs. Re-
Regarding the more political argument (that it would not be acceptable to include these types of costs in economic evaluations) they acknowledged that it may seem that including unrelated medical costs will aggrieve the elderly, but argue that these costs are still real. Moreover, inclusion of these real costs does not automatically result in these interventions no longer being cost effective (unlike the suggestion of Mushlin and Fintor[37]). On the other hand, exclusion of these costs may lead to favouring interventions that extend life among the elderly, rather than improve quality. Furthermore, Meltzer and Johannesson[10] use a simple example to demonstrate that, even if the assumption of conditional independence holds (so that future spending, conditional on survival, is not influenced by current medical consumption) and interventions are aimed at patients of one single age (so that the stream of future costs is identical across interventions), excluding unrelated future costs still leads to biased outcomes in practice, simply because one QALY can be gained by prolonging life for 1 year in perfect health or for 2 years in a health state valued at 0.5 QALY, resulting in 1 or rather 2 years of unrelated future spending.

In a reaction to Meltzer and Johannesson,[10] Garber[44] argued that “…the Panel in effect recommended calculating such costs except when they are known to be small or equal among the various alternatives …” and therefore the criticism about the recommendation regarding unrelated medical costs seems only “… a matter of emphasis.” Finally, important arguments have been presented against the point raised by the Panel that if one includes unrelated medical costs in gained life-years, one should include all costs in gained life-years, including costs of other consumption, the inclusion of which would be in line with Meltzer’s argument.[18]

In a recent influential paper, Nyman[11] proposed a different set of inclusion principles that should serve as a practical guide for deciding which costs and effects to include in cost-effectiveness analysis. According to Nyman,[11] the set of criteria should result in cost-utility ratios that are – what has been labelled by van Baal et al.[13] – internally consistent, also sometimes referred to as a symmetry rule.[36] One of the crucial criteria put forward by Nyman is “Include in the analysis the costs of those resources that directly produce the utility that is being measured in the denominator of the cost utility ratio.”[8] Application of this criterion has important implications, particularly regarding the issue of how to account for unrelated medical costs in life-years gained. In respect of this, Nyman wrote, “The key, as Meltzer (1997) correctly points out, is recognising that these unrelated medical costs are not simply costs, but costs that are incurred to obtain an expected real benefit in the form of an increase in the probability of survival or an increase in health-related QOL. In many existing cost-utility analyses, these benefits are already accounted for in the QALYs, therefore, including their costs in the numerator would only be consistent.” Thus, application of Nyman’s criteria[11] leads to inclusion of unrelated medical costs in life-years gained whenever the associated health gain is captured as projected health gains.

The internal consistency argument put forward by Nyman[11] was subsequently used by van Baal et

---

7 We do not wish to dwell on this point here, but some have argued that since the utility measure (i.e. the QALY) in health economic evaluations is defined rather narrowly, one might argue that a more narrow consideration of costs would also be appropriate.[11,36]  

8 The two other criteria are “Exclude the costs of those resources that produce utility that is not being measured in the denominator, even though the costs are causally associated with the intervention” and “Include the costs of those resources consumed that are causally related to the intervention, but that have no countervailing utility gains.”[11]
al.\cite{13} to discuss its practical implications in the example of a cost-effectiveness analysis of a smoking-cessation programme. They estimated four cost-utility ratios, each differing as to what costs and effects were included. These four ratios were assessed not only in terms of internal consistency but also in terms of the implicit underlying objectives and budget responsibilities of the decision maker. van Baal et al.\cite{13} also addressed the difficulty of separating related and unrelated future costs and showed that the line between what is considered related and what is not becomes fuzzy and hazy, especially for primary prevention aimed at risk factors affecting many chronic diseases. More importantly, they pointed out that it is unclear why a distinction between related and unrelated medical costs would be relevant at all for a decision maker trying to make optimal use of scarce healthcare resources. In the following section, we outline the four different strategies used to account for costs and effects in the analyses used by van Baal et al.\cite{13} to highlight the importance of the current debates.

However, before doing so, a very recent paper deserves mention. Contrary to much of the recent literature, Lee\cite{20} argued against inclusion of all unrelated future costs. His article followed directly on from the theoretical work of Garber and Phelps\cite{9} and Meltzer.\cite{18} Using Meltzer’s model,\cite{18} but with a different budget constraint, Lee\cite{20} reached the conclusion that unrelated medical costs should not be accounted for in economic evaluations. The debate in this area will surely continue, if only because the budget constraint used by Lee,\cite{20} which is pivotal for reaching his conclusions, has been criticized before in the literature\cite{47} and basically assumed that prolonging life is always associated with increasing income. Thus it avoids allocation decisions related to (increased) scarcity – the very reason to perform economic evaluations.\cite{9} While this may be a reasonable assumption from the individual perspective, it is highly questionable whether it is also reasonable to assume this from a societal or healthcare perspective.

### 4. Different Cost-Utility Ratios: an Example

van Baal et al.\cite{13} demonstrated the consequences of different practices of including costs and effects using an example of increased implementation of a smoking-cessation intervention.\cite{48} They estimated the cost effectiveness of a smoking-cessation programme in two scenarios, which only differed with regard to the age groups to which the intervention was offered: in scenario 1, smokers aged between 25 and 44 years were targeted; in scenario 2, smokers aged between 45 and 64 years were targeted. In both intervention scenarios, 25% of the smokers received minimal counselling by a GP and/or a GP’s assistant in combination with nicotine replacement treatment for 1 year (see Silagy et al.\cite{49} for more details). For ease of interpretation, non-medical costs and savings were excluded from the cost-utility ratios, adopting a healthcare perspective.\cite{10} Related diseases were distinguished from unrelated diseases in that the intervention changed only the prognosis and/or the (age- and sex-specific) incidence rate of related diseases. This example considered a preventive intervention targeted at the risk factor ‘smoking’, thus only those diseases for which smoking was a risk (relative risk >1) were considered.

Table I presents the four cost-utility ratios related to the smoking-cessation intervention, as well as the appraisal by van Baal et al.\cite{13} The appraisal of each cost-utility ratio is based on whether the ratio com-

---

\footnotesize{
9 As was pointed out by one of the anonymous reviewers of our article, scarcity may also enter the equation via the use of utility functions including leisure.\cite{19}
10 As mentioned previously (section 1), the issue of inclusion of unrelated medical costs is largely independent of the perspective adopted.
}
Table I. Overview of different cost-utility ratios and results of smoking-cessation intervention scenarios 1 and 2

<table>
<thead>
<tr>
<th>Cost-utility ratio</th>
<th>What does it include?</th>
<th>Formula</th>
<th>Result (€ per QALY)</th>
<th>Appraisal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Only intervention costs divided by all QALY gains</td>
<td>$\frac{C_1}{QALY_r + QALY_u}$</td>
<td>3500</td>
<td>3400</td>
</tr>
<tr>
<td>2</td>
<td>Intervention costs and related costs only divided by all QALY gains</td>
<td>$\frac{C_1 + C_r}{QALY_r + QALY_u}$</td>
<td>1500</td>
<td>900</td>
</tr>
<tr>
<td>3</td>
<td>Intervention costs and related costs only divided by QALY gains only attributable to related costs</td>
<td>$\frac{C_1 + C_r}{QALY_r - QALY_u}$</td>
<td>2900</td>
<td>1900</td>
</tr>
<tr>
<td>4</td>
<td>Intervention costs, related costs, and unrelated costs divided by all QALY gains</td>
<td>$\frac{C_1 + C_r + C_u}{QALY_r + QALY_u}$</td>
<td>4400</td>
<td>6600</td>
</tr>
</tbody>
</table>

a Compared with current practice: a combination of all current initiatives to stop smoking and willpower alone.

$C_1$ = intervention costs; $C_r$ = healthcare costs of medical care of related diseases; $C_u$ = healthcare costs of medical care of unrelated diseases; $QALY_r$ = QALYs gained due to related diseases; $QALY_u$ = QALYs gained due to unrelated diseases; $S_1$ = intervention scenario 1; $S_2$ = intervention scenario 2.

Unrelated Medical Costs in Economic Evaluations

The first cost-utility ratio in table I is internally inconsistent because the numerator of the ratio only accounts for the intervention costs, while the projected effects in the denominator result from both related and unrelated medical care. Moreover, this ratio is incomplete and implies that the decision maker aims to maximize health effects given a disease-specific budget constraint that only includes the costs of the intervention. This limitation is clearly not very realistic, since the ratio ignores both related and unrelated future medical costs associated with the intervention, while these would normally be accounted for when such a disease-specific budget is applied.

The second cost-utility ratio is also internally inconsistent. Costs of related medical care are now included together with the intervention costs, but possible costs incurred due to unrelated medical care in added life-years are ignored. However, the denominator comprises all future effects, due to both related and unrelated medical care. This practice is currently implicitly recommended in the Dutch guidelines (as distinguishing effects from related care as opposed to unrelated care is uncommon and difficult). The fact that lower ratios were found for cost-utility ratio 2 compared with cost-utility ratio 1 can be explained by savings in terms of related costs ($C_r$) due to a lower incidence of smoking-related diseases (i.e. $C_r$ in cost-utility ratio 2 is negative). In terms of the underlying budget allocation problem, this ratio is still fairly limited since the available budget is only intended for care related to a specific disease, risk factor or preventive programme. Therefore, results from this ratio appear to be meaningful only for healthcare decision makers with a very narrow focus.

In cost-utility ratio 3, the costs of the intervention and related medical care are included in the numerator, while the denominator contains only the QALY gains produced by the costs listed in the numerator (i.e. effects caused by unrelated medical spending in gained life-years are subtracted from the total gains). In absence of empirical data regarding exact-
ly how unrelated medical care affects health, the assumption was made that 50% of the gains were attributable to related costs. This yielded cost-utility ratios of €2 900 and €1 900 per QALY gained for scenario 1 and 2, respectively.\textsuperscript{11} Cost-utility ratio 3 is internally consistent, since only those effects that are a direct consequence of the included costs are considered. However, this approach requires knowledge on the length and QOL of patient groups when not receiving any unrelated medical care during gained life-years. Such knowledge is hard to come by and its informative value for the decision maker has been questioned,\textsuperscript{13} although in some instances such a breakdown might still be interesting (e.g. in the case of disease-specific budgets). More specifically, excluding health gains due to unrelated medical care in gained life-years from the objective function of a decision maker seems rather nonsensical and, therefore, the underlying budget allocation problem of this ratio cannot be considered very meaningful.

Finally, the fourth cost-utility ratio is similar to the second ratio, except that now, besides the costs of the intervention and related medical care, the costs of unrelated medical care (Cu) are also included in the numerator. Although costs are saved because of a decrease in smoking-related diseases (so that Cr is negative), these savings are offset by Cu because of costs of unrelated medical care consumption in the gained life-years. As a consequence, the fourth cost-utility ratio yields the least favourable results. This ratio is not only internally consistent but also relates to a meaningful budget allocation problem since the denominator and objective function comprise the total health effects and the numerator and budget constraint consider all healthcare costs.\textsuperscript{13}

The four cost-utility ratios highlight the different options to deal with related and unrelated future medical costs. It is not surprising that including all costs and effects in the cost-utility ratios as compared with only some costs and all effects results in a higher estimate of the costs per QALY gained. The reason for this is simple. Living longer brings about competing diseases that result in healthcare expenditures (which in turn yield health). Thus, when unrelated medical costs in life-years gained are included in the analysis, interventions aimed primarily at improving QOL become relatively attractive compared with life-prolonging interventions. Moreover, the smoking-cessation example showed that treatment among the younger group of smokers proved more cost effective than treatment among the older group when costs of unrelated medical care were included. van Baal et al.\textsuperscript{13} argued that this can be explained because, for the younger smokers, “… the high costs of unrelated medical care occur farther away in the future and are, thus, more heavily discounted.”\textsuperscript{12} Besides this example of a smoking-cessation intervention, several other studies have demonstrated empirically that the cost-effectiveness ratios significantly change when the costs of unrelated medical care are included.\textsuperscript{50,51} Gyrd-Hansen et al.\textsuperscript{50} concluded that including the costs of unrelated medical care – as a function of age – favours intervention among relatively younger groups. These findings suggest that different practices of accounting for costs and effects lead to significant differences in cost-effectiveness ratios. If we assume that the results of such economic evaluations are consid-

\textsuperscript{11} Note that the percentage of QALYs gained attributable to unrelated medical care depends on both the intervention and population.\textsuperscript{13} Varying this percentage from 25% to 75% resulted in ratios ranging from €5 800 to €1 900 per QALY gained in scenario 1 and from €3 700 to €1 200 per QALY gained in scenario 2.

\textsuperscript{12} In the smoking-cessation example, discounting was applied according to the recommendations in the Dutch guidelines for pharmacoeconomic research;\textsuperscript{14} costs were discounted at a rate of 4%, while effects – in terms of QALYs – were discounted at 1.5%.
ered by healthcare policy makers in their decisions, the inclusion or exclusion of these costs may directly affect the allocation of healthcare resources.

5. Time to Change the Guidelines?

In summary of the above, we conclude that the dominant argument in the recent literature appears to be that unrelated future (medical) costs need to be included in economic evaluations of healthcare programmes. Therefore, it is immediately clear that most current economic evaluations do not give a full picture of the cost effectiveness of healthcare interventions. While the controversy is not (yet) completely resolved,\(^{20,45,46}\) it seems that a majority of recent papers\(^{10,11,13,46}\) argue in favour of including these costs. From a decision maker’s viewpoint, would it not make perfect sense to look at all costs brought about by an intervention, directly and indirectly, knowing that the projected health effects would not be realised without incurring these costs?

In that context, two types of consistency are important in economic evaluations: (i) internal consistency as put forward by Nyman;\(^{11}\) and (ii) what we will label *external consistency*. The former is clear: as long as the projections of gained QALYs (implicitly) incorporate normal medical care consumption during added life-years, it is inconsistent not to include the costs of this care consumption. While it is theoretically possible to produce cost-effectiveness estimates that exclude both the costs and the effects of this unrelated care, in practice it proves very difficult to accurately estimate health gains in absence of this unrelated care consumption. van Baal et al.\(^{13}\) had to make rather bold simplifying assumptions in order to estimate the health effects of the intervention without any further unrelated future care consumption. Equally important, the result of such an exercise is not very meaningful to decision makers.\(^{29}\) Therefore, the *external consistency* requires that the results of an economic evaluation address a meaningful decision-maker’s budget allocation problem (i.e. are consistent with the problem addressed). Economic evaluations can only serve their purpose if that requirement is met. As van Baal et al.\(^{13}\) argued, without using the term external consistency, only the fourth cost-utility ratio appears to satisfy both types of consistency.

We therefore claim that the preferred ratio is that which includes both the unrelated costs and the associated projected health gains. In terms of consequences, ignoring costs of unrelated medical care in life-years gained improves the cost-effectiveness ratio of interventions such as smoking-cessation programmes that increase length of life relative to interventions that primarily increase QOL. When interventions are judged on the basis of their ICER, and a more or less fixed threshold is used as benchmark, including more costs may then cause an intervention to become a less likely candidate for funding. Moreover, when economic evidence is used to formulate more precise medical practice guidelines, including the costs of unrelated medical care can lead to other treatment profiles and to including or excluding certain age groups from treatment, as the smoking-cessation example highlights.\(^{13}\) But note that excluding costs also has distributional consequences! Apart from this, systematic exclusion of unrelated medical costs does not necessarily enhance the comparability of economic evaluation results because “The costs to include in a CUA [cost-utility analysis] ... is often determined by the type of intervention and the target disease(s) of the intervention, because this determines what medical care is related,”\(^{13}\) which can result in strange and arbitrary lines between related and unrelated medical care. Therefore, the matter is not trivial and a reconsideration regarding the exclusion of these costs seems warranted.

In view of the current debates in the literature, we feel that a strong case can be made for including unrelated future medical costs in gained life-years. The major problem in advocating this may perhaps
be the practical issue of how to find reliable estimates of these costs. But not only do we agree with Meltzer and Johannesson\cite{10} that a reasonable although imprecise estimate is better than an unreasonable estimate of precisely zero, the progress in this area allows more reasonable estimates to be made. Gandjour and Lauterbach\cite{12} highlight the use of estimates of future medical care corrected for time to death in economic evaluations, and van Baal et al. (unpublished data) recently proposed a further refinement of these estimates. Moreover, a clear directive to start including these costs will undoubtedly increase developments in this area. In terms of standardization and facilitation of inclusion, national institutes such as NICE in the UK or the Dutch Health Care Insurance Board (CVZ) can play an important role. For example, the Swedish LFN offers analysts assistance in this respect. Therefore, it seems timely to move forward by including these costs, whenever relevant, in economic evaluations of healthcare technologies, and to include recommendations for their inclusion in guidelines for economic evaluations.

We acknowledge that this raises the question of where to draw the line. Unsurprisingly, the article by Nyman\cite{11} provoked a (new) round of debate regarding whether survivor non-medical costs (consumption and production) should also be included in an economic evaluation.\cite{32-35} If a healthcare perspective is taken, these consumption and production costs are deemed irrelevant by definition, while the inclusion of future medical costs would still very much be relevant. This makes it relatively easy to advocate inclusion of these costs (without further debate) from a healthcare perspective. However, from a societal perspective, matters are more complex. It has been convincingly argued that, from a welfare economic perspective, such survivor consumption costs should not be ignored.\cite{18} Others have argued that the choice for a narrow (i.e. health-related) outcome measure on the outcome side, would also limit the costs to be included on the cost side of an analysis.\cite{11,29,36} The adoption of a complete societal perspective therefore requires capturing all other relevant costs and effects separately, including broader utility gains. It seems crucial to recognize the tension between the perspective of the decision maker that the economic evaluations is designed to assist and the broad societal perspective. We would argue that, although the healthcare perspective may be a natural one from the viewpoint of a healthcare decision maker, “… completely losing sight of the societal perspective is undesirable as well.“\cite{13} Ignoring certain costs or effects can not only result in suboptimal allocation of resources, but it also seems unlikely that a healthcare decision maker would want to be left completely ignorant of the non-healthcare consequences of his/her spending. Lifting such ignorance requires the calculation and reporting of different costs, but does not necessarily require the equal treatment of all costs.\cite{13} Regardless, it seems inappropriate to exclude future medical costs from economic evaluations just because we have not yet resolved the question of whether non-medical costs should be included. This particularly holds since this controversy seems to be strongly related to the broader controversy regarding the appropriate perspective, which, one could say, is even at the heart of how we design our evaluations. For instance, even the choice of the narrow health-related utility outcomes appears a deliberate attempt to stay close to the perspective of the relevant decision maker.\cite{52}

6. Conclusions

The inclusion of unrelated medical costs in life-years gained in economic evaluations has long been

---

\textsuperscript{13} This issue is part of a broader discussion about the welfare theoretic foundation of QALYs, i.e. whether QALYs can be interpreted as utilities.\textsuperscript{16}
controversial. However, there appears to be growing support for incorporating all future medical costs. While some controversy may remain, it must be noted that this also holds for many other cost categories (e.g. productivity costs) and methodologies (e.g. discounting). If we only included those things for which a complete consensus exists, it is likely that no economic evaluation would be performed. Strong theoretical arguments, regarding both optimality,10,18 and internal and external consistency,11,13 all point towards the inclusion of unrelated future medical costs. The practical possibilities for estimating these costs appropriately are also increasing (van Baal et al., unpublished data),12 facilitating their inclusion.

This is not a trivial issue, as the impact of including these costs can be substantial. Furthermore, inclusion is consistent with both a healthcare and a societal perspective. Good practice in economic evaluations is essential in order to provide policy makers with valuable and meaningful information about the relative efficiency of health technologies. It seems timely that the inclusion of unrelated medical care in added life-years in economic evaluations, whenever relevant, should be the new standard. Guidelines should recommend including these costs rather than excluding them.

Acknowledgements

This study was part of the project ‘Living longer in good health’, which was financially supported by Netspar. The opinions expressed in the paper are those of the authors.

The authors have no conflicts of interest that are directly relevant to the content of this article.

References


34. Richardson JRJ, Olsen JA. In defence of societal sovereignty: a comment on Nyman ‘the inclusion of survivor consumption in CUA’. Health Econ 2006; 15: 311-3