THE RELEVANCE OF FUTURE, UNRELATED HEALTH COSTS IN ECONOMIC EVALUATION IN NICE APPRAISALS

REPORT BY THE DECISION SUPPORT UNIT

26 February 2020

Benjamin Kearns¹

¹School of Health and Related Research, University of Sheffield
ABOUT THE DECISION SUPPORT UNIT

The Decision Support Unit (DSU) External Assessment Centre is based at the University of Sheffield with members at York, Bristol, Leicester and the London School of Hygiene and Tropical Medicine. The DSU is commissioned by The National Institute for Health and Care Excellence (NICE) to provide a research and training resource to support the Institute’s Centre for Health Technology Evaluation Programmes. Please see our website for further information www.nicedsu.org.uk.

The production of this document was funded by the National Institute for Health and Care Excellence (NICE) through its Decision Support Unit. The views, and any errors or omissions, expressed in this document are of the authors only. NICE may take account of part or all of this document if it considers it appropriate, but it is not bound to do so.

This report should be referenced as follows:
EXECUTIVE SUMMARY

Life-extending treatments will result in additional medical expenditure during the period of extra survival. These additional costs can be categorised as either related or unrelated to the treatment of interest. Related costs would include the cost of the treatment as well as the costs of treating any adverse events. An example of unrelated costs is the cost of treating unrelated diseases. Currently the NICE methods guidance governing TA, HST and Medical technologies states that all unrelated costs should be excluded. As these unrelated costs represent real costs that would be incurred by a healthcare system, there is a strong argument that these costs should be included in technology appraisals. The aim of this report is to highlight the issues to consider if there was a change to the NICE guidance on unrelated medical costs in the future.

This report begins with an overview of current arguments in the literature, both for and against the inclusion of unrelated costs. The primary argument for including unrelated medical costs is that if they occur, they represent healthcare expenditure that cannot be spent to improve health elsewhere. Hence, to ensure decisions are made consistent with maximizing health, these unrelated costs should be included. In addition, it is also argued that unrelated health effects (on quality and quantity of life) are typically included in technology appraisals. As such, the inclusion of unrelated costs would be required to ensure a consistent approach. Arguments against the inclusion of unrelated costs appeal to practical and ethical issues. At a practical level, there is currently no standard approach (or reference case source) that can be used to generate estimates of unrelated costs. This raises the potential for the inconsistent inclusion of unrelated costs across appraisals. Ethical arguments question if it is acceptable to deny access to a life-extending treatment because other (unrelated) treatments that may be consumed in the period of additional life extension are themselves not cost-effective. Other arguments against the inclusion of unrelated costs note that they may increase health inequalities, and it is unclear how to define the boundary of which future medical costs should be included. Further, it is not clear if NICE appraisals consistently include unrelated health benefits, which weakens one of the arguments for including unrelated costs.

NICE guidance was compared to guidance provided by other HTA bodies. Including NICE, 40 guidelines were identified from HTA bodies across the world. Of these, only five explicitly
recommended the inclusion of unrelated costs. In contrast, 15 guidelines recommended that such costs be excluded. A further two guidelines noted the importance of being consistent in the decision to include or exclude unrelated costs without specifying if they should be included or not. In general, there was little justification provided for why unrelated costs should be included or excluded. Where this was discussed, the practical and ethical issues were provided to support exclusion. Of the five guidelines recommending the inclusion of unrelated costs, two also recommended using age-adjusted utility values (although this was not explicitly linked to the recommendation to include unrelated costs).

One of the five HTA bodies to recommend the inclusion of unrelated costs was the Dutch National Health Care Institute. To support this recommendation, an accompanying interactive toolkit for estimating unrelated medical costs is publically available. This toolkit allows users to specify the diseases of interest, and then calculates unrelated disease costs by age, sex, and time to death. There is currently no comparable UK-specific tool. Estimates of unrelated disease costs are typically derived from estimates of per-capita healthcare expenditure, adjusted for the costs of related diseases. A number of UK-based studies were identified which have estimated per-capita healthcare expenditure via a bottom-up costing approach. This requires combining data on expenditure from a number of different sources, such as hospital admissions, mental health services and primary care. A drawback of this bottom-up approach is that there is no guarantee that all relevant expenditure will be included. An alternative approach is top-down costing, where overall healthcare expenditure is broken-down by disease and patient characteristics. An approach using English data has been described, and has promise for use as a standard source of unrelated costs. The main limitation with this approach is that it does include the increase in healthcare expenditure near the end of life. Including this is important as end of life costs are significant, and are delayed by life-extending treatments.

If unrelated medical costs were included in NICE appraisals, then life-extending treatments would become more costly, and hence less cost-effective. The degree to which any life-extending treatment would be affected by this change will depend on the age-profile of the treated patients. Older patients incur higher mean healthcare expenditure, and so higher unrelated costs. Hence, life-extending treatments targeting the elderly would be particularly affected by the inclusion of unrelated medical costs in NICE appraisals. Appraisals that use a cost-consequence approach, such as those of medical technologies, will be particularly
affected by the inclusion of unrelated costs. In these situations, the more effective a treatment is at prolonging survival, the greater the degree of costs incurred which will not be offset by the health benefits of additional survival. In general, as healthcare expenditure is unequally distributed across the population, including unrelated costs may have equity issues for NICE appraisals. Their inclusion would also lead to a different threshold for cost-effective, although it is currently unclear what this threshold should be.

If unrelated costs were included in NICE appraisals, it would be important to ensure the consistent modelling of both these unrelated costs and also unrelated benefits. For example, end of life costs can be substantial. As they can confound the relationship between age and unrelated costs, it is important that they are included. Committees would also want assurances that the impact of unrelated treatments and diseases on both survival and health-related quality of life has been appropriately quantified within an appraisal.

There are a number of areas that require further work to provide consensus and guidance about if and how unrelated costs should be included in NICE appraisals. In particular, is it possible to develop a reference source of unrelated costs that is of sufficient quality for use in appraisals. Such a reference source would ideally follow the approach of the Netherlands toolkit in allowing estimates broken-down by age, sex and time to death, with costs for related diseases removed. There are also likely to be ethical objections to the inclusion of unrelated costs – a thorough consideration of these would be required, to provide assurances that ethical issues are being addressed. If unrelated costs are included, different thresholds for estimating cost-effectiveness would be required. Further, guidance on the inclusion of unrelated health benefits may also be necessary, to ensure consistency of approaches. Further, there may be treatments which do not extend life but do enable patients to subsequently receive unrelated treatments that they would not otherwise receive. Thought should be given to if these unrelated costs should also be included.

Life-extending treatments will result in unrelated costs during periods of extra survival. These costs are real costs, once incurred they cannot be spent elsewhere to improve health. As such, the inclusion of these unrelated costs in NICE appraisals is expected to lead to improved decision-making. There are however practical limitations and ethical concerns that impede the immediate adoption of unrelated costs. Future research should address these
limitations and concerns to support discussions about the inclusion or exclusion of unrelated costs in NICE appraisals.
# ABBREVIATIONS AND DEFINITIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CADTH</td>
<td>Canadian Agency for Drugs and Technologies in Health</td>
</tr>
<tr>
<td>CES</td>
<td>French Health Economists Association</td>
</tr>
<tr>
<td>CPRD</td>
<td>Clinical Practice Research Datalink</td>
</tr>
<tr>
<td>CUAs</td>
<td>Cost-utility analyses</td>
</tr>
<tr>
<td>DSU</td>
<td>Decision Support Unit</td>
</tr>
<tr>
<td>HAS</td>
<td>French Haute Autorité de santé</td>
</tr>
<tr>
<td>HES</td>
<td>Hospital episode statistics</td>
</tr>
<tr>
<td>HTAs</td>
<td>Health technology assessments</td>
</tr>
<tr>
<td>ICER</td>
<td>United States Institute for Clinical and Economic Review</td>
</tr>
<tr>
<td>ISPOR</td>
<td>International Society for Pharmacoeconomics and Outcomes Research</td>
</tr>
<tr>
<td>KCE</td>
<td>Belgian Health Care Knowledge Centre</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
<tr>
<td>NICE</td>
<td>National Institute for Health and Care Excellence</td>
</tr>
<tr>
<td>NOMA</td>
<td>Norwegian Medicines Agency</td>
</tr>
<tr>
<td>PAID</td>
<td>Practical Application to Include future Disease costs</td>
</tr>
<tr>
<td>PHARMAC</td>
<td>New Zealand Pharmaceutical Management Agency</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
<tr>
<td>ZIN</td>
<td>Dutch National Health Care Institute</td>
</tr>
</tbody>
</table>
2. INTRODUCTION

2.1. BACKGROUND

In health technology assessments (HTAs), future unrelated medical costs are costs which are incurred by a healthcare system, but which are not associated with treatment of the disease of interest (1, 2). For example, in an appraisal of a cancer treatment, the costs of treating hypertension would be viewed as unrelated (assuming that hypertension is not an adverse event). These unrelated costs stem from both an unrelated disease (in the example, hypertension), and the unrelated treatments associated with the unrelated disease. When comparing two or more treatments, if there are no differences in survival it is usual to assume that all unrelated costs will be the same for each treatment, and so can be ignored, as they will not affect estimates of cost-effectiveness. Hence, there may also be unrelated costs due to the disease of interest, if these costs are not altered by the treatment of interest. If instead there are differences in survival between treatments then the life-extending treatment may be associated with additional future unrelated healthcare costs. These are resources that would be used for treating other conditions that arise during the extended survival period. For example, a “cure” for an otherwise fatal childhood condition will see survivors consume healthcare resources that they otherwise would not have across the rest of their lives.

Current guidance from the National Institute for Health and Care Excellence (NICE), formalised in their 2013 guide to the methods of technology appraisal, is to exclude these unrelated costs (3). Section 5.5.7 of the NICE reference case states:

“Costs that are considered to be unrelated to the condition or technology of interest should be excluded.”

The question of whether unrelated costs should be included in HTAs has been debated for a long-time in the literature, with no consensus (4-9). This is reflected in a 2018 European Delphi exercise on the use of costs in health economic evaluations (10). This exercise included 26 experts from 17 countries. Across two Delphi rounds there was a lack of consensus on whether future medical costs should be included, with 53% and 50% preferring their inclusion in the first and second rounds, respectively. A lack of consensus was also cited as a reason for the influential first Panel on Cost-Effectiveness in Health and Medicine not recommending the inclusion of unrelated costs in 1996 (11). Recommendations of the second panel, published in 2018, do explicitly include unrelated costs (12).
This review

The aim of this report was to inform potential future changes to the NICE guide to the methods of technology appraisal. This report had the following objectives:

• Identify current arguments in the literature for and against the inclusion of unrelated costs.

• Provide an overview of how HTA bodies other than NICE handle unrelated costs in appraisals.

• Describe the methods available for estimating unrelated costs and including them within appraisals.

• Discuss the potential implications of including or excluding unrelated costs within NICE appraisals.

Unrelated costs may be categorised as either unrelated medical costs (costs borne by healthcare systems), or unrelated non-medical costs (productivity and non-medical consumption). The focus of this report is unrelated medical costs, also referred to as unrelated healthcare costs or indirect medical costs(6, 13). Unrelated non-medical costs are not considered, as the cost perspective of this report is the same as that specified in the NICE reference case: the National Health Service (NHS) and personal social services. For brevity, future unrelated medical costs are referred to as “unrelated costs” in this report. Where it is important to distinguish between unrelated medical and non-medical costs, this shall be explicitly stated.
3. EXISTING LITERATURE ON UNRELATED COSTS

3.1. ARGUMENTS FOR THE INCLUSION OF UNRELATED COSTS

There are two main arguments for the inclusion of unrelated costs in HTA. One is internal consistency; if unrelated health effects are included, then unrelated healthcare costs should also be included. The second, more important, argument is that the inclusion of unrelated outcomes (health effects and costs) will result in more accurate estimates of the true costs of a treatment, and hence more accurate decisions than the exclusion of unrelated outcomes. This second argument is also referred to as that of external consistency (14). As an illustrative example, re-consider the example of a cancer treatment provided in the introduction. Suppose that without the cancer treatment patients would live on average to the age of 60, whilst with the treatment they live on average to the age of 65. During these five years of extra survival, the direct costs associated with cancer treatment are included in a HTA. Some patients would also develop (or already have) hypertension during the years of extra survival; If during these five years the individual develops hypertension, the resulting (unrelated) treatment costs for this unrelated disease are typically not included in a HTA. To exclude these unrelated costs ignores the fact that, under a fixed healthcare system, the unrelated costs cannot be spent elsewhere to produce health benefits. Therefore, unrelated costs have real opportunity costs, and ignoring these will lead to sub-optimal funding decisions for a fixed budget (1, 2). The two arguments of internal and external consistency are discussed in more detail.

Unrelated costs are incurred by the (unrelated) treatment of unrelated diseases. These unrelated treatments and unrelated diseases will have unrelated health effects. One argument is that these unrelated health effects are often included in HTAs (1, 5). Examples include:

- Basing estimates of long-term survival on national life-tables or disease registries. In either case, the survival experience will be influenced by any unrelated treatments received by the population.
- Assuming that utility values are the same as the general-population. As with survival, the utility of the general population will be affected by unrelated diseases and treatments.
Applying age-related disutilities. These disutilities may be applied to baseline utilities, which may come from the general population or be disease-specific. The disutilities are not solely due to the process of ageing; instead, they will be influenced by the presence of morbidities, which increase as people age.

To include unrelated health effects but not unrelated healthcare costs is inconsistent, and will make life-extending treatments appear more cost-effective than they actually are. This intuitive argument was formalized by Nyman, who suggested that costs should be “included if they represent resources that directly produce the utility that is being measured in the denominator of the cost–utility ratio” (15). The internal consistency argument will also be satisfied by not including unrelated health effects. However, estimating only direct (related) health effects may be difficult, as it would require knowledge on long-term survival and utility values in the absence of any unrelated diseases or treatments. More importantly, the consistent exclusion of unrelated outcomes would lead to an incomplete estimate of the opportunity cost associated with a life-extending treatment (1). During the years of additional life, any unrelated treatments received by a patient will require funding by the healthcare system. With a fixed budget, these funds cannot then be allocated elsewhere. Therefore, if the purpose of HTA is to aid with decisions about how best to allocate a fixed budget to maximise patient outcomes, all unrelated outcomes should be included. Following the approach of Morton and colleagues, the potential options for including or excluding unrelated outcomes is summarised in Table 1 (5). Arguments in favour of including unrelated costs contend that the current approach in HTA is typically quadrant C; where unrelated health effects are modelled, but not unrelated costs. This approach lacks internal consistency (as would quadrant B, although there is no evidence that appraisals take this approach). Quadrants A and D are both internally consistent; of the two it is argued that use of quadrant A (including all unrelated outcomes) is preferred. This is because it is likely to be more feasible to obtain the required evidence, and it is externally consistent, so more likely to result in improved decision-making.
Table 1: Approaches to the inclusion or exclusion of unrelated costs and health benefits

<table>
<thead>
<tr>
<th>Unrelated costs</th>
<th>Unrelated health effects</th>
<th>Included</th>
<th>Excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Included</td>
<td>A</td>
<td></td>
<td>B</td>
</tr>
<tr>
<td>Excluded</td>
<td>C</td>
<td></td>
<td>D</td>
</tr>
</tbody>
</table>

There have also been a number of theoretical studies, which attempt to quantify the impact of including or excluding unrelated costs on optimal decision-making. Results from these studies are mixed, although an overview of these theoretical discussions suggests that studies that recommend excluding unrelated costs had significant flaws in their assumptions and so, on a theoretical level, unrelated costs should be included(1). This finding is supported by a recent study that considers HTA from a welfare-economic perspective(16). Results from the theoretical literature may be of limited practical relevance, as they depend strongly on the assumptions used and their relevance to applied appraisals is unclear. For example, Van Baal and colleagues note that in the theoretical literature, unrelated costs are assumed to depend solely on age and nothing else (conditional on survival), but that this assumption is typically violated as costs are known to vary based on proximity to death(17).

2.2 ARGUMENTS AGAINST THE INCLUSION OF UNRELATED COSTS

The two main arguments against including unrelated costs are first practical - how to adequately estimate unrelated outcomes - and secondly ethical - should a treatment be penalised for keeping people alive longer?

The primary practical issue is how to estimate consistently unrelated costs and health effects. Ideally, estimates of unrelated cost should be granular enough to be able to identify and remove costs due to the related disease(s), to avoid double counting. It should also be able to separate out unrelated costs due to end-of-life care (decedent costs) from other costs (survivor costs), as it is known that health-care expenditure increases with proximity to death(18). Failure to account for either double counting or decedent costs will lead to biased estimates of unrelated costs(17). In addition, different HTA decision problems will have
different patient case-mix (such as different age distributions, and male to female ratios), so for accuracy the estimation of unrelated costs should be available separately for key patient characteristics, such as age and sex, and potentially the number of comorbidities (19, 20).

A further practical issue is how to estimate unrelated costs due to unrelated comorbidities. Unrelated comorbidities are defined as morbidities which are more common amongst people with the disease of interest when compared to the general population, but which are not altered by the treatment of interest. For example, suppose that a treatment influences cancer survival, but does not affect the prevalence of stroke amongst people with cancer. The costs of stroke are unrelated as during the period without life-extension there is no cost difference between treatments. During the period of life-extension these costs would need estimating if unrelated costs are included. This estimation would need to reflect that the prevalence of stroke in the United Kingdom (UK) is between three to five times more common amongst people with cancer compared to the U.K. general population (21-23). Accounting for unrelated comorbidities could be achieved by weighting per capita healthcare expenditure for these diseases by their prevalence amongst the disease of interest, although this would require estimates of disease-specific per capita costs. For accurate estimates, all unrelated comorbidities should be included when estimating unrelated medical costs. If the unrelated costs are calculated separately for different combinations of patient characteristics (such as by age, sex, and time to death), then this would require information on the prevalence of comorbidities for each combination, which may be prohibitively time-consuming to obtain and incorporate within a health economic model. One option would be to model these comorbidities as related diseases. This removes the need to weight per capita costs by the prevalence of the comorbidities, but it will introduce additional decision uncertainty (24). It will not address the issue of how to identify which diseases should be classified as co-morbidities, and if this classification should vary by combinations of patient characteristics.

Ideally, there would be a single reference case source for unrelated costs, to avoid the potential for inconsistent estimates across appraisals and for reasons of efficiency. A lack of consistency could arise from the use of different data sources to estimate unrelated costs, or the inclusion of different diseases. In either situation, there is the possibility that cost-effectiveness results would be sensitive to the choice taken, with a lack of comparability across HTAs. For the same reasons, the inclusion of unrelated health effects should also be consistent. For utilities this suggests estimating the impact of unrelated diseases and
treatments on utility from age-related decrements from the general population (using a standard reference source), and applying these to utility values for the population of interest. Good-practice on using utility data in HTA is provided by the Decision Support Unit (DSU) utilities technical support documents series and an International Society for Pharmacoeconomics and Outcomes Research (ISPOR) task force report(25, 26), both of which recommend age-adjustment of utilities. The extent to which age-adjusted utilities are used in NICE appraisals is unclear. One review found that this adjustment was only performed in six of 71 submissions (8.5%). However, this review only focused on NICE appraisals between July 2005 and June 2008, so the incorporation of age-adjusted utilities in more recent appraisals is unknown. For survival, a similar approach to that for utility values could be used. This would combine general population life-tables (to capture survival due to unrelated diseases and treatments) with disease-specific survival. However, there is no consensus on how to combine survival data; a recent review article provides suggestions, but no definitive approach(27). This review article also noted that, of 214 UK HTAs published between January 2004 and October 2010, only 24 (11%) incorporated data that was external to the clinical trial - suggesting that any unrelated survival benefits may not be routinely captured in HTAs. As a further practical concern, within a health economic model it may not always be possible to prospectively identify periods of life-extension. If so, unrelated costs would be included during periods when there is no survival gain. This will lead to increased decision uncertainty and so increase the probability of making the incorrect decision(11, 24).

The primary motivation for ethical objections to the inclusion of unrelated costs is that life-extending treatments are no longer judged on their own merits. Instead, estimates of cost-effectiveness are dependent upon the cost-effectiveness of any unrelated treatments. A stark example is that of treatments for patients with chronic kidney disease who require dialysis, where the novel treatment of interest, such as phosphate binders, would not affect rates of dialysis in the absence of a survival benefit(4). The dialysis costs may then be viewed as unrelated. Dialysis is costly, and including this unrelated cost for the additional survival period means that the treatment of interest would appear less cost effective than would be the case had they been excluded from the analysis. At the extreme, it is possible that the new technology would not be considered cost effective even if it commanded a zero price (7). The implication is that the novel treatment, despite being clinically effective and offering an extension to life, would not be funded because an unrelated treatment is itself not cost-
effective at a traditional cost-effectiveness willingness-to-pay threshold (such as the range £20,000 to £30,000 employed by NICE). This is ethically difficult as the implication is that life-extending treatments would not be funded because the additional survival in these patients is not cost-effective given the additional health resources required when compared to the benefits that would be generated by other patients with those resources. An alternative approach would be to note that there has already been a commitment to fund dialysis, despite it not being cost-effective at a conventional willingness-to-pay threshold. Hence, if the unrelated costs of dialysis are included, this suggests that during the phase of life-extension a higher weight should be given to health benefits. Identifying what these weights should be is not obvious. A potential approach is to assume that because dialysis is already funded, its inclusion should not in itself alter any estimates of cost-effectiveness. This approach proposes that, in the case of dialysis, unrelated costs can be excluded. However, to ensure consistency it would be important also to exclude any unrelated health effects due to dialysis; identifying these may not be straightforward. A different approach would be to review the decision to fund dialysis. For example, observational evidence has suggested that not everyone will benefit from dialysis, and within the UK there is an on-going trial comparing dialysis with responsive management amongst the frail elderly (ISRCTN17133653). However, if it decided that the unrelated treatment is not cost-effective, withdrawing it when it is already available as current practice may be challenging.

There are also some more subtle points associated with the inclusion of unrelated costs. Where treatments offer a survival benefit, the NICE reference case stipulates that a lifetime time horizon is appropriate to ensure that all relevant differences in outcomes and costs are included. Hence, if females receive a life-extending treatment, then all of the unrelated costs that they would incur during the remainder of their life should be included. Should these unrelated costs include costs associated with giving birth? These are real costs incurred by the healthcare system, so based on the argument of including all opportunity costs to make optimal funding decisions (which is a key argument in favour of including unrelated costs; see the previous Section) they should be included. However, if the costs are included, should not the costs and health effects of any children (and their subsequent lives, which could also include further births) also be included? This difficulty in defining the scope of unrelated costs has been labelled by Culyer as the “ad nauseam problem of infinite regress” (30), with the author further noting:
“Indeed, if all unrelated costs are, as a matter of principle, to be considered, no matter how remote in time or geography, then there appears to be no limit to the costs to be considered other than those imposed by discounting and time preference on the one hand and the sheer cost of making the necessary estimates on the other”

Under this view, to account fully for all the relevant unrelated costs and health would be unfeasible. Yet to exclude the unrelated costs and health outcomes of any future children raises questions about whether any other unrelated costs and health outcomes should be also excluded. For example, delaying a death means that carers (whose health-related quality of life is included in the NICE reference case perspective on outcomes) would be able to spend additional time with a loved one, delaying the utility loss of bereavement. Ultimately, a pragmatic decision about the boundaries of unrelated costs and health effects would be required(30).

The inclusion of unrelated costs may also have equity impacts. The estimate of unrelated costs is known to vary by socio-demographic factors such as age and sex(31), so “the impact of including unrelated medical costs will be unequally distributed across the population”(32). Ideally, this impact would be estimated prior to implementation, so that decision makers can identify if including unrelated costs would lead to an increase of inequalities, and if so what steps could be taken to mitigate against these. This is of particular importance for NICE appraisals, as a guiding NICE principle is to reduce health inequalities(33).

2.3 SUMMARY

Arguments in favour of including unrelated costs in HTAs have been described, as have arguments against their inclusion. There are arguments of both internal and external validity in favour of including unrelated costs. There are also strong practical and ethical objections. Collectively, these arguments demonstrate a lack of consensus in the academic literature about the inclusion of unrelated costs. The next section considers if there is any consensus in the guidelines provided by HTA bodies.
4. GUIDELINES ON UNRELATED COSTS

4.1. HTA GUIDELINES

Guidance from NICE explicitly recommends the exclusion of unrelated medical costs in HTAs. This guidance was compared with the recommendations of other bodies or organisations involved in HTA. As a starting point, the guidelines available via the ISPOR tool were used, where English language versions could be obtained(34). This resulted in 34 guidelines (including NICE guidance). These were supplemented by six additional guidelines; see Appendix A.1 for full details on these guidelines. These guidelines varied from official recommendations to support country-specific HTA submissions (such as reference cases), to position statements published by experts. There was a lack of consistency in the terminology used in the guidelines evaluated. As unrelated costs are not directly related to the treatment of interest, they are a type of indirect healthcare cost. However, many guidelines did not distinguish between indirect healthcare and indirect non-healthcare costs. In these situations, the term “indirect costs” was typically used to refer to non-healthcare costs, mainly productivity gains and losses(35). If a guideline gave a recommendation on indirect costs, but only discussed non-healthcare indirect costs, it was assumed that the guideline did not make any specific recommendation on unrelated medical costs.

Of the 40 guidelines, five (12.5%) explicitly recommended the inclusion of unrelated medical costs and 15 (37.5%) recommended their exclusion. A further two guidelines were not prescriptive, but stressed the importance of being consistent in the approach taken. An overview of these guidelines is provided in Table 2.
## Table 2: Overview of recommendations on unrelated costs in HTA guidelines.

<table>
<thead>
<tr>
<th>Guidelines recommending inclusion</th>
<th>Guidelines focusing on consistency</th>
<th>Guidelines recommending exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Dutch National Health Care Institute (ZIN)</td>
<td>The Canadian Agency for Drugs and Technologies in Health (CADTH)</td>
<td>NICE guide to the methods of technology appraisal 2013</td>
</tr>
<tr>
<td>The state of Israel Ministry of Health</td>
<td>The World Health Organisation (WHO)</td>
<td>The All Wales Medicine Strategy Group</td>
</tr>
<tr>
<td>The Medical Association of Thailand.</td>
<td></td>
<td>The Belgian Health Care Knowledge Centre (KCE)</td>
</tr>
<tr>
<td>The Second Panel on Cost-Effectiveness in Health and Medicine (Second panel)</td>
<td>The French Health Economists Association (Collège des Économistes de la Santé; CES)</td>
<td></td>
</tr>
<tr>
<td>The United States Institute for Clinical and Economic Review (ICER)</td>
<td>The French Haute Autorité de santé (HAS)</td>
<td>The German Institute for Quality and Efficiency in Health Care (IQWiG)</td>
</tr>
<tr>
<td></td>
<td>The Norwegian Medicines Agency (NOMA)</td>
<td>The Portuguese Ministry of Health</td>
</tr>
<tr>
<td></td>
<td>The Czech State Institute for Drug Control</td>
<td>The Hungarian Institute of Pharmacy and Nutrition</td>
</tr>
<tr>
<td></td>
<td>The New Zealand Pharmaceutical Management Agency (PHARMAC)</td>
<td>The Taiwan Society for Pharmacoeconomics and Outcomes Research</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The Japanese Central Social Insurance Medical Council</td>
</tr>
<tr>
<td></td>
<td></td>
<td>World Health Organisation guide for standardisation of economic evaluations of immunization programmes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The Scottish Medicines Consortium</td>
</tr>
</tbody>
</table>
Of the five guidelines that recommended inclusion, three were official guidelines supporting HTA submissions published by: ZIN, the state of Israel Ministry of Health, and the Medical Association of Thailand. The remaining two guidelines were published by the Second panel, and the United States ICER. ICER guidance notes that the inclusion of unrelated medical costs could result in a clinically effective treatment not being cost-effective even if it has a zero price. In this situation it is recommended that a scenario analysis be conducted which removes unrelated costs. The Second panel noted that their recommendation to include unrelated costs was a change from the First panel, which suggested analysts use their discretion due to a lack of consensus amongst the panel members. This previous lack of consensus is attributed to the lack of a theoretical rationale. Of the three official guidelines, the ZIN provides a supplementary costing manual(36), along with a tool for calculating unrelated medical costs (discussed in the following Section of this report). The authors note that “it cannot be predicted which unrelated diseases a patient will suffer from in the future” so recommended deriving future unrelated costs from per capita healthcare use, not costs per patient.

Of these five guidelines, only two provided an explicit recommendation to model utilities by age and sex. The Israel Ministry of Health advice that utilities “should be age and gender-adjusted”, with adjustment values provided by age (single years between 0 and 99 inclusive) and sex (male, female, either). Whilst not stated, these adjustment values appear to be multiplicative disutilities. The First panel recommends using age and sex specific estimates of utility. It further notes that variation in cost-effectiveness by sociodemographic demographics “may be ethically problematical”, and so recommends using sensitivity analyses to identify the drivers of cost-effectiveness results. The remaining three guidelines recommended the inclusion of unrelated costs, but did not discuss capturing the effects of unrelated diseases and treatments on utility values. None of the guidelines discussed capturing the effects of unrelated diseases and treatments on survival. Guidance on the implementation of unrelated costs could not be found for any of the other guidelines.

Guidelines from both CADTH and WHO highlight the lack of consensus on whether unrelated medical costs should be included. As such, they emphasise the importance of taking a consistent approach to the inclusion or exclusion of unrelated costs. In addition, CADTH guidance recommends a scenario analysis which includes unrelated costs and unrelated health benefits.
Table 2 lists the 15 guidelines that recommend exclusion of unrelated medical costs. Of these, all bar one explicitly discuss unrelated costs. The one exception is the Scottish Medicines Consortium which states that the NICE reference case “is appropriate for use in a submission”. Hence, it is assumed to implicitly recommend the exclusion of unrelated costs. The inclusion of unrelated costs in scenario analyses was mentioned as an option in five guidelines (PHARMAC, KCE, IQWiG, Czech, Hungary).

The majority of guidelines do not provide an explicit rationale for the exclusion of unrelated costs. A notable exception is the PHARMAC guidance, which highlights the practical issues in estimating both unrelated costs and benefits, as well as the ethical implications of penalising life-extending interventions in cost-utility analyses (CUAs):

“A key concern with including these costs in CUAs is that it would result in life-saving (or life-extending) treatments potentially being less cost-effective, hence biasing against those treatments that extend life. This is a particularly important issue when CUA results are used in the relative setting (ie where life-saving treatments need to be directly compared with treatments that improve quality of life). These costs are also very difficult to calculate and are associated with a significant amount of uncertainty. In most cases, limited data are available on these costs, and obtaining data may be time consuming. Further, future interventions may also be associated with health gains that would need to be taken into account in the analysis, significantly increasing the complexity of the analysis (and hence risk of error).”

The unresolved practical and ethical issues arising due to the inclusion of unrelated costs is also cited by the WHO as a reason for excluding unrelated costs in economic evaluations of immunization programmes. The Portugal guidance also justifies the exclusion of unrelated costs due to the difficulty in identifying them “…it is not, therefore, possible to relate the subsequent use of care with the treatment”. A different type of justification was provided by CES, who took a wider societal perspective and assume that unrelated medical costs would be offset by productivity gains during the years of additional life. Hence, they could be excluded unless there were evidence to the contrary:
“The following point is currently under debate: should not the future health care spending of survivors be deducted from the income expected from survival? The proposition seems logical. And yet, in a universal health insurance system within which intergenerational solidarity operates, it may be assumed that the discounted amount of lifelong contributions is equal to the discounted amount of health care spending. The future health care spending of survivors would only be taken into account if this assumption could not be verified.”

4.2. OTHER GUIDELINES RELEVANT TO ENGLAND

Whilst not specific to HTA, the HM Treasury is the British government department responsible for economic policy. It has published a guidance document, known as the Green Book(37). This does not explicitly discuss related or unrelated costs. Instead, it states that, where possible “All relevant costs and benefits which may arise from an intervention should be valued and included”. The definition of what is relevant is not explicitly provided, and so is open to interpretation. Those in favour of including unrelated costs have argued that these are relevant costs(5), whilst those who advocate for their exclusion do so on the basis that they are not relevant(4).

4.3. SUMMARY

Amongst the identified guidelines, recommendations to exclude unrelated costs were three times more common than recommendations to include. However, the majority of the guidelines either did not explicitly discuss unrelated costs, or did not provide a justification for their recommended approach. Reasons for exclusion typically focused on both the practical and ethical issues involved with the appropriate inclusion of unrelated costs. Possible solutions to these are discussed in the next two sections.
5. METHODS FOR INCORPORATING UNRELATED COSTS

Whilst there are compelling arguments for the inclusion of unrelated costs within technology appraisals, a practical barrier is being able to quantify these costs. This section discusses existing approaches. It begins with a discussion of a Dutch costing tool that has been developed specifically to estimate unrelated future costs for use in economic evaluations. Other approaches to estimating unrelated costs are then presented and discussed. For these, there is a focus on English examples, as these will be the most relevant to future NICE appraisals. Examples from other countries are also provided where they help to highlight key points.

5.1. EXISTING METHODS TO SUPPORT APPRAISALS

As noted in the previous section, Dutch guidance from ZIN explicitly recommends the inclusion of unrelated costs. To support this, the Erasmus School of Health Policy and Management, in collaboration with the Dutch National Institute for Public Health and the Environment, has developed a toolkit which estimates per capita annual health care expenditure(20). The toolkit, known as the ‘Practical Application to Include future Disease costs’ (PAID) estimates unrelated disease costs and provides results by age and sex, and are provided separately for the last year of life and other years. It is possible to define the related diseases of interest, to ensure that they are not included in the estimates of unrelated disease cost (and hence avoiding double counting). Examples of economic evaluations that have used PAID are:

1. A childhood influenza vaccination program(38).
2. Treatment of chronic heart failure(39).
3. Treatment of relapsing-remitting multiple sclerosis(40).

The impact of including unrelated costs varied in these three examples. For the first, the vaccination program changed from being cheaper (than no vaccination program) to being more expensive. For the second example, the incremental cost-effectiveness ratio increased by approximately 50%. In the third example, the inclusion of indirect costs had minimal impact on estimates of cost-effectiveness.
The general idea behind PAID is that of a bottom-up costing. Evidence is obtained on the annual per capita costs of disease, for 99 different diseases (in the current version, PAID 3.0). The current version of PAID is available as an online interactive application. Previous versions used Excel. As mentioned, these costs are stratified by the patient characteristics of age (single year, ages 0 to 99 inclusive) and sex, as well as proximity to death. Costs are also estimated separately by type of health care provider (hospital, nursing and residential care, other providers of medical goods including retail, ambulatory care, and ‘other’ health care providers). As such, PAID offers very detailed granular evidence on how and where healthcare costs are incurred. A user may then define the diseases and healthcare providers that contribute to unrelated costs for their decision problem, and obtain per capita costs for use in their economic evaluation. Two limitations acknowledged by the PAID authors are first that the modelling of decedent costs could be more granular (using more than one time period), and secondly that cross-sectional data are being interpreted as longitudinal data. The methodological framework for PAID is discussed in Appendix A.2.

From the perspective of a NICE appraisal, it would be possible to convert the estimates of unrelated costs from PAID into English costs, for example using purchasing power parities. However, this approach would be potentially limited as it is unclear how generalisable costs are across countries. Ideally, country-specific versions of the evidence used in PAID would be available. In the absence of this, other approaches to estimating unrelated costs are described.

5.2. UK-SPECIFIC APPROACHES

Within the UK, detailed evidence on hospital resource use is available from hospital episode statistics (HES) data. This data may be combined with unit costs (healthcare resource groups, which collates hospital stays with similar resource use) to provide estimates of disease-specific costs. It is possible to refine further these costs by accounting for longer than usual stays and specific high-cost treatments(41). Due to the granularity of HES data, it is possible to obtain these estimates by patient characteristics such as age, sex, number of comorbidities, and deprivation. HES data may also be linked to mortality data to incorporate the impact of time to death on hospital costs. For example, Seshamania and Gray categorised time to death by quarters, and found that costs in the last quarter were
over ten times that of 20 quarters to death (£1,698 compared to £148)(42). Research into hospital expenditure that used HES has provided a number of useful findings that are relevant to this report:

- Patients from more deprived areas have higher expenditure than patients from areas of less deprivation. Further, for the general population aged under 65, over 30% of annual medical spending is incurred by 1% of the population(43). This suggests that if unrelated costs are included there may be important distributional impacts on inequality.

- Annual per capita expenditure has been increasing over-time(44). If future unrelated costs are modelled, it may be worth modelling the trends in these costs. For consistency, trends in overall survival and utility would also need to be modelled.

- Expenditure increases with the number of morbidities (health conditions). When combined with evidence on time to death, there was little additional impact of age on expenditure, with an average cost of $9,000 (2014 US$) in the last year of life, compared with an average cost of $3,500 two years before death(44).

The feasibility of using HES data to derive healthcare expenditure by age, sex, time to death and disease has been demonstrated in a study, which used 260 mutually exclusive disease categories(41). This could be used to derive per capita costs and so be used in the same was as PAID. An advantage of using HES data, compared with PAID, is that the former is longitudinal, and so the impact of time to death on healthcare expenditure will be more accurately captured. The main drawback with this approach is that it is limited to hospital expenditure, and so does not provide a complete picture of per capita healthcare costs. For example, using Dutch data, hospital expenditure represented 44% and 35% of total healthcare costs for decedents and survivors, respectively(20). Due to these limitations, a better approach would be to use HES data in combination with other national sources of healthcare expenditure. Van Baal, Meltzer and Brouwer describe combining HES data with Office for Health Economics data, but it is unclear which source(s) were used, and if they varied by age(45). An alternative analysis by Asaria combined HES data with a nationally representative GP dataset (Q Research)(31). This allowed for an analysis of how costs varied by age, sex, and level of deprivation. The combined dataset covered hospital expenditure and some primary care expenditure; costs incurred by dental and ophthalmology services
were not included in Q Research, and so were taken from summary data, with the assumption that costs were uniformly distributed across the population. Another assumption was required for outpatient costs: in contrast to inpatient stays, it was not possible to derive specific costs for each outpatient episode, so a single average cost was used. Walker and colleagues describe a similar approach to estimating healthcare costs, this time focusing on patients with stable coronary artery disease(46). Some differences from the Asaria study are that the Clinical Practice Research Datalink (CPRD) was used instead of Q Research, and the authors did not attempt to estimate outpatient costs. By focusing on a specific disease (stable coronary artery disease), the data from Walker and colleagues demonstrate how it would be possible to separate-out related and unrelated disease costs. For example, in the first year of follow-up 38% of the cohort experienced a hospitalisation, but only 54% of these hospitalisations were for the related disease.

Whilst not discussed in either study, both analyses appear to omit some categories of healthcare expenditure, such as on mental health and social services. A more complete estimate of English healthcare expenditure is provided by Ride and colleagues, who linked HES data with both the CPRD and the Mental Health Services Data Set (previously known as the ‘Mental Health and Learning Disabilities Data Set’ and the ‘Mental Health Minimum Dataset’) (47). The resulting integrated dataset provides very granular data - for example, the authors are able to demonstrate that a number of variables affect overall healthcare expenditure, including time since diagnosis, patient age, level of deprivation and ethnicity. Drawbacks of this approach are that a non-trivial amount of data-cleaning was required, and the integrated dataset is still incomplete, as the authors note that it omits community prescribing and community social care.

A detailed description of an alternative approach to estimating age-sex specific unrelated health-care costs is described by Briggs, Scarborough and Wolstenholme (hereafter, ‘the Briggs approach’)(13). Whilst previous approaches used micro-costing (bottom-up), the Briggs approach uses gross-costing (top-down). A particular strength of the Briggs approach is that it covers all NHS expenditure. The key data-source used is NHS England cost curves(48). These provide estimates of relative healthcare expenditure by age and sex, broken down by four categories of care: general and acute, mental health, prescribing, and primary care. This does not cover all of the NHS healthcare spending, as maternity services
and specialised services are not included; data on the overall expenditure is available for these two services, but not by age. To obtain an age-sex breakdown for specialised services, it is assumed that 46% of the expenditure follows the cost curve for mental health care, and the remaining 54% follows the cost curve for general and acute care (this ratio was based on 2012/13 data, when a breakdown of costs by care category was available for specialised services). Maternity costs were allocated by age based on the distribution of live births in England and Wales. Hence, age-sex breakdowns of relative healthcare expenditure are obtained for five categories of care (general and acute, mental health, prescribing, primary care, and maternity). To estimate absolute costs by age and sex, the total healthcare expenditure by category of care is required. Total NHS healthcare expenditure is obtained from three (mutually exclusive and exhaustive) data-sources:

1. Programme budgeting data, which provides estimates of expenditure, broken down by 56 disease categories (including maternity services and primary care prescribing).
2. Primary care costs (excluding prescribing costs).
3. Specialised services

Of the primary care costs, 90% are allocated to primary care, with the remaining 10% allocated to mental health (based on the fact that 10% of primary care prescribing costs are spent on mental health). The cost of specialised services were divided between mental health, and general and acute care, as previously described. Details on the allocation of programme budgeting costs to the categories of care depends on the specific disease (see (13) for more details). To summarise, the Briggs methodology is as follows.

- Subtract related disease costs from the overall NHS healthcare budget. This provides an estimate of overall expenditure on unrelated diseases.
- Allocate this overall unrelated expenditure to the five categories of care.
- Apply estimates of relative healthcare expenditure by age, sex and category of care to the overall expenditure (by category) to obtain estimates of spending on unrelated diseases by these characteristics.

When compared with the micro-costing approaches, a drawback of the Briggs approach is the increased work required to identify related medical costs. These costs may have been
estimated separately (for example, by measuring resource use during a trial and applying unit costs to these), but the authors also describe an approach for estimating related medical costs that is consistent with the estimation of unrelated costs. The main idea is that the broad programme budgeting categories (and their costs) are decomposed into the costs for individual diseases. This may be achieved using examples from the literature, or failing this, using HES data. Examples provided by the authors include:

- Costs for the programme budgeting category “04a Diabetes” are decomposed so that type two diabetes receives 90% of the costs, and type one diabetes receives the remaining costs. This is based on a published estimate of the relative NHS expenditure on these two diseases(49).

- Stomach cancer is a part of the programme budgeting category “02b Cancer, upper GI”. HES data was used to estimate the proportion of admissions for upper GI cancers that were due to stomach cancer. This proportion was then applied to the costs for “02b Cancer, upper GI” to derive stomach cancer costs.

There are three points to note with the authors’ approach:

1. The estimation of unrelated costs is based on a ‘top-down’ approach, as opposed to the ‘bottom-up’ approach used in the HES-based studies. That is, in the Briggs approach, related costs are subtracted from total expenditure, and the result is then broken-down by age and sex. In contrast, with the HES-based studies, costs of unrelated disease by age and sex are summed to obtain unrelated costs by age and sex. The bottom-up approach avoids the assumption (used in the Briggs approach) that the related disease costs has the same age and sex distribution as the unrelated diseases. This is an acknowledged limitation of the Briggs approach(13).

2. Unlike PAID, the Briggs approach does not make an adjustment for end of life costs (or more generally, time to death). As life-extending treatments will delay the occurrence of end of life costs, it is important to explicitly include these when estimating unrelated costs(20). Estimates of annual healthcare costs separated by survivors and decedents may be obtained for the hospital setting using HES data, as previously mentioned. Overall expenditure on other healthcare would require decomposing into survivor and decedent costs. One option, described in an English
study examining the impact of physical activity on dementia(50), is to use decedent costs from a Nuffield Trust report(51). This report used linked data on 73,243 people from a sample of English regions to estimate costs in the last year of life by age, sex and 26 different diseases. Costs included hospital care (inpatient, outpatient and accident and emergency) and social care (residential and nursing, home care and other care). These decedent costs can be combined with overall costs and the probability of death to derive survivor costs, using the approach in PAID (see Equation 1 in Appendix A.2). A limitation with the Nuffield report is the granularity of the reported data: the 26 disease categories is a lot less than the number of categories in PAID, and disease-specific estimates are not broken down by either age or sex. An alternative approach is to use the ratio of decedent to survivor costs that has been estimated for the PAID tool.

3. In the Briggs approach, three programme budgeting categories were excluded when estimating unrelated costs: ‘social care needs’, ‘healthy individuals’, and ‘other’. For use in NICE appraisals, social care needs should be included. As the remaining two categories represent healthcare spend, they may need to be included to account appropriately for opportunity costs. However, it is unclear if these may be interpreted as unrelated disease costs.

As noted, the key methodological difference between the Briggs approach and the HES case-studies is that the former begins with an estimate of overall healthcare expenditure and decomposes this to obtain estimates by disease, age and sex. An advantage of this top-down costing approach is that it ensures that all healthcare expenditure is included. A disadvantage is the assumptions (and potentially additional data sources) required to decompose overall expenditure. Two other studies were identified that also decomposed overall healthcare expenditure. The first was the Dutch cost-of-illness study that is used as the primary data source in PAID(20). The second was an Australian disease costs study(52). Both studies used a combination of national registry data and survey data. For example, the Australian study used national data covering hospital care, medical services, and prescriptions. This was supplemented by a national health survey and a survey of ‘morbidity and treatment in general practice’, and a survey of ‘disability, ageing and carers’. The alternative bottom-up costing approach requires integrated datasets. Examples using national linked datasets have been described for Denmark and New Zealand(53, 54). The
Danish study also considered wider societal costs (production and non-healthcare consumption), but not time to death. The New Zealand study focused on medical costs, which were broken down by age, sex and time to death (last six months, last year, others).

The previously mentioned approaches all used national data sources. Two English examples using local data sources have also been described. These both used a bottom-up costing approach. The first is the Kent Integrated Dataset(55), which includes all residents of Kent, and the majority of health and social care expenditure. Some notable evidence gaps were primary care prescribing data and hospice costs. Linkage was based on patients’ NHS number, with the linked dataset provided evidence on resource use. This was combined with reference costs to obtain healthcare expenditure; the reference costs included both locally obtained costs and national unit costs, such as the healthcare resource groups that are applied to HES data. In a published study, the Kent Integrated Dataset was used to examine the association between deprivation and healthcare expenditure amongst the elderly (those aged 55 and older). Data on costs were available for 63% of the eligible population (323,401 of 512,120), and showed a strong effect of deprivation; with an increase in per capita costs of over £400 in the most deprived quintile compared with the least deprived. Results were independent of the effects of age and gender. The second example is integrated data from South Somerset, described as the Symphony Project(56). The published data are for 2012, and cover 114,874 people. Diseases are classified into 49 categories, and costs are provided by eight care settings: primary care (including prescribing), inpatient and day-cases, outpatients, accident and emergency, mental health, community care, social care, and continuing care. Regression analyses considered differences by age, sex and the number of diseases but not by the time to death. Results suggested that age had little impact. Details on linkage and derivation of costs are not provided, but are assumed to be similar to those of the Kent Integrated Dataset. The data were used in a study that modelled the impact of physical activity and dementia - costs from the Symphony Project were combined with end of life costs from the previously described Nuffield Report(50, 51).

For these two local integrated datasets, it would be beneficial to identify how representative the estimates of per capita healthcare expenditure are for use at a national level. European examples of using local integrated datasets to estimate healthcare costs were identified for Baix Emporda in Catalonia, and for Copenhagen(19, 57). In the former case- study, both
proximity to death and the number of morbidities were found to be important drivers of healthcare costs. The Copenhagen data were used as part of a HTA assessing interventions to reduce alcohol consumption, with little comment on the data used, beyond noting that it may underestimate costs as it excluded long-term costs due to injuries.

No other UK examples of estimating per capita healthcare costs were identified. One other approach, used primarily by American studies, was the use of national survey data, which contain evidence on healthcare expenditure. One example is the Consumer Expenditure Survey published by the Bureau of Labor Statistics(8, 58). This does not disaggregate results by disease, so estimates of expenditure on related diseases should ideally be removed to avoid the effects of double counting. An alternative source is the Medical Expenditure Panel Survey(59), although a noted limitation of this survey is that it does not include the institutional population(8). None of the identified case studies included an adjustment for the time to death. Within England, there is some survey data, which combines limited evidence on the prevalence of diseases with healthcare resource use, and so could be combined with unit costs to obtain estimates of healthcare expenditure by disease. This includes the British Household Panel Survey and Understanding Society(60, 61). Local surveys may also be useful, for example the Yorkshire Health Study collects evidence on 13 named diseases (with a free-text option for additional diseases), as well as the frequency of visiting 25 named health-care providers by four broad categories: hospital, other healthcarer (such as GP or dentist), other carer (such as care worker and health visitor), and alternative therapist(62). A primary limitation with the use of survey data is the potential inaccuracies of self-report, which would include recall of healthcare resource use within a given time-period(63).

5.3. SUMMARY

A number of approaches and potential data sources for estimating per capita healthcare expenditure, and from this estimates of unrelated medical costs, have been identified. Of these, the Briggs approach appears to be the most promising for obtaining estimates specific to the UK. Unlike the other UK examples, all relevant healthcare expenditure is guaranteed to be included by the top-down costing approach employed. Further, estimates by age, sex and healthcare category can be obtained using publically available data with relatively little need for data cleaning. In theory, it would be possible to use the Briggs approach to attribute
the overall expenditure by age and sex to disease-specific expenditures. The general idea for this approach and some examples have previously been described. A further adjustment to obtain separate survivor and decedent costs would also be possible, for example using the Nuffield Report as in an existing case study(51) or an alternative suitable source. A recent systematic review of palliative care costs in the UK identified ten relevant studies, although it noted that due to variation it was not possible to synthesise results across studies(64). The resulting estimates would then be an English equivalent to the inputs used in the PAID toolkit. The main practical difficulty is in deriving consistent disease-specific expenditures from overall healthcare category expenditures. For hospital categories, HES data could be used to obtain the relative costs of different diseases, with a constraint that their sum (by age and sex) matches the overall age- sex matched costs. For the other care categories, it may be possible to use the other datasets used in the integrated HES-based approaches. For example, the Mental Health Services Data Set could be used to attribute overall mental health expenditure to specific diseases, whilst a nationally representative primary care dataset (such as Q Research, CPRD, The Health Improvement Network, or Research One) could be used for the attribution of primary care costs to specific diseases. Alternatively, the use of local datasets or local costings, such as those used in the Kent Integrated Dataset or the Symphony Project may be possible. Another approach would be to obtain disease-specific estimates via the use of survey data, although this would be limited by the number of diseases that can be reported in a survey, in addition to the potential biases associated with self-report(63).

It would also be beneficial to compare the estimates from the top-down Briggs approach with those from the various bottom-up costing approaches. As both approaches are attempting to measure the same per capita healthcare costs (by patient characteristics) they should provide similar estimates; any areas where they differ may provide useful insights into their relative strengths and limitations. Of the case studies identified, two examined the impact of the number of morbidities on healthcare expenditure(19, 44). These both found an important effect. In particular, when the number of morbidities and time to death were modelled, there was little residual impact of age on healthcare expenditure. This suggests that future improvements may be to incorporate the effects of multimorbidity. To do so would further increase the number of individual categories for which estimates are required, as would more granular modeling of time to death (beyond a binary yes/no for the last year of life) and the inclusion of any other potentially important patient characteristics such as
ethnicity(47). It is unlikely that sufficient data would be observed to populate directly costs for every individual category. The majority of the identified case studies employed regression models to derive estimates; there is the potential that different approaches to regression modelling would lead to different estimates(65, 66). The sensitivity of results to the choice of regression model should also be explored in future research.
6. POTENTIAL IMPACT ON NICE APPRAISALS

If the current NICE guidance on unrelated costs does not change, and they remain excluded from appraisals, then the argument of Section 2.1 that the appraisals lack internal consistency would remain. As noted in Section 2.2, it is unclear how many appraisals actually incorporate unrelated health effects, for future appraisals it would be beneficial to explicitly describe the assumptions that are made about unrelated health effects, and how these are included or excluded in the appraisal. In addition to a potential lack of internal consistency, the continued exclusion of unrelated costs would have the effect of under-estimating the true impact on a healthcare budget of a life-extending treatment.

Assuming that unrelated health benefits are appropriately included, then including unrelated costs would make life-extending treatments more costly and hence less cost-effective. Unrelated costs increase with age, whilst general population utilities decrease with age. As such, including unrelated costs would lead to higher increases in cost for treatments of older populations. Younger populations would incur unrelated costs for a longer time period, but the effect of increasing unrelated costs with age would be lessened by the impact of dampening. Based on the current definition of unrelated costs, their inclusion would not impact on treatments which improve quality of life without extending it. However, such treatments may improve patient characteristics (such as physical functioning) to a level where patients are able to receive unrelated treatments that they would not have received in the absence of the treatment. If unrelated costs due to increased quantity of life are included in NICE appraisals, then unrelated costs due to increased quality of life should also be considered.

In NICE appraisals of medical technologies, a cost-consequence approach is used(67). In these situations, life-extending treatments will be associated with increased unrelated costs, but no health benefits. This creates a perverse incentive to not demonstrate survival gains for medical technologies.

Decision-making in NICE appraisals is informed in part by comparing the incremental costs per incremental QALYs gained against a willingness to pay threshold(3). This threshold is notionally based on the cost effectiveness of what is displaced(68). As some of the potentially
displaced treatments include life-extending treatments, their cost-effectiveness would be affected by the inclusion of unrelated costs. As such, if including unrelated costs in NICE appraisals, different thresholds would be required to those currently used; future research would be required to identify these thresholds.

The inclusion of unrelated costs within NICE appraisals leads to ethical questions of equity, as outlined in Section 2.2. It is acknowledged that these issues should be explicitly addressed, although there is currently no consensus on how best to do so. For example, in a manuscript entitled “Future unrelated medical costs need to be considered in cost effectiveness analysis”(2) the authors note “We share the moral intuition that denying access to life-extending treatments to dialysis patients specifically, is ethically challenging.” If NICE did not have an explicit process in-place to address equity concerns then appraisal committees would be required to make value judgments about when equity issues are more important than cost-effectiveness on a case-by-case basis, which opens the possibility for subjective and inconsistent decision-making. There is the related concern that introducing unrelated costs may increase inequalities, as their impact would be unequally distributed across the patient population. This would be contrary to one of NICE’s guiding principles to reduce health inequalities: “If possible, our guidance aims to reduce and not increase identified health inequalities”(33).

NICE appraisals would also require data sources to estimate unrelated costs. A number of potential approaches and data sources have been described in Section 4, but there is currently no standardised source of unrelated costs for use in a NICE appraisal. Even if a suitable evidence source were identified, there are unresolved methodological issues about which covariates to include (for example, should the number of morbidities be a factor?), and the approach to statistical modelling of cost data. The absence of both a suitable dataset and methodological consensus raises the potential for inconsistent approaches to the inclusion of unrelated costs. This in turn would hamper comparisons across appraisals, and so increase the potential for making sub-optimal decisions(13). For internal consistency, the inclusion of unrelated costs would also need to be accompanied with the consistent inclusion of indirect health effects. One approach would be to model the utility values and survival of the population of interest relative to that of the general population. For example, a disease-specific disutility could be applied relative to age-gender general population values. For survival, it is possible to decompose the overall (observed) survival into survival of the general population (expected survival), and survival due to the disease(69). For example, if it is
assumed that observed survival is the product of expected survival and disease-specific survival, then the methods of relative survival may be used. Alternatively, it may be assumed that a certain proportion of the patient population will experience survival equal to the general population, so cure models may be employed. For both relative survival models and cure models it should however be noted that there is no consensus on the best approach to use(27).

The inclusion of unrelated medical costs would broaden the set of costs that are included in NICE appraisals. A related question is then unrelated non-medical costs should be included (and so a broader cost perspective used by NICE). This may ameliorate some of the impact of including unrelated medical costs, as individuals who live for longer have greater opportunity to contribute to society.

Ideally, the aforementioned issues and limitations would be addressed prior to the incorporation of unrelated costs in NICE appraisals. This would require time to implement. The alternative is that unrelated costs are included at an earlier point, and any issues are addressed in a reactive manner, as they arise during appraisals. An advantage of this is that the key practical issues that arise in NICE appraisals would be explicitly identified, with efforts focused on resolving these. The disadvantage of this reactive approach is two-fold. First, if the inclusion of unrelated costs is sub-optimal, this will have tangible impacts on decisions, which may in turn be sub-optimal. The inclusion of unrelated costs in HTAs is advocated as addressing existing sub-optimal decision making, but if unrelated costs are not appropriately included, there is no guarantee that they will improve decision making. The second disadvantage of a reactive approach is that initial NICE appraisals which include unrelated costs, and are approved by NICE committees, would set a precedent. Hence, even if it is found that the initial approach to incorporating unrelated costs was flawed or could be improved, it may be difficult to justify not using the initial approach when it has been previously approved. For these reasons, the proactive approach is preferred, with the practical and ethical issues associated with including indirect costs addressed prior to their inclusion in NICE appraisals.
7. CONCLUSIONS

There are compelling arguments for the inclusion of unrelated costs; both to ensure both internal consistency with the inclusion of any unrelated health effects, and to provide a more accurate estimate of the costs associated with a life-extending treatment. Arguments against the inclusion of unrelated costs focus on the practical and ethical issues of implementation: suitable evidence sources and methods for estimating unrelated costs are required, and the inclusion of such costs would raise ethical concerns for which there is currently no clear solution. These practical issues may be a factor behind the majority of identified HTA guidelines either not discussing the use of unrelated medical costs, or explicitly recommending their exclusion. Reasons for exclusion were rarely provided, but when they were, they referred to practical and ethical issues in implementation. The inclusion of unrelated costs was only identified in the guidance of five HTA bodies; of these, only the Dutch ZIN provided practical guidance on how to incorporate these costs. This includes the publically available interactive PAID tool, which allows for granular estimates of unrelated cost that also remove costs due to pre-defined related diseases. A key feature of PAID is the use of per capita healthcare expenditure to derive unrelated costs. A number of English approaches to estimating per capita healthcare expenditure were identified, including top-down and bottom-up costing exercises, as well as the use of national and local datasets. None of the identified English approaches had the same level of detail or ease of use (particularly for defining related diseases) as PAID. Of the existing English approaches, the Briggs approach appears to be the most promising for further refinement as it was the only approach to use a similar top-down costing approach to the Dutch cost-of-illness study that was the primary evidence source for PAID.

There are a number of areas with a current lack of either consensus or clarity. These would require further research if NICE were to consider making changes to the methods guide. These include:

- Is it possible to create a toolkit for estimating unrelated costs, similar to that used in the Netherlands (PAID)? If so, which methods and data sources would be the most appropriate to use?
- If a toolkit is used, is the PAID approach sufficient, or should any other factors (such as the number of morbidities) be accounted for?
Can the distributional impacts of incorporating unrelated costs be anticipated in advance? In particular, would including unrelated costs lead to increased inequalities in health outcomes?

Ethical objections to the inclusion of unrelated costs have been raised. Would the process of NICE decision-making need adjustment to place greater emphasis on equity concerns? For example, in the Netherlands, the proportional shortfall approach is used to combine equity concerns with estimates of cost-effectiveness (70).

How will the inclusion of unrelated costs impact on the NICE cost effectiveness threshold?

If unrelated costs are included, should the NICE methods guide be more prescriptive about the inclusion of unrelated health effects? In particular, should the methods guide recommend the inclusion of age-adjusted utilities?

When extrapolating survival gains, what methods should be used to provide assurances that the impact of unrelated diseases and unrelated treatments is accounted for?

Unrelated costs could occur in situations when there is no life-extension (if treatment allows for additional future unrelated treatment). Should these unrelated costs also be included?

As unrelated costs represent real costs incurred by a healthcare system they should ideally be accounted for in appraisals of life-extending treatments. However, to do so is not straightforward, and may be the reason why the majority of identified guidelines from HTA bodies did not recommend their inclusion. This report provides an overview of the key considerations and highlights areas requiring further research.
8. REFERENCES


# APPENDIX

## A.1 List of Guidelines by HTA Organisations

<table>
<thead>
<tr>
<th>Country</th>
<th>Guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>Commonwealth of Australia as represented by the Department of Health.</td>
</tr>
<tr>
<td>Country</td>
<td>Source</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Israel</td>
<td>State of Israel Ministry of Health. Guidelines for the submission of a request to include a pharmaceutical product in the national list of health services (The National Health Insurance Law, 1994). 1994.</td>
</tr>
<tr>
<td>Slovenia</td>
<td>Health Insurance Institute of Slovenia. Rules on the classification of medicinal products on the list. 2013.</td>
</tr>
</tbody>
</table>
### A.2 Methodological Framework for PAID

Let ‘lhc’, ‘sc’ and ‘dc’ denote lifetime healthcare costs, survivor costs, and decadent costs, respectively. These costs are then related by the formula:

\[
lhc(g) = \sum_{a}^{n-1} \sum_{i} sc_i(a, g) + \sum_{i} dc_i(n, g)
\]

where \(a, g, i, n\) represent age, sex, disease, and age at death, respectively. With this bottom-up approach to estimating overall healthcare costs, the lifetime costs due to unrelated diseases (unrelated healthcare costs; ‘uhc’) may simply be defined as Equation 1 excluding costs for all related diseases, which are denoted by \(Z\). This gives:

\[
uhc(g) = \sum_{a}^{n-1} \sum_{i \notin Z} sc_i(a, g) + \sum_{i \notin Z} dc_i(n, g)
\]

To estimate the inputs to PAID (the \(sc_i\) and \(dc_i\)), the main data-source used was a Dutch cost of illness study(20), with the 2017 version used in PAID 3.0. The starting point of the study was total annual medical expenditure in the Netherlands, which was decomposed into estimates of healthcare cost for 107 diseases by age, sex and healthcare provider. For age, 21 age classes were used; interpolation to single years of age (for use in PAID) was obtained using
cubic splines. The cost of illness study did not provide separate costs by survivors and deecedents; instead, it gave evidence on overall (average) per capita healthcare costs ($ac_i$). Hence, for use in PAID, the authors assumed that the $ac_i$ was a weighted average of $sc_i$ and $dc_i$ with weights given by the mortality rate ($m$):

$$ac_i = (1 - m) \times sc_i + m \times dc_i$$  \hspace{1cm} (3)

$$= (1 - m) \times sc_i + m \times r_i \times sc_i$$  \hspace{1cm} (4)

where $r_i$ is the disease-specific ratio of decedent to survivor costs. All calculations were stratified by age and sex. For hospital expenditure, estimates of $r_i$ could be obtained directly. For the remaining healthcare providers, it was either assumed that $r_i = 1$, or data was available on the overall value of $r$ (averaged over diseases), so those overall ratios were assumed to follow a similar disease-specific decomposition to that observed for hospital expenditure (see (20) for more details). One assumption in Equation 2, which is not explicitly discussed, is that having a given disease will not influence the costs of unrelated diseases compared with the general population. That is, conditional on the related diseases, costs of unrelated diseases are the same as per capita costs. The accuracy of this assumption is unclear.