Background study

A background study on the 'cost-effectiveness' package principle for the benefit of the appraisal phase in package management

Publication number 291

Publisher	College voor zorgverzekeringen P.O. Box 320 1110 AH Diemen Fax (020) 797 85 00 E-mail info@cvz.nl Internet www.cvz.nl
Series number	29079523
Department	Care Advice
Authors	Prof. Dr. J.J. van Busschbach (Erasmus MC) en Dr. G.O. Delwel (CVZ)
Direct line	Tel. +31 (0) 20 797 85 44
Orders	Extra copies can be ordered via our website (www.cvz.nl) or by calling the service desk via telephone number +31 (0) 20 797 88 88.

Contents:

page

1 1. Summary
3 2. Introduction
6 3. Cost-effectiveness
7 4. The lack of a ceiling value
7 4.a. Other arguments in addition to costs per QALY
7 4.b. Eliciting strategic behaviour
8 4.c. The lack of a normative framework
10 5. A bandwidth
10 5.a. Other criteria
13 6. Separating assessment and appraisal
15 7. The limited alternatives to QALYs
18 8. When is cost-effectiveness relevant?
18 8.a. Cost-effectiveness is a criterion of a higher order
19 8.b. Limited budget
20 9. Why apply a threshold value to cost-effectiveness?
21 10. The size of the threshold value
22 11. The interaction between additional arguments and cost-effectiveness
23 12. The bandwidth between reasonable and dubious
25 13. How reasonable is a threshold value if it leads to a person's death?
26 14. Aren't the reasonable threshold values actually much too high?
29 15. The domain of cost-effectiveness
29 15.a. Cure, care and prevention
30 15.b. Efficiency, cost-effectiveness and appropriate use
31 15.c. Confusion surrounding necessity
34 16. Additional criteria
36 17. Criteria that increase the clemency of the cost-effectiveness

requirement
36 17.a Burden of disease
38 17.b. Rarity of the disease (orphan drugs)
39 17.c. Informal care
39 17.d. Public health care risks
41 18. Criteria that increase the strictness of the cost-effectiveness requirements
41 18.a. Limited overlap with the health care domain
41 18.b. Budget impact
42 18.c. Future medical costs not included
43 18.d. Unsuited to insurance due to high prevalence
43 18.e. Unsuited to insurance due to patient having a lot of influence on dose of treatment
44 18.f. Uncertainty about appropriate use of the intervention
45 19. Criteria that should not be allowed to count
45 19.a. 'Lifestyle'/high-risk behaviour
47 19.b. Age, gender, ethnicity, sexual preference and social-economic status
49 20. References

Appendix

List of abbreviations

1. Summary

The *College voor zorgverzekeringen* (CVZ, Health Care Insurance Board) has created a committee that will compare the outcomes of the four package principles, necessity, effectiveness, cost-effectiveness and feasibility with one another, per intervention and also with other possible arguments. This appraisal committee, the Package Advice Committee (ACP), will provide weighted advice on whether an intervention is eligible for inclusion in the basic package of health care provisions. This background study will first examine the meaning of cost-effectiveness in relation to appraisal and then weigh this package principle in relation to the other three package principles and other possible arguments.

This background study, though it is not binding, was written as guidance for the ACP as it is a specific elaboration of the 'cost-effectiveness' package principle seen from the point of view of an appraisal committee such as the ACP.

CVZ's guidelines for pharmacoeconomic research define cost-effectiveness as the costs per Quality-Adjusted Life-Year (QALY). Three decades of scientific research have shown that QALYs are the most useful and valid means of expressing the cost-effectiveness of health care.

This cost-effectiveness in terms of costs per QALY has been 'weighed' against the three other package principles and possible other arguments. In the past CVZ had previously suggested operationalising the 'necessity' package principle via the concept burden of disease. This idea was recently confirmed by advice issued by the Council for Public Health and Health Care (RVZ). By weighting cost-effectiveness according to necessity (burden of disease), instead of defining a fixed limit for costs per QALY, CVZ is working according to a bandwidth. This band ranges from €10,000 for a limited burden of disease up to €80,000 for an extremely severe burden of disease. In other words: the evaluation of an intervention's cost-effectiveness is partly determined by the

burden of the disease. This is a way of creating solidarity with patients who suffer a high burden of disease.

In addition to necessity, the evaluation of cost-effectiveness is also determined by the two remaining package principles (effectiveness and feasibility) and a number of other arguments. Effectiveness can be interpreted as the certainty that an intervention actually does what it is expected to do. A high effectiveness represents a high degree of certainty, which will reduce the weight given to cost-effectiveness. The same applies to efficient feasibility. Other arguments permitting a higher cost-effectiveness are the rarity of a disorder (orphan drugs), a positive effect on informal care-givers and a reduction in risks to others (e.g., due to a reduced chance of contagiousness). Arguments that increase the weighting are limited overlap with health care, budget impact, non-insurability due to high prevalence, or due to patients having a lot of influence on the treatment dose, uncertainty about the appropriateness of the intervention (the degree to which the right patients are being treated). 'Lifestyle'/high-risk behaviour is normally excluded as an argument because these concepts also apply (in part) to most ordinary diseases and accidents. Also excluded as evaluation factors for cost-effectiveness are age, gender, ethnicity, sexual preference and social economic status.

This report provides CVZ for the first time with a fairly exhaustive list of factors that can be used for weighting the cost-effectiveness of a given treatment. This is a refined method for an appraisal committee such as the ACP for allowing the cost-effectiveness argument to play a role in decisions on whether or not to include interventions in the basic package.

2. Introduction

Reason

Since 1st January 2005 cost-effectiveness has played a role in the assessment of reimbursement applications for new unique medicines under the Dutch system for reimbursing medicines (GVS). This involves CVZ issuing advice on a medicine's therapeutic value, the consequences for the pharmaceutical budget and the substantiation of its cost-effectiveness (appropriateness).
In its role as package supervisor, CVZ has established four package principles: necessity, effectiveness, cost-effectiveness and feasibility. Package decisions are examined according to these criteria in the appraisal phase. It is important that CVZ's advice is not merely a statement on the substantiation of the cost-effectiveness, but also an evaluation of that that cost-effectiveness.

Questions

This report on "the cost-effectiveness package principle for the benefit of the appraisal phase in package management" is a background study that addresses two main questions.

1) What is the value of the cost-effectiveness, and
2) What weight does the cost-effectiveness carry in relation to the other package principles and other arguments that may play a role in the appraisal phase.

Method of work

These questions have been elaborated upon by Prof. Dr. J.J. van Busschbach (psychologist and professor at the Erasmus MC) and Dr. G.O. Delwel (advisor on Package Advice, CVZ). Content-related discussions took place within CVZ with J. Zwaap (advisor and secretary to the ACP) and Dr. A. Boer (Executive Board and Chairman of the ACP). The draft document was subsequently discussed twice by the ACP (October and December 2008). This background study is an appendix to CVZ's report 'Package management in practice 2' that was sent to the Ministry of VWS on 2nd June 2009.

Background to Package Management in Practice 2

Insured Package

CVZ has set up a committee that weighs up – per intervention – the outcomes of the four package principles, necessity, effectiveness, cost-effectiveness and feasibility in relation to one another and other possible arguments. This appraisal

Advisory Committee

committee, known as the Insured Package Advisory Committee (*Advies Commissie Pakket*, ACP), must provide a considered opinion on whether an intervention is eligible for inclusion in the basic package of health provisions, and is expected to pay particular attention to aspects relating to society. This background study first examines the actual evaluation of the cost-effectiveness (when is cost-effectiveness good and when is it poor) and then addresses the weighting of this package principle in relation to the other three package principles and other possible arguments.

Society-related aspects

Dealing with arguments and considerations

Although this background study is not binding, it was written as guidance for the ACP as it is a specific elaboration of the package principle 'cost-effectiveness' seen from the point of view of appraisal. The committee will have to weigh up the various arguments in relation to one another, for example 'cost-effectiveness' in relation to 'burden of disease'. This report describes what is said in the literature about such considerations, and in an earlier phase the ACP actually referred to it as "a toolbox of arguments and considerations". The report does not dictate what the outcomes of such considerations should be: after all, the actual weighing up is the *raison d'être* of a committee such as the ACP. One exception to this rule is where certain outcomes are known to be inconsistent or where it is clear that a given factor will lead to unexpected, undesired outcomes.

Conclusions

CVZ does not apply a ceiling value to cost-effectiveness. The cost-effectiveness of many interventions lies within a bandwidth with a median value of €40,000/QALY.
Various criteria influence the assessment of cost-effectiveness in the appraisal phase. These criteria may increase the clemency of the cost-effectiveness requirement, e.g., as in the case of burden of disease and rarity of a disease, or they may increase the strictness of the cost-effectiveness requirement, as is the case with lack of certainty regarding the appropriateness of the intervention. Some criteria are prohibited from playing a role in this evaluation.

'Toolbox'

The criteria elaborated upon in this background study, the so-

called ‘toolbox’, facilitate a transparent assessment of cost-effectiveness for package decisions.

3. Cost-effectiveness

CVZ regards cost-effectiveness as a package principle that is used when taking decisions about whether or not to include a treatment in the basic package. CVZ currently has two decades of experience in determining cost-effectiveness in health care. This experience was described in guidelines for pharmacoeconomic research which were drawn up in 1999 and updated in 2006 (CVZ, 2006). According to these guidelines, cost-effectiveness should preferably be expressed in costs per Quality-Adjusted Life-Year (QALY). CVZ's guidelines describe in detail which costs should be included in cost calculations. In addition, the guidelines for pharmacoeconomic research describe how benefits to health should be measured using QALYs. Compared with this *clarity* regarding the measurement of cost-effectiveness, there is still a *lack of clarity* about its use as a criterion in compiling the package. The big question that still remains is: Where do we draw the line distinguishing between an intervention that is cost-effective and an intervention that is not?

On the basis of the guidelines for pharmacoeconomic research, this ceiling is expressed as the maximum number of Euros per QALY that can still be regarded as cost-effective budget expenditure. For example: €20,000 per QALY is – only just – acceptable, whilst €25,000 per QALY is not.

4. The lack of a ceiling value

Governmental organisations usually avoid explicit reference to a ceiling for costs per QALY. There are several reasons for this, which are described in brief below. One of them, the existence of other criteria in addition to cost-effectiveness, is elaborated upon in more detail below.

4.a. Other arguments in addition to costs per QALY

Firstly, avoiding any reference to an explicit limit an indication that cost-effectiveness is not the only principle on which the composition of the package is based. Although 'costs per QALY' do already cover a number of criteria (survival, quality of life and costs), it is quite conceivable that other arguments also exist that are used when compiling the package. A well-known example of such a criterion is 'burden of disease', which is also referred to as 'necessity'. If other criteria are important in addition to cost-effectiveness, then it would be impossible to optimise the decision-making process by allowing everything to depend on costs per QALY. On this basis, there is no room for an absolute ceiling for costs per QALY. The other criteria that exist and how these correlate with the interpretation of cost-effectiveness are discussed in more detail below.

4.b. Eliciting strategic behaviour

A second important reason for caution in stipulating an absolute 'ceiling for costs per QALY' is that this could provoke undesirable strategic behaviour on the part of those who provide care. For example, with a fixed ceiling of €20,000 per QALY, it is conceivable that the price of each new medicine would be driven up to €19,999 per QALY. Equally conceivable therefore is that the representatives of those who finance health care, such as CVZ, refuse to give up the negotiating position they have.

When offering interventions just under the ceiling, people should realise that the ceiling is an upper limit and not simply the ‘average of all that is available within the package’. Everything that touches the ceiling or is just below it belongs to the least cost-effective interventions that are only just acceptable. The present comprehensive package exists thanks to the fact that most interventions are much cheaper per QALY. Continually including expensive interventions that are just under the ceiling will lead either to maximum acceleration in increased costs of care or to the repression of older interventions that are much more cost-effective by newer interventions that are less cost-effective (Claxton et al., 2008).

4.c. The lack of a normative framework

A third reason for not having a fixed ceiling is that it has proven to be extremely difficult to ask people for a standard price for ‘costs per QALY’ or even ‘costs per life-year gained’. To start with, there are practical problems. For example, a survey on ‘willingness-to-pay’ in relation to holidays, cars and jars of peanut butter is quite conceivable, because people have a pretty good idea of the value they attach to these products. Furthermore, the public are used to making decisions about the prices of such products on a daily basis. That does not apply to most forms of health care, with the possible exception of dental care as patients in the Netherlands do make their own cost appraisals. It is inconceivable that any patients have ever weighed up whether a life-year gained – let alone a QALY – is worth €20,000. Patients and the general public simply lack any form of ‘price anchor’, i.e., a clear normative framework in which they can place the costs of health care.

Apart from this practical problem, the question is whether – theoretically speaking – there is a valid method for obtaining such a statement from the public (Gyrd-Hansen, 2008). The 'willingness-to-pay' approach is part of the so-called 'prospect theory’, whilst 'costs per QALY' or 'costs per life-year' are part of the so-called 'cumulative prospect theory’. For example, in the first theory, 'willingness-to-pay' is emphatically a function

of income, whilst in the second theory, these differences are ignored because we feel that basic health care should be available for all citizens, irrespective of their income. In other words, not only practical complications, but also theoretical complications, have ensured that as yet attempts to arrive at a convincing upper limit for ‘costs per QALY’ on the basis of interviews with patients or the general public have been unsuccessful.

5. A bandwidth

In fact it is impossible to arrive at a clear upper limit for 'costs per QALY' due to 1) the presence of other arguments than cost-effectiveness, 2) the attempts to avoid strategic behaviour on the part of care-providers, and 3) the lack of a normative framework. The lack of a clear limit has resulted in the current practice of working with a 'bandwidth'. The breadth of this band is generally determined according to previous decisions. A frequently quoted example of such a decision in the Netherlands was the 40,000 Guilders limit per QALY that was extrapolated from the clinical guidelines on cholesterol-reducing products dating from 1998 (Casparie, Van Hout, Simoons, 1998). In a recent study, the Council for Public Health and Health Care [*De Raad voor de Volksgezondheid en Zorg,* RVZ] claimed that the bandwidth ranges from €10,000 for a disease with a low burden to €80,000 for a disease with a high burden (RVZ, 2006). The upper limit of this bandwidth in particular received a lot of attention from the media. The Council based the upper limit on a number of observations, including:

- The annual costs per patient for a nursing home (€60,000)
- From international studies involving various interventions (bandwidth €12,000 to €73,000)
- The WHO norm 'three times the national product per head of the population' (€90,000)
- Devlin & Parkin study (2004) of decisions made by NICE (National Institute for Health and Clinical Excellence) in England (€79,000)
- Day's (1999) meta-analysis of estimates of the value of a statistical life, based on 17 American publications, which supplied a value of about €5,600.00 for a statistical life. Assuming an average life expectancy of 79 years, this amounts to €71,000 per year.

The Council finally concluded: "*… that society regards €80,000 as a reasonable upper limit for a QALY."*

5.a. Other criteria

As indicated above, this upper limit of €80,000 received a

great deal of attention in the media. Less attention was given to the Council's emphatic statement that this sum only applied for patients with an illness that is associated with a high burden of disease. In other words, this limit only applies to diseases whereby patients would die immediately if given no treatment or whereby patients without treatment would have a very low quality of life. Where this is not the case, then a lower sum per QALY should apply as the upper limit (RVZ, 2006).

In fact, the RVZ did not actually propose an upper limit, but defined the bandwidth. The proposed bandwidth shows that the RVZ feels that cost-effectiveness is not the all-embracing criterion for compiling the basic package. The Council states that, though it is an important criterion, it 'must be weighted by burden of disease'. This means that we should show greater solidarity with (i.e.: "We are prepared to spend more money on …") patients with a heavy burden of disease, than on patients with a lower burden of disease. To put it differently again: patients with a low burden of disease show solidarity towards patients with a high burden of disease.

In cases where burden of disease affects how we interpret a treatment's cost-effectiveness, it is conceivable that other matters may also play a role. Rare diseases form a widely cited example. Because the turnover of medicines for this group is – by definition – low, the price per patient of newly developed medicines is often high. For example, the RVZ states that: "[patients] *may not be allowed to suffer from the fact that their disorder, illness or handicap is rare …"* (RVZ, 2007, page 22). In other words: patients with common ailments should show solidarity with patients with rare diseases. A disease with a high burden and a low frequency could form a ground for accepting more costs per QALY than would otherwise be the case.

In addition to burden of disease and disease frequency, other matters could also conceivably lead to such a line of thought. For this reason CVZ has set up a committee, the ACP, to examine, per intervention, whether the value of the package principle 'cost-effectiveness' counterbalances the three other

package principles (necessity, effectiveness and feasibility) and other possible arguments. In other words, this committee does not determine the cost-effectiveness, but assesses it with respect to all relevant arguments.

6. Separating assessment and appraisal

By separating the measurement of the package principles and their assessment relative to one another, CVZ is following an international trend to separate the assessment phase from the appraisal phase (RVZ, 2007; Technology Appraisal Committee, see also: www.NICE.org.uk; Dear et al., 2007). Noticeably, the appraisal phase, unlike the assessment phase, is not characterised by strict methodological guidelines. In its advice on the appraisal phase, for example, the RVZ only elaborates upon the interaction between burden of disease and cost-effectiveness (RVZ, 2006, 2007). The explanation provided by the RVZ on other arguments, such as the above-mentioned infrequency of a disease, is limited to a few paragraphs (RVZ, 2007, page 22). The number of arguments raised is also limited. According to the Council, arguments allowed to play a role in assessing cost-effectiveness are:

- Burden of disease
- Personal responsibility,
- Incidental effects on society,
- The temporary nature of the interventions
- The rarity of a disease

The Council actually explicitly excludes some criteria:

- Age
- Gender
- Ethnicity
- Sexual preference
- Social-economic status
- Geography
- 'lifestyle' and (high-risk) behaviour

According to the Council, these explicitly excluded criteria can actually be taken into account when assessing the effectiveness of an intervention. This means that an intervention may be more effective on women than on men, and that will and can lead to a different decision for men than for women. The document subsequently elaborated upon this argumentation.

This report is an attempt to increase the number of arguments and elaborate upon them in order to provide CVZ's appraisal committee (ACP) with a list of the most plausible arguments for assessing cost-effectiveness in relation to the other package principles (necessity, effectiveness and feasibility).

7. The limited alternatives to QALYs

QALYs are an attempt to arrive at an all-embracing standard for health. Another possibility is to take a more limited aspect of health, for example life-years or blood pressure. Thus, instead of 'costs per QALY', cost-effectiveness is limited to 'costs per life-year gained', or 'costs per reduced millimetre of mercury'.

The guidelines for pharmacoeconomic research permit use of the above method for expressing cost-effectiveness, though preference goes out to costs per QALYs. This is based on two important arguments.

First of all, it provides a more limited view of effects on health comparisons between interventions. For example, how does one compare 'costs per life-year' with 'costs of a reduced millimetre of mercury'? The impossibility of making such comparisons makes it difficult to arrive at an assessment of the price per effect. For example, is it worthwhile to pay €1,000 per patient, per year, in order to reduce blood pressure by one millimetre? In other words, this does not circumvent but rather complicates the crucial discussion about just what is an acceptable price for health, because the acceptable price will have to be determined per effect.

A second argument for using QALYs as outcome measure, instead of focusing on a single effect, is that the latter approach may result in important effects being missed. The best-known example of this is the focus on life-years gained in oncology: this neglects the possibility of significant effects on quality of life.

In the restricted sense of the word, cost-effectiveness may be useful when examination is limited to a single field of treatment. For example, when examining two blood pressure-lowering products with little to differentiate them, then the cost-price per reduced millimetre of mercury can be sufficient. During such a comparison the representatives of the most expensive intervention typically come up with a claim that

their intervention has better properties in addition to reducing the blood pressure, for example fewer side effects. This is their way of suggesting that more should be taken into account than just the main effect. If that is the case, then one should also be looking for a broader effect parameter that also reflects side effects and this will inevitably lead to using QALYs.

In spite of the above, researchers in Germany are still searching for methods to work with limited cost-effective analyses that make no use of QALYs. The IQWiG (Institute for Quality and Efficiency in Health Care) in Germany suggests making use of a 'limited cost-effectiveness ratio'. In order to determine the maximum cost-price per effect, they propose examining the present cost-price per effect, i.e., the cost-price per effect that currently exists in practice. For example, the current medication can reduce blood pressure on average by 3 millimetres of mercury for €1,500. This is equal to €500 per millimetre. In that case, a new medicine that achieves another millimetre reduction is allowed to cost a maximum of €500 per patient. A new medicine that reduces blood pressure by 5 millimetres would be allowed to cost a maximum of €2,500. There have been vehement protests about several aspects of the logic of this line of reasoning.

Firstly, it is not clear whether achieving one extra millimetre reduction in mercury is the best way to spend €500, because it is not clear what that millimetre actually does to our health, e.g., in terms of survival and quality of life. In other words: "what good is it to the patient?" This is the same as the argument mentioned above: the problem surrounding the interpretation of the parameter "costs per QALY". Using the more limited cost-effectiveness analyses does not solve the problem, but instead splinters it into a large number of new interpretation-related problems per specific effectiveness measure. It may be the case that 'millimetres mercurial pressure' are more familiar than 'QALYs', but there is no shred of scientific or normative (ethical) substantiation of the concept 'cost-price per millimetre of mercurial pressure', whilst during the past three decades 'cost-price per QALY' has

proved to be a useful concept in the ethical debate surrounding cost-effectiveness in health care.

A second often-quoted argument against using the more limited cost-effectiveness analysis is that it assumes current treatment practice is 'optimal'. After all, the cost-effectiveness of the current practice is what determines the cost-price per effect, e.g., the €500 per millimetre mercurial pressure. In health fields in which money tended to flow like water in the past, these 'historical cost-effectiveness ratios' perpetuate the idea that this will continue to be the case in the future. Cost-effectiveness ratios are keener in fields that were less well-off in the past and the purse strings tend to be drawn more tightly. It is a typical case of "the rich getting richer whilst the poor get poorer".

The scientific community is keeping a watchful eye on developments in Germany, because of inconsistencies in evaluating cost-effectiveness ratios and the fact that the status quo is inherently maintained (Krauth et al., 2008; Jönsson, 2008).

8. When is cost-effectiveness relevant?

8.a. Cost-effectiveness is a criterion of a higher order

Measuring cost-effectiveness in terms of 'costs per QALY' is not always relevant, because the cost-effectiveness criterion is a compendium of other criteria (costs, survival, quality of life), which are themselves capable of providing sufficient information. In other words, cost-effectiveness is a criterion of a higher order. The fact is that CVZ only assesses the cost-effectiveness package principle for new medicines in situations where:

- the medicine is not mutually replaceable;
- the medicine has a therapeutic added value;
- the medicine leads to added costs.

This means that cost-effectiveness only becomes important when the new medicine has more effects and will cost more than the old alternative. In these cases the Medicinal Products Reimbursement Committee (CFH) assesses the substantiation of the cost-effectiveness analysis. In principle, this committee assesses not whether the cost-effectiveness is high or low, but only whether the method of determination was valid. To do this the CFH uses the above-mentioned guidelines as laid down in the Guidelines for Pharmacoeconomic Research (CVZ, 2006). These guidelines stipulate that cost-effectiveness must be estimated in terms of a cost-effectiveness analysis in which costs represent costs to society and effects are expressed as Quality-Adjusted Life-Years (QALYs).

Thus, a cost-effectiveness analysis only becomes relevant after the assessment has already involved a number of other steps. When weighing up the 'cost-effectiveness' package principle, the appraisal committee can therefore rightly assume that:

- the medicine is unique (one that is not mutually replaceable);
- it has a statistically significant and clinically relevant effect (the medicine has a therapeutic added value)
- using the product will require making funds available

(substantial added costs are involved)
- the cost-effectiveness analysis complies with CVZ's guidelines.

8.b. Limited budget

Assessing the cost-effectiveness of a treatment only makes sense if the budget is limited. If the budget is not limited, then carrying out a cost-effectiveness analysis would be a waste of effort and an effect study would suffice. There is a lack of clarity about the degree to which the health care budget is limited. The fact that budget growth is limited by numerous policy measures is evident enough. Nevertheless, up till now the budget has always grown, relatively as well as literally. Clearly, a limited budget is not an irrefutable fact. This means that how the appraisal committee interprets cost-effectiveness analyses will be largely determined by the degree to which the committee itself is assuming that the budget is limited. Cost-effectiveness could play a more prominent role if the committee feels that the budget is limited, than if the committee feels that budget growth is still possible. In the latter case, other arguments than cost-effectiveness will gain the upper hand.

9. Why apply a threshold value to cost-effectiveness?

If the appraisal committee assumes a limited budget, then it should also assume that granting funds for a new treatment will always go hand-in-hand with funds being withdrawn for another treatment that is already in the package. After all, a limited budget implies that accepting something into the package will mean saying farewell to another reimbursement. If this is to be done properly, the committee will need to know not only the cost-effectiveness of the new treatment but also the cost-effectiveness of the treatment that can now no longer be financed. Preferably, therefore, the appraisal committee should be provided with two cost-effectiveness analyses:
1) a cost-effectiveness analysis of the new treatment; and
2) a cost-effectiveness analysis of the treatment that will have to be excluded from the package (Buxton, 2007).

However, the latter analysis will hardly ever be placed in the hands of the committee. This is partly due to the large number of treatments in the package, which makes it difficult to determine exactly where budget has been withdrawn. Even if a number of possible candidates were known, we would still need to know the cost-effectiveness data of those candidates at the moment that the new candidate is being appraised. In practice, therefore, a pre-determined threshold value for cost-effectiveness is used, or the above-mentioned bandwidth.

10. The size of the threshold value

The threshold value or bandwidth for cost-effectiveness is generally determined on the basis of treatments that are already included in the package. The highest costs per QALY found in the package are usually taken into consideration. Assuming a limited budget, this would inevitably lead to interventions with a better cost-effectiveness being excluded from the package. This would be illogical as it would mean the basic package as a whole would be providing less health care than before the new treatment was included. In other words: the inclusion of a new treatment has increased suffering instead of reducing it. This is one of the reasons why most of the government authorities involved in cost-effectiveness in the various countries cite a stricter cost-effectiveness threshold than the highest values found in the national package. The highest cost-effectiveness values that are barely acceptable are thus reserved for interventions that need extra arguments, such as interventions for patients with a high burden of disease. This report also sums up the other arguments that are related to accepting an inferior cost-effectiveness.

The way in which NICE refers to its threshold values is illustrative of the above. NICE claims that treatment that costs less than £20,000 per QALY is cost-effective and must therefore be reimbursed. If the cost-effectiveness is between £20,000 and £30,000, then NICE requires additional arguments before issuing positive advice. Above the £30,000, everything depends on those additional arguments, because the cost-effectiveness is interpreted as being low (House of Commons, 2007; NICE, 2003; Culyer et al., 2007). Nevertheless, Devlin (2004) found that NICE was permitting interventions up to £55,000. Apparently, NICE quotes a stricter threshold value than the one they actually use in practice. In keeping with what was said in the previous paragraph, there is clearly room for additional arguments that permit a higher cost price per QALY.

11. The interaction between additional arguments and cost-effectiveness

For the ACP appraisal committee, the key question with regard to cost-effectiveness is: “Exactly what are the additional arguments and how do they relate to the threshold value of cost-effectiveness?”

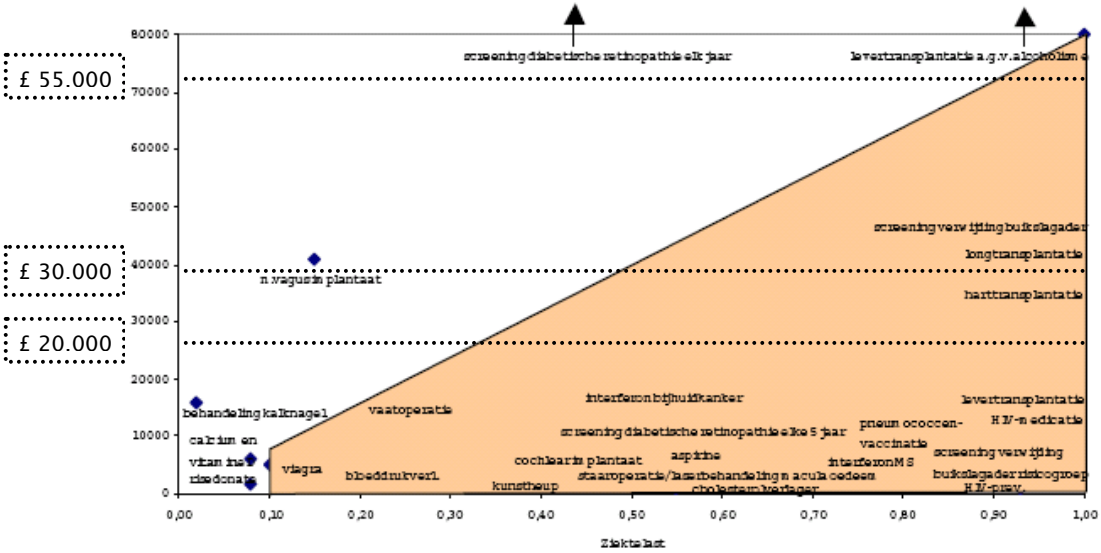
The most elaborate reply to this question was provided by the Netherlands. In 2006, in the above-mentioned (RVZ) report, the RVZ elaborated upon the burden of disease criterion in relation to threshold values. Basing themselves in part on earlier research carried out by CVZ, the RVZ claimed the existence of consensus within society that patients with a mild burden of disease must show solidarity towards patients with a severe burden of disease. In that case, it is justifiable that a new treatment with high costs per QALY forces a treatment with low costs per QALY out of the package, as long as the new treatment focuses on patients with a severe burden of disease and the other treatment was for patients with a mild burden of disease. This can be interpreted as accepting high costs per QALY for patients with a severe burden of disease and being stricter when assessing treatments for patients with a mild burden of disease. The Scottish Medicines Consortium (SMC) works along similar lines: there is no fixed threshold, but additional arguments become increasingly important as the cost-price per QALY rises (Cairns, 2006).

12. The bandwidth between reasonable and dubious

The RVZ has established an €80,000 per QALY threshold for patients with maximum burden of disease. This threshold drops to €10,000 for the category with the lowest burden of disease. The €80,000 upper limit is the same as the highest threshold value that NICE uses in practice: £55,000 (the value of the British pound has since fallen dramatically). The midpoint of the bandwidth is the same as the bandwidth quoted by NICE, ranging from £20,000 to £30,000. The same applies to the SMC (Dear et al., 2007). The most natural choice would therefore be to quote €40,0000 per QALY as a guideline for indicating the limit at which reasonable cost-effectiveness ends and dubious cost-effectiveness begins. These matters are illustrated in figure 1 taken from the RVZ. It shows that the RVZ model works with an interaction between cost-effectiveness (costs per QALY) and burden of disease, whilst NICE works with bandwidths. It also shows that the cost-effectiveness of most interventions lies within the proposed limits.

Having said this: does this mean that most of the ACP's work has been regulated as far as cost-effectiveness is concerned? This is not the case. The committee will probably use the midpoint of the bandwidth, €40,000. The limits of the band have not as yet been determined. Indeed: in all probability, these limits will be influenced by external developments. For example, increases or economies in the total budget are crucial because these determine the basic assumption of a fixed budget. This is why a number of authors describe appraisal committees such as the ACP not as committees that apply a threshold value, but rather as committees in search of the threshold value (Buxton, 2007, Cuyler et al., 2007).

Fig. 1. Costs per QALY according to severity of the disorder



Source: RVZ, 2006

For the rest, the curve has not yet taken on a fixed shape. Nevertheless, few statements have been made and little research done on shifts around the midpoint. The most obvious is to presume a monotonous tendency to rise, in which case a linear relationship usually has a powerful predictive capacity. This is exactly what the RVZ report shows. One might suppose that in a few years time it will be possible to sketch the curve based on decisions made by the ACP.

13. How reasonable is a threshold value if it leads to a person's death?

Is the said threshold value of €40,000 society's 'firm cut-off point'? Is this the same as saying: this far but no further? This is unlikely. One cannot imagine a society that would consider it acceptable to refuse treatment for a patient in a life-threatening situation simply because that treatment is more expensive than, for example, €40,000 per QALY. The crux of this observation lies in the phrase "a life-threatening situation". The value of a life-year or a QALY is not the same for everyone: for example, if a patient has only a few years left (the life-threatening situation), then we are prepared to reach deeper into our purses. In other words: the greater the 'need' to treat, the more we are prepared to pay. This is why necessity is often translated into a relatively large loss of life-years and/or quality of life, i.e., a heavy burden of disease. This is exactly what CVZ discovered in 2001 and the RVZ in 2006 (CVZ, 2001, RVZ, 2006).

The fact that we are prepared to go far in acute situations that are clearly life-threatening also explains why the RVZ chose the sum of €80,000 as the highest threshold value. It also explains why values exceeding €100,000 per QALY are often found in the various studies that attempted to estimate the value of a life (Hirth, 2000). Such cases always involve the acute expectation of death, that is, a very severe burden of disease. Reason enough to differentiate between the 'common or garden variety' of health care and attempts to prevent excessive acute suffering. In other words: an important issue in assessing interventions is the burden of disease one is trying to avoid.

14. Aren't the reasonable threshold values actually much too high?

One often hears, particularly from patients and the pharmaceutical industry, that the English organisation, NICE, is too strict in its assessment of the cost-effectiveness of new interventions. In other words: the NICE threshold of about €40.000 per QALY is unfair. Expectations during the past few years were therefore that the threshold value would eventually be raised. Surprisingly enough, this turned out not to be the case. As it happens, a recent development in this discussion is that the local NHS implementation organisations, the so-called Primary Care Trusts (PCTs), claim that NICE is actually too lenient. The PCTs are expected to realise the NICE recommendations within the existing budgetary frameworks. These representatives of actual practice claim that authorising expensive new interventions is putting pressure on normal, but highly effective forms of aid (House of Commons, 2007; Martin, Rice, Smith, 2007; NICE 2007; Buxton, 2007). They claim that NICE should bring the threshold values more into line with current practice, because common, effective forms of treatment will otherwise be pushed out of the package by expensive new developments that are much less effective. These matters have given a new twist to the discussions surrounding the height of the threshold (Towse, Raftey, 2009). As a result, the House of Commons has initiated research that will examine whether the NICE threshold value should be lowered to bring it more into line with daily practice (House of Commons, 2007, page 6).

The observation made by local English parties, that the €40,000 per QALY threshold is everything but 'normal', is in keeping with an observation made in the Netherlands by Meerding et al. (2007), that the current package contains many interventions that are much more cost-effective than the threshold value of €40,000 (see table 1).

For example: the costs of interventions per QALY for infectious diseases and heart and vascular disease currently range between €2,000 and €5,000. The costs for oncological

interventions, at €16,000 to €18,000 per QALY gained, are considerably higher, but still far under the €80,000 maximum RVZ limit which should apply to this group with a high burden of disease. Similar statistics exist for England, with a cost-price of £19,100 per QALY for oncology and £12,000 for heart and vascular diseases (House of Commons, 2007, page 60).

Table 1. Costs per QALY for interventions already included in the package Meerding et al . (2007).

Diseases	€ / QALY
Infectious diseases where antibiotics are the most important intervention	
Gastrointestinal infections	2,771
Tuberculosis	71
Pneumonia	6,049
Sepsis	11,164
Infectious diseases where vaccination is the most important intervention	
Meningitis	-1,015
Meningococci	4,208
Diphtheria	257
Whooping cough	442
Polio	-22,268
Measles	991
Cancer	
Lung cancer	18,618
Colorectal cancer	10,893
Breast cancer	2,387
Prostate cancer	30,095
Testicular cancer	692
Non-Hodgkin	6,980
Hodgkin	1,077
Heart and vascular diseases	
Coronary heart diseases	3,531
Strokes	-3,428

A number of interventions, e.g., with a case of polio, lead to a negative cost-price per QALY. This is because the economies achieved by these interventions exceed the costs.
In practice, individual treatments, or even sub-groups of treatments, may be much more expensive than the average

values shown in table 1. Nevertheless, these average prices show that everything above €40,000 per QALY should be categorised as having a relatively poor cost-effectiveness. Poor cost-effectiveness is only justifiable if there are other important arguments that justify reimbursement, such as those elaborated upon by CVZ and the RVZ for burden of disease.

15. The domain of cost-effectiveness

15.a. Cure, care and prevention

Cost-effectiveness analyses are widely used in the field of cure, which is where they were largely developed. Cost-effectiveness analyses are also often used in prevention. One difference between cure and prevention is that the cost-effectiveness requirement in prevention is usually stricter. This can be explained by a number of factors.

Firstly, prevention involves – by definition – a future patient, and it is a known fact that matters in the future do not carry as much weight as matters in the present. Secondly, prevention involves patients who are – again by definition – as yet unknown. Unlike patients of flesh and blood, these statistical victims in the future do not elicit immediate solidarity, and this results in a lower priority. It is also conceivable that the uncertainty surrounding calculations is overestimated for prevention and less so for cure. The stricter interpretation of the cost-effectiveness of prevention is yet another illustration of the existence of arguments that influence how cost-effectiveness is interpreted.

Doubt is often expressed about the value of using standard cost-effectiveness analyses in the field of care because it is often difficult to translate the research questions into costs per QALY analyses. For example, it is hard to imagine how to translate 'more privacy' or 'improved visiting arrangement' into something as generic as a QALY. The obvious answer is to develop different measuring instruments for this type of parameters that relate to comfort and quality.

A number of new concepts and measuring instruments are currently being developed for classifying what we know about the cost-effectiveness of care. One promising example is an elaboration by Flynn et al. (2008), who have linked QALYs with the capabilities approach of Sen (Verkerk, Busschbach, Karssing, 2001) by means of discrete choice experiments. One wonders where this development will eventually end. Quite

conceivably, we might even return to a 'pure QALY'. This is only possible, when the description of the alternative form of care is sufficiently clear. For example, if nursing homes did not exist, then not only would the quality of life of patients be lower, one might also expect mortality to increase. If that is the case, then the QALY is a sensitive measure for care. This is probably the reason why the RVZ based their standard of €80,000 per QALY in part on the cost-price of a year in a nursing home (RVZ, 2006, page 87). The discussion about using the QALY, or any other measure of cost-effectiveness in care, is expected to go on for some time.

15.b. Efficiency, cost-effectiveness and appropriate use

One matter of confusion is that the concepts of efficiency, cost-effectiveness and appropriateness are often used simultaneously, whilst they actually refer to different matters. This applies in particularly to the use of efficiency and cost-effectiveness. In science, the concepts of efficiency and cost-effectiveness are synonymous, but that is not always the case with respect to policy and politics. In matters of policy, the concept of efficiency is used in particular to indicate that the new intervention:

- Has the same effects, but costs less
- Has better effects, but costs less
- Has better effects, but costs just as much

The situation ‘better effects, more costs’ is left out of the equation. All four of the above-mentioned situations are covered within the concept cost-effectiveness, and in fact it is ‘better effects, more costs’ that leads to most discussion. In fact the three previously mentioned situations form an ‘open-and-shut-case’: no-one would be opposed to the reimbursement of an intervention that is just as good but costs less. In fact, the concept of efficiency, as defined in the three above-mentioned situations, has nothing to contribute to what we would already have done. We recommend therefore using the concept of cost-effectiveness instead of the incomplete concept of efficiency.

Appropriateness is a concept for indicating whether a medicine is actually reaching the target group in practice and it is often referred to as 'appropriate use'. It reflects the difference between scientific research and clinical practice, i.e., the difference in efficacy and effectiveness in practice. Due to the fact that cost-effectiveness analyses focus on the situation in practice ('real costs' and 'real effects'), the degree of appropriateness is often included in a cost-effectiveness analysis. See also the paragraph on 'Uncertainty about appropriate use of the intervention'.

15.c. Confusion surrounding necessity

The four package principles of CVZ were developed from the four criteria of the Funnel of Dunning, one of which was the concept 'necessity'. Dunning defined necessity upon introduction of the criteria, as a situation in which:
1) the disorder leads to early mortality, or
2) the disorder hinders – or prevents – normal participation in society.

In particular, using the concept 'normal participation in society' presented many problems. As a result, over the course of time the necessity criterion has developed into the concept of burden of disease, as this is easier to determine than 'normal participation in society' (CVZ, 2001).

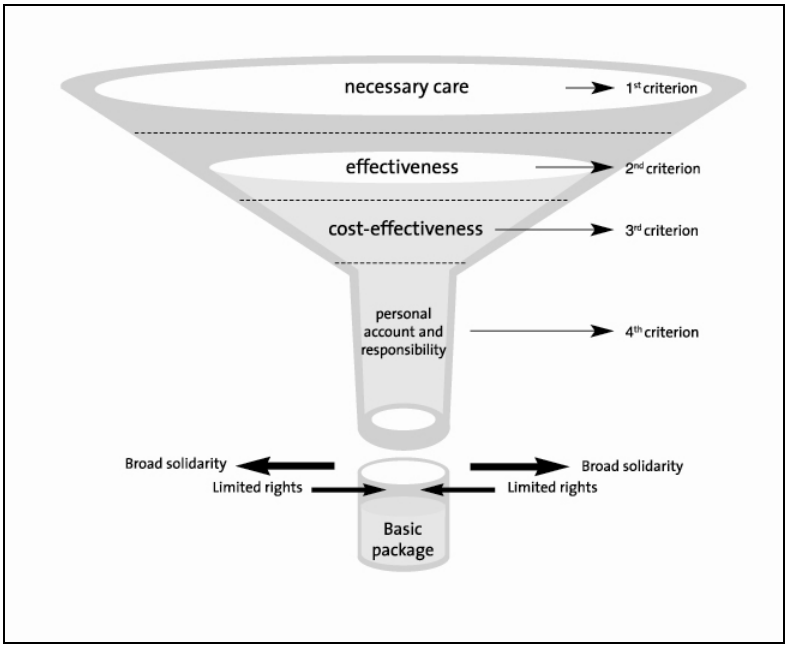
After thus clarifying the definition of necessity, CVZ recently proposed renewed blurring of the definition of necessity: "*Does the disease or the necessary care, given the cultural context, justify a claim on solidarity*" (CVZ, 2006).
According to CVZ (2006), necessity covers not only 'burden of disease' and 'care requirement', but also 'considerations relating to personal responsibility' (CVZ 2006, page 36).

This means that Dunning's 4th criterion, 'personal accountability and responsibility', has been incorporated into the first criterion 'necessity'. For a number of reasons this merger between Dunning's first and fourth criteria was

received with some hesitation.

1. Firstly, necessity can no longer be assessed unequivocally, but will have to be subjected to multidimensional assessment: burden of disease, care requirement and not-at-one's-own-expense.
2. Secondly, the statement "*Does the disease or the necessary care justify a claim to solidarity ….*" is a description of the outcomes of an assessment, and not a criterion.
3. Thirdly, the addition of "*given the cultural context*" has brought us back to the concept 'normal', a concept that had proved to be so sadly unproductive in the Funnel of Dunning.
4. Fourthly, there is no elaboration of what is meant by the concept 'care requirement'.

Fig. 2. Dunning's funnel



We therefore recommend a critical examination of the proposal. It would probably be better to move some of the 'criteria' that have been placed under the heading 'necessity' and re-locate them under 'feasibility'. Feasibility can then assess typical insurance matters such as the predictability of patients' costs, the effect of patients' behaviour on costs and the possibility that they bear the costs themselves. Spectacles are a good example of this. The need for spectacles is great, because not having spectacles would lead to a high burden of disease. However, spectacles are unsuited to inclusion in a social insurance package. This is partly because people would no longer be careful with their spectacles (their own responsibility would no longer be reinforced). Furthermore, insurance is not really necessary because a new pair of spectacles is usually affordable. We therefore recommend, when elaborating in more detail on the concept of necessity, shifting 'considerations relating to personal accountability' to the fourth criterion 'feasibility'.

16. Additional criteria

Quite conceivably, in addition to the RVZ criterion 'burden of disease', there may also be other criteria which – similarly to burden of disease – should influence how we assess the cost-effectiveness of a form of treatment. For example, the rarity of the treatment, budget impact and treatments that have not been validly measured via QALY analyses. Unfortunately, none of these additional criteria have been elaborated upon as was done by the RVZ for the burden of disease criterion. Below is a list of many of these additional criteria. Most of these criteria would increase the threshold value, i.e., the treatments to which the criteria relate are more likely to be cost-effective. The opposite applies to some of the criteria; if these criteria apply, then the cost-effectiveness should be examined more critically. Although some criteria are mentioned often, their inclusion is highly questionable. For example, being guilty of causing the disorder, as with smoking or skiing accidents. Table 2 lists these criteria. The rest of this document describes the criteria.

Table 2. Criteria that play a role in assessing cost-effectiveness.

Criteria that affect the **interpretation van cost-effectiveness**	**Increase the clemency** of the cost-effectiveness requirement	Make the cost-effectiveness requirement **stricter**	Should **not** be included
• High burden of disease • Rareness • Lots of Informal care • Public health risks			
• Little overlap with health care domain • High budget impact • Future medical costs have not been included • Unsuited to insurance due to high prevalence • Unsuited to insurance due to excessive patient influence on the treatment dose • Uncertainty about the appropriateness of the intervention			
• 'lifestyle'/high-risk behaviour • Age, gender, ethnicity, sexual preference and social-economic status			

17. Criteria that increase the clemency of the cost-effectiveness requirement

17.a. Burden of disease

There is a reason why the RVZ started by elaborating upon the burden of disease criterion as moderator for cost-effectiveness: burden of disease has been cropping up in the literature for many years, in various disguises, as the most important distribution principle, alongside effectiveness and cost-effectiveness. Economic literature refers to the 'equity debate', the dissemination of arguments about how burden of disease can be interwoven with arguments relating to cost-effectiveness. In the Netherlands this led to a similar political discussion about the Funnel of Dunning, in which the criterion 'necessity' refers to burden of disease.

It is important to realise that there are two sides to elaborating upon the burden of disease criterion: disorders with a severe burden of disease are more likely to be eligible for reimbursement, but this must imply that disorders with a mild burden of disease will only be reimbursed if they have a very good cost-effectiveness ratio. This clearly shows that the RVZ is well aware that the funds for financing the treatment of disorders with a severe burden of disease have to be found somewhere. It also shows that the greater the generosity of the appraisal committee ACP towards disorders with a severe burden of disease, the more they will have to refuse highly cost-effective treatments for another target group.

In operationalising the concept of burden of disease, the RVZ suggests complete non-reimbursement for diseases with a 10% burden of disease. One might wonder whether disorders with such a low burden of disease are in fact disorders and whether their treatment actually belongs in the field of health care (see below). On the other hand, a 10% threshold does seem rather high.

The RVZ defined burden of disease by means of 'proportional

shortfall', a concept that was introduced by André Ament of the University of Maastricht and elaborated upon further in Elly Stolk's dissertation (2005). Proportional shortfall takes into account age and duration of the disorder and it is a compromise between more radical definitions of burden of disease, such as 'the rule of rescue' and 'fair innings'. CVZ will shortly be elaborating upon the definition and assessment of necessity for the benefit of appraisals by the ACP.

Patients and other laymen often point out the importance of burden of disease, in addition to cost-effectiveness. A good example can be found in the reports of the Citizens' Council of NICE. This Citizens' Council has had elaborate discussions about whether and how burden of disease should be weighted in connection with cost-effectiveness. In January 2008 for example, the Citizens' Council of NICE spent two and a half days discussing the subject:
Should NICE and its advisory bodies take into account the severity of a disease when making decisions? If yes, should the advisory committees: 1) take severity "into consideration" alongside the cost and clinical effectiveness evidence; 2) or should severity be included in the calculation of the QALY?
(NICE, 2008 Page 4).

Note that, as far as burden of disease is concerned, the first part of the question is exactly the same question posed in this report. The second part of the question is about operationalising the first question. The conclusions of the Citizens' Council for the first part of the question were as follows:
The Citizens' Council concluded, by 24 to 2, that NICE and its advisory bodies should indeed take the severity of a disease into account when making decisions. Among the 24 of us who took this view there was unanimity that rather than do so by including severity in the calculation of the QALY, it should be taken "into consideration" alongside the cost and clinical effectiveness evidence"
We are not calling for the questionnaire [they are referring here to the EuroQoL EQ-5D] *or the QALY to be abandoned; rather we are suggesting that, in the light of experience so far,*

it is time they were subjected to a thoughtful and penetrating review (NICE 2008, page 4).

It is important to point out that, in spite of its critical attitude, the Citizens' Council of NICE does not dismiss the QALY paradigm, but calls for consideration of burden of disease, alongside cost-effectiveness. This is in keeping with CVZ's reports over recent years, and the RVZ report and is characteristically fitting for to an appraisal committee such as the ACP.

17.b. Rarity of the disease (orphan drugs)

In their report *'Rechtvaardige en Duurzame zorg'*, the RVZ argues that, in addition to burden of disease, attention should also be given to the size of the patient group. If the group is small, then the turnover of medicines for this group will be low. This means that the pharmaceutical industry will find it difficult to recoup their development costs. This would be easier if a higher price per patient were allowed for newly developed medicines. The RVZ puts it as follows: these patients may *"not fall victim of the fact that their disorder, disease or handicap is rare"* (RVZ, 2007, page 22). Thus, according to the RVZ, rarity of a disease is a reason for increasing the leniency of the cost-effectiveness threshold.

Broad support exists for the RVZ's line of thinking. CVZ has even designed separate assessment procedures in order to accommodate it. Although the assessment procedures are not fully operational, it seems that these procedures will amount to authorising interventions with much higher costs per QALY than the upper limit of €80,000 referred to by the RVZ. This idea has not been received without some criticism. Opponents point out that up till now no-one has succeeded in providing conclusive theoretical substantiation for this line of thought. This makes it difficult to persist in the idea that giving undue preference to orphan drugs is 'ethically' justified (McCabe, Claxton, Tsuchiya, 2005). An examination of the consequences makes it clear that such conclusive substantiation is lacking: how does one justify refusing to treat people with a common

or garden ailment, when we are willing to pay for people with an exotic disease (under circumstances that are otherwise equal ….)?

Despite the above-mentioned criticism, broad support does seem to exist for assessing the cost-effectiveness of rare diseases more mildly. This may be because orphan drugs are often used in situations where other favourable arguments also exist, such as a high burden of disease, recognisable victims when care is denied and low budget impact. The broad basis of societal support, the lack of a conclusive theoretical framework and the presence/link with other favourable arguments, in combination, make the assessment of orphan drugs an important challenge for the ACP.

17.c. Informal care

CVZ's Guidelines for pharmacoeconomic research state that cost-effectiveness analyses should be limited to the effects of treatment on a patient. This means that 'care-giver effects' are not included. It may be important to nevertheless include these in an assessment of diseases such as Alzheimer, borderline personality disorders and handicapped children. Where clear 'care-giver effects' do exist, they may actually be magnified by including these patients in the cost-effectiveness ratio.

17.d. Public health care risks

We argue below that there is no need to insure many frequently occurring disorders that are treated with relatively cheap interventions. For example, pain-killers and the contraceptive pill for adults. There will be exceptions to this rule. Should a patient neglect his disease, as a result of which the patient becomes a nuisance to himself or his surroundings, it may make sense to remove financial obstacles. Examples are vaccinations, the contraceptive pill for minors, care of drug addicts, etc. In these cases, it is important to keep the threshold as low as possible because

the seriousness of the consequences are greater than patients may suspect. Some cost-effectiveness analyses already work along this line of thought by including in the analysis the extra costs and effects in the patients' surroundings. An extra value for effects is unnecessary in such cases, because all effects have already been included in the analysis. However, the cost-effectiveness analysis may have been limited to effects on patients. In that case, it is advisable to include an extra weighting for the costs and effects external to the patients.

18. Criteria that increase the strictness of the cost-effectiveness requirements

18.a. Limited overlap with the health care domain

QALY-analyses can be used to assess the cost-effectiveness of just about all health care activities. However, one cannot simply reverse the order of this reasoning: it is not the case that all treatments that can be assessed by means of QALY-analyses actually belong within the domain of health care. For instance, the cost-effectiveness ratio for cosmetic surgery may be impressive. The same can be said of remedial teaching (dyslexia) and assistance with bringing up children. This does not mean to say that these three treatments clearly fit into the domain of health care. With respect to cosmetic care, this is regarded as health care that actually exceeds the natural average limit, whilst health care should focus on patients who are below the standard. Stolk, Brouwer and Busschbach (2005) elaborated upon this as the principle of 'pleasure-seeking' versus 'pain-avoidance' during the discussion surrounding the reimbursement of Viagra. Here also there is a relationship with burden of disease: a burden of disease does not exist until one falls below normal values, and this is why a social insurance package has no room for health care that enables patients to transcend the norm.

18.b. Budget impact

It is often claimed that the size of the required budget should not be an issue in discussions about reimbursement (RVZ, 2007, blz. 23). After all, budget impact is emphatically excluded from the four package principles: necessity, effectiveness, cost-effectiveness and feasibility. The question is whether budget impact is not being involved, informally, in reimbursement discussions. For example, a positive instance of this is the mildness that exists towards the reimbursement of rare interventions (see above). The importance of considering budget impact is apparent from the fact that in

almost all countries reimbursement requests must always be accompanied by an estimate of the required budget. Budget impact seems in particular to play a role where there is increasing uncertainty about use of the health care intervention and its (cost-) effectiveness. One might even say that a high budget magnifies uncertainty: an error with a small budget results in a small problem, whilst an error with a large budget results in a large problem. In this sense, it would be rational to be more restrictive with a high budget. There is growing interest in elaborating upon the rationality of the budget impact argument in more detail (Cohen, Stolk, Niezen, 2008). An improved description of this argument is expected to become available in the near future.

18.c. Future medical costs not included

Discussions are currently taking place on how to include patients' future medical costs in cost-effectiveness analyses. The Guidelines for pharmacoeconomic research state: "The evaluation may not include medical costs that are not related to the treatment of a disorder" (CVZ, 2006). This means that if a cured cancer patient dies three years later of a heart disorder, the costs of the heart complaint cannot be included in the analysis. They are included if the patient dies from a recurrence of the cancer. The reasoning behind this method of work is that life-extending treatments will otherwise be 'expensive' due to unrelated diseases. Nevertheless, there is a growing body of opinion that all future health care costs should be included. This plays a role for example in anti-smoking campaigns: contrary to what many people expect, health care costs will actually increase because ex-smokers will become older and therefore consume more care. In this case the cost-effectiveness actually remains favourable, because a lot of extra QALYs have been gained, so that the cost-price per QALY remains favourable. The discussion about how to deal with future health care has not yet been resolved. Until it has, we recommend examining whether such costs can be relevant, and whether they can be included in the analyses. The cost-effectiveness ratio should be assessed more strictly if they are not included, than if they are included.

18.d. Unsuited to insurance due to high prevalence

Interventions for diseases with a high prevalence and low costs, such as pain-killers for normal usage, plasters, etc., are not really suited to insurance, because everyone is confronted with these on a regular basis and because (almost) everyone can cope with these costs. To insure these would make this daily shopping unnecessarily expensive: after all, we do not take out 'bread insurance'. In other words: We take out health insurance in order to protect ourselves from 'catastrophic costs', and interventions for illnesses with a very high prevalence and low costs do not represent 'catastrophic costs'. Dental care is a good example of this. Everyday dental care can only be insured voluntarily, but dental care for damage due to accidents is reimbursed. The care required after a catastrophic event is insured, whilst everyday dental care is not. The above is sometimes referred to under the title of 'personal responsibility', as is the following criterion.

18.e. Unsuited to insurance due to patient having a lot of influence on dose of treatment

Insurance only works via the mediation of independent claims experts, that is, doctors in the case of health care. Transforming the recipients into claims expert would reduce the efficiency of insurance due to a lack of incentives towards efficiency, as well as increasing improper use (abuse). For an analogy, see the inefficiency of travel insurance for the loss of personal property. Insurance is clearly inappropriate for products where patients have a lot of influence on the dose of treatment, as is the case with sleeping pills, spectacles, dental care and contraceptive pills. The same applies to the reimbursement of interventions with low therapy compliance, such as self-help in giving up smoking (Zyban) and losing weight. This criterion is also referred to as the 'personal responsibility' criterion.

18.f. Uncertainty about appropriate use of the intervention

Appropriateness of use is about the degree to which the right patients are treated. If there is a great deal of uncertainty about prescribing an intervention outside the established diagnosis (a low level of appropriateness), then the cost-effectiveness should be scrutinised more critically. After all, in these cases costs will increase rapidly and the effects will be reduced. This applies in particular when the budget impact has already been estimated at a high level. An example in which this emphatically occurs is with expensive interventions for rare diseases. The costs rise rapidly and the cost-effectiveness is reduced when these expensive interventions are used not only for the rare diseases. This is an important reason why continually following appropriateness over time plays such an important role in the conditional reimbursement of orphan drugs.

19. Criteria that should not be allowed to count

19.a. 'Lifestyle'/high-risk behaviour

The definition of the concept ‘lifestyle’ has a laborious history, as can be seen from the literature. The term is often used loosely, mostly with a negative connotation in reimbursement discussions. Examples of things to which ‘lifestyle’ can refer are: luxury health care, high-risk behaviour (skiing, unsafe sex, smoking), eccentric preferences (sex changes, plastic surgery, extreme sports). This report adheres to Gilbert's definition (1999). He categorised a number of the ways in which people use the term ‘lifestyle drug’:

1) A medicine for a problem that is not really a health-related problem. In this respect, people even speak of a medicine that makes one more than normally healthy. An often-quoted example is Viagra, but it could also apply to anti-ageing cosmetics. Using the term ‘lifestyle’ in this way falls under the criterion ‘burden of disease’ and demarcation of the domain of health care (see above).
2) References to interventions that can be deployed for problems that can be put down (in part) as a personal accountability, such as obesity, stomach ulcer and smoking.

It is tempting to suppose that scarcity in health care would benefit if we were to stop reimbursing interventions required after an accident or a disease that is due to a patient's own behaviour. Often cited examples are smoking and skiing accidents. However, a practical complication to this line of thinking is that one can find many health problems that are, at least in part, related to people’s behaviour. This applies to skiing accidents, but also to falling off a step-ladder in the kitchen and to diabetes. In the debate, an extremely vulnerable aspect of the argument has turned out to be the lack of a clear division between high-risk behaviour and daily accidents and illnesses. For example, proponents of the ‘lifestyle’ and high-

risk behaviour argument are told that the consistent application of this line of reasoning would mean, for example, that a person would only be eligible for reimbursement in relation to an accident if they had been pushed. It is due to these complications that this often cited 'lifestyle' and high-risk behaviour approach has never been implemented.

The 'lifestyle' and high-risk behaviour argument does sometimes play a role during treatment. Treatment may be terminated if a patient does not cooperate sufficiently. A well-known example is that liver transplants are not carried out for patients who are still addicted to alcohol. Even in this case, however, one could say that this is not so much a 'lifestyle' argument, but rather that lack of treatment effect is the deciding factor. The 'lifestyle' argument becomes even more feeble if one examines the Multidisciplinary Guidelines for Cardiovascular Risk-Management (2006). Although – according to these guidelines – smoking should be discouraged, in the meantime medicine is *more likely* to be prescribed to smokers than to non-smokers: the reverse of what was intended by the 'lifestyle' argument. Apparently, even during treatment, the 'lifestyle' criterion often turns out not to be the powerful argument it was expected to be.

The discussion about 'lifestyle' could profit from developments surrounding the concept of capabilities. The concept of 'capabilities' refers to the possibilities a person has, irrespective of whether he or she takes advantage of them (Verkerk, Busschbach, Karssing, 2001). 'Societal guarantees' for capabilities are often offered that are linked to user characteristics: for example, the opportunity to attend further education up to a certain age. Older people may carry on studying, but they have to pay more. Seen in this light, Viagra (a discussion in which the concept of 'lifestyle' is often cited) for young diabetic patients is not quite the same as Viagra for the elderly. Another concept that may be helpful in interpreting 'lifestyle' is 'fair innings': health is guaranteed by society up to the level of what is regarded as 'normal health' (Stolk, van Donselaar, Brouwer, Busschbach, 2004).

19.b. Age, gender, ethnicity, sexual preference and social-economic status

There is a broad consensus that age, gender, ethnicity, sexual preference and economic status may not be used primarily to regulate rights to care. Cost-effectiveness analyses were developed within an environment that has traditionally been meticulous in ensuring the exclusion of these factors. This resulted in the credo:

"A QALY is a QALY, no matter who gets it" (Alan Williams, 2001)

One exception to this exists in situations in which scientific research has demonstrated that the efficacy of an intervention – due to one of these characteristics – is greater or smaller than in that of other groups in relation to that characteristic and that these differences cannot be explained by other factors (RVZ, 2007, page 19). The RVZ illustrated this with the example:
In 2005 the American Food and Drug Administration (FDA) registered BiDil as a medicine for heart failure in negro people, after research proved that the number of deaths among these heart failure patients fell by 43% and the number of hospital admissions fells by 39% as a result of taking this medicine. The medicine had little effect on other ethnic groups. As a result the medicine is only reimbursed when it is prescribed to people of the negroid race (RVZ, 2007, pages 19-20).

The RVZ example happens to involve an outcome whereby no advantage was to be gained by the disadvantaged groups. The question remains whether this argument would still stand if such advantages did exist, however limited they may be. A well-known example is that of older patients. One might expect the outcomes of interventions in older people to be less effective, even if only because an older patient has fewer remaining years left. Nevertheless, explicitly referring to age as a criterion causes a great deal of unrest. For example, the discussions towards the end of the nineteen-eighties about imposing a 55-year age limit for heart transplants. Imposing such an age limit would make it possible to allocate the scarce

donor hearts to patients who would be able to live longest with them. Schuyt (1990) predicted that this would lead to a vehement discussion and he was right. Age as an explicit criterion eventually disappeared from the agenda, although research had shown that being older is clearly a risk factor (Simoons and Weimar, 1990).

The heart transplant example illustrates the unsuitability of explicitly using age as a criterion. However, when embedded in other medical factors, age is interpreted in the way indicated by the RVZ above. This is illustrated by the cholesterol guidelines, whereby the elderly are excluded and whereby different criteria are used for men and women on the basis of effectiveness criteria.
"The working group recommended not to prescribe statins to men older than 70 and women older than 75 years. This should not be interpreted as age discrimination" (Casparie, Van Hout, Simoons, 1998; page 2076).

In other words it is possible to allow age, gender, etc. to play a role in the cost-effectiveness of an intervention and it leads to acceptable differences in treatment. Where no relationship exists with (cost-)effectiveness, then the use of these criteria would not seem to be in keeping with the egalitarian nature of a social health insurance system. These facts are also clearly supported by the Citizens' council of NICE, as can be seen from the statement:
There is much debate over whether, or how, age should be taken into account when allocating healthcare resources. The Citizens' Council decided that health should not be valued more highly in some age groups than in others and that social roles at different ages should not affect decisions about cost effectiveness. They said, though, that where age is an indicator of benefit or risk, it can be taken into account.
(NICE, 2007).

20. References

Buxton, M. Looking for willingness-to-pay (WTP) threshold for a QALY – does it make sense? A practical view. ISPOR Connections 2007;13:9-11

Cairns J. Providing guidance to the NHS: The Scottish Medicines Consortium and the National Institute for Clinical Excellence compared. Health Policy. 2006 Apr;76(2):134-43

Casparie AF, Van Hout BA, Simoons ML. Guidelines en kosten. Nederlands Tijdschrift voor de Geneeskunde 1998;142:2075-77

Claxton K, Briggs A, Buxton MJ, Culyer AJ, McCabe C, Walker S, Sculpher MJ. Value based pricing for NHS drugs: an opportunity not to be missed? BMJ. 2008 Feb;336(7638):251-4

Cohen JP, Stolk E, Niezen M. Role of budget impact in drug reimbursement decisions. J Health Polit Policy Law. 2008 Apr;33(2):225-47.

Culyer AJ, McCabe C, Briggs A, Claxton K, Buxton M, Akehurst R, et al. Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. J Health Serv Res Pol 2007;12:56-8

CVZ. College voor Zorgverzekeringen. Toenders WGM. Breedte Geneesmiddelenpakket. Uitwerking criteria noodzakelijkheid, eigen rekening en verantwoording en lifestyle. Publicatienummer 01/05, 27 april 2001, College voor Zorgverzekeringen, Amstelveen.

CVZ. College voor Zorgverzekeringen. Rapport pakketbeheer in de praktijk. Mastenbroek, C, G., Van der Meer, FM, Zwaap, J, Rikken, F, Polman, P. Diemem: College voor Zorgverzekeringen, 2006.

CVZ. College voor Zorgverzekeringen. Guidelines voor pharmacoeconomic research, geactualiseerde versie. Geldig vanaf 1 april 2006. Volgnummer 25001605. College voor zorgverzekeringen, Diemen Maart 2006

Day B. A meta-analysis of wage-risk estimates of the value of statistical life. Centre for Social and Economic Research on the Global Environment. London: University College London, 1999.

Dear, J, O'Dowd, C, Timoney, A, Paterson, KR, Walker, A, J., WD. Scottish Medicines Consortium: an Overview of Rapid New Drug Assessment in Scotland. Scottish Medical Journal 2007;52:20-26.

Devlin, Parkin. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Econ 2004;13(5): 437-52

Flynn TN, Louviere JJ, Marley AA, Coast J, Peters TJ. Rescaling quality of life values from discrete choice experiments for use as QALYs: a cautionary tale. Popul Health Metr. 2008 Oct 22;6(1):6.

Gyrd-Hansen, D. Looking for willingness-to-pay (WTP) threshold for a QALY – Does it make sense? A critical view. ISPOR Connections, 2007 Aug;13(4):5-8
Gilbert D. Lifestyle drugs: Who will pay? Scrip report. London: PJB Publications; 1999.

Hirth RA, Chernew ME, Miller E, Fendrick AM, Weissert WG. Willingness to pay for a quality-adjusted life year: in search of a standard. Med Decis Making. 2000 Jul-Sep;20(3):332-42.

House of Commons, HC. National Institute for Health and Clinical Excellence (NICE). First Report of Session 2007–08. 2007.

Jönsson B. IQWiG: an opportunity lost? Eur J Health Econ. 2008 Aug;9(3):205-7.

Krauth C, John J, Aidelsburger P, Brüggenjürgen B, Hansmeier T, Hessel F, Kohlmann T, Moock J, Rothgang H, Schweikert B, Seitz R, Wasem J. AG Methoden der gesundheitsökonomischen Evaluation. Comments by the Working Group for Methods of Economic Evaluation in Health Care (AG MEG) to IQWiG's Draft Guidelines "Methods for Assessment of the Relation of Benefits to Costs in the German Statutory Health Care System" [Article in German] Gesundheitswesen. 2008 Jun;70(6):e1-16.

Martin S, Rice N, Smith PC. The link between health care spending and health outcomes: evidence from English programme budgeting data. University of York: Centre for Health Economics, 2007. (Research Paper 24.)

McCabe C, Claxton K, Tsuchiya A. "Orphan drugs and the NHS: should we value rarity?" BMJ 331(7523): 1016-9.

Meerding WJ, Polder JJ, Hollander AEM de, Mackenbach JP. Hoe gezond zijn de zorguitgaven?. Zorg voor euro's - 6. De kosten en opbrengsten van gezondheidszorg bij infectieziekten, kankers en hart- en vaatziekten. Zorg voor Euro's – 6. RIVM Rapport 270091002. ISBN: 978-90-6960-174-8, 2007

Multidisciplinaire richtlijn Cardiovasculair risicomanagement 2006. Kwaliteitsinstituut voor de Gezondheidszorg CBO en Nederlands Huisartsen Genootschap. ISBN-10: 90-8523-131-0. www.cbo.nl

NICE: National Institute for Health and Clinical Excellence. Guide to the Methods of Technology Appraisal. London: NICE, 2003

NICE, 2007. Social value judgements. Principles for the development of nice guidance. Second draft: National Institute for Health and Clinical Excellence (NICE), 2007 November 2007.http://www.nice.org.uk/media/998/50/SVJ2ForPublicConsultation.pdf

NICE, 2008. Report on NICE Citizens Council meeting Quality Adjusted Life Years (QALYs) and the severity of illness 31 January – 2 February 2008. 16 July 2008. http://www.nice.org.uk/media/2C3/31/CCReportOnQALYsAndSeverity.pdf.

Raad voor de Volksgezondheid en Zorg. Zinnige en duurzame zorg. Raad voor de Volksgezondheid en Zorg, Den Haag. 2006

Raad voor de Volksgezondheid en Zorg. Rechtvaardige en duurzame zorg. Publicatienummer 07/04. ISBN-13: 978-90-5732-183-2, Raad voor de Volksgezondheid en Zorg, Den Haag. 2007

Schuyt, C. J. M. (1990). "Tragische keuze." Volkskrant (Open Forum) 14/2/1990

Simoons, ML, Weimar W. (1990). Indicatiestelling en selectie bij harttransplantatie. Bulletin Nedelandse Transplantatie Vereniging, 1, nr. 3, december 1990

Stolk EA. Equity and Efficiency in Health Care Priority Setting: how to Get the Balance Right? Doctoral thesis, Erasmus University Rotterdam, 2005

Stolk EA, Brouwer WDF, Busschbach JJV. Rationalising rationing: economic and other considerations in the debate about funding of Viagra. Health Policy 2002;59:53-63

Stolk EA, van Donselaar G, Brouwer WB, Busschbach JJ. Reconciliation of economic concerns and health policy : illustration of an equity adjustment procedure using proportional shortfall. Pharmacoeconomics. 2004;22(17):1097-107.

Towse A, Raftey J. Should NICE's threshold range for cost per QALY be raised? Yes / No BMJ. 2009;338:268-9.

Verkerk MA, Busschbach JJV, Karssing ED. Health-related quality of life research and the capability approach of Amartya Sen. Quality of life Research 2001;10:49-55

Williams, A. The 'fair innings argument' deserves a fairer hearing! Comments by Alan Williams on Nord and Johannesson. Health Econ 2001;10:583-585.

APPENDIX

List of abbreviations

ACP *Adviescommissie Pakket* [Package Advice Committee]
CFH *Commissie Farmaceutische Hulp* [Medicinal Products Reimbursement Committee]
CVZ *College voor zorgverzekeringen* [Health Care Insurance Board]
FDA Food and Drug Administration
IQWiG Institute for Quality and Efficiency in Health Care
NICE National Institute for Health and Clinical Excellence
PCT Primary Care Trust
QALY Quality-adjusted life-year
RvB *Raad van Bestuur* [Executive Board]
RVZ *Raad voor de Volksgezondheid en Zorg* [Council for Public and Health Care]
SMC Scottish Medicines Consortium
VWS *Volksgezondheid, Welzijn en Sport* [Ministry of Health, Welfare and Sport]
WHO World Health Organisation