

# A missing cornerstone in the Norwegian Priority Commission's weighting scheme

– Sub-treatment balancedness is a necessary property for priority setting criteria

MATHIAS BARRA <sup>1</sup>  
KIM RAND-HENDRIKSEN <sup>1, 2</sup>

<sup>1</sup> Health Services Research Center, Akershus University Hospital, Lørenskog, Norway

<sup>2</sup> Department of Health Management and Health Economics, Faculty of Medicine,  
University of Oslo, Oslo, Norway

**Abstract:** The Norwegian government recently put in place a priority commission tasked with suggesting a set of explicit criteria for priority setting in the health care sector. The commission suggested three criteria, the first two of which equate to cost-effectiveness, where, essentially, the gain is measured in terms of Quality Adjusted Life Years (QALYs). The third criteria specifies that the number of QALYs be multiplied by a factor depending on the total health loss – also measured in QALYs – without the treatment in question. In this paper, we will show that the suggested weighting scheme creates situations in which the priority of treatment programmes will change based on arbitrary bundling (where two or more treatments are combined into one) or sub-divisions (where a treatment is split up into two or more components.) We show that these types of problems can be avoided or ameliorated if the QALY weighting scheme satisfies a property which we call sub-treatment balanced – informally, that the total weighted QALY gain is preserved when treatments are bundled or sub-divided. To our best knowledge, this property has not previously been discussed in the priority setting literature. We demonstrate that sub-treatment balance can easily be achieved in general, and in particular we show how to adapt the weighting scheme suggested by the Norwegian priority commission in order to satisfy this sub-treatment balance. Finally, we argue that any weighting scheme used in health care priority setting should be sub-treatment balanced with respect to any other attribute of a treatment which policy makers would want to take into account when making their decisions. At the time of writing, the Norwegian government has yet to conclude on a final set of criteria for prioritization, and a task-group, led by professor Jon Magnussen, is re-evaluating the severity criterion suggested by the priority commission. However, sub-treatment balance is still relevant, as it should be required of any weighing scheme, and is crucial given that (i) the criterion results in weighting QALYs, and (ii) if the selected measure of severity is affected by the administered treatment.

**JEL Classification:** I1, H51

**Key words:** priority setting, cost-utility, QALY, criteria, sub-treatment balanced

---

\* *Correspondence to:* Mathias Barra, Health Services Research Center, Akershus University Hospital, Postboks 1000, 1478 Lørenskog, Norway. E-mail: Mathias.Barra@ahus.no

## 1 Introduction

Imagine two patients, Alice and Beth. At the age of 20, Alice was diagnosed with a condition that would be debilitating without treatment. The condition could be kept at bay through the continuous use of a costly treatment regimen, in which case Alice would remain asymptomatic. However, due to its high cost, the treatment for the condition did not reach the threshold for cost-effectiveness in terms of cost per quality-adjusted life year (QALY – see section 1.1).

Imagine that the criteria for priority setting specified that prior to computing the cost-effectiveness of a treatment, the expected health gain (in terms of QALYs gained) was to be multiplied by a severity factor based on the *ex ante* health loss in the absence of treatment, i.e. the health-loss (also measured in QALYs) that would be incurred if no treatment was administered<sup>1</sup>. Without the treatment, the expected health loss for Alice, beginning at the age of 20 and stretching to a life expectancy of more than 80 years, would be substantial. After applying this severity weight, the costly treatment for the condition was then considered cost-effective for Alice, and she was offered the treatment for the remainder of her life.

Fast-forward forty years. Alice and Beth are both 60 years old. Alice has been receiving treatment for her condition continuously, so that she has not experienced any health loss. Until now, Beth has been healthy, and has also had no health loss. Now, at the age of 60, Beth is diagnosed with the same condition, and the unweighted cost per QALY is still prohibitive. Without treatment, the *ex ante* health loss for Beth is substantial, but much less so than it was for Alice forty years previous. Due to this smaller expected health-loss, the health loss-weighted cost per QALY is still too high, and therefore Beth is not offered the treatment.

Importantly, at the age of 60, Beth and Alice have had the same experienced health-loss so far (none), and have the same future prospects with and without treatment. However, while Alice will continue to get treatment, Beth will remain untreated. This apparent difference occurs because the remaining QALY-gain from the treatment for Alice is still multiplied by the severity factor first calculated on the basis of her expected health loss at the age of 20, while the identical expected QALY-gain for Beth is multiplied by a severity factor that is calculated on the basis of her expected health loss at the age of 60.

This example is hypothetical. However, a set of priority criteria exhibiting this kind of problem have recently been suggested by a commission on priority settings in health care in Norway. The suggested implementation of the weighting scheme aimed at prioritizing treatment based on ‘severity’ implies that the priority of any specific treatment will change, often dramatically, if it is split into sub-treatments or if it is bundled with other treatments. In this particular case, the proposed scheme means that combining any two treatments will always result in a higher priority than those same treatments viewed separately, so that e.g. in the example, Alice gets treatment also after the age of 60, when her case is indistinguishable from Beth’s, because, in the example, we have bundled together her life-long treatment. If Alice’s case were to be sub-divided into two distinct treatments considered separately, one before the age of 60 and one after, she would, like Beth, not receive treatment after the age of 60.

In this paper, we explore the properties and conditions under which severity weighting schemes result in such inconsistencies, and describe how such problems can be mitigated or avoided.

---

<sup>1</sup> *Ex ante health loss* is a key concept in this work. It means the expected health loss incurred in the absence of a treatment. In contrast *ex post health loss* refers to the remaining health loss after a treatment that does not fully restore health.

### 1.1 Background – Priority in health care

The legitimacy of the publicly financed health care systems of Northern Europe is under pressure: there is an ever-increasing gap between the medical frontier and budgetary constraints, and new beneficial treatments are not necessarily cost-effectiveness. To ameliorate this situation, clear policy practises for the priority setting of treatments are needed. Furthermore, these must both be acceptable and be perceived as fair by the general public.

In June 2013, a body of experts, politicians and representatives of other stake holders' organizations were appointed by the Norwegian government to deliver a so-called *Norwegian Official Report*<sup>2</sup> (NOU). This body – called *the Priority Commission*<sup>3</sup> – had as its mandate to clarify the existing (and, where needed, to chisel out new) guidelines for priority setting in the Norwegian public health care system. The work behind the NOU was led by Professor Ole Frithjof Norheim, and the resulting NOU *Transparent and Fair – Priority Setting in the health services* (Norheim et al., 2014), hereafter referred to as *The NOU*, was submitted to the Norwegian Ministry of Health on November 12<sup>th</sup> 2014.

The final legal status of a Norwegian Official Report is not determined until the policy process of which it is part has concluded; it may result in a new legislation (enactment), or it may be abandoned (Ringard et al., 2012, p. 24). When an NOU is enacted, it is an important source source for interpreting the resulting act – in particular during an initial period, when the intention behind the law is to be interpreted by bureaucrats, lawyers, and judges.

The NOU discusses a range of subjects and issues related to priority setting in the health care sector. The mandate of the commission was to describe and set out principles of, criteria for, and the means of implementing priority settings within the Norwegian health care system.

The commission was also tasked with revising the criteria for the priority settings currently in place. The criteria originate from the last commission on priority setting, the *Lønningutvalget II*'s report (Lønning et al., 1997). These criteria were, to put it bluntly, *cost-effectiveness with a glance towards 'severity'*.

The NOU recommended substituting the existing severity criterion with an explicit weighting of the severity of a condition in terms of the *health loss* the condition incurs.

Both gains and losses are measured in *quality adjusted life years (QALYs)*<sup>4</sup>. For the work presented here, only a superficial understanding of the QALY-paradigm is needed: the QALY integrates longevity and Health Related Quality of Life into a single numerical measure of health. The QALY thus reflects both improvements in survival time and improvements in health status. The interested reader may see e.g. Pliskin et al. (1980) or Weinstein et al. (2009) for an introduction to the topic.

Unfortunately, the way the NOU has suggested to implement the severity weighting will lead to inconsistencies of the type presented in the initial example of Alice and Beth. In this paper we explain how and why these problems occur, and present a simple solution.

<sup>2</sup> In Norwegian: Norsk offentlig utredning.

<sup>3</sup> In Norwegian: Prioriteringsutvalget.

<sup>4</sup> The NOU actually use the slightly more general term Good Life Years (GLYs); see section 4 (Discussion and Conclusion) for more on this distinction. However, for the work presented here, the distinction is of no importance.

### 1.1.1 The Priority Commission's criteria for priority setting

Chapter seven of the NOU details the suggested criteria for priority setting. The NOU proposes three criteria for setting priorities for different treatments<sup>5</sup>  $T_i$  offered by the Norwegian health care system:

- Resources: the cost  $C_i$  of treatment  $T_i$ ;
- QALY gain: the expected health gain  $q_i$  associated with treatment  $T_i$ ; and
- QALY loss: the expected *ex ante* QALY loss  $\ell_i$  suffered by the recipient of treatment  $T_i$ .

Taken together, the first two items are standard cost-effectiveness: in the choice between otherwise equally resource-intensive treatments, the one which procures the greatest QALY gain should be preferred.

The third criterion states that *some* weight should be attributed to the magnitude of the QALY loss suffered by the recipient of the QALY gain. This means that if two individuals stand to obtain equal QALY gains  $q$ , then the one with the largest QALY loss should be given priority.

### 1.1.2 Health loss – the devil is in the details

In the following section, we first set down some notation to use in discussing this issue, and summarize what the NOU writes about it.

The NOU's chapter on the health loss criterion (Norheim et al., 2014, Chap. 7.5 pp. 93–96), discusses several ways of approaching the health loss issue. The NOU decided on health losses being defined as *the life-long expected health loss, ex ante the treatment being considered; ex post treatment already received*.

The NOU (Norheim et al., 2014, section 7.6.2) thus recommends that in evaluating a treatment  $T_i$  in a priority setting context, the standard CUA-fraction<sup>6</sup>  $\frac{q_i}{C_i}$  should be replaced with an expression of the form  $\frac{\phi(\ell_i, q_i)}{C_i}$  which incorporates the health loss-parameter  $\ell_i$  in addition to the health gain-parameter  $q_i$  to yield a health loss-weighted QALY-gain  $\phi(\ell_i, q_i)$ . The health loss  $\ell_i$  is composed of two parts:  $\ell_i = \ell_i^0 + \ell_i^1$  where  $\ell_i^0$  represent the health loss already experienced in the past, while  $\ell_i^1$  represents the expected future health loss *ex ante* the treatment  $T_i$ . For example, the total health loss  $\ell$  associated with a kidney-transplant operation, is the sum of the estimated health loss  $\ell_0$  already experienced at the time of the operation, plus the expected future health loss  $\ell_1$  without (i.e. *ex ante*) the operation.

The NOU suggests that the function  $\phi(\ell_i, q_i)$  should take the form  $W(\ell_i) \cdot q_i$  for a suitable weighting-function  $W$ . The function  $W$  outlined in the NOU is a stepwise increasing function, where<sup>7</sup>  $W(\ell)$  is 1 for  $\ell < 15$ ;  $W(\ell) = 2$  when  $15 < \ell < 30$ , and  $W(\ell) = 3$  when  $30 < \ell$  (Norheim et al. 2014, p. 98).

In all fairness, the NOU does hint at problems with *ex ante* health loss, and suggests that ‘*the most precise way [of weighting health gains] consists in weighting each unit of health gain (e.g., a good life year) by the size of the health loss at the time the beneficiary*

<sup>5</sup> We will use the generic term treatment throughout the paper to refer to any treatment, procedure, intervention, medicine regime, rehabilitation, etc. which consumes resources and which is aimed at providing QALY gains for patients.

<sup>6</sup> We thus express cost-effectiveness as *QALYs per unit of cost* here: in the priority setting arena, this means that a larger value means higher priority.

<sup>7</sup> That is  $W(\ell) = \left\lceil \frac{\ell}{15} \right\rceil$  where  $\lceil \cdot \rceil$  rounds up to the next integer.

receives the health gain'<sup>8</sup>, but asserts that to weight the health gains directly *on the margin* would be demanding<sup>9</sup> (Norheim et al., 2014, p. 97). In this work, we show that this is not the case.

### 1.1.3 The introduction formalized

**Definition 1 (QALY weighting function)** Let  $T$  be some available treatment, which is expected to provide a QALY gain  $q$  to a patient with expected ex ante QALY loss  $\ell$ . We always assume  $\ell \geq q$ , and both  $q$  and  $\ell$  refer to, respectively, QALYs gained with the treatment, and total, life-long QALYs lost, ex-ante the treatment.

A QALY weighting function (QWF) is a function  $u(\ell, q)$  which satisfies:

1.  $u(\ell, 0) = 0$  zero gain – zero weighted gain;
2.  $q > q' \Rightarrow u(\ell, q) > u(\ell, q')$  greater gain – greater weighted gain;
3.  $\ell \geq \ell' \Rightarrow u(\ell, q) \geq u(\ell', q)$  at least equal loss – at least equal priority.

A QWF is trivial if  $u(\ell, q) = u(\ell', q)$  for all  $q, \ell$  and  $\ell'$ .

This weighted QALY gain is to be used in connection with cost-effectiveness analysis. More specifically, it is common to operate with a monetary threshold value, representing the maximal (*societal*) willingness-to-pay (WTP) for one additional QALY gained. For example, in Norway, an unofficial threshold value of NOK 500 000 is often used; thus, a gain of 2 QALYs would be considered cost-effective below a total cost of NOK 1 000 000.

The interpretation of the terms in Definition 1 is that if  $T$  is an available treatment, so that  $T$  yields an unweighted QALY gain of  $q$  and is such that the expected life-long ex ante QALY loss of beneficiaries of  $T$  is  $\ell$ , then the loss-weighted QALY gain is  $u(\ell, q)$ . Furthermore, given a WTP-value for one additional *loss-weighted* QALY, and if we adopt the convention of expressing the cost  $C$  of treatment  $T$  in units of the societal WTP threshold value per QALY, then  $T$  is considered to be cost-effective when  $u(\ell, q) \geq C$ , since this inequality obtains exactly when  $1 \geq C/u(\ell, q)$ , i.e. when the cost per weighted QALY is less than the threshold WTP. For example, using the unofficial Norwegian threshold value of NOK 500 000 as the unit for the cost  $C$ , a treatment  $T$  costing NOK 2.5 million (equals 5 times the WTP-threshold, i.e.  $C = 5$ ) and providing a weighted gain of 3 QALYs, we have  $3 < 5$ , and  $T$  would *not* be considered cost-effective. If  $T$  is instead provided at the cost of NOK 1.25 million ( $C = 2.5$ ), we see by  $3 \geq 2.5$  that  $T$  achieves cost-effective relative to the NOK 500 000 WTP threshold.

### 1.1.4 Main aim and outline of this work

It seems as if a false dichotomy has been set up by the NOU, in which one must either compute QALY loss ex ante *or* ex post: by pointing to problems associated with computing ex post, they conclude that ex ante is the better solution. It is not; the better solution is computing losses on the margin. This means that weight assigned to the health gain changes continuously because the remaining ex ante health loss is reduced as health-improvement is achieved. Another way to view this is to say that the treatment is subdivided into

<sup>8</sup> 'Den mest presise måten består i å vekte hver enkelt gevinstenhet (for eksempel et godt leveår) etter størrelsen på helsetapet hos gevinstmottakeren på det tidspunktet hun eller han mottar den enheten.'

<sup>9</sup> 'Direkte vekting av hver gevinstenhet er krevende.'

infinitesimal improvements, each of which is weighted by the ex ante health loss of that infinitesimal ‘slice’ of the treatment. Contrary to the NOU’s comments, we will show that this calculation is neither difficult nor impractical using simple integrals.

In the rest of this paper, we first give some examples demonstrating serious problems with the QALY loss weighting scheme as proposed in the NOU. We will show that unless the weighting scheme satisfies a simple property – to be dubbed *sub-treatment balanced* – the consequences will be undesirable: health care providers will have perverse incentives, and the absence of this property will imply what we regard as unfair priority setting policies. Next we will show that if the weighting scheme is a function of health loss alone,  $u(\ell, q) = W(\ell) \cdot q$ , it cannot be balanced. We will then provide a very simple, albeit general, solution to some of these problems.

The relationship between the loss-weighting in the NOU and the solutions developed here is described, and we also prove a simple result about cost-effective solutions under any weighting scheme.

We have not aimed to review the severity literature in any detail. The NOU has been our primary source of references, and we will not go beyond that in this paper. However, although we have not read all of the works cited in the NOU (the reference list covers 16 pages), we have not been able to find our solution presented in any other work from this literature.

## 2 Potentially problematic properties of QALY weighting functions

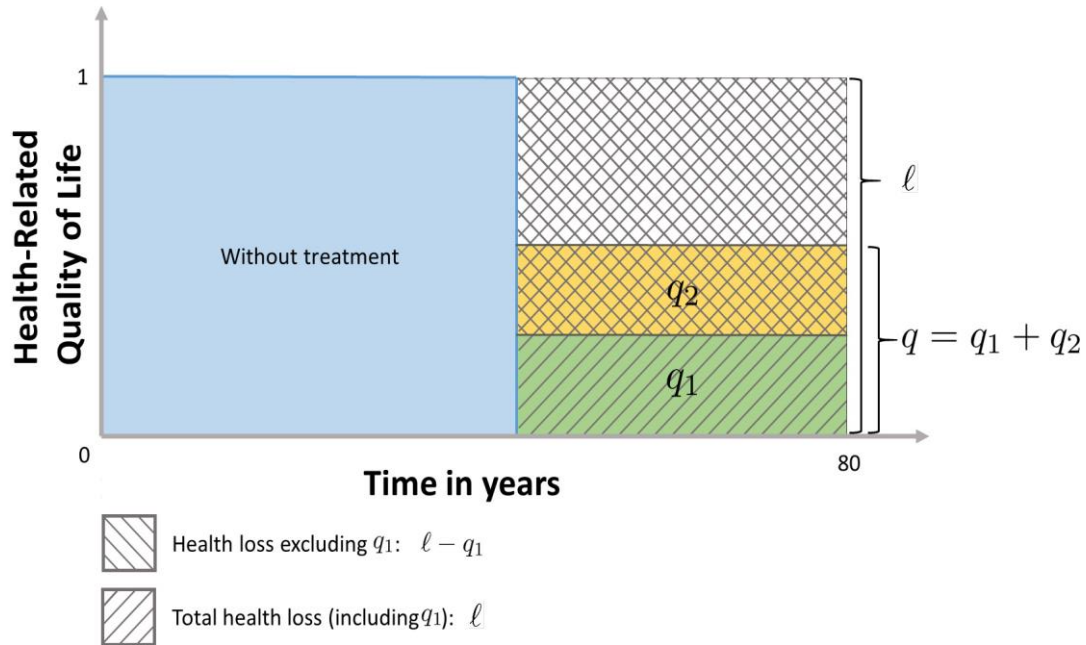
In this section, we start by providing two short examples illustrating problems with the weighting scheme suggested in the NOU, before we describe the problem more formally.

**Example 1** Consider treatments  $T_1$  and  $T_2$  targeting a patient group providing unweighted QALY gains of  $q_1 = 15$  and  $q_2 = 15$ . For this example, let us require that  $T_2$  cannot be administered unless  $T_1$  has been provided first. Figure 1 illustrates this example. The initial ex ante QALY loss for the patient group is  $\ell = 30$ . Using the weighting scheme  $W$  suggested in the NOU, treatment  $T_1$  will provide  $15 \cdot W(30) = 15 \cdot 2 = 30$  weighted QALYs. After  $T_1$  have been performed, the remaining QALY loss (ex ante for  $T_2$ ) is only 15, and  $T_2$  will provide  $15 \cdot W(15) = 15 \cdot 1 = 15$  weighted QALYs. Let the WTP per weighted QALY be 1, and let the cost for each of the treatments be 25. Since  $T_1$  provides 30 weighted QALYs, the cost per weighted QALY falls below 1, and it is considered cost-effective.  $T_2$  provides only 15 weighted QALYs, each at a cost of more than 1, and it is therefore not considered cost-effective.

Consider what happens if we bundle  $T_1$  and  $T_2$  into one larger treatment unit,  $T$ , providing 30 unweighted QALYs at a combined cost of 50. With an ex ante QALY loss of 30, the weighted QALY benefit from  $T$  is  $30 \cdot W(30) = 30 \cdot 2 = 60$ . Since the cost for each of the weighted QALYs is below 1,  $T$  is considered to be cost-effective. Thus, by the mere act of bundling  $T_1$  and  $T_2$ , the sub-treatment  $T_2$  has suddenly become cost-effective. As a second observation, consider the situation if  $T_1$  had a total

cost of 35 and  $T_2$  had a total cost of 25. Considered separately, neither treatment would reach the threshold value for cost-effectiveness, but bundled into  $T$ , thereby providing 60 weighted QALYs, both would be accepted.

**Figure 1: Graphical illustration of the sub-treatment problem**



Note: Graphical illustration of the sub-treatment problem: Prior to any treatment, the health loss ( $\ell$ ) is the whole right area of the figure, including  $q_1$  and  $q_2$ . Following sub-treatment  $q_1$ , the remaining health loss is reduced to  $\ell - q_1$ . If  $q$  is considered as a whole, the benefit of the two sub-treatments will be weighted by  $\ell$ , but if considered separately,  $q_1$  will be weighted by  $\ell$  and  $q_2$  will be weighted by  $\ell - q_1$ .

The first observation is problematic, while the second borders on the absurd. Clearly these phenomena provide perverse incentives for private health care providers.

The problem with assigning weights to QALY gains based on the ex ante magnitude of a non-health gain quality, has also been noted by Hope et al. (2010) in a paper focusing on various strategies for the weighting of a primary health outcome by orthogonal considerations (Hope et al. (2010) studies resource allocations in a context of *needs*), and they specifically point to the problem of value-boosting by bundling treatments. It is worth noting that the present work can also be adapted to a solution to this problem in the context of Hope et al. (2010).

This kind of weighting scheme would result in over-investment in treatments when bundled together and under-investments in treatments targeting patients with lower health-losses. These effects may be amplified by not loss-weighting on the margin. However, the situations described above can be ameliorated if the QWF satisfy what we call *sub-treatment balancedness*.

## 2.1 Formal examples and description of the problem

Recall that the weighting of a QALY gain obtained through a treatment should be weighted with the QALY loss ex ante this treatment. This means that if the patient stands to lose  $\ell$  QALYs without treatment, and to gain  $q \leq \ell$  QALYs with the treatment, then the weighted

QALY gain become  $u(\ell, q)$ . By extension, if a patient stands to gain, say 3 and 9 QALYs from two subsequent treatments  $T_1$  and  $T_2$ , and stands to lose 20 QALYs in total without any treatment, the weighted QALY gain should be

$$\underbrace{u(20,3)}_{T_1 \text{ with ex ante QALY loss } 20} + \underbrace{u(17,9)}_{T_2 \text{ with ex ante QALY loss } 20-3=17} \quad (1)$$

We also see that if the two treatments were bundled into one treatment  $T$  – with a total unweighted QALY gain of  $12 = 3 + 9$ , then the weighted QALY gain would be computed as  $u(20,12)$ . In this example the total unweighted QALY gain is  $q = 12$ , and the two partial gains are  $0.25q$  and  $0.75q$ , respectively.

More generally, if two treatments  $T_1$  and  $T_2$  provide QALY gains of  $pq$  and  $(1-p)q$  respectively (for  $0 < p < 1$ ), yielding a total QALY gain of  $q = pq + (1-p)q$ , then it is interesting to compare the various total *weighted* QALY gains relative to some QWF  $u$ .

$$\text{Combined treatment: } u(\ell, q) \quad (2)$$

$$\text{First } T_1, \text{ then } T_2: u(\ell, pq) + u(\ell - pq, (1-p)q) \quad (3)$$

Since both values (2) and (3) in a sense represent the same total QALY gain relative to the same ex ante QALY loss, one could argue that the weighted QALY totals should be equal. If this is not the case, then health care providers could manipulate the willingness to pay by arbitrarily bundling (combining treatments), splitting up treatments, or changing the order in which the treatments are administered.

The most simple QWF is  $u(\ell, q) = q$ . This QWF depends only on  $q$  (it is trivial), and it is clear that with this QWF the three different weighted totals above are all equal. In fact when the unweighted QALY gain is used,  $u(\ell, pq) = pq = p \cdot u(\ell, q)$ . We next briefly discuss some consequences for situations when equality between (2) and (3) does not hold.

Assume that a treatment  $T$  provides a QALY gain of  $q$  QALYs, to someone with an ex ante QALY loss of  $\ell \geq q$ . This implies that  $T$  could be priced at  $C_T = u(\ell, q)$  and therefore be adopted. Assume next that  $u(\ell - (1-p)q, pq) < p \cdot u(\ell, q)$  for some  $p \in [0,1]$ . Say  $p = 0.5$ ; then if  $T$  is split up into treatments  $T_1$  and  $T_2$  providing subsequent gains of  $0.5q$  QALYs each, then  $T_2$  would *not* be considered cost-effective if priced at  $0.5C_T$ . This might be acceptable; indeed, all the weighting schemes we consider here will have this property. However, we may hope that if  $T$  is cost-effective at cost  $C_T$ , then for any  $p \in [0,1]$ , if  $T$  is split up into two subsequent treatments  $T_1$  and  $T_2$  providing QALY gains of  $pq$  and  $(1-p)q$ , then the cost  $C_T$  can be allocated between  $T_1$  and  $T_2$  so that both sub-treatments remain cost-effective.

Assume next that  $u(\ell - (1-p)q, pq) \geq p \cdot u(\ell, q)$  for all  $\ell, q$  and  $p \in [0,1]$ . By non-triviality, we may then choose  $\ell, q$  and  $p$  such that  $u(\ell, pq) > u(\ell - (1-p)q, pq)$ . Then, we see that

$$u(\ell, pq) + u(\ell - pq, (1-p)q) > p \cdot u(\ell, q) + (1-p) \cdot u(\ell, q) = u(\ell, q).$$

In other words, one may increase the total weighted QALY gain of a treatment by splitting it into sub-treatments. This seems like an even more undesirable property than the one discussed immediately above. It should be noted that none of the QWFs proposed either below or in the NOU have this property.



### 3 Solution to the problem

In this section we first describe a criterion for QWFs which ensure that some of the problems discussed in Section 2 are either solved or ameliorated. We next describe how QWFs satisfying this criterion are constructed.

#### 3.1 The sub-treatment balanced priority setting criterion

We always assume that  $0 \leq q \leq \ell$ ; the QALY loss ex ante  $\ell$  is at least as great as the potential QALY gain  $q$  of the treatment under consideration.

**Definition 2 (Sub-treatment balanced)** A QWF  $u(\ell, q)$  is sub-treatment balanced when, for all  $p \in [0, 1]$ :

$$u(\ell, q) = u(\ell, pq) + u(\ell - pq, (1 - p)q) \quad (4)$$

For example,  $T_1$  and  $T_2$  may be two doses of a medicine, taken at two future points in time (e.g., pain killers for rheumatic disorder), or they could be two surgical interventions performed at the same time independently of each other (e.g., mastectomy  $T_1$  possibly followed by reconstructive surgery  $T_2$ ). If  $t_1$  and  $T_2$  provide unweighted QALY gains of  $pq$  and  $(1 - p)q$ , respectively, and therefore offer  $q$  QALYs in total, then equation 4 now becomes:

$$\underbrace{u(\ell, q)}_{\text{Combination of } T_1 \text{ and } T_2} = \underbrace{u(\ell, pq)}_{T_1 \text{ with ex ante health loss}} + \underbrace{u(\ell - pq, (1 - p)q)}_{T_2 \text{ with ex ante health loss}} \quad (5)$$

If the QWF is balanced, it is impossible to make a treatment more valuable by bundling. It is also always possible to take a cost-effective treatment  $T$  and relocate the costs between any two sub-treatments  $T_1$  and  $T_2$  in such a way that both become cost-effective. Indeed, when  $u$  is sub-treatment balanced, then

$$u(\ell - (1 - p)q, pq) \leq p \cdot u(\ell, q) \leq u(\ell, pq). \quad (6)$$

In less mathematical terms, balance is the requirement that any treatment that can be construed as consisting of more than one sub-treatment will have the same overall priority as the combined priority of its constituent sub-treatments. Without this property, arbitrary grouping or sub-division of treatment options will have a separate impact on priority setting. We have not been able to come up with any set of ethical guidelines that are congruent with the notion that bundling is in itself an ethically relevant factor to consider for priority setting.

The next result shows that any weighting-scheme like the one proposed by the NOU fails to be sub-treatment balanced. This both illustrates a serious shortcoming of the proposed health-loss criterion, and justifies the development of a family of balanced QWFs.

**Proposition 1** Let  $W$  be a positive monotonically increasing function, so that

$$u(\ell, q) = W(\ell) \cdot q \quad (7)$$

is a QWF. Then  $u$  is either not sub-treatment balanced, or  $W$  is constant.

*Proof:* Let  $0 \leq q \leq \ell$ , and assume  $u(\ell, q) = W(\ell) \cdot q$  is sub-treatment balanced. Then

$$W(\ell) \cdot q = W(\ell) \cdot pq + W(\ell - pq) \cdot (1 - p)q$$

which after some simple algebra yields  $W(\ell) = W(\ell - pq)$ . Because  $p \in [0, 1]$  and  $0 \leq q \leq \ell$  were all arbitrary, it follows that  $W$  is constant. QED

If  $W$  is constant, then  $W(\ell) \cdot q$  is a trivial QWF. Thus, Proposition 1 above means that no multiplicative decomposition of a balanced QWF  $u(\ell, q)$  into a purely ‘health loss-weighted health gain’ exists.

We end this section with a theorem which shows that a sub-treatment balanced QWF deals with the problems from Example 1 insofar as possible: (i) if a treatment is cost-effective, then it remains so when it is split up into sub-treatments – possibly after distributing the costs according to weight – and, (ii) there is no cost-effectiveness bundling of cost-ineffective treatments,

**Theorem 1** *Let  $u$  be a sub-treatment balanced QWF. Then*

- i. Let  $T$ , at cost  $C$ , with gain  $q$  from an initial ex ante health loss  $\ell$  be sub-divisible into  $T_p$  and  $T_{1-p}$  to provide subsequent gains of  $pq$  and  $(1 - p)q$ . Then, if  $u(\ell, q) \geq C$ , one may find  $C_p$  and  $C_{1-p}$  such that  $C_p + C_{1-p} = C$  and such that  $u(\ell, pq) \geq C_p$  and  $u(\ell - pq, (1 - p)q) \geq C_{1-p}$ .
- ii. Let  $T_p$  and  $T_{1-p}$ , at costs  $C_p$  and  $C_{1-p}$  provide subsequent gains of  $pq$  and  $(1 - p)q$  from an initial ex ante health loss  $\ell$ . Then, if  $u(\ell, pq) < C_p$  and  $u(\ell - pq, (1 - p)q) < C_{1-p}$ , then  $u(\ell, q) < C_p + C_{1-p}$ .

*Proof:* Since  $u$  is balanced  $C \leq u(\ell, q) = u(\ell, pq) + u(\ell - pq, (1 - p)q)$ . Setting

$$C_p = \frac{u(\ell, pq) \cdot C}{u(\ell, q)} \quad \text{and} \quad C_{1-p} = \frac{u(\ell - pq, (1 - p)q) \cdot C}{u(\ell, q)}$$

thus ensures  $C_p + C_{1-p} = C$ ,  $C_p \leq u(\ell, pq)$  and  $C_{1-p} \leq u(\ell - pq, (1 - p)q)$ ; proving (i). Part (ii) is immediate from the definition of sub-treatment balancedness. QED

When  $u$  is balanced, then the order of treatments is also irrelevant for the total weighted QALY gain.

### 3.2 A family of sub-treatment balanced QALY weighting functions

Thus far we have shown that balancedness removes or ameliorates problems, and that any weighting scheme as proposed in the NOU cannot be balanced. In this sub-section we therefore define a family of QWFs  $u_w(\ell, q)$ , which, for any suitable (to be defined later) function  $w$ , satisfies the following criteria:

1.  $u_w$  incorporates health-loss ex ante at the margin of health gains.

2.  $u_w$  is as such independent of the QALY paradigm; any ratio-scale utility assigned to ‘good life years’ can be used to measure the health loss and the health gains<sup>10</sup>.
3.  $u_w$  is sub-treatment balanced.

**Definition 3 (Sub-treatment balanced QALY  $w'$ -weighted function)**

Let  $w'$  be any positive and integrable function, so that  $w$  is well-defined<sup>11</sup> by:

$$w(x) = \int_0^x w'(y) dy$$

We call  $w'$  a *loss-weight function (LWF)*. We next define, uniformly in  $w'$ , a function  $u_w(\ell, q)$  by

$$u_w(\ell, q) = \int_0^q w'(\ell - x) dx \quad (8)$$

It is immediate from the definition that

$$u_w(\ell, q) = w(\ell) - w(\ell - q) \quad (9)$$

and thus that for  $p \in [0,1]$  we have  $u_w(\ell, q) = u_w(\ell, pq) + u_w(\ell - pq, (1-p)q)$ . Also,  $u_w(\ell, 0) = 0$  is obvious. Since the LWF  $w'$  is required to be positive,  $w$  is strictly increasing in  $q$  when  $q \leq \ell$ . If  $w'$  is non-negative, then  $u_w$  is also monotonic in  $\ell$ , whence  $u_w$  is a QWF as per Definition 1. We have also shown

**Proposition 2** Let  $u_w$  be defined by Equation 8; then, if the LWF  $w'$  is an increasing function, then  $u_w$  is a sub-treatment balanced QWF. QED

The exact properties of the QWF will obviously depend strongly on which  $w'$  is chosen as the LWF; consequently, finding the ‘right’ LWF  $w'$  will be a matter of some importance. Given point estimates for  $u_w(\ell, q)$  for different values of  $\ell$  and  $q$ , it will be a rather straightforward task to fit an acceptable  $w'$  to such observations.

The requirement that  $w' > 0$  only means that a greater QALY gain always means a larger loss-weighted gain. Similarly, if the weighting function  $w'$  is also non-decreasing – i.e. never gives less importance to greater ex ante QALY losses – then it is also increasing in  $\ell$ .

But are such QWFs – as asserted in the NOU – overly complicated to understand and difficult to apply? Quite the contrary, as we show in the next section.

### 3.3 Examples

The simplest LWF is  $w' \equiv \alpha$ ; this gives equal weight – and thus *no weight* – to all losses. Because  $w(x) = \alpha x$  we obtain

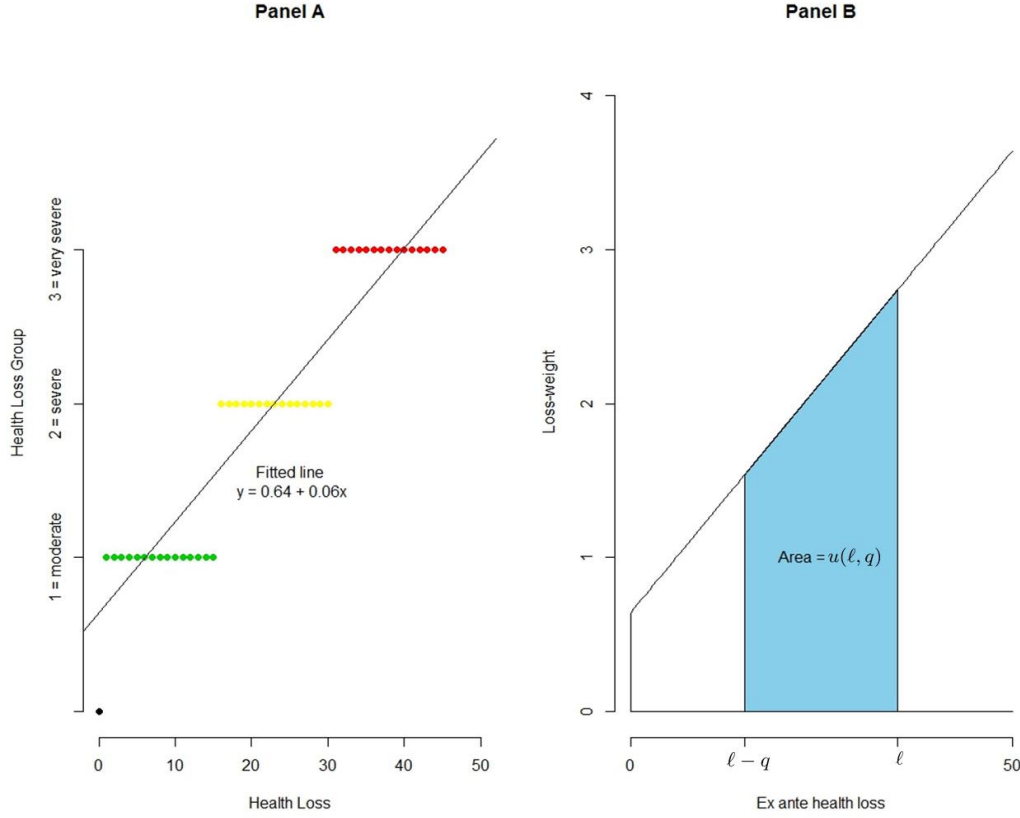
$$u_w(\ell, q) = \alpha(\ell - (\ell - q)) = \alpha q$$

<sup>10</sup> See footnote 4 (p. 3) and section 4 (Discussion and Conclusion) for more on this issue.

<sup>11</sup> Hence  $\frac{dw}{dx} = w'$  and  $w$  is an anti-derivative of  $w'$ .

so that the standard cost-utility expression reappears. This  $w'$  coincides with the only type possible where the QWF does have a multiplicative decomposition:  $u(\ell, q) = W(\ell) \cdot q$  and  $W$  is identically 1.

**Figure 2: Graphical illustration of step-wise (NOU) vs. balanced weighting**



Note: Panel A shows the step-wise weights of the QALY loss groups as outlined in the NOU. Furthermore, a regression line has been fitted. In panel B, the fitted line is used as the LWF as in Equation 10 with  $\alpha = 0.64$  and  $\beta = 0.06$ . Hence,  $u_w(\ell, q)$  is interpretable as the area  $A$  under the curve (shaded), which is bounded by the QALY loss and the QALY loss minus the health gain. It is evident that the the same gain  $q$  leads to a larger total utility when the loss  $\ell$  increases, as long as  $w'$  is monotonically increasing.

For  $w'(x) = \alpha + \beta x$  we obtain  $w(x) = \alpha x + \frac{\beta x^2}{2}$ . This form gives the first unit of gain a weight proportional to the remaining QALY loss in addition to a fixed weight proportional to the gain.

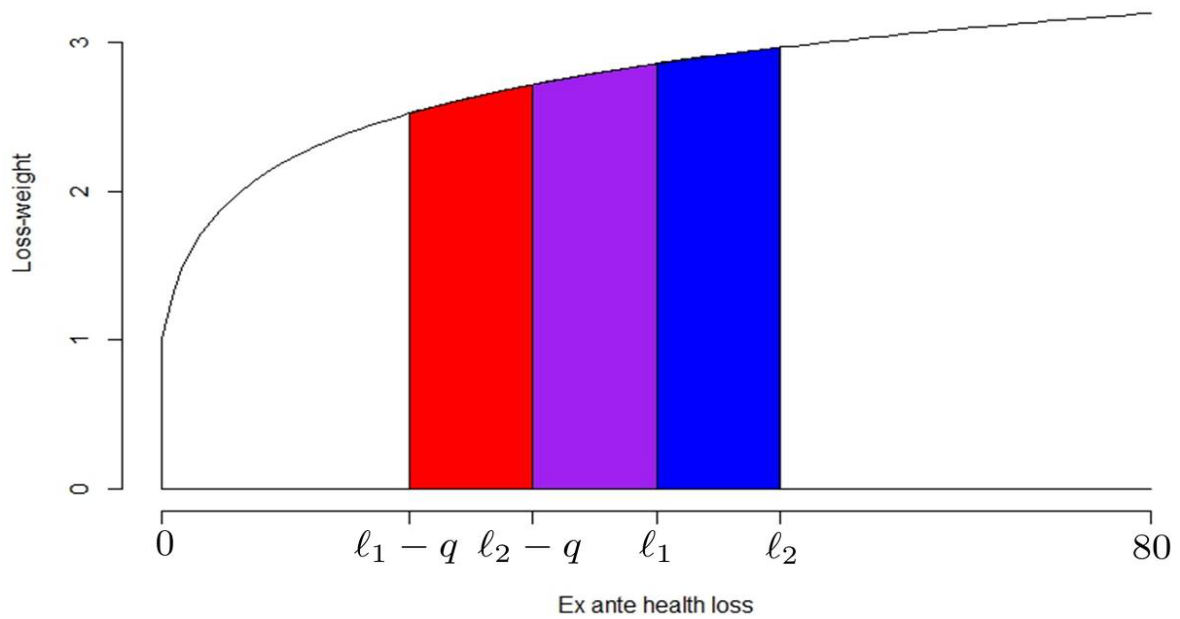
$$u_w(\ell, q) = \alpha \ell + \frac{\beta \ell^2}{2} - \alpha(\ell - q) - \frac{\beta(\ell - q)^2}{2} = \alpha q + \frac{\beta}{2}(2\ell q - q^2) \quad (10)$$

For  $\alpha \simeq 0.64$  and  $\beta \simeq 0.06$  this LWF is – *at the margin* – the weighting scheme suggested in the NOU. The value  $u_w(\ell, q)$  corresponds to the area under the graph of  $w$  between  $\ell - q$  and  $q$ ; see Figure 2, panel B.

According to the NOU's suggested algorithm, any treatment targeting a patient group with a very high expected QALY loss – say, of 35 QALYs – would imply a societal WTP threshold three times higher than for treatments targeting patients with an expected QALY loss of below 15 QALYs. A treatment expected to restore these patients to full health would justify costs up to 6.2 million Euro (assuming base rate WTP of 0.060 million Euro

per QALY). Under the LWF above, the WTP with marginal weighting yields a WTP for the full restoration of health of 3.5 million Euro. If a treatment for the same group has an expected QALY gain of only 5 QALY, the WTP would be 0.88 million Euro with the NOU's weights, and 0.76 million Euro with the marginal weighting above instead. As expected, the difference is much smaller for the latter example of restoring only 5 of 35, since most of that QALY gain is weighted with a substantial health loss.

**Figure 3: Illustration of log-type WLF**



Note: Illustration of log-type WLF (as in Equation 11 with  $\alpha = 2$ ) is interpretable as the areas under the curve. The left plus the middle areas correspond to  $u_w(\ell_2, q)$ , while the middle plus the right areas correspond to  $u_w(\ell_1, q)$ ; even if the gain is equal, the health losses are different, resulting in greater weight to the first quantity.

Looking at panel B of Figure 2, it is clear that the weighting is linear in loss. This may not reflect public preferences. Another possible  $w'$  is  $\alpha + \beta \cdot f(x)$ , which when  $\frac{df}{dx} > 0$  result in an  $u_w$  which has marginal decreasing priority in the QALY loss when  $\frac{d^2f}{dx^2} < 0$ , and which has marginal increasing priority in the health loss when  $\frac{d^2f}{dx^2} > 0$ . For example,  $\alpha = \beta = 1$  and  $f(x) = \log(1 + x)$  to obtain<sup>12</sup> a less steep LWF as seen in Figure 3 with decreasing marginal weighting of losses. The QWF is also readily computable (set  $\ell' = \ell + 1$ ) as

$$u_w(\ell, q) = (\alpha - \beta)q + \beta(\ell' \cdot \log(\ell') - (\ell' - q) \cdot \log(\ell' - q)); \quad (11)$$

<sup>12</sup> The addition of the constant 1 is merely to ensure that  $f(x) > 0$  for positive  $x$ .

a somewhat involved expression, but no match for a standard spreadsheet.

#### 4 Discussion and conclusion

The NOU suggests prioritizing health care interventions by weighting their cost-effectiveness using a function of the patients' QALY losses. Unfortunately, the suggested operationalization means that bundling and splitting up treatments can have a direct influence on priority setting – a breach of what we call sub-treatment balancedness. Furthermore, we have shown that any multiplicative decomposition such as the one suggested by the NOU fails to be balanced, but that adjusting the scheme to allow weighting on the margin is simple, practical, and results in a sub-treatment balanced QWF. Finally, we have provided an algorithm for constructing sub-treatment balanced QWFs.

Note that the discussion of sub-treatment balance is independent of the merits (or lack thereof) of the prioritization criteria suggested in the NOU. This is a technical property which we believe should be required of *any* weighting scheme applied to QALYs.

It is well worth noting that in the criteria suggested in the NOU, the health gains and the health losses are operationalized with respect to the same measure: QALYs. The operationalization of severity via (ex ante) health loss – *also measured in QALYs* – by design cannot distinguish health loss due to relatively shorter periods in very poor, painful and degrading health conditions from health loss due to long periods of relatively high HRQoL. While we do not support the use of instantaneous HRQoL as the basis for severity weighting of QALYs, as promoted by e.g. Nord (1993,2005); Nord and Johansen (2014) we do agree that this operationalization appears to be much closer to a layman interpretation of *severity* than the health loss criterion from the NOU. Curiously, this issue is not discussed *at all* either in the NOU's chapter 8 on 'Other suggested criteria' (pp. 101–112), or in the NOU's Appendix 4 (pp. 196–199), which further elaborates on the technicalities of health losses. If a criterion similar to the one proposed by Nord (1993,2005); Nord and Johansen (2014) were to be implemented, it should be operationalized in such a way as to ensure sub-treatment balance. (Consider that the instantaneous HRQoL of patients will change under successful treatment. Thus, if treatment options were to be weighted on the basis of ex ante HRQoL loss, inconsistencies similar to those described in this paper will occur.)

We should point out that the NOU does not directly use QALYs as their basic outcome measure; rather they abstract away from this particular model of HRQoL to one in which so-called *good life years* play the role of outcome measure. The NOU does, however, state that the QALY framework is the best available representation of such good life years. QALYs are the best measure of good life years<sup>13</sup>. For the purposes of clarity, we have used the term QALYs throughout this manuscript. All the statements in this paper will hold, and sub-treatment balance should still be a requirement as long as some measure which assigns numerical utilities to different lives in given health states (e.g. QALYs) is employed.

The method of weighting proposed in this paper is applicable to other such weighting schemes, and we posit that sub-treatment balance should be required for any suggested operationalization of a priority setting system. As such, sub-treatment balancedness is not a priority setting criterion in itself; it is a criterion for priority setting criteria.

There are other technical issues arising from the suggested health-loss criterion that are not fully addressed by ensuring sub-treatment balance. For example, both with and without sub-treatment balance, group-level cost-effectiveness weighted by health loss will yield different results if individual health-gains are weighted by individual health losses and

<sup>13</sup> 'Til tross for sine metodeproblemer framstår kvalitetsjusterte leveår (QALYs) som en godt egnet måleenhet for å uttrykke helsegevinsters ulike størrelser[.]' (Norheim et al., 2014, p. 141)

then averaged, and if average health gain is weighted by average health loss – a problem akin to the ecological fallacy.

#### 4.1 Directions for further research

If the framework of health-loss weighting QALY-gains proposed in the NOU is adopted, we encourage further efforts aimed at identifying the LWFs which yield the QWFs that most closely reflect the social welfare function of the adopting jurisdiction. This should be effected through elicitations from experts and/or the public. In general, for any weighting scheme in which the property or properties on which the weights are calculated will be altered by the act of treatment, adjustments similar to the ones discussed here will be required to ensure sub-treatment balance.

#### 4.2 Conclusion

If it is decided that ex ante health loss should be used to weight gains in health economic cost-utility calculations for priority setting, as suggested by the NOU, our solution (Equation 10) can be used directly for cost-utility analysis. If the health-loss notion of severity is abandoned, other severity weighting of the QALY should also adhere to the sub-treatment balanced property; suitably rephrased, it is not an approximation to the loss-weight-on-the-margin problem only, but an exact and easy solution to other types of QALY-weighting schemes.

Since this paper was first drafted, it has become clear that the ex ante health-loss as it was proposed by the priority commission will not be adopted in its suggested form. The Norwegian government has appointed a special task-group, lead by professor Jon Magnussen, to suggest a new operationalization of severity for use in a priority setting<sup>14</sup>. Severity weighting of QALYs can take a number of different forms. However, as the title of this paper indicates, we suggest that *any* operationalization of a priority-weighting scheme by a measure of severity (or other similar concepts) should satisfy sub-treatment balancedness. Any weighting scheme using a measure of severity that is influenced by the treatment in question will have the potential for lack of sub-treatment balance, as well as problems such as the ones discussed by Hope and colleagues (2010). These problems can be ameliorated by employing the methods presented here – that is, by use of a suitable integral.

#### Acknowledgements

We would like to thank the diligent reviewers. Their input has substantially improved both the quality and structure of the paper.

#### References

- Hope, T., Østerdal, L. P., and Hasman, A. (2010). An inquiry into the principles of needs-based allocation of health care. *Bioethics*, 24(9) 470–480.
- Lønning, (1997) Prioritering på ny: Gjennomgang av retningslinjer for prioritering innen norsk helsetjeneste, volume 18 of Norges Offentlige Utredninger. Statens forvaltningstjeneste, 1997.
- Nord, E. (1993). The trade-off between severity of illness and treatment effect in cost-value analysis of health care. *Social Science & Medicine*, 23, 227–238.

---

<sup>14</sup> See <https://www.regjeringen.no/no/aktuelt/regjeringen-vil-ikke-innfore-helsetapkriterium--vil-arbeide-videre-med-kriterium-for-alvorlighet/id2422928/> for the Norwegian Government's own explanation and the Mandate of the task-group (Norwegian only)

- Nord, E. (2005). Concern for the worse off: Fair innings versus severity. *Social Science & Medicine*, 60, 257–263.
- Nord, E. and Johansen, R. (2014) Concerns for severity in priority setting in health care: A review of trade-off data in preference studies and implications for societal willingness to pay for a qaly. *Health Policy*, 116, 281–288.
- Norheim, O. F., Allgott, B., Aschim, B., Førde, R., Gjøl, G. K., Gundersen, T., Kakad, M., Kjelleve, A., Kvinnsland, S., Melberg, H. O., Moen, A., Mæland, Ø., Olsen, J. A. and Sjøli, S. (2014). Åpent og rettferdig – prioriteringer i helsetjenesten. Norges Offentlige Utredninger. The 2013–2014 Royally Appointed Priority Commission, 2014.
- Pliskin, J., Shepard, D. S., and Weinstein, M. C. (1980) Utility functions for life years and health status. *Operations Research*, 1 206–224.
- Ringard, Å., Sagen, A., Saunes, I. S. and Lindahl, A. K. (2012) Norway: health system review. *Health systems in transition*, 15(8) 1–162.
- Weinstein, M. C., Torrance, G., and Alistair McGuire, A. (2009) Qalys: The basics. *Value in Health*, 12, S5–S9.