# COMPARISON AND EVALUATION OF STATES OF HEALTH *

## Carmen Herrero
Universidad de Alicante & IVIE

## Jose Luis Pinto
Universitat Pompeu Fabra

**Abstract**

Starting from a finite or countable set of states of health, and assuming the existence of an objective transitive preference relation on that set, we propose a way of performing interpersonal comparisons of states of health. In so doing, we first consider the population divided into types, and consider that two individuals of a different type have a comparable state of health whenever they sit at the same centile of their respective type. A way of comparing and evaluating states of health for different groups is then proposed and rationalized. This can be viewed as both an alternative and an extension of the traditional QALY approach.

**Address for correspondence:** Carmen Herrero, Departamento de Fundamentos del Análisis Económico, Universidad de Alicante, 03071 ALICANTE, Spain. Phone: (34)65903618, Fax: (34)65903685, E-mail: Carmen. Herrero@aitana.ua.es

1

# 1    Introduction

In tackling the problem of how to properly evaluate the state of health of a group of people, we may either consider the individualistic point of view, which generally involves problems of aggregation, or the point of view of a social planner.

The so called QALY approach has so far proved to be the most popular one. The justice principle behind the computation of the aggregate QALYs for a group of people is the *principle of impartiality*, clearly reflected in the sentence:"a QALY Is a QALY Is a QALY" [cf. Weinstein (1988)].

Notice that health is a very particular commodity. From consumers point of view, by means of a certain amount of *resources (health services)*, every individual is able to produce a certain *state of health*, depending upon his/her personal characteristics. Whereas it is easy to measure the amount of money devoted to some patient, it is not so easy to evaluate the health improvement he gains, and it is yet more difficult to properly compare gains for two different individuals. The traditional QALY approach does not take into account differences in characteristics across individuals in the aggregation procedure. Nevertheless, it has been argued that some characteristics, in particular age, family responsibilities, etc, have to play a role in weighing QALYs [see Williams (1988) or Murray & Lopez (1996)]. In any case, previous proposals have not yet managed to achieve a general consensus on the best method.

In this paper we try to offer an alternative proposal. First, we shall consider that there is a set of well-defined *states of health*, $S$. Furthermore, we shall assume that there is a preference relation on $S$ satisfying completeness and transitivity. The *best* state of health, $b$, and the *worst* state of health, $w$, are well specified, and are the best and worst elements of the aforementioned ordering.

Then, we introduce a way of *making interpersonal comparisons of states of health*. We claim that identical states of health are not equally socially valuable, irrespective of individual characteristics. A proper way of grouping characteristics induces a classification of the population in *types*. Then, we tackle the problem of socially evaluating the profiles of health states profiles for a group of individuals, from the point of view of a social planner. Asking social planner's preferences to satisfy the traditional VNM assumptions, plus additive independence across individuals, enables us to represent these preferences in an additive way. Then, we introduce the comparability criterion

using types. If the population is made out of a single type, we end up with a formulation which is compatible with the traditional QALY approach. Otherwise, we obtain a type-dependent way of measuring health in an aggregated way. An outline of potential applications, and an example of the way the theory applies in the evaluation of health care technologies, more specifically Total Hip Arthoplasty, illustrate the possibilities of this approach. Final remarks on the difficulties and the advantages of this framework, as well as future lines of research close the paper.

## 2   States of health.   Social ordering on the states of health.

Let us consider a set $S$ of states of health. A particular state of health is defined by means of a vector of characteristics [see, for instance, the EUROQOL questionnarie, or alternatively, we may think of a set of functionings, a la Sen, see Sen (1985), Pereira (1993) or Herrero (1996); that set of functionings convey to a certain *capability set*, associated to a particular state of health]. Suppose, for the sake of simplicity, that $S$ is a finite or countable set.

Assume now that there exists a preference relation $\Pi$, defined over $S$ . For any two states of health, $x, y \in S$, $x\Pi y$ means that state $x$ is considered to be better than state $y$. If, for two states $x, y \in S$, it is not true that $x\Pi y$, we say that $y$ is at least as good as $x$, and we write $y\Xi x$. If, simultaneously, $x\Xi y$ and $y\Xi x$, then we say that $x$ and $y$ are *similar or equally good,* and write $x\Upsilon y$.

We will also ask $\Pi$ to satisfy some additional requirements:

(i) *Preference*, namely that $\Pi$ is asymmetric and negatively transitive, i.e., for any $x \in S$, it is not true that $x\Pi x$, and for any $x, y, z \in S$, if $x\Xi y$ and $y\Xi z$, then $x\Xi z$

(ii) *Existence of extremes*, namely there are two states, $w, b \in S$, such that for any $x \in S$, $b\Xi x\Xi w$.

Previous requirements indicate, (i) that $\Pi$ is an *ordering,* and (ii) that this ordering has a minimum and a maximum, that is, there exists a particular state of health (maybe several, equally good ones) which is better than any other, the *perfect state of health, best,* and there is another state of health (maybe several, equally good ones), which is worse than any other, the *worst*

3

*state of health*, possibly death.

It is important to stress that so far $\Pi$ is purely ordinal, namely, for two given states of health, $x, y \in S$, we may say $x$ is better than $y$ ($x\Pi y$), or $y$ is better than $x$ ($y\Pi x$), or they are equally good ($x\Upsilon y$), but there are no cardinal valuations. So, if we consider four states of health, and it turns out that $x\Pi y$, and $z\Pi t$, we cannot measure the increase in going from $y$ to $x$ in relation with the increase in going from $t$ to $z$.

Previous requirements guarantee that $\Pi$ can be represented by means of a *utility function* $v : S \to R$, such that $x\Pi y$ iff $v(x) > v(y)$. Furthermore, any monotone transformation of $v$ is also a utility representation of $\Pi$. In consequence, we may choose a particular representation such that $v(b) = 1$, and $v(w) = 0$, and therefore, any state of health, $x \in S$, will be associated with a number $v(x)$, such that $0 \leq v(x) \leq 1$. Again, it has to be noticed that those numbers have only ordinal significance [cf. Kreps (1988), Chapter 3].

# 3 The interpersonal comparison problem. A solution linked to particular populations.

Consider now a particular population, $N$, in a precise time. For that population, we have the *state of health function*, $h : N \longrightarrow S$, namely, for every individual $a \in N$, $h(a) \in S$ indicates individual $a's$ state of health. We face the problem of socially comparing the state of health of two individuals, $a, a' \in N$. Suppose that $h(a) = x$, whereas $h(a') = y$.

We may behave naively and only consider $x, y$, forgetting about the particular individuals to whom these states of health are associated with. Nonetheless, it is clear that it is not a sensible way of making comparisons. It may be the case that one of the individuals is 20 years old and the other is 80, and if $x\Upsilon y$, for instance, both need a wheelchair, we cannot say that they have a similar state of health. Thus, some personal characteristics have to be taken into account in order to properly compare them.

Assume that we have the population divided into types, $N = N_1 \cup N_2 \cup ... \cup N_T$, in such a way that individuals belong to one and only one type. Types can be defined by using the characteristics which society considers to be relevant (e.g., age, gender, race, income level, etc.).

If two individuals belong to the same type, they are considered to be *socially similar.* Individuals belonging to different types are considered to be

*socially different.* Let $\tau = \{1, ..., T\}$ the set of types.

Now, for every $k \in \tau$, consider the following functions:

$f_k : S \longrightarrow N$, such that $f_k(x) = \# \{i \in N_k \mid h(i) = x\}$, and

$F_k : S \longrightarrow N$, such that $F_k(x) = \# \{i \in N_k \mid x \Xi h(i)\}$

Namely, $f_k(x)$ is the number of individuals of type $k$ whose state of health is $x$, and $F_k(x)$ stands for the number of individuals of type $k$ whose state of health is worse than or as good as $x$.

Notice now that $F_k$ is a *utility function for* $\Pi$, for every $k \in \tau$. It can be understood as a *cardinal utility function.* In such a case, it turns out that, by considering

$g_k : S \longrightarrow R$, such that $g_k(x) = \frac{f_k(x)}{\# N_k}$ and

$G_k : S \longrightarrow R$, such that $G_k(x) = \frac{F_k(x)}{\# N_k}$,

$G_k$ is also a utility function for $\Pi$, and $F_k$ and $G_k$ represent identical *cardinal preferences,* since $G_k = \lambda F_k(x)$, where $\lambda = [\# N_k]^{-1}$. Notice, nevertheless, that $G_k$ and $G_j$ for $k \neq j$, $k, j \in \tau$, represent different cardinal preferences in spite of the fact that they represent identical ordinal preferences ($\Pi$). Suppose that we consider two different types, $i, j$, and states of health $x, y$ such that $G_i(x) = G_j(y)$. This means that *the proportion of people of type i having a state of health worse than or as good as x coincides with the proportion of people of type j having a state of health worse than or as good as y.* In other words, if we attach to every type its respective cumulative distribution of states of health, *state of health x for type i, and state of health y for type j, they correspond to the same centile in their respective cumulative distributions.*

$g_k(x)$ can also be interpreted as the probability that an individual in type $N_k$ has a state of health $x$.

Let us now consider the function: $t : N \rightarrow \tau$, attaching to every individual in the population $N$ her type, namely $t(a) = i$ means that individual $a \in N_i$.

Consider now the following definition [cf. Roemer (1993), (1996)].

**Definition 1** - *We shall say that two individuals $a, a' \in N$ have a comparable state of health whenever $G_{t(a)}[h(a)] = G_{t(a')}[h(a')]$*

The above definition indicates that we consider two individuals belonging to different types as having a *comparable state of health* whenever they sit at the same centile of their types. This idea can be interpreted as saying that, by means of the utility functions $G_k$, $k \in \tau$, we associate cardinal numbers to states of health, *in a type-dependent way,* by using the distributions of states of health in population $N$.

5

Notice that if $t(a) = t(a')$, then $a, a'$ have a comparable state of health iff $h(a) \Upsilon h(a')$. In consequence, our criterion is an extension of the usual valuation for individuals of the same type.

If state $w$ is defined in such a way that $F_k(w) = 0, \forall k \in \tau$, then $G_k(w) = 0, G_k(b) = 1, \forall k \in \tau$.

# 4 Ranking profiles of states of health for a group of agents.

Let us now consider a group of agents, $A \subset N$, and consider the problem of ranking profiles of states of health for such a group of individuals. Notice that since population $N$ was divided into types, every agent $a \in A$ belongs to one and only one type. That is, we may consider the restriction of function $t$ to group $A$, and by slightly abusing language, call it also $t$.

A profile of health for $A$ is a mapping $s : A \to S$, where $s(a)$ indicates the state of health of individual $a$ in profile $s$. Let us call $S^A$ the set of all possible profiles of health for group $A$. Call $\Omega$ the set of lotteries over $S^A$, namely, a lottery $L \in \Omega$ is a mapping $L : S^A \to [0, 1]$ with finite support, such that $\sum_{s \in S^A} L(s) = 1$.

If $L, M \in \Omega$, and $\lambda \in [0, 1]$, define $[\lambda L + (1 - \lambda)M](s) = \lambda L(s) + (1 - \lambda)M(s)$. Thus, $[\lambda L + (1 - \lambda)M] \in \Omega$.

For a profile $s : A \to S$, and a state of health $x \in S$, denote by $(s^{-a}, x) = s' \in S^A$ the profile such that $s'(a') = s(a')$, whenever $a' \neq a$, $s'(a) = x$. That is, $(s^{-a}, x)$ coincides with $s$ in all agents but agent $a$, and the state of health of agent $a$ in profile $s'$ is $x$.

We shall now consider the existence of a binary relation $P$ defined over $\Omega$, understood as a *strict preference relation*, in such a way that $R$ and $I$ are, respectively, the *weak preference relation* and the *indifference relation* associated to $P$. That is, for any $L, M \in \Omega$, $LRM$ iff it is not true that $MPL$, and $LIM$ iff simultaneously, both $LRM$ and $MRL$.

Notice that $P$ also induces a binary relation on $S^A$, since any profile $s \in S^A$ can also be interpreted as a degenerated lottery in $\Omega$, where $s(s) = 1$, and $s(s') = 0$, for all $s' \neq s$.

Let us now consider the following assumptions:

**Preference.-** *P is a preference relation* on $\Omega$, *namely, it is asymmetric and negatively transitive, i.e., $\forall L, M \in \Omega$, if $LPM$, then it is not true that*

6

$MPL$, and $\forall\ L, M, N \in \Omega$, if both $LRM$ and $MRN$, then $LRN$.

**Independence.-** *For all $L, M, N \in \Omega$, and for all $\lambda \in (0, 1]$, if $LPM$, then $[\lambda L + (1 - \lambda)N]P[\lambda M + (1 - \lambda)N]$.*

**Continuity.-** *For all $L, M, N \in \Omega$, if $LPMPN$, then there exist $\lambda$, $\mu \in (0, 1)$, such that $[\lambda L + (1 - \lambda)N]PMP[\mu L + (1 - \mu)N]$*

**Additive Independence.-** *For any $s, r \in S^A$, any agent $a \in A$, if we call $q = (s^{-a}, r(a))$, $p = (r^{-a}, s(a))$, and $L(s) = L(r) = \frac{1}{2}$, $M(q) = M(p) = \frac{1}{2}$, then $LIM$.*

Preference, Independence and Continuity are the basic assumptions in the Von Neumann-Morgenstern expected utility theory. In this particular case, they say that preference $P$ on $\Omega$ is a complete ordering, that common chances are ignored in the valuation of lotteries over profiles of health for $A$, and that by properly combining probabilities we may fill all valuation gaps.

Additive Independence asks for states of health for the different agents in $A$ to be additive independent, namely preferences depend only on the marginal probability distribution and not on the joint distribution.

Then, we obtain the following results:

**Proposition 1** .- *Under Preference, Independence and Continuity, there exists a function $u : S^A \to R$ such that $LPM$ iff $\sum_{s \in S^A} L(s)u(s) > \sum_{s \in S^A} M(s)u(s)$. Then we say $P$ admits an expected utility representation, and that $u$ is a utility function of $P$ over $S^A$. Furthermore, $u$ is unique up to positive linear transformations, namely, if an alternative $u'$ also represents $P$ then there exist real numbers $\lambda > 0$ and $\mu$, such that $u'(s) = \lambda u(s) + \mu$, for any $s \in S^A$.*

PROOF: It is a direct consequence of Von Neumann-Morgenstern expected utility theorem. Cf. Kreps (1988, Theorem 5.4).♠

Call $W \in S^A$ the profile such that $W(a) = w$ for all $a \in A$, and $B \in S^A$ the profile such that $B(a) = b$ for all $a \in A$.

**Proposition 2** .- *Under Preference, Independence, Continuity and Additive Independence, there exist $u_a : S \to R$, $a \in A$, such that: (i) $u(s) = \sum_{a \in A} \Gamma_a u_a[s(a)]$; (ii) $u_a$, $a \in A$, are normalized so that $u_a(w) = 0$, $u_a(b) = 1$; (iii) $\Gamma_a = u(W^{-a}, b)$; (iv) $u$ is normalized so that $u(W) = 0$, $u(B) = \#A$.*

PROOF: Additive Independence indicates that attributes $\{s(a)\}_{a \in A}$ are additively independent, since it implies that preferences over lotteries on them depend only on their marginal probability distributions and not on their joint probability distribution. Thus, we may apply Keeney and Raiffa

(1976, Theorems 5.1 and 6.4), in order to obtain the representation result in (i) [cf. Bleichrodt, Theorem 3.2]. We are free to normalize $u_a$ and $u$ as we wish. Then, the values of $\Gamma_k$ can be obtained as follows:

$u(B) = \sum_{a \in A} \Gamma_a u_a(b) = \sum_{a \in A} \Gamma_a$
$u(W^{-a}, b) = \Gamma_a u_a(b) = \Gamma_a$. ♠

Consider now the following assumption:

**Neutrality.-** *For any two profiles, $s = (W^{-a}, x)$, $r = (W^{-a'}, y) \in S^A$, if state $x$ enjoyed by agent $a$ is comparable to state $y$ enjoyed by agent $a'$, then $sIr$.*

Neutrality means the following: suppose that agent $a \in N_i$, whereas agent $a' \in N_j$. Then, agent $a$ state of health in $s$, $x$, is comparable to agent $a'$ state of health in $r$, $y$ whenever $G_i(x) = G_j(y)$. In such a case, $sIr$.

**Proposition 3** .- *Under Preference, Independence, Continuity, Additive Independence and Neutrality, $u(s) = \sum_{a \in A} G_{t(a)}[s(a)]$.*

PROOF: Notice first that, since $u(W^{-a}, b) = \Gamma_a$ and for any $a, a' \in A$, it turns out that $G_{t(a)}(b) = G_{t(a')}(b) = 1$, by Neutrality, $u(W^{-a}, b) = u(W^{-a'}, b)$, for all $a, a' \in A$. In consequence, $\Gamma_a = \Gamma_{a'}$, for all $a, a' \in A$. Furthermore, since $u(B) = \sum_{a \in A} \Gamma_a = \#A$, it follows that $\Gamma_a = 1$ for all $a \in A$. In consequence, $u(s) = \sum_{a \in A} u_a[s(x)]$. Suppose now that we consider profiles $s = (W^{-a}, x)$, $r = (W^{-b}, y) \in S^A$ such that $G_{t(a)}(x) = G_{t(a')}(y)$. Under Neutrality, $sIr$, namely, $u(s) = u(r)$. That is, $u(s) = u_a[s(a)] = u_a(x) = u(r) = u_{a'}[r(a')] = u_{a'}(y)$. Notice that furthermore, previous identities are fulfilled iff $G_{t(a)}(x) = G_{t(a')}(y)$. In consequence, $G_{t(a)}$ and $u_a$ represent identical preferences, and thus, they are related by a positive affine transformation. Furthermore, $G_{t(a)}(w) = u_a(w) = 0$, and $G_{t(a)}(b) = u_a(b) = 1$, for all $a \in A$, and thus, $G_{t(a)} = u_a$. In consequence, $u(s) = \sum_{a \in A} G_{t(a)}[s(a)]$.♠

Assume now that all individuals in $A$ belong to the same type. Then Neutrality implies that for any state of health $x \in S$, and for any two individuals $a, a' \in A$, if we consider the profiles $s, r : A \to S$, where $s(a) = x, r(a') = x$, and $s(a") = w$, for all $a" \neq a$, $r(a") = w$, for all $a" \neq a'$, then $sIr$. That is, if all individuals but one are at state $w$, and the remaining individual is at state $x$, the planner is indifferent about the particular individual outside $w$. Thus, Neutrality with a single type is an instance of impartiality. In such a case, in Proposition 4.3, it follows that $u_a(x) = u_{a'}(x)$, for all $x \in S$, and for all $a \in A$. In consequence, we obtain a formulation compatible with the traditional QALY aggregation procedure.

8

# 5 An outline of possible applications.

The above framework can be useful as a tool in the analysis of different types of decisions dealing with health policy. We mention some potential examples.

## 5.1 Measuring benefits in health on an individual level.

Consider two individuals, $a$, $a'$, and the possibility of treating them. Before treatment, the state of health of individual $a$ is $x$, and after the treatment, her state of health is $y$. As for individual $a'$, her state of health before treatment is $x'$, and after treatment it is $y'$. We may consider four different profiles of health for group $A = \{a, a'\}$:

$s_1$, where $s_1(a) = x$, $s_1(a') = x'$, namely $s_1$ represents the "no treatment" situation

$s_2$, where $s_2(a) = y$, $s_2(a') = y'$, namely $s_2$ represents treatment for both individuals.

$s_3$, where $s_3(a) = y$, $s_3(a') = x'$, $s_3$ indicates treatment for $a$, no treatment for $a'$

$s_4$, where $s_4(a) = x$, $s_4(a') = y'$, $s_4$ is treatment for $a'$, no treatment for $a$.

We can cardinally compare the previous alternatives, in particular $s_3$ and $s_4$, and thus decide which individual will most enjoy treatment, in case there is no possibility of treating them both. Notice that $u(s_1) = G_{t(a)}(x) + G_{t(a')}(x')$; $u(s_2) = G_{t(a)}(y) + G_{t(a')}(y')$; $u(s_3) = G_{t(a)}(y) + G_{t(a')}(x')$; $u(s_4) = G_{t(a)}(x) + G_{t(a')}(y')$.

If we compare alternatives $s_3$, $s_4$ versus $s_1$, we may also compare the gains in going from $s_1$ to $s_3$ with the gains in going from $s_1$ to $s_4$. Notice that these correspond with the gain of individual $a$ (respectively of $a'$) from treatment. That is, we may cardinally compare individual gains.

## 5.2 Evaluation of alternative policies over two different groups.

Suppose now that we consider the social impact of two different proposals, as, for example, a campaign of prevention of breast cancer for women between 45 and 55 years old, or the administration of a new drug in order to improve the situation of terminal AIDS patients. We may then consider the group made out of $A \cup B$, $A$ being the first subgroup, and $B$ the second subgroup.

If, in the absence of any policy, the expected health profile for group $A$ is $s : A \to S$, and the expected health profile for group $B$ is $s' : B \to S$, and by means of policy I, the expected health profile of group $A$ will move to $r : A \to S$, whereas the expected health profile of group $B$ will be $r' : B \to S$, under policy II, we may construct four profiles for $A \cup B$, corresponding to "no policy"; "policy I"; "policy II", and "policies I and II simultaneously". Then, we may compare the gains of policy I versus no policy, and that of policy II versus no policy, in order to make a decision.

## 5.3  Comparison of states of health within different communities

Suppose that $N$ represents the population in a country, and we have defined types according to $N$ are divided in classes $N_i$, $i = 1, ..., T$. Suppose now that the population is also divided by political reasons into other groups. For instance, we may consider two different regions, $A$ and $B$. We may ask ourselves about the relative situation of health within these two regions. As an example, suppose that we consider individuals in the same type, say $k$, both in $A$ and $B$. Then, we may consider $A_k = A \cap N_k$, and $B_k = B \cap N_k$, and the corresponding cumulative distributions, $G_{A_k}$ and $G_{B_k}$ respectively. If it turns out that $G_{A_k}(x) \leq G_{B_k}(x)$ for all $x \in S$, then we may say that *for type k, population A is in a better state of health than population B*. Things are not that clear if no such domination relation exists, but even then, we may suggest compareing the median value of both $G_{A_k}$ and $G_{B_k}$. Suppose that $G_{A_k}(x) = G_{B_k}(y) = .5$. If $x \Pi y$, we may also say that the median state of health of type $k$ in population $A$ is better than the median state of health of type $k$ in population $B$. Notice, nevertheless, that this comparison has only ordinal significance.

## 5.4  Equity considerations

The measure $G_k$ can be considered as a *socially comparable measure of the state of health of an individual*. In consequence, it can be used in order to perform interpersonal comparisons of states of health across people. Thus, by combining it with different equity criteria, we may obtain rules on how to allocate resources in order to properly *equalize health*.

# 6    An Example

We devote this section to analyzing the performance of our model in the evaluation of health care technologies, more specifically Total Hip Arthoplasty (THA).

In order to implement the model, we need information about:

(a) *The set  S of states of health.* We shall use the set of health states decribed by the EuroQol [see Brooks (1996), and also Appendix 1]. This is a health profile with five dimensions and three items per dimension defining 243 possible health states. Unconcious and dead are also valued.

(b) *A Preference relation $\Pi$ defined over $S$.* We will use the valuation estimated by the MVH group [see Williams (1995)]. These values were estimated in a survey of 2997 people using the Time Trade-Off (TTO) technique in 1993-94. It has to be noticed that by means of such a procedure we obtain a cardinal valuation of the elements of $S$, even though our model only asks for an ordinal relation on the states of health.

(c) *A population of reference, $N$.* We shall use the population in Catalonia as the reference population. Because of the Catalan Health Survey (CHS), we have information on the distribution of the states of health in this population, by using the EuroQol instrument.[1]

(d) *A set of types $\tau$ such that individuals each belong to one and only one type.* As a first step, we choose types related to age. In principle, we divide the population into four age groups, namely, 16-40; 41-60; 61-80; and over 80.

At a second step, we consider also gender in order to classify individuals. Thus, we finally will look at two different scenarios: either we consider four different types (age related), or eight different types (age and gender related).

Figure 1 shows the mean value of health status by age (EuroQol), and also by age and gender. It clearly shows that health worsens as people age. It also shows that for each age, the health state for men is better than the health state for women. This fact somehow justifies the selection of types we made.

(e) *Information on the distribution functions $G_k$, for all $k \in \tau$.* Again, this information is obtained from the Catalan Health Survey (CHS). Figure

---

[1]The Catalan Health Survey (CHS) was comissioned by the Catalan Government. A total of 15.000 people were interviewed during 1994. One of the questions included in it was a description of each individual's own health state using the EuroQol instrument.

2 shows the distribution of health by age groups. Figures 3 to 6 show the cumulative distribution for the different types.

(f) *A group of people A extracted from population N.* We consider a group of 213 patients suffering an intervention of THA in 1994 in seven catalan hospitals. For each patien we have information about age, gender, health state both the day before the intervention and six months later. Each patient described his/her own state of health in each of this two moments of time by using the EuroQol classification system.

## 6.1    Comparison of states of health for two populations

Notice that the state of health deteriorates with age (Figure 1), and we may say that men have a better state of health than women (Figures 3 to 6).

## 6.2    Comparison of states of health for two individuals

As an example of the way our approach works, consider the case of an 85 year-old woman in EuroQol health state 23221 (TTO value 0.2). She is in percentile 23.6 of her type (see Figure 6). Consider now a 35 year-old man in the same health state. He is in percentile 0.8 of his population (see Figure 3). Even though they have exactly the same EuroQol value, our model considers that the 35 year-old man is in a worse state than the 85 year-old woman.

Following with the previous example, assume that our 85 year-old woman and our 35 year-old man have the same final state of health after THA, for example, state 12111 (TTO value 0.815). We notice that she is now in percentile 75.4 of her population whereas he is now in percentile 11 of his population. Again the state of health of both individuals is very different.

Notice that the traditional approach estimates that both individuals were in the same situation before and after the intervention.

## 6.3    Computing improvements in health on an individual level

In our previous example, the 85 year-old woman has an improvement of $75.4 - 23.6 = 51.8$, whereas the 35 year-old man has an improvement of $11.1 - 0.8 = 10.3$. Again, we are facing results extremely different than those obtained using the traditional approach, under which both individuals show an identical improvement, namely an improvement of 0.615.

## 6.4  Average Improvements due to THA in a group

By considering group $A$, made out of our 213 patients, and collecting the information of their state of health after and before THA, we may compute the group improvement by using proposition 4. Tables 1 and 2 show the benefits of the intervention if types are related only to age. Table 2 illustrates the differences between our approach and the traditional one. Tables 3 to 6 show the benefits of the intervention if types are related both to age and gender.

# 7  Final Remarks

This paper attempts to provide an alternative way of evaluating states of health for either individuals or groups of people. Our construction relies on two main assumptions: (1) The existence of a (finite or countable) set of *states of health, $S$,* such that every element in $S$ is well defined and independent of personal characteristics; (2) The existence of a complete ordering on $S$.

An important aspect of our contribution is the idea to clasify the population into types, which may deserve an asymmetrical treatment in the aggregation procedure. Thus, our approach can be viewed as an extension of the traditional QALY aggregation procedure (in which all individuals belong to the same type).

When the population is divided into more than one type, our approach provides a different way of evaluating states of health to that given by the traditional one. In our example, it is clear that some characteristics of the individuals, being age the most obvious one, are closely related to their state of health, and should also be related to their valuation. Our model seems to provide a sensible way of so doing. Nevertheless, it also suffers from limitations.

Some of these limitations, which become apparent in the example provided in Section 6, refer to the distribution of health among different types. In the example, the distribution of states of health for young people in the general population is quite good. Only 18% of those under 40 declare themselves to suffer any kind of health problem. Although this may be due to the lack of sensitivity of the instrument used (EuroQol), we think that this is a problem which will appear whenever types are related to age. Apparently

high health improvements may actually produce a very small jump in the percentile where the patient is before and after treatment. Let us illustrate this with an example. Assume a 35 year-old woman in state 33333 (a terrible situation) has a TTO value of -0.59, and is in percentile 0.1. If she is treated and reaches health state 12221, then she is in percentile 3.5. If she now jumps to state 12111, then she is in percentile 16. In our example, the improvement from 33333 to 12221 has a value of 3.4, and the improvement from 12221 to 12111 has a value of 12.5, for a 35 year-old woman. In the traditional approach, previous improvements are of 1 and 0.35, irrespective of type.

Notice that our approach only requires ordinal information on the way states of health, $S$, are ordered. Nevertheless, it is robust enough to be applied under cardinal information. In such a situation (as it is under the ordering used in the example of Section 6), our model is open to further ways of refining our approach, in order to obtain more accurate results. We may think, for instance, of providing with additional statistically significant information in order to avoid some of the problems previously mentioned.

Nevertheless, our framework provides a useful information due to the way we understand the comparability criteria. If types are correctly chosen, and if it is agreed that comparability across types is correctly made, it is legitimate to attach different weights to similar improvements, in a type-dependent way. Thus, grate care must be taken over the way in which types are selected. In the example of Section 6, we first use a criterion that seems to us normatively appealing (age), and another one that seems to us to be more controversial (gender). The election of types may, somehow lead to controversial decisions. Suppose, for instance, a 65 year-old man whose health state has a TTO value of 0.8. He is in percentile 37.3 of his type. A woman of the same age and with the same state of health is at percentile 57.5. If this is interpreted as the man being in a worse condition than the woman, he may have priority for treatment, against intuition.

An interesting contribution of this model is that it can provide an answer to one of the main ethical problems that have been raised against the cost-utility analysis, the *double jeopardy* argument [cf. Hadorn (1992)], that is, discrimination against the disabled. In fact, Oregon's attempt to ration Medicaid was initially rejected because it was believed to violate the Americans With Disabilities Act [see Sullivan (1992)]. The argument is that if somebody suffers a chronic condition, then his/her upper bound in the scale of TTO values is smaller than 1. Suppose that an individual in a chronic con-

14

dition suffers an illness unrelated to the previous chronic condition, and that another person (without the chronic condition) suffers an identical illness. The benefits of treatment for these two people in the traditional analysis are different: the disabled person's is lower. This problem can be avoided in our framework by considering types taking into account chronic problems. Thus, health improvements of the disabled individuals are evaluated in relation to other disabled people.

So far, our formulation is done in a static framework. Introducing time, life streams and ways of comparing them in our framework is left for future research.

# 8   Appendix

THE EUROQOL CLASSIFICATION SYSTEM

MOBILITY
1. No problems in walking about
2. Some problems in walking about
3. Confined to bed.

SELF-CARE
1. No problem with self-care
2. Some problems washing or dressing self
3. Unable to wash or dress self

USUAL ACTIVITIES
1. No problems with performing usual activities (e.g., work, study, housework, family or leisure activities)
2. Some problems with performing usual activities
3. Unable to perform usual activities

PAIN/DISCOMFORT
1. No pain or discomfort
2. Moderate pain or discomfort
3. Extreme pain or discomfort

ANXIETY/DEPRESSION
1. Not anxious or depressed
2. Moderately anxious or depressed
3. Extremely anxious or depressed

Note: For convenience each composite health state has a five digit code number relating to the relevant level of each dimension, with the dimensions always listed in the order given above. Thus, 11223 means:

1 No problems in walking about
1 No problems with self care
2 Some problems with performing usual activities
2 Moderate pain or discomfort
3 Extremely anxious or depressed

# 9 References

Bleichrodt, H., 1995, Applications of utility theory in the economic evaluation of health care, Thesis Erasmus University, Rotterdam.

Brooks, R. with the EuroQol group ,1996, EuroQol: the current state of play. *Health Policy* 37, 53-72.

Fishburn, P.C., 1965, Independence in utility theory with whole product sets, *Operations Research* 13, 28-45.

Fishburn, P.C., 1970, *Utility theory for decision making* (Wiley, New York).

Hadorn, D.C., 1992, The problem of discrimination in health care priority setting, *JAMA* 268, Vol 11, 1454-59.

Herrero, C., 1996, Capabilities and Utilities, *Economic Design* 2, Vol 1, 69-88.

Keeney,R. & H. Raiffa, 1976, *Decisions with multiple objectives* (Wiley, New York).

Kreps, D.M., 1988, *Notes on the theory of choice (*Westview Press, Boulder, Colorado).

Murray, C.L.J. & A.D. Lopez, eds, 1996, *The Global Burden of Disease and Injury Series,* Vol 1: The global burden of disease, (Harvard University Press, Cambridge, MA).

Pereira, J., 1993, What does equality in health mean?, *Journal of Social Policy* 22, 19-48.

Roemer, J.E., 1993, A pragmatic theory of responsibility for the egalitarian planner, *Philosophy & Public Affairs* 22, 146-166.

Roemer, J.E., 1996, *Theories of Distributive Justice (*Harvard University Press, Cambridge, MA).

Sen, A.K., 1985, *Commodities and Capabilities* (North-Holland, Amsterdam).

Sullivan, L.W., 1992, Letter to Gov. Barbara Roberts. Department of Health and Human Services.

Weinstein, M., 1988, A QALY Is a QALY Is a QALY- or Is it?, *Journal of Health Economics* 7, 289-290.

Williams, A., 1988, Ethics and efficiency in the provision of health care, in Bell, J.M. & Mendus, S (eds), *Philosophy and Medical Welfare* (Cambridge University Press).

Williams, A., 1995, The measurement and valuation of health: a chronicle. Discussion working paper 136, Center for Health Economics, The Uni-

versity of York.

# 10  Tables and Figures

| Table 1: Comparison of absolute increments | | | | |
|---|---|---|---|---|
| | 16-40 | 41-60 | 61-80 | >80 |
| EuroQol-B | 0.21 | 0.15 | 0.17 | 0.26 |
| EuroQol-A | 0.70 | 0.75 | 0.78 | 0.79 |
| $\Delta EuroQol$ | 0.49 | 0.60 | 0.61 | 0.52 |
| Percentile-B | 1.49 | 6.14 | 10.8 | 18.52 |
| Percentile-A | 19.11 | 45.79 | 54.58 | 59.19 |
| $\Delta Percentile$ | 17.63 | 39.65 | 43.78 | 40.67 |

| Table 2: Comparison of relative increments | | | | |
|---|---|---|---|---|
| | 16-40 | 41-60 | 61-80 | >80 |
| EuroQol | 80 | 98.3 | 100 | 85.2 |
| Percentile | 40.2 | 90.56 | 100 | 92.9 |

| Table 3: Comparison of absolute increments (women) | | | | |
|---|---|---|---|---|
| | 16-40 | 41-60 | 61-80 | >80 |
| EuroQol-B-W | 0.12 | 0.09 | 0.11 | 0.32 |
| EuroQol-A-W | 0.47 | 0.81 | 0.75 | 0.80 |
| $\Delta EuroQol - W$ | 0.35 | 0.72 | 0.64 | 0.48 |
| Percentile-B-W | 0.83 | 3.32 | 10.91 | 28.54 |
| Percentile-A-W | 2.77 | 48.79 | 54.15 | 71.40 |
| $\Delta Percentile - W$ | 1.94 | 45.47 | 43.24 | 42.86 |


| Table 4: Comparison of relative increments (women) | | | | |
|---|---|---|---|---|
| | 16-40 | 41-60 | 61-80 | >80 |
| EuroQol | 49 | 100 | 89 | 67 |
| Percentile | 4 | 100 | 95 | 94 |

| Table 5: Comparison of absolute increments (men) | | | | |
|---|---|---|---|---|
| | 16-40 | 41-60 | 61-80 | >80 |
| EuroQol-B-M | 0.26 | 0.19 | 0.28 | 0.19 |
| EuroQol-A-M | 0.83 | 0.71 | 0.85 | 0.77 |
| $\Delta EuroQol - M$ | 0.57 | 0.52 | 0.57 | 0.58 |
| Percentile-B-M | 1.54 | 5.79 | 9.32 | 11.25 |
| Percentile-A-M | 27.26 | 40.51 | 56.03 | 49.43 |
| $\Delta Percentile - M$ | 25.72 | 34.72 | 46.71 | 38.16 |

| Table 6: Comparison of relative increments (men) | | | | |
|---|---|---|---|---|
| | 16-40 | 41-60 | 61-80 | >80 |
| EuroQol | 98 | 90 | 98 | 100 |
| Percentile | 55 | 74 | 100 | 82 |