



Open Access Full Text Article

RESEARCH ARTICLE

## État des lieux sur les méthodes d'élicitation du QALY

### [Overview of QALY elicitation methods]

Vickie Fauteux<sup>1</sup>  
Thomas G. Poder<sup>2,3</sup>

<sup>1</sup>Département d'Économique,  
École de Gestion, Université de  
Sherbrooke

<sup>2</sup>UETMIS, CIUSSS de l'Estrie –  
CHUS

<sup>3</sup>Centre de recherche du  
CHUS, CIUSSS de l'Estrie –  
CHUS

Correspondence:  
Thomas G. Poder, UETMIS,  
CIUSSS de l'Estrie – CHUS,  
580 rue Bowen Sud,  
Sherbrooke (QC), J1G2E8,  
Canada

Email:  
[tpoder.chus@ssss.gouv.qc.ca](mailto:tpoder.chus@ssss.gouv.qc.ca)

Article received:  
29 September 2017

First response:  
17 October 2017

Article accepted:  
22 October 2017

**Résumé :** Le concept d'année de vie ajustée par la qualité (AVAQ ou QALY) est de plus en plus utilisé pour la prise de décision en santé. Ce concept développé dans les années soixante-dix a connu de nombreux développements méthodologiques et de nombreuses méthodes d'élicitation des utilités en santé ont ainsi vu le jour. Cet article propose un tour d'horizon des différentes méthodes d'élicitation du QALY (Quality Adjusted Life Year). Ce faisant, les fondements théoriques du QALY sont exposés et une comparaison des avantages et inconvénients des différentes méthodes d'élicitation est effectuée. Parmi les méthodes d'élicitation directes, le pari ordinaire et l'arbitrage temporel sont les plus utilisées, alors qu'au sein des méthodes indirectes on constate que les instruments d'utilité multi-attributs les plus utilisés sont le EQ-5D et le SF-6D. Aucun consensus n'existe actuellement sur les méthodes les plus robustes ou les plus appropriées. De plus, l'utilisation du QALY reste limitée par un certain nombre de problèmes tant théoriques qu'empiriques. Il demeure donc nécessaire de poursuivre les recherches en ce domaine et de rester prudent sur son utilisation dans le domaine de la prise de décision.

**Mots clés :** AVAQ, qualité de vie reliée à la santé, utilité, préférence en santé.

**Abstract:** The quality adjusted life year concept (QALY) is increasingly being used in healthcare decision-making. This concept developed in the seventies has undergone numerous methodological developments and numerous methods of elicitation of health utilities have thus emerged. This article provides an overview of the different QALY elicitation methods. In doing so, the theoretical foundations of the QALY are presented and a comparison of the advantages and disadvantages of the different elicitation methods is carried out. Among direct elicitation methods, standard gamble and time trade-off are the most commonly used, whereas in indirect methods it is found that the most widely used multi-attribute utility instruments are EQ-5D and SF-6D. There is currently no consensus on the most robust or appropriate methods. Moreover, the use of QALY is limited by a number of theoretical and empirical problems. It is therefore necessary to continue research in this area and to remain cautious about its use in the area of decision-making.

**Keywords:** QALY, health-related quality of life, utility, health preference.

## Introduction

En santé comme ailleurs, le cadre de toute évaluation économique se caractérise par des besoins illimités et des ressources limitées. Pour obtenir les meilleurs résultats possibles, un impératif nécessaire est d'être sans cesse à la recherche de solutions permettant à la fois de minimiser les ressources utilisées et de maximiser la santé de la population. Si les ressources sont généralement mesurées en termes d'unités monétaires, la santé ne dispose pas d'une unité de mesure commune aussi évidente. Cette difficulté réside dans la nature même de ce qu'est la santé. Selon l'Organisation Mondiale de la Santé, celle-ci se définit comme « un état de complet bien-être physique, mental et social et ne consiste pas seulement en une absence de maladie ou d'infirmité » [1]. Suite à ce constat, il est très rapidement apparu que la santé ne pouvait se limiter à la simple utilisation d'indicateurs cliniques. Pour mesurer la santé et les différentes dimensions qu'elle recouvre, le concept de qualité de vie reliée à la santé (QVRS ou HRQoL en anglais : health-related quality of life) a ainsi émergé dès la fin des années 1960 aux États-Unis [2]. De nombreux instruments psychométriques de QVRS ont depuis lors été créés.

L'objet principal de ces instruments de mesure est d'établir quelles interventions sont les plus efficaces pour améliorer la QVRS et contribuer à prendre de meilleures décisions. Certains de ces instruments sont génériques comme le SF-36 [3] ou le Nottingham Health profile [4] alors que d'autres ciblent une pathologie particulière comme le QLQ-C30 [5] ou le diabetes QoL [6]. Ces instruments ont pour la plupart été validés et sont fortement corrélés avec des indicateurs cliniques d'évolution de l'état de santé [7-9]. Toutefois, pour de multiples raisons, ils ne sont généralement pas recommandés pour effectuer des analyses de coût-efficacité [10]. En particulier, ces instruments de QVRS négligent deux aspects importants dans le cadre d'une telle analyse. Tout d'abord, ils n'évaluent pas

nécessairement la durée associée à différents états de santé et surtout, ces instruments ne se basent pas sur les préférences des individus [10-11]. Les préférences en santé sont définies comme une mesure de l'utilité, c'est-à-dire du bien-être ou de la satisfaction que retire un individu de la consommation d'un bien ou d'un service, ici un programme ou une intervention en santé [12]. Il est donc problématique de faire des choix d'allocations de ressources à l'aide d'instruments psychométriques ne possédant pas ces propriétés, en particulier lorsque l'on doit comparer un programme plus cher et plus efficace à un autre moins cher et moins efficace. En effet, deux interventions différentes n'affectent pas forcément les mêmes dimensions de la santé (e.g. mobilité, douleur physique, état mental) et si l'on ignore les préférences des individus pour ces domaines, il devient impossible de bien évaluer quelle intervention possède la plus grande valeur en termes de coût-efficacité.

Afin de résoudre ce problème, des instruments basés sur les préférences et pouvant être uniformisés en une mesure standard ont été développés [13]. Le quality adjusted life year (QALY ou AVAQ en français : année de vie ajustée par la qualité) fait partie de ces instruments [14-15]. Son développement a entraîné un changement de paradigme, puisque le QALY ne représente pas seulement la notion d'efficacité, mais également celle d'utilité, c'est pourquoi lorsqu'on l'utilise, on parle désormais d'analyse coût-utilité. De fait, le QALY combine la qualité de vie ainsi que la longévité en un score unique tout en considérant les préférences des individus pour différentes dimensions de la santé. Il représente ainsi une mesure standard qui facilite l'allocation optimale des ressources disponibles dans le domaine de la santé [16].

La mesure du QALY peut être effectuée de façon directe ou indirecte. Parmi les méthodes d'élicitations directes, les plus fréquemment utilisées sont l'arbitrage temporel (TTO ou time trade off) et le pari

ordinaire (SG ou standard gamble). Ces deux méthodes, ainsi que d'autres, seront présentées plus bas. On peut également mesurer un QALY à l'aide d'instruments utilisant un système de notation préconçu, basé sur les préférences d'individus représentatifs de la population générale. Les instruments de ce type permettent une mesure indirecte du QALY et correspondent à des questionnaires standardisés. Ces instruments font partie de la classe des MAUI (multi-attribute utility instrument) dont les plus connus sont le EQ-5D, le SF-6D et le HUI [16]. À noter que certains de ces MAUI correspondent à des versions écourtées de questionnaires traditionnels de QVRS. Par exemple, le SF-6D a été conçu à partir du SF-36 dans le but de générer une valeur d'utilité qui puisse être dérivée des réponses obtenues au SF-36 [17]. Dans cet article, nous présentons les fondements conceptuels du QALY, ses différentes méthodes d'élicitation ainsi que les principaux MAUI utilisés.

### Fondements conceptuels du QALY

Le QALY est un instrument de mesure basé sur la théorie de l'utilité espérée [18]. Le terme QALY a été utilisé la première fois en 1976 par Zeckhauser et Shepard [19] afin de décrire la combinaison d'une unité de

temps (un an) et de la qualité de vie dans une mesure de résultat de santé. Toutefois, le concept avait été abordé un peu plus tôt lors du développement d'un « health status index » par Fanshel et Bush [20], Torrance [21] et Torrance et al. [22]. Les hypothèses sous-jacentes au modèle ont été énoncées par Pilskin et al. [23] au début des années 80 et pendant le reste de la décennie on discuta des hypothèses et fondations du QALY. Ces dernières sont exposées plus bas. En 1992, on dénombrait 51 études utilisant le QALY comme mesure de résultats d'évaluation économique en santé [24]. Deux décennies plus tard, le QALY est considéré un comme outil de référence dans l'analyse coût-utilité avec plusieurs milliers de références par année [24].

L'utilisation de deux dimensions, qualité de vie et longévité, est basée sur le principe qu'un individu passera à travers différents états de santé tout au long de sa vie. Chaque état de santé est associé à une valeur; un état plus désirable se voyant accorder une valeur plus élevée. On évalue l'utilité de cet état de santé à travers les préférences des individus au moyen d'enquêtes populationnelles. Le QALY est donc la somme, sur l'ensemble de la vie espérée, des années de vie pondérées par leur qualité [15]. Son calcul est le suivant :

$$QALYs = \sum_{t=1}^{t=T} \frac{q_t}{(1+r)^t}$$

Lorsque vient le temps de répartir les ressources de façon optimale, le QALY permet aux décideurs de faire un arbitrage entre le coût et l'efficacité d'un programme ou d'une technologie en santé. En maximisant le nombre de QALY, on cherche ainsi à maximiser les bénéfices liés à la santé des individus. Les décideurs doivent toutefois prendre en considération que le QALY n'évalue pas l'équité entre différentes situations [15]. En effet, un QALY est considéré comme identique, quel que soit l'âge ou la situation des individus. Le QALY représente ainsi la qualité de vie auto-

T : espérance de vie

$q_t$  : la qualité de vie de l'année t ( $t \in [0, T]$ )

r : taux d'actualisation

perçue d'un individu et sa valeur varie généralement entre 0 et 1 [25]. Une année complète en parfaite santé équivaut ainsi à 1 et 0 indique la mort. Notons cependant qu'un état de santé considéré comme étant pire que la mort aura une valeur négative [25]. Aussi, on considère qu'un gain de santé aura la même valeur pour un individu passant de 0,5 à 0,7 que pour un individu passant de 0,2 à 0,4 [15].

Le QALY présente de nombreux avantages. Tout d'abord, il est intuitif et simple à représenter. Ensuite, il permet de combiner plusieurs dimensions de la santé

en un score unique. Il est également reconnu comme universel en plus d'être sans ambiguïté (e.g.  $1 > 0,2$ ) et est recommandé par le National Institute for Health and Care Excellence (NICE) et d'autres agences nationales d'évaluation reconnues dans le domaine de la santé. Surtout, le QALY est basé sur les préférences des individus, ce qui est essentiel pour l'analyse économique [10]. En l'utilisant on accepte les hypothèses suivantes [15] :

- 1- Une décision liée à l'allocation des ressources doit être prise;
- 2- Les résultats des alternatives peuvent être spécifiés en termes d'état de santé, de changement ou de durée;
- 3- Les ressources sont limitées et toute alternative implique de consommer des ressources (coût);
- 4- Un objectif majeur du décideur est de maximiser la santé de la population sous contraintes de ressources;
- 5- La santé est définie en termes d'année de vie pondérée par sa qualité;
- 6- La valeur est mesurée en termes de préférence entre différents états de santé;
- 7- Chaque individu est neutre au risque et possède une fonction d'utilité additive à travers le temps;
- 8- Les préférences individuelles peuvent être agrégées et utilisées pour représenter les préférences d'un groupe;
- 9- Un QALY à la même valeur pour tout individu.

Selon Weinstein et al. [15], les six premières conditions ne posent pas de restriction sur l'évaluation des états de santé utilisés pour calculer un QALY. Par contre, les trois dernières conditions sont controversées et restrictives. La septième condition pose ainsi une hypothèse très forte et simplifie considérablement la réalité. Elle est toutefois nécessaire pour que le QALY puisse représenter la fonction d'utilité en santé des individus. La neutralité face au risque permet en effet de calculer l'espérance de vie ajustée selon la qualité et

l'additivité de la fonction d'utilité permet d'évaluer un état de santé à un moment  $t$  indépendamment d'où ce moment est situé sur une séquence de temps [15]. La huitième condition est fort utile pour l'évaluation économique, car elle permet, sur la base des préférences individuelles, de prendre des décisions collectives. Toutefois, on doit s'interroger à savoir si les préférences individuelles représentent adéquatement les préférences sociales. Plusieurs preuves empiriques semblent en effet démontrer que les individus n'accordent pas les mêmes valeurs à un état de santé [26-29]. De plus, les fonctions d'utilité doivent être cardinales et comparables entre les individus, ce qui n'est pas toujours démontré [30]. Quant à la dernière condition, celle-ci soulève une question de justice morale. En supposant qu'un QALY possède la même valeur pour tous on prend en considération seulement le gain ou la perte totale de QALY, peu importe les caractéristiques personnelles des individus, soit l'âge, la sévérité de la maladie, la classe sociale, etc. ce qui peut être moralement contestable et n'est pas sans poser problème dans le processus d'allocation des ressources en santé où divers intérêts entrent en compétition [16].

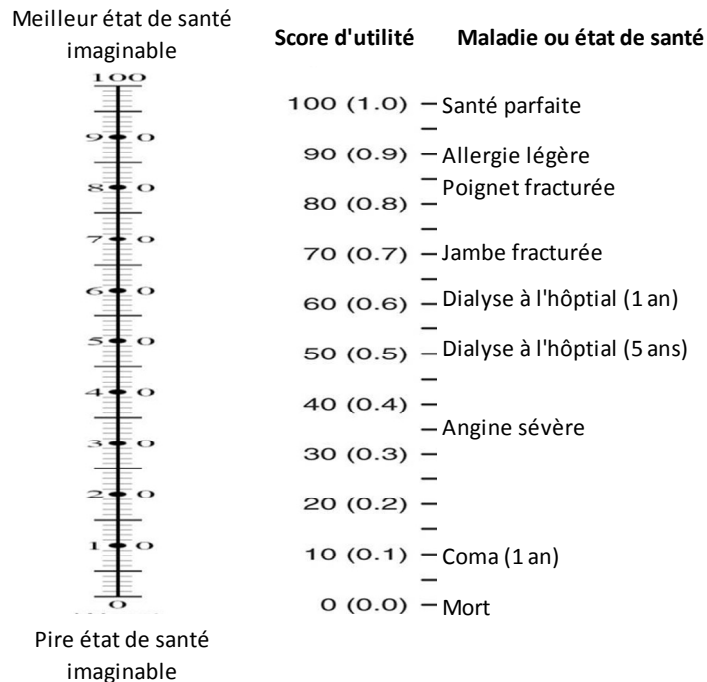
## **Méthodes d'élicitation du QALY**

### ***Mesures directes***

Afin de mesurer un QALY, plusieurs méthodes d'élicitation directe sont disponibles. La plus simple de ces méthodes utilise une échelle de notation (RS ou rating scale). Une valeur d'ancrage égale à 0 est choisie pour représenter la mort ou le pire état de santé possible, alors que la santé parfaite prend la valeur de 1 (ou 100). On demande ensuite aux répondants d'ordonner sur une échelle linéaire différents états de santé. Il devient alors simple d'observer la valeur que prend un état de santé et de le comparer à d'autres. On peut aussi en déduire combien cela coûterait aux individus de passer d'un état de santé à un autre. La valeur attribuée à l'état de santé est ici égale à sa valeur de QALY. Cette méthode est toutefois critiquée

pour ne pas être une bonne représentation des préférences. On y retrouve en effet un biais de fin d'échelle et un biais d'espacement. De plus, les individus n'effectuent aucun arbitrage entre les différents états de santé proposés, ce qui est très important en économie [31]. Ce à quoi s'ajoute le fait que les notions de temps (i.e. durée) et de risque (i.e. probabilité de survenue de l'état de santé)

sont également absentes. Dans la famille des échelles de notations, les échelles visuelles analogues (VAS ou visual analog scale) sont les plus fréquemment utilisées. Un autre exemple d'échelle de notation est l'échelle modifiée de Karnofsky qui permet d'évaluer un état de santé par rapport à d'autres dont les valeurs sont prédéfinies (Figure 1).



**Figure 1 :** Échelle visuelle analogue (à gauche) et échelle modifiée de Karnofsky (à droite)

En comparaison aux échelles de notation, la méthode de l'arbitrage temporel (TTO ou time trade-off) est une mesure d'utilité qui représente mieux les préférences des individus. Cette méthode est actuellement la plus utilisée pour calculer le QALY [31]. Les participants à un TTO se voient généralement présenter 2 scénarios, soient X nombres d'années de vie en parfaite santé et X+Z années de vie dans un état de santé intermédiaire (i.e. un état de santé moins enviable). Si le participant choisit l'état de santé parfait, on lui propose un nouvel arbitrage avec une durée de vie réduite pour l'état de santé parfait. Si le participant choisit l'état de santé intermédiaire, on augmente la durée de vie de l'état de santé parfait. Le processus est

répété jusqu'à ce que le répondant soit indifférent entre les 2 états de santé proposés. Le QALY est alors obtenu en faisant le ratio entre le nombre d'années de vie en santé parfaite et le nombre d'années de vie avec l'état de santé intermédiaire [31]. La Figure 2 permet de visualiser le principe du TTO en prenant comme exemple une situation de diabète [16]. La Figure 3 illustre quant à elle le choix qui peut être demandé à un individu en utilisant un état de santé intermédiaire décrit par le SF-6D.

Le pari ordinaire (SG ou standard gamble) permet également d'éliciter les préférences des répondants pour différents états de santé. Le SG est d'ailleurs considéré comme l'étalon-or en ce qui a trait aux

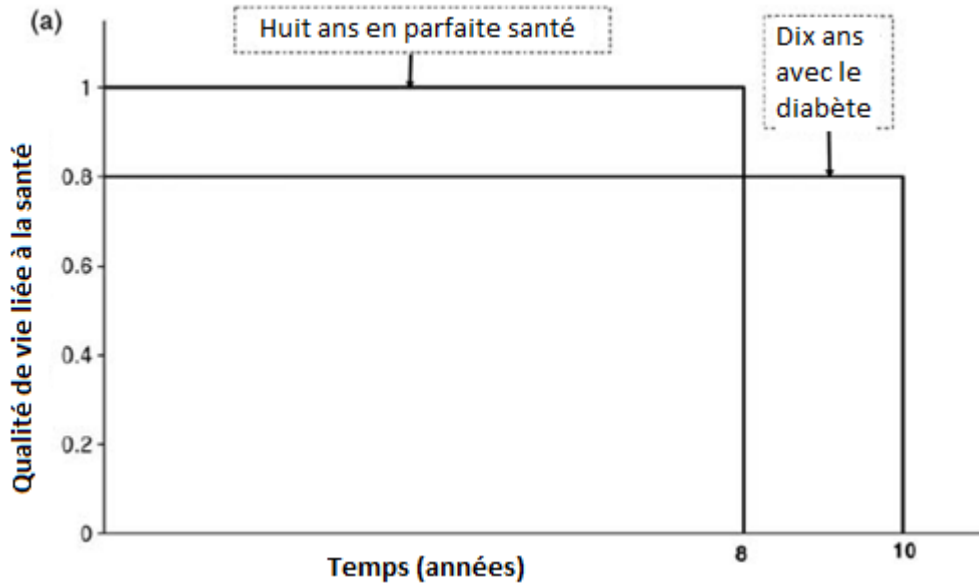


Figure 2 : Arbitrage temporel (TTO ou time trade-off). Adapté de Whitehead et Ali [16].

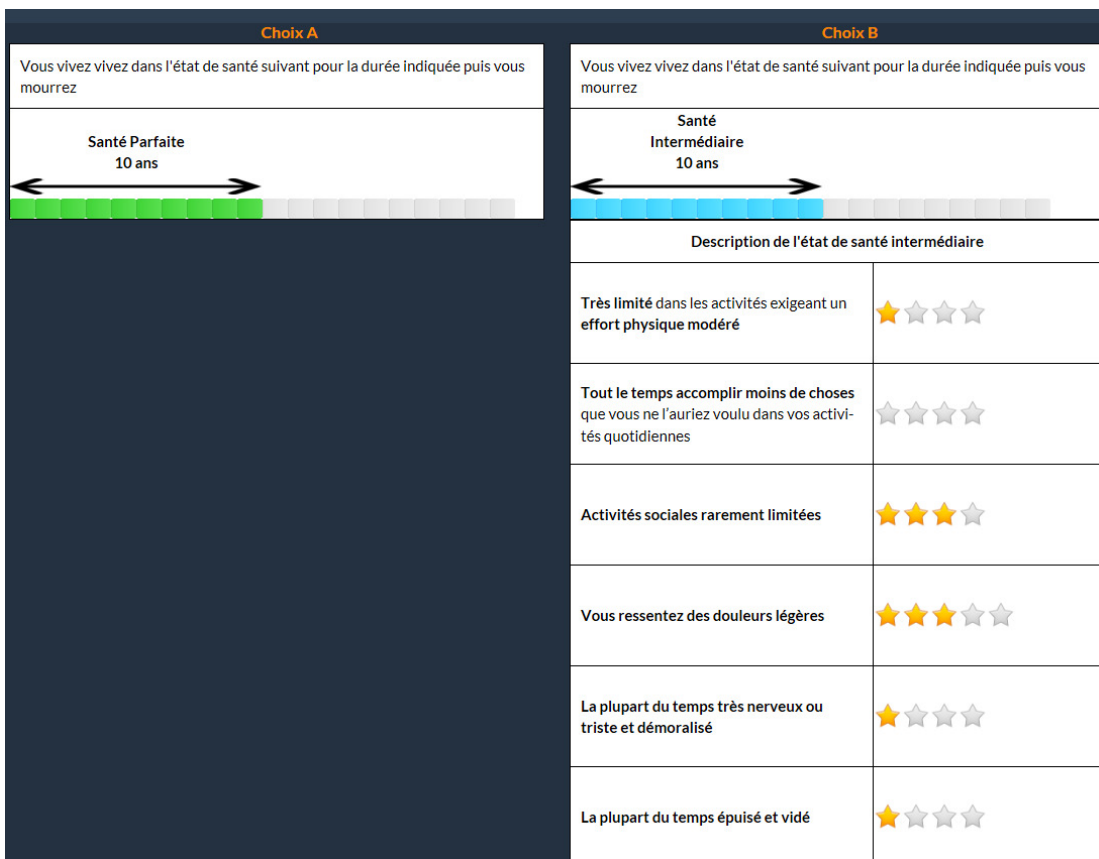


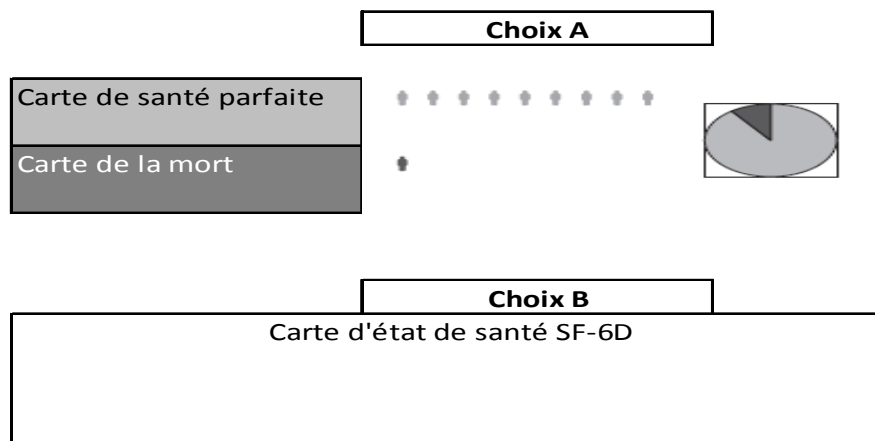
Figure 3 : Arbitrage temporel (TTO) utilisant un état de santé décrit par le SF-6D

techniques d'élicitation de QALY puisqu'il se fonde sur la théorie de l'utilité espérée [32]. La technique respecte ainsi les axiomes de préférences en situation d'incertitude, soit

la transitivité, l'indépendance et la continuité. Cette méthode possède aussi des propriétés d'échelle cohérentes avec la notion de QALY. C'est-à-dire, qu'un saut

entre 0,2 et 0,4 équivaut à un saut entre 0,6 et 0,8. Toutefois, elle comporte un biais cognitif potentiel dans la mesure où elle recourt à la notion de probabilité qui peut être difficile à interpréter par les répondants [15]. De fait, il est demandé de faire un choix entre 2 scénarios ou loteries où le résultat de l'un est certain et l'autre non. Chaque participant doit ainsi faire un choix entre : vivre X années dans un état de santé intermédiaire (i.e. situation certaine) ou bénéficier d'un traitement médical avec un pourcentage de succès Y (i.e. situation incertaine). Si le traitement réussit, le répondant vit alors X années en parfaite santé. Par contre, si le traitement échoue, c'est la mort immédiate ou vivre pendant X années dans un état de santé considéré comme étant le pire possible. Le choix de la mort immédiate ou du pire état de santé dans le scénario avec traitement est déterminé par le chercheur selon qu'il souhaite ancrer le choix du répondant par

rapport à l'une ou l'autre de ces 2 situations. Le processus est répété en faisant varier la probabilité de succès du traitement jusqu'à ce que les répondants soient indifférents entre l'état de santé intermédiaire, mais certain, et le traitement au résultat incertain. Le QALY obtenu pour l'état de santé intermédiaire correspond ainsi à la probabilité d'indifférence entre cet état de santé et le traitement proposé. À noter ici que le type de traitement proposé n'est pas nécessairement précisé. De la réponse de l'ensemble des répondants, on tire un score de préférence basé sur la moyenne des scores individuels. Cela nous donne l'utilité sociale accordée aux états de santé, qu'on transforme ensuite en QALY [12]. La Figure 4 illustre le principe de cette méthode d'élicitation où le choix A correspond au traitement médical avec résultat incertain et le choix B à un état de santé intermédiaire mais certain.



**Figure 4 :** Pari ordinaire (SG ou standard gamble)

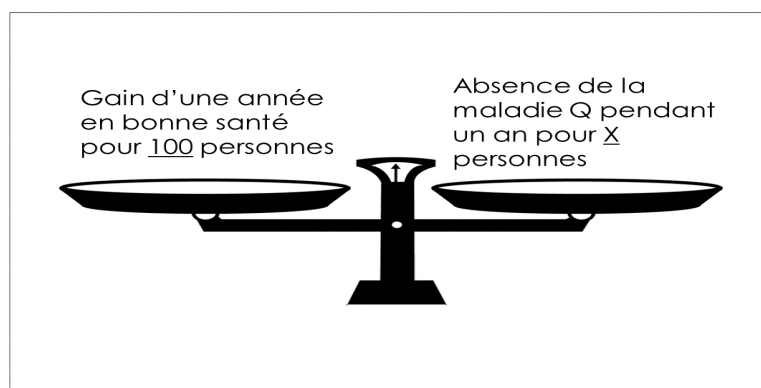
Ces trois techniques d'élicitation (i.e. RS, TTO et SG) sont les plus fréquemment utilisées pour calculer un QALY. L'avantage d'utiliser une échelle de notation comparativement au TTO ou au SG réside dans la facilité d'utilisation de cette approche. Cette technique est peu coûteuse, plus rapide et plus facile à comprendre que les autres techniques. Toutefois, il est difficile d'y évaluer la mort et plus important encore sa validité est souvent remise en question puisque les

préférences y sont évaluées dans un contexte de certitude, sans arbitrage et prise en compte de la dimension temporelle [32]. De leur côté, le SG et le TTO sont des techniques basées sur des choix contraints. Elles sont donc une mesure de la force des préférences. Le SG est considéré comme le « golden standard », car sa base théorique est plus forte que celle du TTO. En effet, le SG est basé sur la théorie de l'utilité espérée tandis que le TTO est lié à l'économie du bien-être par le concept de

choix, mais n'est pas soutenu par une théorie précise si ce n'est celle dite de l'utilité aléatoire. Cependant, le SG incorpore la notion de probabilité qui est parfois difficile à comprendre pour les répondants. De plus, la validité de la technique est également contestée, car les individus ne respectent pas toujours les axiomes liés à la théorie de l'utilité espérée [32]. Le TTO est considéré comme plus facile à comprendre que le SG par les répondants, mais la technique suppose que les choix sont constamment proportionnels, c'est-à-dire qu'un individu désirant sacrifier 1 an sur un horizon temporel de 5 ans pour avoir une santé parfaite voudrait également sacrifier 2 ans sur un horizon de 10 ans ou encore 10 ans sur un laps de 50 ans. Cela nuit à sa validité, car certains travaux avancent que la valeur accordée aux états de santé varie en fonction du nombre d'années que l'individu passera dans cet état [32-33]. En règle générale, le SG donne

des valeurs de QALY plus élevées que pour le TTO car certains individus sont averses au risque et que la technique du SG fait explicitement mention du risque associé au traitement lors des exercices de choix.

En dehors de ces méthodes couramment utilisées, d'autres sont également utilisées pour calculer les QALY, telles que l'arbitrage populationnel (PTO ou person trade-off) ou les choix expérimentaux discrets (DCE ou discrete choice experiment). Le PTO permet d'incorporer une certaine forme d'équité dans l'évaluation du QALY. Les répondants indiquent s'ils préfèrent un gain d'une année en bonne santé pour 100 personnes ou l'absence de maladie pour X individus pendant 1 an. Le processus est itéré jusqu'à indifférence [15]. Peu d'études empiriques ont été conduites avec cette technique, ce qui ne nous permet pas d'établir avec assurance sa validité [32]. La Figure 5 illustre cette méthode.



**Figure 5 :** Arbitrage populationnel (PTO ou person trade-off)

Finalement, on s'intéresse de plus en plus à la technique des choix expérimentaux discrets (DCE ou discrete choice experiment) afin de mesurer l'utilité associée à différents états de santé [34]. Si le DCE n'est pas à proprement parler une technique d'élicitation distincte du SG ou du TTO dans la mesure où elle en reprend les principes, elle utilise cependant un format novateur. De fait, un DCE sera dit TTO ou SG selon la façon dont il élicitera les préférences des individus en faisant varier les probabilités ou les durées de survie. Le

DCE est un format de plus en plus utilisé car il facilite la compréhension auprès du grand public de ce que peut être un arbitrage entre différents états de santé en demandant directement aux répondants de choisir entre 2 ou 3 états de santé lequel il préfère. Dans le cas général, deux cartes représentant des états de santé intermédiaires sont présentées au répondant. Ces cartes représentent les différentes dimensions (ou attributs) d'un état de santé qui sont associées à des niveaux et à un certain nombre d'années de



vie. Dans ce cas général, il s’agit d’un DCE de type TTO. Il est cependant possible de remplacer le nombre d’années de vie dans l’état de santé décrit par une probabilité de décès (i.e. DCE de type SG). Plus récemment, on a également combiné un DCE sans attribut de durée ou de probabilité avec un exercice de TTO afin d’ancrer le DCE avec la mort et la santé parfaite [35]. À noter que le nombre d’états de santé présenté dans un DCE n’est pas

limité à 2 ou 3 et que dans le cas d’un « best-worst choice » on peut aussi ajouter un état de santé représentant la mort immédiate. En présentant différentes cartes de choix de manière aléatoire, on peut ainsi évaluer l’ensemble des préférences des individus pour de nombreux états de santé. La Figure 6 donne un exemple de DCE de type TTO en utilisant des cartes de santé issues du SF-6D.

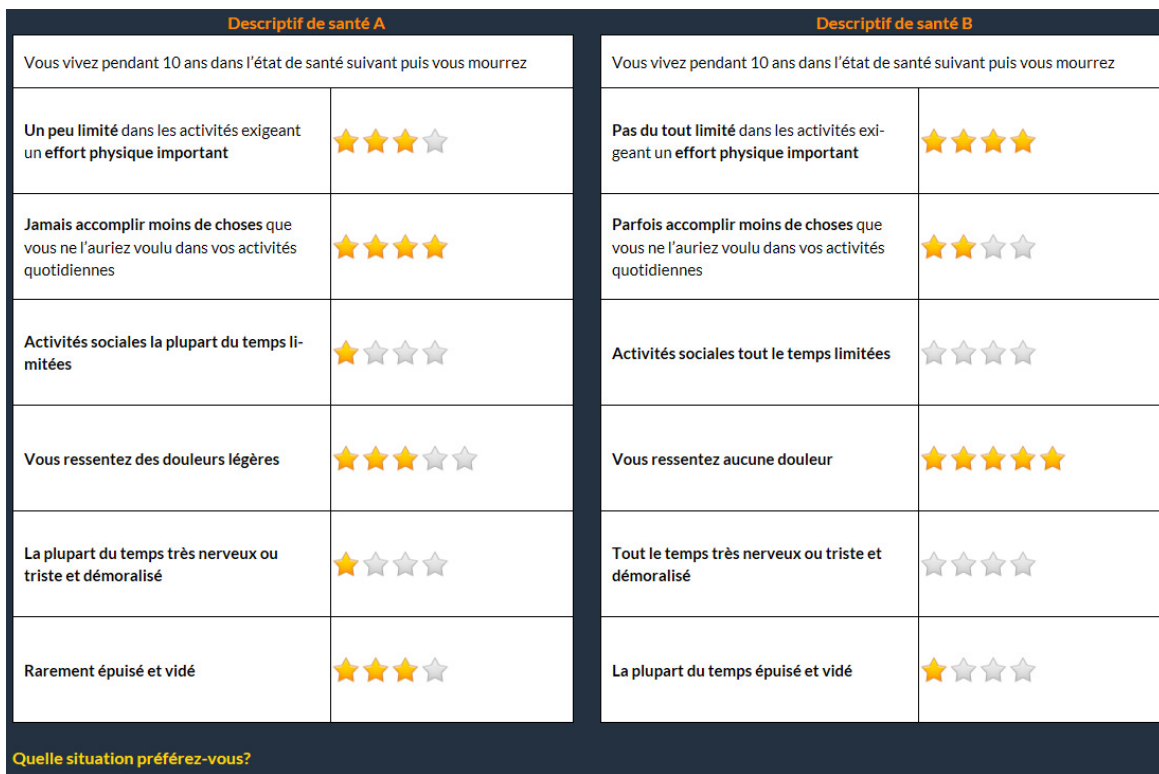


Figure 6 : Choix expérimental discret (DCE) décrit par le SF-6D

**Mesures indirectes**

Un autre chemin possible pour mesurer un QALY sans avoir à utiliser les méthodes plus ou moins sophistiquées décrites plus haut est d’utiliser un instrument d’utilité multi-attribut (MAUI ou multi-attribute utility instrument) [12,16]. Il s’agit là d’une méthode indirecte qui utilise les préférences d’un groupe de population dit représentatif pour en déduire des valeurs de QALY pour un individu ou un groupe spécifique. À noter ici que si les mesures directes se font le plus souvent dans un contexte spécifique d’une maladie et d’un

traitement, les mesures indirectes ont quant à elles été élaborées pour sortir de ces cas particuliers en utilisant des descriptions génériques d’états de santé.

Un MAUI se décompose en 2 parties : le questionnaire et son algorithme de conversion en QALY. Trois étapes importantes entrent dans son utilisation. Tout d’abord, on doit choisir un questionnaire de MAUI en se basant sur les différentes dimensions de la santé qui sont importantes pour les individus ciblés. Des exemples fréquents de questionnaires de MAUI sont le QWB, 15D, EQ-5D (3L ou 5L),

HUI (2 ou 3), SF-6D, AqoL-8D [12] et le CHU-9D [36]. Ensuite, on demande aux individus ciblés de remplir le questionnaire. Finalement, on utilise un algorithme de calcul de QALY afin d'attacher un score d'utilité aux états de santé décrits par les répondants [12]. Les algorithmes utilisés sont issus d'études ayant utilisé l'une ou l'autre des méthodes décrites ci-dessus (i.e. principalement le TTO et le SG) en se servant des mêmes dimensions et niveaux de santé que ceux décrits dans les questionnaires de MAUI. Les scores d'utilité calculés grâce à ces algorithmes de conversion peuvent alors être utilisés pour réaliser des évaluations économiques coûts-utilité. Les avantages des MAUI par rapport aux méthodes d'élicitation directes sont un faible coût d'administration, une facilité d'utilisation accrue et une durée d'administration réduite. En dehors du fait que les MAUI n'élicitent pas directement les préférences des individus ciblés, leur principal désavantage réside dans la difficulté à établir une comparaison fiable de la qualité des différents instruments utilisés. Les instruments mentionnés plus haut possèdent ainsi entre 5 et 15 dimensions. Ces dimensions intègrent chacune un item pour le HUI, 15D, EQ-5D, SF-6D et CHU-9D, tandis que le AqoL-8D inclut en moyenne 4 items par dimension. Le nombre d'états de santé définis par ces différents questionnaires varie ainsi entre 243 pour le EQ-5D-3L et  $2.37 \times 10^{23}$  pour le AqoL-8D [12]. Un tableau comparatif des dimensions et du contenu des principaux MAUI pour une population adulte peut être trouvé dans l'article de Richardson et al. [12]. Pour une clientèle pédiatrique, on se référera davantage à l'article de Ungar [37].

Les MAUI les plus populaires sont le EQ-5D, HUI3 et SF-6D [12,16]. On observe un effet très marqué de loyauté locale dans le choix d'un MAUI, c'est-à-dire qu'un instrument développé dans un pays donné est grandement utilisé dans ce pays [12]. Ce résultat est peu surprenant dans la mesure où les algorithmes de conversion en QALY pour ces MAUI sont généralement disponibles pour les seuls pays dont ils sont

issus, même si cela est de moins en moins le cas. De plus, certaines agences nationales d'évaluation, comme la Haute Autorité de Santé en France recommandent de n'utiliser des MAUI que lorsque qu'il existe un algorithme de conversion propre au pays concerné [38]. Il est cependant à noter que de plus en plus de travaux visent à développer des algorithmes de conversion d'un même MAUI pour plusieurs pays. En ce sens le EQ-5D et le SF-6D sont les MAUI possédant le plus grand nombre d'algorithmes correspondant à différents pays. Plusieurs critères s'appliquent cependant à la sélection d'un bon MAUI. Il se doit ainsi d'être pratique, fiable, court et que l'ensemble de ses composantes, c'est-à-dire, le système descriptif, la méthode d'évaluation et le type de modèle, soient valides [12]. De plus, à ces critères internes s'ajoutent des critères externes qui varient selon le but de la recherche [12]: évalue-t-on ex-ante ou ex-post, distribue-t-on le questionnaire aux patients ou à la population générale et cherche-t-on une mesure individuelle ou sociale ? Le choix de ces critères externes doit s'appliquer au contexte de la recherche. Par exemple, si l'on veut évaluer un programme spécifique, il serait approprié d'interroger ex-post ceux qui sont touchés par le programme, c'est-à-dire les patients. En effet, dans une évaluation ex-post, on vise à connaître les états de santé que vivent les patients et ces états sont valorisés par les valeurs obtenues auprès de la population générale. Par contre, si l'on désire optimiser l'allocation des ressources en santé, il faudrait interroger de façon ex-ante la population générale [15].

Comme nous avons pu le voir plus haut, les valeurs d'utilités des MAUI sont basées sur les préférences de la population générale, ce qui peut constituer une limite par rapport aux mesures directes. En effet, l'avantage d'une mesure directe, tel que souligné par Whitehead et Ali [16] et Versteegh et Brouwer [29], est que les répondants sont généralement informés sur la maladie et les états de santé associés, ce qui devrait permettre l'élicitation d'une

valeur d'utilité en relation avec le vécu de cet état de santé. Si le répondant est un membre du public, son élicitation est basée sur sa connaissance hypothétique de l'état de santé. S'il s'agit d'un patient, cette élicitation est basée sur son expérience. On voit donc ici qu'utiliser des MAUI basés sur les préférences de la population générale conduit potentiellement à utiliser des tables d'utilité ne reflétant pas les véritables utilités vécues par les individus étant réellement dans ces états de santé. Toutefois, on s'inquiète aussi du biais que peut créer la capacité d'adaptation des patients à leurs propres états de santé. Empiriquement, les effets de l'adaptation sont contradictoires et peuvent parfois avantager les patients et parfois leur nuire [29]. Actuellement, l'argument principal pour utiliser les MAUI dans le cadre d'évaluations ex-ante se rapporte au fait que les préférences de la population générale représentent la perspective sociétale [29]. L'idée étant ici que puisque l'on s'interroge sur l'allocation de ressources publiques on se doit d'utiliser les préférences de la population générale [16]. De plus, l'ensemble des citoyens n'ont souvent pas d'intérêt particulier dans l'évaluation des états de santé, ils offrent donc en théorie un regard non biaisé (i.e. voile d'ignorance). Toutefois, ils sont bien moins informés que les patients, ce qui pose le problème d'un biais hypothétique dans l'évaluation des valeurs de QALY [16]. Bien que ces questions soient non résolues dans la littérature actuelle, Versteegh et Brouwer [29] émettent deux pistes de solutions, soit : informer la population générale sur les différents états de santé tels qu'expérimentés par les patients ou encore reconnaître la pluralité des points de vue et sonder à la fois la population générale et les patients.

## Conclusion

Le QALY est un instrument de plus en plus utilisé par les décideurs afin de procéder à des choix en santé et d'optimiser l'utilisation des ressources disponibles. Si cet outil présente de nombreux avantages,

notamment celui d'être facile et intuitif à comprendre et de permettre une comparaison entre plusieurs programmes ou interventions en santé pour différentes pathologies ou problématiques, celui-ci montre également certaines faiblesses tant au niveau théorique que pratique. En l'absence d'alternative crédible au QALY, les travaux visant à résoudre les problèmes posés par l'utilisation des QALY constituent toutefois l'approche à privilégier. Dans cet article, nous avons présentés les fondements du QALY et les différentes méthodes utilisées pour éliciter les valeurs d'utilité associées à différents états de santé. Chacune de ces méthodes présente ses avantages et inconvénients et aucun accord n'existe dans la littérature sur celles étant les plus robustes ou les plus appropriées [39]. Un constat qui s'étend également au choix des MAUI et à la source des préférences à utiliser pour ces derniers (i.e. préférences de la population générale ou des patients). En outre, le débat reste ouvert à savoir si « a QALY is a QALY is a QALY » ou encore si l'agrégation des préférences individuelles représente nécessairement les préférences sociétales [40]. On constate ainsi que si le QALY a fait l'objet d'un très fort développement méthodologique au cours des quarante dernières années, une étape importante reste encore à franchir pour passer d'un outil conceptuel puissant à celui d'un outil reconnu et validé. Il est ainsi nécessaire de garder à l'esprit que le QALY doit être utilisé avec précaution et être considéré pour ce qu'il est dans l'aide à la prise de décision, c'est-à-dire un élément à considérer parmi d'autres et non la pierre angulaire [39].

## Remerciement

Nous remercions un évaluateur anonyme pour ses commentaires avisés.

## Financement

Ce travail a été financé grâce à un fonds de recherche du Conseil de Recherche en Sciences Humaines (CRSH) du Canada, Programme de Subventions de développement Savoir 2015-2017, numéro de dossier 430-2015-00712.

## Conflit d'intérêt

Aucun.

## Références

- [1] OMS. Constitution de l'Organisation mondiale de la Santé. Organisation mondiale de la santé: Genève; 1946. Disponible au : <http://apps.who.int/gb/bd/PDF/bd47/FR/constitution-fr.pdf?ua=1>
- [2] Elkinton JR. Medicine and the quality of life. *Ann Intern Med* 1966;64:711-714.
- [3] Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. *Medical Care* 1992;30(6):473-483.
- [4] Hunt S, McEwen J, McKenna S. Measuring health status: a new tool for clinicians and epidemiologists. *J R Coll Gen Pract* 1985;35(273):185-188.
- [5] Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, sw Haes JC et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85(5):365-76.
- [6] The DCCT Research group. Reliability and Validity of a Diabetes Quality-of-Life Measure for the Diabetes Control and Complications Trial (DCCT). *Diabetes Care* 1988;11(9):725-732.
- [7] Mchorney A, Ware J, Lu R, Sherbourne C. The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of Data Quality, Scaling Assumptions and Reliability across Diverse Patient Groups. *Medical Care* 1994;32(1):40-66.
- [8] Salaffi F, Sarzi-Puttini P, Girolimetti R, Gasparini S, Atzeni F, Grassi W. Development and validation of the self-administered Fibromyalgia Assessment Status: a disease-specific composite measure for evaluating treatment effect. *Arthritis Res Ther* 2009;11(4):R125.
- [9] Watkins K, Connell CM. Measurement of health-related QOL in diabetes mellitus. *Pharmacoeconomics* 2004;22(17):1109-1126.
- [10] Brazier J. The Short-Form 36 (SF-36) Health Survey and Its Use in Pharmacoeconomic Evaluation. *Pharmacoeconomics* 1995;7(5):403-415.
- [11] Fayers P, Hays R. Assessing quality of life in clinical trials. Second edition. Oxford University Press: New York; 2005.
- [12] Richardson J, McKie J, Bariola E. Multiattribute Utility Instruments and Their Use. *Encyclopedia of Health Economics* 2014;2:341-357.
- [13] Glied S, Smith PC. *The Oxford Handbook of Health economics*. Oxford University Press: New York; 2013.
- [14] Weinstein MC, Stason W. Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med* 1977(13);296:716-721.
- [15] Weinstein MC, Torrance G, McGuire A. QALYs: The Basics. *Value Health* 2009;12:S5-S9.
- [16] Whitehead A, Ali S. Health outcome in economic evaluation: the QALY and utilities. *Br Med Bull* 2010;96:5-21.
- [17] Brazier J, Usherwood T, Harper R, Thomas K. Deriving a Preference-Based Single Index from the UK SF-36 Health Survey. *J Clin Epidemiol* 1998;51(11):1115-1128.
- [18] McFadden D. The Measurement of Urban Travel Demand. *Journal of Public Economics* 1974;3:303-328.
- [19] Zeckhauser R, Shepard D. Where now for savings lives? *Law and Contemporary Problems* 1976;40:5-45.
- [20] Fanshel S, Bush JW. A health-status index and its application to health-services outcomes. *Operations Research* 1970;18:1021-66.
- [21] Torrance G. A generalized cost-effectiveness model for the evaluation of health programs. Research report 101. McMaster University, Faculty of Business: Hamilton, Ontario; 1970.
- [22] Torrance G, Thomas W, Sackett D. A Utility Maximization Model for Evaluation of Health Care Programs. *Health Serv Res* 1972;7(2):118-131.
- [23] Pilskin JS, Shepard DS, Weinstein MC. Utility functions for life years and health status. *Operations Research* 1980;28:206-224.
- [24] Sassi F. How to do (or not to do)... Calculating QALYs, comparing QALY and DALY calculations. Oxford University Press in association with The London School of Hygiene and Tropical Medicine; 2006.
- [25] Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. *Health Econ* 2005;14(2):197-208.
- [26] Joannesson M, Johansson P. Quality of life and the WTP for an increased expectancy at an advanced age. *Journal of Public Economics* 1997;65:219-228.
- [27] Loomes G, McKenzie L. The use of QALYs in health care decision-making. *Soc Sci Med* 1989;28:209-308.
- [28] Weinstein MC. Economic Techniques for Technology Assessment. *The Economics of Medical Technology*. Proceedings of an International Conference on Economics of Medical Technology, 1988;44-53. Disponible au : [https://link.springer.com/chapter/10.1007/978-3-642-72785-6\\_7](https://link.springer.com/chapter/10.1007/978-3-642-72785-6_7)
- [29] Versteegh M, Brouwer W. Patient and general public preferences for health states: A call to reconsider current guidelines. *Soc Sci Med* 2016;165:66-74.
- [30] Liljas B, Lindgren B. On Individual Preferences and Aggregation in Economic Evaluation in Healthcare. *Pharmacoeconomics* 2001;19(4):323-335.
- [31] Garza A, Wyrwich K. Health Utility Measures and the Standard Gamble. *Acad Emerg Med* 2003;10(4):360-363.
- [32] Ryan M, Scott D, Reeves C, Bate A, Teijlingen E, Russell E, Napper M, Robb C. Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technology Assessment* 2001;5(5):17-38.
- [33] Jonker MF. Non-linear time preferences in health state valuations; time for time-corrected QALY tariffs? Presented at the 6th Meeting of the International Academy of Health Preference Research, 6-7 July 2017, Boston, USA. Disponible au : <http://iahpr.org/wordpress/wp-content/uploads/2016/09/Program170709.pdf>

- [34] Lancsar E, Louviere J, Donaldson C, Currie G, Burgess L. Best worst discrete choice experiments in health: Methods and an application. *Soc Sci Med* 2013;76(1):74-82.
- [35] Bailey H, Stolk E, Kind P. Toward Explicit Prioritization for the Caribbean: An EQ-5D Value Set for Trinidad and Tobago. *Value Health Reg Issues* 2016;11:60-67.
- [36] Stevens K. Valuation of the Child Health Utility 9D Index. *PharmacoEconomics* 2012;30(8):729-747.
- [37] Ungar WJ. Challenge in Health State Valuation in Paediatric Economic Evaluation – Are QALYs contraindicated? *PharmacoEconomics* 2011;29(8):641-52.
- [38] Haute Autorité de Santé. Choix méthodologiques pour l'évaluation économique à la HAS. 2011, Paris : HAS. Disponible au : [https://www.has-sante.fr/portail/upload/docs/application/pdf/2011-11/guide\\_methodo\\_vf.pdf](https://www.has-sante.fr/portail/upload/docs/application/pdf/2011-11/guide_methodo_vf.pdf)
- [39] Drummond M, Brixner D, Gold M, Kind P, McGuire A, Nord E. Toward a Consensus on the QALY. *Value in Health* 2009;12(1):S31-S35.
- [40] Nord E, Daniels N, Kamlet M. QALYs: Some Challenges. *Value in Health* 2009;12(1):S10-S15.