

QALYs or not QALYs : that is the question ? *QALYs or not QALYs : that is the question ?*

J.P. MOATTI⁽¹⁾, P. AUQUIER⁽¹⁻²⁾, A.-G. LE COROLLER⁽¹⁾, G. MACQUART-MOULIN⁽¹⁾

Revue Epidém. et Santé Publ., 1995, 43: 573-583

- (1) INSERM Unité 379 « Epidémiologie et sciences sociales appliquées à l'innovation médicale », Institut Paoli-Calmettes, 232 Bld Sainte-Marguerite, 13273 Marseille Cedex 9
- (2) Laboratoire de Santé Publique, Faculté de Médecine, Marseille

The article discusses the proposal of some health economists to use the «cost per QALY (quality-adjusted-life year)» ratio as an universal indicator for economic assessment of medical interventions, in the so-called «cost-utility» analyses. Authors argue that QALYs are not a straightforward application of expected utility theory, which is the standard economic model of individual behaviours toward risk and uncertainty. Indeed, QALYs are compatible with economic utility theory only if individuals' preferences regarding health states satisfy certain very restrictive properties : utility independence between length of life and quality of life, constancy of the proportional trade-off between quality of life and length of life, risk neutrality towards health states, constancy through time of the utility associated with each health state. Aggregation of individual QALYs to obtain an indicator for patient groups at the societal level also raises complex equity problems. Last but not least, from the epistemological point of view, QALYs are based on the hypothesis that health interventions only affect the health of the individual and not any other aspects of his well-being. The authors conclude that the « cost per QALY » approach should be abandoned in order to avoid ambiguities that could impede the development of health economics in the medical field.

Quality-adjusted-life-years (QALYs). Economic Assessment. Expected Utility Theory. Cost-utility analysis.

L'article discute la proposition de certains économistes de la santé de faire d'un ratio de coût par QALY (année de vie ajustée sur la qualité) un indicateur universel pour la comparaison économique de stratégies médicales par des analyses dites « coût-utilité ». Contrairement aux affirmations des promoteurs de cette approche, les QALYs ne sont pas une application directe de la théorie de l'utilité espérée qui est effectivement le modèle économique de référence en matière de comportements des agents individuels en situation de risque et d'incertitude. La construction d'un indicateur QALY suppose des hypothèses restrictives sur les préférences individuelles qui sont peu conformes aux comportements réels : stricte indépendance des arbitrages sur la survie et ceux sur la qualité de vie, constance de la proportionnalité de ces arbitrages, « neutralité » face au risque, constance au cours du temps de l'utilité associée à chaque état de santé. L'agrégation des QALYs individuels pour obtenir une mesure collective dans des groupes de populations soulève également de délicats problèmes d'équité. Enfin, du point de vue épistémologique, le QALY présuppose de façon discutable qu'une intervention médicale n'a d'influence que sur l'état de santé stricto-sensu et n'affecte en rien les autres éléments qui font qu'un temps de vie apporte de l'utilité à l'individu. L'article conclut à la nécessité d'abandonner purement et simplement les approches de « coût par QALY », sous peine d'engendrer des confusions multiples qui seraient en définitive préjudiciables à la diffusion des méthodes d'analyse économique appliquées à la santé.

Année de vie ajustée sur la qualité. Evaluation économique. Théorie de l'utilité. Analyse coût-utilité.

Depuis plus d'une dizaine d'années, des analyses économiques dites « coût-utilité » qui s'efforcent de comparer des stratégies médicales et sanitaires sur un critère de coût par « *année de vie ajustée sur la qualité* » (QALY) ont été développées dans la littérature internationale en économie de la santé¹. Tout en reconnaissant les hypothèses fortes qui sous-tendent la construction d'un indicateur de type QALY, certains économistes ont même proposé de se fonder sur ce critère de coût par QALY pour orienter les choix de la collectivité en matière d'allocation des ressources entre les différents secteurs de la médecine et de la santé²⁻³. La Revue d'Epidémiologie et de Santé Publique s'est récemment faite l'écho d'un tel point de vue⁴. Les arguments avancés en faveur de cette approche insistent sur le caractère opérationnel et « synthétique » d'un indicateur QALY censé « mettre en évidence les relations existantes entre les coûts additionnels (d'une intervention médicale) et l'amélioration observée du service médical rendu » ; mais, ils invoquent également le fait que la notion de QALY reposerait sur un fondement scientifique incontestable : la théorie de l'utilité espérée. Cette théorie constitue en effet la référence fondamentale des approches micro-économiques visant à modéliser les comportements des agents confrontés à des choix en situation de risque ou d'incertitude⁵.

Ce volet théorique de l'argumentation des promoteurs des approches « coût-utilité » et QALY a, en fait, suscité d'intenses débats parmi les économistes de la santé, une partie d'entre eux ayant d'emblée contesté la cohérence revendiquée entre une démarche QALY et les fondements de la théorie économique de la décision en régime d'incertitude⁶. Au-delà des préoccupations théoriques des économistes qui se réclament de « l'économie du bien-être » (welfare economics)⁷, ces débats ont des conséquences pratiques immédiates pour la mise en œuvre des évaluations économiques de stratégies et pratiques médicales⁸.

Dans cet article, nous voudrions montrer que selon nous, il est illégitime d'assimiler un indicateur QALY à une mesure de l'utilité individuelle, au sens de la théorie économique. Nous en déduisons que la terminologie d'analyse « coût-utilité » est en elle-même fallacieuse. Nous nous efforcerons d'étayer la recommandation concrète pour les économistes de la santé d'abandonner purement et simplement les approches de « *coût par QALY* », sous peine d'engendrer des confusions multiples qui seraient en définitive préjudiciables à la diffusion des méthodes d'analyse économique appliquées dans le champ de la santé.

DE L'UTILITÉ ESPÉRÉE AUX QAL Ys

Les économistes disposent effectivement d'un cadre théorique pour rendre compte de la rationalité des agents confrontés à des choix en incertitude, dite théorie de l'utilité espérée dont on fait remonter la formalisation à Von Neumann et Morgenstern⁹. Confronté au choix entre différentes actions impliquant différents résultats, l'individu rationnel est censé choisir celle qui maximise l'utilité espérée du résultat. Le calcul d'espérance d'utilité nécessite la définition préalable d'un ensemble « d'états de la nature » Θ et des probabilités qui leur sont associées $P(\theta_j)$, ainsi que la détermination d'un indice d'utilité U associé subjectivement par l'individu aux conséquences. L'action optimale parmi l'ensemble A des actions (a_i) envisageables est alors celle qui maximise la quantité suivante :

$$E [a_i] = \sum_{j=1}^n U(O_{ij}) \times P(\theta_j),$$

avec O_{ij} la conséquence associée au choix de l'action a_i , si se produit l'état de la nature θ_j .

Un exemple concret permettra d'illustrer schématiquement ce mode de raisonnement dont on pourra trouver ailleurs une présentation détaillée et rigoureuse (voir notamment¹⁰ et la note (1) de l'appendice). Supposons qu'un patient (ou un clinicien devant agir pour le meilleur intérêt de son patient) soit confronté au choix entre deux traitements d'un eczéma dont les coûts financiers sont identiques. Le traitement 1 (T1) guérit l'eczéma avec certitude en 2 semaines et sans effet secondaire. Le second traitement (T2) a un résultat moins prévisible : dans 80 % des cas, l'eczéma est guéri en un jour sans effet secondaire ; dans 15 % des cas, il est guéri avec certitude en 2 semaines et sans effet secondaire (situation équivalente au résultat de T1) ; enfin, dans 5 % des cas le traitement agit en 2 semaines mais s'accompagne d'une forte réaction fébrile qui oblige à garder le lit. Comment choisir entre T1 et T2 ? L'utilité espérée est censée fournir une solution « rationnelle » en reformulant les termes du choix.

Classons les trois résultats de T1 et T2 par ordre de préférence décroissant : A, B et C. En premier lieu, l'individu préférera être guéri en un jour sans effet secondaire (A) ; en second lieu, il préférera être guéri en 2 semaines sans effet secondaire (B) ; enfin, être guéri en 2 semaines au prix d'une grave réaction secondaire (C) constituera certainement l'état le moins souhaitable pour l'individu. Affectons une valeur maximale (par ex. 1) à la valeur (l'utilité) que l'individu accorde au résultat le meilleur $U(A)=1$ et une valeur minimale (0) à celle qu'il accorde au pire $U(C)=0$. Le problème peut alors être reformulé comme un choix pour l'individu entre l'alternative suivante : soit bénéficier avec certitude du résultat intermédiaire (B), soit jouer une « loterie » qui lui fera bénéficier du meilleur résultat avec une probabilité (p) ou au contraire du pire résultat avec une probabilité (1-p). Il en découle immédiatement que l'utilité associée par l'individu au résultat intermédiaire $U(B)$ est égal à la probabilité (p) qui laisse l'individu « indifférent » entre le fait de jouer la loterie ou celui d'obtenir (B) avec certitude.

Dans cet exemple, l'utilité espérée du traitement 1 est égale à $U(B)$ puisque l'occurrence de (B) est certaine $E[T1] = U(B)$. Celle du traitement 2 est égale à la somme des utilités associées à chacune des conséquences possibles (A, B, C) pondérées par les probabilités d'occurrence de chacune d'elles :

$$\begin{aligned} E[T2] &= [0,8 \times U(A)] + [0,15 \times U(B)] + [0,05 \times U(C)] \\ &= (0,8 \times 1) + (0,15 \times U(B)) + (0,05 \times 0) \end{aligned}$$

L'individu préférera le traitement 1 au traitement 2 si son utilité espérée est supérieure, c'est-à-dire si :

$$E[T1] > E[T2], \text{ soit si : } U(B) > 0,8 + 0,15 U(B) \text{ soit } U(B) > 0,941.$$

Les analyses « coût-efficacité » sont les plus fréquentes dans la littérature d'évaluation médico-économique visant à comparer des stratégies médicales. Ces analyses rapportent les ressources consommées par chaque stratégie à un indicateur d'efficacité médicale exprimé en unités physiques (par ex., un pourcentage de guérisons, ou un nombre de décès évités, ou un nombre d'années de vie gagnées ou enfin l'amélioration d'un indicateur clinique ou biologique dont la valeur pronostique est démontrée). En principe, le passage des analyses « coût-efficacité » aux analyses « coût-utilité » consiste simplement à modifier l'indicateur de résultat en se basant sur un autre indicateur, exprimant cette fois l'utilité espérée, pour le patient, du résultat de chaque stratégie (T_i). Si $\Delta U(s_k)$ est le gain (ou la perte) d'utilité ressenti par le patient du fait du passage de son état de santé initial (s_0) à un autre état de santé (s_k) [$\Delta U(s_k) = U(s_k) - U(s_0)$], l'utilité espérée associée au traitement T_i sera :

$$\sum_k P_{ki} \Delta U(s_k), \text{ avec } (P_{ki})_{k \in K}$$

la distribution de probabilité sur les conséquences possibles qui représente l'incertitude sur les résultats de T_i . Et, la comparaison entre deux traitements T_i et T_j s'effectuera logiquement sur la base :

$$T_i \geq T_j \text{ si } \sum_k P_{ki} \Delta U(s_k) \geq \sum_k P_{kj} \Delta U(s_k)$$

Ce n'est pas ici le lieu de discuter, de façon approfondie, les limites générales de validité des analyses « coût-efficacité » ou « coût-utilité » (Appendice 2)¹¹⁻¹²⁻¹³. Nous nous concentrerons sur la question que soulève l'élaboration des indicateurs empiriques censés « représenter » en pratique l'utilité associée par les individus aux résultats d'une intervention médicale améliorant leur état de santé.

La solution la plus couramment proposée pour la mesure de l'utilité des états de santé est celle du recours à un indicateur QALY (année de vie ajustée sur la qualité). Weinstein & Fineberg¹⁴ définissent les QALYs comme « les années de vie en parfaite santé qui seront considérées comme équivalentes aux années effectivement vécues dans un état de santé donné », ce qui revient donc à proposer un arbitrage entre survie et qualité de cette survie. La forme la plus connue du QALY est celle de la matrice proposée par Kind *et al.*¹⁵, qui est illustrée dans le *tableau 1*. Elle associe des états de santé, définis par la combinaison d'une échelle d'incapacité physique et d'une échelle de détresse psychologique, avec une pondération « subjective » sur la valeur (utilité) accordée par les individus à un temps donné passé dans chaque état. Cette pondération varie de 0 « décès » à 1 « état de parfaite santé », certaines situations pouvant même se voir associées des pondérations de valeur négative car considérées comme « pires que la mort » (voir les cases VII D et VIII A du *tableau 1*). La matrice QALY est censée permettre une comparaison quantitative (« cardinale » disent les économistes) de l'utilité associée à différents états de santé. Par exemple, dans le *tableau 1*, un an dans la situation VI B (mobilité impossible sans aide, associée à des troubles psychologiques bénins) est considéré comme équivalent à 84,5 % d'une année en parfaite santé.

L'indicateur QALY peut apparaître comme une solution universelle à tout problème de comparaison entre des stratégies médicales. Ces stratégies peuvent être incompatibles c'est-à-dire qu'elles visent une même indication pathologique et sont alternatives les unes aux autres (par ex. comparaison de différents protocoles thérapeutiques alternatifs dans les cancers du sein métastatiques). Mais, il peut également s'agir de stratégies compatibles entre elles, par exemple pour arbitrer l'allocation des ressources entre différents secteurs de la médecine, comme cela est suggéré par les « league tables » qui présentent les ratios coût marginal par QALY d'interventions sanitaires diverses allant de la prévention du tabagisme à la dialyse⁴.

TABLEAU I. — Matrice d'évaluation de la qualité de vie.

Echelle d'incapacité (a)	Echelle de détresse psychologique			
	A Aucune	B Bénigne	C Modérée	D Sévère
I	1,000	0,995	0,990	0,967
II	0,990	0,986	0,973	0,932
III	0,980	0,972	0,956	0,912
IV	0,964	0,956	0,942	0,870
V	0,946	0,935	0,900	0,700
VI	0,875	0,845	0,680	0,000
VII	0,677	0,564	0,000	- 1,486 (**)
VIII	- 1,028		Non applicable	

Parfaite santé = 1 ; Décès = 0

Source : [10].

(a) Echelle d'incapacité

I = Aucune

II = Incapacité sociale limitée

III = Incapacité sociale sévère et/ou réduction de la performance au travail ; capable de réaliser des tâches ménagères

IV = Performance professionnelle sévèrement limitée ; tâches ménagères limitées mais possibilité de se déplacer pour faire les courses soi-même

V = Impossibilité de travailler : personnes âgées confinées chez elles

VI = Mobilité impossible sans aide

VII = Confiné au lit

VIII = Etat d'inconscience.

Le QALY est censé fournir une valeur numérique agrégée du résultat de différentes stratégies médicales, particulièrement commode pour la comparaison économique de ces stratégies :

- Le calcul des QALYs tient compte de l'évolution au cours du temps, pour chaque patient, de l'utilité qu'il attribue à son état de santé. Si l'intervention T0 produit une première année valant 0,4 QALY, une seconde 0,2 et une troisième 0,7, tandis qu'une intervention T1 produit pour le même patient respectivement 0,9, 0,6 puis 0,9, alors on pourra en déduire que le traitement T1 permet un gain additionnel de 1,1 (2,4 - 1,3) QALYs par rapport à T0. En termes généraux, si l'on compare une stratégie médicale i par rapport à une stratégie initiale j pour l'individu n pendant une durée de T années (définie du moment présent 0 à T moment du futur où les conséquences sur la santé de l'individu n des deux interventions i et j peuvent par hypothèse être considérées comme identiques, par exemple du fait du décès de n, on aura :

$$\Delta \text{QALY} (i,j,n) = \sum_{t=0}^{T-1} U_n [s_{it}] - \sum_{t=0}^{T-1} U_n [s_{jt}]$$

avec s_{it} l'état de santé de l'individu n, lorsque la stratégie T_i a été mise en œuvre, durant la période $[t, t+1]$ (cet état étant supposé constant sur la période d'une année) ;

- Le calcul des QALYs tient compte de l'incertitude et du caractère probabiliste des résultats de la plupart des traitements et interventions en pondérant les QALYs par leur probabilité d'occurrence. Par exemple, si T_1 a 50 % de chance de mettre le patient pendant un an dans un état de santé correspondant à 0,9 QALY et 50 % dans un état moins favorable de 0,4, le résultat de T_1 pour la première année sera valorisé à 0,65 QALY; avec s_{it} l'état de santé de l'individu n, lorsque la stratégie T_i a été mise en œuvre, durant la période $[t, t+1]$ (cet état étant supposé constant sur la période d'une année) ;
- Enfin, les QALYs permettent d'agréger les résultats individuels pour comparer des groupes (y compris très différents de patients) ; si l'on considère un groupe de X individus, on posera :

$$\Delta \text{QALY} (i,j, X) = \sum_{n=1}^X \Delta \text{QALY} (i,j,n)$$

Cette commodité d'utilisation apparente explique l'engouement suscité par l'indicateur de coût par QALY auprès de nombreux auteurs. Malheureusement, elle repose sur un certain nombre de présupposés implicites. Sans prétendre rendre compte de la vaste littérature sur ce sujet, nous évoquerons successivement à propos des QALYs : 1) les hypothèses réductrices sur la forme des préférences individuelles ; 2) les problèmes liés à la prise en compte du temps ; 3) ceux liés à l'agrégation des préférences individuelles ; 4) les hypothèses réductrices sur la relation santé-bien-être global de l'individu. Ces hypothèses conduisent à mettre en doute la prétendue cohérence théorique entre les QALYs et la mesure économique de l'utilité et établissent même, selon nous, un fossé quasi infranchissable entre les deux.

QUATRE RAISONS D'UN FOSSÉ INFRANCHISSABLE

Les hypothèses sur les préférences individuelles

Les QALYs peuvent-ils légitimement être considérés comme une mesure appropriée de l'utilité espérée, au sens de la théorie économique ? Une abondante littérature théorique (voir par ex.⁶⁻¹⁶) a démontré qu'une réponse positive ne pouvait être donnée à cette question qu'à la condition que les préférences des individus concernant les différents états de santé obéissent à des hypothèses très particulières, dont on peut raisonnablement penser qu'un grand nombre d'individus ne s'y conforment pas. On peut repérer au moins trois hypothèses sous-jacentes pour que les QALYs puissent être considérés comme une mesure adéquate de l'utilité, hypothèses dont nous allons souligner les contraintes et les conséquences.

Une première hypothèse est que les arbitrages que l'individu fait en matière de durée de vie demeurent strictement indépendants de ceux qu'il effectue sur la qualité de vie. Si l'individu est « indifférent » entre une survie de 14 ans avec un handicap majeur et une « loterie » avec une chance sur deux de vivre 10 ans ou 20 ans avec ce même handicap, il devra l'être « automatiquement » s'il est confronté au même arbitrage sur sa durée de vie dans un autre état de santé : il devra donc par exemple être indifférent entre 14 ans en parfaite santé et une loterie avec une chance sur deux de vivre 10 ou 20 ans en parfaite santé. Inversement, s'il est indifférent entre la certitude de vivre 20 ans avec un handicap majeur et une loterie avec une chance sur deux de vivre 20 ans avec ou sans ce handicap majeur, il devra maintenir un arbitrage identique si le choix de vivre avec ou sans handicap concerne une durée différente (10 ans ou 30 ans). Cette hypothèse d'indépendance des utilités implique également que les préférences ne sont pas modifiées en fonction de la séquence temporelle où différents états de santé se présentent : l'individu sera censé être totalement indifférent entre « se casser la jambe dans un an et avoir une rage de dents dans une semaine » et « se casser la jambe dans une semaine et avoir une rage de dents dans un an ».

Une seconde hypothèse pour que les QALYs correspondent aux principes de mesure de l'utilité individuelle est moins fréquemment relevée dans la littérature théorique sur le sujet. Elle concerne la constance dans la proportionnalité des arbitrages entre survie et qualité de vie. Cette hypothèse peut être illustrée par l'exemple suivant: si un individu est indifférent entre 40 années de vie avec un handicap et 32 années sans ce handicap, cela signifie qu'il est prêt à « sacrifier » 1/5 de son espérance de vie en échange d'une meilleure qualité. Une mesure QALY présuppose que l'individu maintiendra toujours cette même proportionnalité dans l'arbitrage, et qu'il sera donc prêt à échanger 30 ans avec ce même handicap contre 24 ans sans celui-ci.

Enfin, une troisième hypothèse sous-jacente aux QALYs est que les individus sont neutres par rapport au risque lorsqu'il s'agit de leur espérance de vie, c'est-à-dire qu'ils sont toujours indifférents entre une loterie portant sur leur survie et la certitude (équivalent certain) d'une survie égale à l'espérance mathématique de cette loterie. Cette hypothèse de neutralité face au risque signifie, par exemple, que tout individu sera censé être indifférent entre la certitude de 5 années de vie supplémentaires et une chance sur deux de décéder immédiatement ou de bénéficier de 10 années supplémentaires.

Or, l'un des intérêts majeurs de la théorie économique de l'utilité espérée est justement de pouvoir tenir compte d'attitudes individuelles différentes face au risque. Confrontés à une même loterie avec des probabilités identiques de gains ou de pertes, certains individus ayant une « aversion pour le risque » seront prêts à payer pour éviter d'avoir à la jouer, alors que d'autres ayant une propension pour le risque exigeraient un dédommagement si on voulait les empêcher d'y participer. Confronté à une chance sur deux de mourir ou de vivre 10 ans en bonne santé, un individu avec une forte

aversion au risque pourra n'accorder au résultat de cette loterie qu'une valeur de 3 QALYs (au lieu des 5 de l'espérance mathématique de résultat) [Appendice (3)].

Il a pu être montré empiriquement, ce qui correspond d'ailleurs à l'expérience intuitive des cliniciens, que les individus confrontés à des loteries ayant des conséquences sur leur survie présentent en général une forte aversion au risque¹⁷, même si cette tendance disparaît lorsqu'il s'agit de risques à faible probabilité individuelle d'occurrence (ce phénomène de renversement des attitudes face au risque pouvant expliquer certaines difficultés d'adoption de comportements de prévention face à des risques « faibles » diffus dans une large population)¹⁸. Malheureusement, les QALYs s'avèrent incapables de prendre en considération des comportements autres que ceux caractérisés par une stricte neutralité face au risque.

Les problèmes liés à la prise en compte du temps

Une mesure QALY correspond par construction à une forme fonctionnelle du type $w(s) \times t$. Cela signifie que l'utilité associée à la qualité d'un état de santé donnée est proportionnelle à la durée de temps que l'on passe dans cet état. En principe, il est bien sûr envisageable de traiter le temps comme une variable continue. On pourrait par exemple exprimer le gain en QALYs d'une stratégie i par rapport à une stratégie j sous la forme :

$$\Delta \text{QALY}(i,j,n) = \int_0^T [U_n [s_i(t)] - U_n [s_j(t)]] dt$$

en faisant l'hypothèse que s_i est une fonction du temps.

Mais en pratique, on recourt toujours à une variable discontinue correspondant à une période de temps donnée (l'année dans la version standard du QALY). Cette simplification a trois conséquences.

La première conséquence est qu'on considère que l'utilité associée par l'individu à un état de santé donné demeure constante au cours de chaque période de temps choisie comme unité. Mais, un tel présupposé est très discutable : l'utilité associée au fait de consacrer son temps à la vision systématique de toutes les retransmissions télévisées de la dernière coupe du monde de football risque, selon les individus, de décroître au fur et à mesure que la répétition rend le spectacle ennuyeux, ou au contraire d'augmenter au fur et à mesure que la finale se rapproche et que l'excitation grandit.

La seconde conséquence est que toutes les atteintes à la « qualité de vie » intervenant sur des périodes plus brèves que l'unité de temps choisie (en pratique, inférieure à l'année) auront tendance à être systématiquement sous-estimées ; or, les effets secondaires, même sur une courte période, sont fréquemment un élément-clé de la comparaison de stratégies médicales. Dans une étude récente¹⁹, nous avons par exemple montré des différences statistiquement significatives des notes mises par les patients sur des échelles de douleur et d'anxiété, selon qu'on collecte des cellules hématopoïétiques aux fins d'autogreffe par un prélèvement de moelle classique (avec anesthésie générale) ou en les prélevant par séances de cytophérèse dans le sang périphérique. Une réduction de ces coûts « intangibles » pour les patients constitue un des avantages de la greffe de cellules-souche du sang périphérique face à la greffe de moelle. L'ampleur croissante des ressources que la collectivité consacre au traitement de la douleur suffit à attester qu'il s'agit d'éléments auxquels une « utilité » ou une « dés utilité » importantes sont attachés et qui auront tendance à être négligés dans une approche de type QALY.

La troisième conséquence est que les QALYs sont, par définition, incapables de rendre compte des situations où la valeur, du point de vue de l'individu, d'un temps de vie supplémentaire tend à devenir infinie, c'est-à-dire des situations fréquentes en médecine où le pronostic vital est immédiatement en jeu¹³. Comme le note Nord²⁰, « la continuation même de sa vie apparaîtra toujours « sans prix » à l'individu concerné. De ce fait, il est logiquement impossible d'obtenir des évaluations individuelles valides arbitrant entre un traitement visant à empêcher le décès immédiat et d'autres états de santé à échéance plus ou moins lointaine pour ce même individu ».

Les problèmes d'agrégation collective des QALYs

L'addition des QALYs obtenus au niveau individuel pour disposer d'une mesure collective du résultat d'un traitement renvoie au problème classique (mais non résolu) en économie du bien-être de l'agrégation des préférences individuelles qui nécessite l'hypothèse discutable de comparaisons interpersonnelles des utilités sur une échelle cardinale unique⁹. La démarche QALY relève de façon restrictive du « simple critère utilitariste qui consiste à juger de la réussite en fonction de la somme totale d'utilité créée, rien d'autre ne se voyant accorder une valeur intrinsèque »²¹. Un nombre croissant d'économistes cherchent précisément à s'affranchir de ce critère et en tout cas à discuter les conséquences sur leur démarche d'autres systèmes éthiques et de valeurs que le seul utilitarisme.

Raisonné en coût par QALY gagné dans un groupe de patients va d'ailleurs plus loin encore que la simplification utilitariste habituelle. L'agrégation des QALYs individuels présuppose de considérer que chaque année de vie a exactement la même valeur, le même poids pour chaque individu. Or, de la même manière que l'utilité marginale du revenu diffère selon les individus et selon leur niveau de revenu initial (1 F supplémentaire n'a pas la même valeur pour un sans domicile fixe que pour un milliardaire), le gain d'une année supplémentaire peut présenter des valeurs marginales très différentes selon les individus (ce problème étant d'ailleurs exacerbé si l'on compare des interventions dans des groupes très différents des jeunes et des vieux par exemple)²².

Les hypothèses réductrices sur la relation entre la santé et le bien-être des individus

Enfin, la notion de QALY débouche sur une interrogation fondamentale quant aux relations entretenues entre la santé d'une part, et le bien-être associé par l'individu à l'existence même d'autre part. En effet, l'utilité qu'un individu dérive d'une extension de sa durée de vie ne dépend pas seulement de la qualité de l'état de santé (S) dont il fait l'expérience, mais également de sa consommation de biens matériels (C), de son temps de loisirs (L) ou consacré à d'autres activités et fonctions sociales (F), de sa vie spirituelle etc. Pour qu'une mesure type QALY de l'utilité associée spécifiquement à l'état de santé soit empiriquement possible, il faut donc faire l'hypothèse que la fonction d'utilité individuelle W_{se} est une fonction « multi-attributs » de forme séparable et additive, c'est-à-dire qu'elle peut être construite par addition des fonctions d'utilité W_s, W_c, W_l, W_f , etc. associées respectivement à chacune de ces classes d'objets (S, C, L, F) :

$$W_{se} = W_s + W_c + W_l + W_f \dots + W_m$$

Concrètement, cela signifie que la mesure QALY présuppose qu'une intervention médicale n'a d'influence que sur l'état de santé stricto-sensu et n'affecte en rien les autres éléments (par exemple la consommation matérielle) qui font qu'un temps de vie apporte de l'utilité à l'individu. Il s'agit d'une hypothèse particulièrement réductrice, par exemple lorsque les coûts indirects (notamment les pertes d'activité liées à la maladie pour le patient et ses proches) sont importants et variables selon les stratégies qu'on compare.

Cette hypothèse conduit surtout à ignorer la dimension humaniste que partagent le meilleur de la tradition clinique et de la science économique. Les études sur la qualité de vie se heurtent souvent au paradoxe apparent d'une absence de différence significative moyenne de la qualité, mesurée sur une même échelle, entre des groupes de patients présentant pourtant des niveaux de gravité très variables (par exemple entre des patients atteints de cancers métastatiques et d'autres avec des cancers de meilleur pronostic, entre des patients atteints de maladies chroniques très invalidantes et la population générale, etc.)²³. A l'inverse, une étude récente sur les bénéficiaires du RMI faisait apparaître qu'ils exprimaient un indicateur de santé perceptuelle significativement dégradé par rapport à toutes les autres catégories de population à groupe d'âge et à niveau d'atteinte pathologique équivalents²⁴. Peut-être ces constats sont-ils justement explicables par des différences dans la façon dont les individus, soumis à différentes situations, estiment les interactions entre leur état de santé et les autres aspects constitutifs de leur existence en tant qu'être humain ?

RENONCER AU GRAAL DE L'INDICATEUR SYNTHÉTIQUE DE QUALITÉ DE VIE

Dans les années 1970 en France, certains économistes avaient popularisé l'idée de "prix de la vie humaine" qui considérait que la valeur d'un individu pouvait être assimilée à la perte de production potentielle, approchée par ses revenus futurs actualisés, qu'il représentait pour la collectivité en cas de décès prématuré. Ils avaient proposé ce critère unique pour fixer les limites de l'allocation des ressources collectives en matière de sécurité. Nous avons montré ailleurs en quoi cette idée relève d'une illusion technocratique qui a beaucoup contribué au rejet des études de Rationalisation des Choix Budgétaires appliquées au domaine de la santé, ce qui a fait prendre un retard considérable à la diffusion des méthodes d'évaluation économique au champ médical dans notre pays²⁵.

Nous espérons avoir montré que, de la même façon, l'approche coût par QALY occulte des présupposés décisifs dans les critères d'allocation des ressources au lieu de les expliciter afin de les soumettre à débat. Il nous semble plus sage que les économistes de la santé résistent à l'illusion de simplicité des choix décisionnels par le recours à des indicateurs uniques de type QALY. Plutôt que de rechercher le Saint Graal d'une mesure de l'utilité de la santé, il est plus raisonnable et réaliste aujourd'hui de systématiquement mieux documenter, dans la comparaison de différentes stratégies médicales, les aspects trop souvent négligés des coûts non monétaires, mais bien réels, subis par le patient. Il existe d'ailleurs pour mesurer l'impact des effets secondaires et des inconvénients des traitements sur le patient d'assez bons outils psychométriques (échelles de douleur, construction d'échelles de gêne occasionnée par les effets secondaires). De même, il est utile de mieux décrire les résultats médicaux en s'efforçant de compléter les mesures de survie par une connaissance plus fine des incapacités et handicaps subis par les patients et leur entourage en utilisant diverses échelles cliniques ou psychométriques. En pratique, il sera toujours possible de présenter l'ensemble des informations recueillies et de réaliser une analyse de sensibilité, c'est-à-dire de discuter dans quelle mesure la modification d'un paramètre ou la prise en compte d'un paramètre supplémentaire (incapacité associée à la survie par exemple) modifie ou non la hiérarchie des stratégies comparées. Enfin, l'approfondissement de la réflexion méthodologique sur la façon de « révéler » les préférences des patients confrontés à différentes alternatives de traitement, est indispensable. Il serait bien sûr souhaitable que les diverses méthodes empiriques proposées dans la littérature économique comme le « jeu de hasard idéalisé » (standard gamble) ou les « temps psychologiquement équivalents » (time-trade-off) connaissent un début d'applications concrètes en France. Une autre piste de recherche pour la révélation des préférences des patients est celle des tentatives de mesure de la « propension individuelle à payer » (willingness-to-pay) une information diagnostique ou un bénéfice thérapeutique²⁶⁻²⁷. Mais, il n'est en rien obligatoire de fondre l'ensemble de cette démarche de révélation des préférences individuelles dans la construction d'indicateurs QALY.

L'idée de comparer les coûts par QALY de différentes interventions médicales peut exercer une séduction certaine sur les « décideurs » en apparaissant comme un moyen commode de dégager des priorités dans les arbitrages financiers entre différents programmes de santé. Il est sans doute souhaitable d'améliorer la transparence politique des choix dans notre système de santé ou dans la gestion des prix du médicament. Mais, il serait à notre sens contradictoire avec ce louable objectif de prétendre y parvenir en imposant un indicateur (le QALY) dont le succès ne pourrait être fondé que sur une fausse simplicité chargée d'ambiguïté et d'obscurité quant à ses conditions d'élaboration ainsi que sur une pseudo rigueur théorique.

REMERCIEMENTS : Cet article n'engage que la responsabilité des auteurs. Nous tenons cependant à remercier Jean-Paul Auray et Gérard Duru (CNRS, GS Santé, Université de Lyon I) et Louis Eeckhoudt (Facultés Catholiques de Mons) pour leurs commentaires éclairants.

APPENDICE

(1) En s'inspirant de [10], on peut généraliser l'exemple du texte dans les termes suivants. Les préférences s'exercent sur différentes classes d'objets :

- Un ensemble D des durées. Soit dans l'exemple $D = \{d_1, d_2\}$ avec $d_1 =$ « un jour », $d_2 =$ « deux semaines » ;
- Un ensemble S des états de santé. Soit $S = \{s_1, s_2\}$ avec $s_1 =$ « sans effet secondaire », $s_2 =$ « avec effet secondaire » ;
- Un ensemble $D \times S$ des couples associant une durée à un état de santé donné (exemple « 13 jours sans eczéma ») ;
- Un ensemble T de traitements. Soit $T = \{T_1, T_2\}$;
- Des conséquences des traitements en termes de couples durée/état de santé (A, B, C dans l'exemple) et des probabilités de survenue de ces couples. Les conséquences des traitements sont donc des distributions de probabilité sur une partie C de $D \times S$. Soit $C = \{(d_1, s_1), (d_2, s_1), (d_2, s_2)\}$ et F , l'ensemble des distributions de probabilité ; $F = \{p_1, p_2, p_3, \sum p_i = 1\}$.

Une façon simple de construire sur un ensemble quelconque E d'objets une relation de préférence est de faire appel aux propriétés de l'ensemble R des nombres réels et de sa relation \geq . Il suffit de construire une application $u(\cdot)$ définie sur E et prenant ses valeurs dans R , puis de construire la préférence $>$ sur X en posant pour tout x et y dans X : $x > y$ si et seulement si $u(x) \geq u(y)$. La fonction $u(\cdot)$ est appelée « fonction d'utilité » et la relation $>$ est alors appelée « préférence sur X associée à l'utilité $u(\cdot)$ ». Dans notre cas, on pourra ainsi définir des préférences sur T, F, C, D et S (notées respectivement $\geq_T, \geq_F, \geq_C, \geq_D, \geq_S$) au travers des fonctions d'utilité associées (respectivement E_T, E_F, U_C, v_D, w_S).

Par hypothèse, on a : $T_1 \geq_T T_2 \iff g(T_1) \geq_F g(T_2)$, donc $E_T = E_F$ à une transformation linéaire près, où g est une application de T dans F qui à tout traitement T_i associe sa conséquence, c'est-à-dire une distribution de probabilité sur C . Utiliser le critère d'utilité espérée (dont on peut montrer qu'il n'est qu'un cas particulier du concept plus général d'utilité neumanienne également appelé « utilité d'intervalles » [10]) pour classer les conséquences revient, quel que soit m_1 et m_2 à comparer deux distributions sur C :

$$\mu_1 \geq_F \mu_2 \iff \int_C U_C d\mu_1 \geq \int_C U_C d\mu_2. \text{ Soit } E_T(T_i) = \int_C U_C d\mu_i$$

Toute la question est alors d'estimer expérimentalement U_C , soit par mesure directe des préférences sur C , c'est-à-dire à l'égard des couples durée/états de santé (c'est stricto-sensu la démarche de l'utilité espérée), soit en construisant U_C à partir des fonctions d'utilité v et w observées respectivement sur D et sur S . On peut montrer que la construction des QALYs relève du point de vue théorique de cette seconde approche qui consiste à considérer que U_C est une fonction d'utilité

« multi-attributs » construite sur $(D \times S)^T$ à partir des fonctions d'utilité V_D sur D , et W_s sur S ; et, à introduire des hypothèses restrictives supplémentaires, dont il sera question plus loin, pour poser que U_C est de la forme $t \times w(s)$ avec t le temps passé dans un état s donné.

(2) En réalité, on peut montrer que ces analyses coût-efficacité et coût-utilité ne sont que des cas particuliers, reposant sur des spécifications implicites très restrictives, des analyses coût-bénéfice (ACB) qui visent à mesurer le surplus social net d'un investissement public par la différence entre ses coûts totaux et ses résultats totaux (ce qui implique inévitablement de mesurer ces deux termes sur une même échelle, une échelle monétaire constituant l'étalon de mesure le plus simple) [11].

(3) Les attitudes individuelles différentes face au risque, entendu comme probabilités d'occurrence d'une conséquence, peuvent aisément être représentées mathématiquement par des coefficients comme celui de Pratt [17] du type :

$$u(x) = x^{1-c}, \quad c < 1$$

avec (x) représentant la durée de vie, $u(x)$ l'utilité associée à cette durée de vie, $u(0)=0$ et $u(1)=1$. Plus, c 'est élevé et plus l'individu exprimera une aversion au risque, la neutralité correspondant à $c=0$.

RÉFÉRENCES

- ¹ Gérard K. Cost-utility in practice: a policy maker's guide to the state of the art. *Hlth Pol* 1992 ; 21 : 249- 79.
- ² Williams A. Economics of coronary artery bypass grafting. *Br Med J* 1985 ; 91 : 326-9.
- ³ Hadorn DC. Setting health care priorities in Oregon. *JAMA* 1991 ; 265 : 2218-25.
- ⁴ Launois R. La prise en compte des préférences des patients dans les choix de santé individuels et collectifs. *Rev Epidém et Santé Publ* 1994 ; 42 : 246- 62.
- ⁵ Willinger R. La rénovation des fondements de l'utilité et du risque. *Revue Economique* 1990 ; 41: 5- 48.
- ⁶ Loomes G., Mc Kenzie L. The use of QALY's in health care decision making. *Soc Sci Med* 1989 ; 28: 299-308.
- ⁷ Culyer AJ. The normative economics of health care finance and provision. *Oxford Rev Econom Policy* 1989 ; 5 : 34-58.
- ⁸ Moatti JP. Evaluation économique: un complément nécessaire de l'évaluation médicale. In Matillon Y, Durieux P, eds. *L'évaluation médicale : du concept à la pratique*, Paris, Coll. Médecine-Sciences, Flammarion, 1994 : 71-80.
- ⁹ Von Neumann I, Morgenstern O. *Theory of games and economic behavior*, Princeton, Princeton University Press, 1947.
- ¹⁰ Auray JP, Duru G. Qualité de vie. In Lily Moto, éd. *Santé & multidisciplinarité : choix et décisions*, Paris, Eds Hermès, 1995 : 82-127.
- ¹¹ Carrère MO, Auray JP, Duru G. La valeur décisionnelle des ratios coût-efficacité et coût-utilité dans l'évaluation économique des programmes de santé. In Actes du Colloque Européen « *De l'analyse économique aux politiques de santé* », Eds CREDES-CES, Paris, 1992, T 3 : 91-100.
- ¹² Henry C. *Economie publique*, Paris, Eds de l'Ecole Polytechnique, 1987.
- ¹³ Moatti JP. *Economie de la sécurité*, Paris, Coll. Analyse et Prospective, Eds INSERM –La Documentation Française, 1989.
- ¹⁴ Weinstein M, Fineberg H. *Clinical decision analysis*, Philadelphia, WB Saunders, 1980.
- ¹⁵ Kind P, Rosser R, Williams A. Valuation of quality of life : some psychometric evidence. In Jones-Lee MW, ed. *The value of life and safety*, Amsterdam, North Holland Pub CO, 1982: 159-70.
- ¹⁶ Culyer AJ. Commodities, characteristics of commodities, characteristics of people, utilities and the quality of life. In Baldwin S, ed. *The quality of life. Perspectives and policies*, London, Rutledge, 1990.
- ¹⁷ Pratt JW. Risk aversion in the small and in the large. *Econometrica* 1964 ; 32 : 122-36.
- ¹⁸ Hellinger FJ. Expected utility theory and risk choices with health outcomes. *Med Care* 1989 ; 27: 273-9.
- ¹⁹ Auquier P, Macquart-Moulin G, Blache JL *et al.* Anxiété et douleur des patients soumis à deux procédures différentes de recueil des cellules souches : "la cytophèrese ou le prélèvement de moelle : *Bull Canc* 1994 ; 6 : 472.
- ²⁰ Nord E. The Qaly. A measure of social value rather than individual utility ? *Hlth Eco* 1994 ; 3 : 89-93.
- ²¹ Sen A. Jugements sur l'économie et la philosophie morale. In A Sen, *Ethique et économie*, Paris, trad fr. PUF, 1993.
- ²² Donaldson C, Atkinson A, Bond J. Should QALY's be programme-specific ? *J Hlth Eco* 1998 ; 7 : 239- 57.
- ²³ De Haes JC, Van Knippenberg FC. Quality of life of cancer patients: review of the literature. *Soc Sci Med* 1985 ; 20: 809-17.
- ²⁴ Obadia Y, Toubiana P, Chanut C, Rotily M. L'état de santé des RMistes dans les Bouches du Rhône. *Rev Fr Aff Soc* 1994 ; 2 : 111-130.
- ²⁵ Moatti JP, Chanut C, Benech JM. Researcher driven versus policy driven economic appraisal of health technologies: the case of France. *Soc Sci Med* 1994; 38: 1625-33.
- ²⁶ Donaldson C. Willingness-to-pay for publicly provided goods : a possible measure of benefit ? *J Hlth Eco* 1990 ; 9: 103-118.
- ²⁷ Moatti JP, Julian C, Seror V, Loundou DA, Le Galès C, Aymé S. Willingness-to-pay for prenatal screening: methodological and empirical issues. In Actes du Colloque Européen" *De l'analyse économique aux politiques de santé* », Eds CREDES-CES, Paris, 1992, T 3 : 71-80.