



Méthodes de mesure de la qualité de vie dans les travaux d'évaluation économique des actions de santé

Dans le champ de l'évaluation économique, deux approches sont développées pour rendre compte de la valeur que les individus accordent à un état de santé. La première consiste à révéler directement, pour chaque situation décisionnelle particulière, les préférences des personnes concernées grâce à des scénarios décrivant des états de santé, en mobilisant une ou plusieurs des méthodes de révélation des préférences (échelle visuelle analogique, méthode de l'équivalent probabiliste, arbitrage temporel). La seconde approche repose sur l'emploi d'index d'états de santé génériques pondérés par les préférences. Les états sont génériques car définis indépendamment d'un contexte clinique ou des particularités d'une population. Les index sont dits pondérés par les préférences car chaque état de santé est associé à un poids obtenu, une fois pour toute, en interrogeant des personnes malades ou des personnes appartenant à la population générale. L'usage qui doit être fait de ces index pondérés par les préférences oriente le choix de la population enquêtée. Ainsi, lorsque les préférences doivent être utilisées pour allouer des ressources collectives, plusieurs groupes d'experts anglo-saxons recommandent d'interroger la population « des citoyens informés ».

m/s n° 5, vol. 16, mai 2000

Alors qu'une communauté scientifique se dessine au travers, notamment, de la constitution d'une société internationale: l'*International Society for Quality of Life* dont le congrès inaugural s'est tenu en 1994, de la parution depuis 1992 d'un journal spécialisé: *Quality of Life Research*, la définition même de l'objet étudié, la qualité de la santé, la qualité de vie, reste une question âprement discutée par des chercheurs aux appartenances disciplinaires variées. La publication de bibliographies, que l'impératif d'exhaustivité transforme en une tâche sans fin, montre que sont regroupés des questionnaires de mesure des performances ou des capacités physiques, d'étude de la santé perçue, d'évaluation du retentissement de la maladie ou de son traitement sur la vie quotidienne des personnes atteintes. Ce regroupement se justifie par l'existence d'un consensus implicite quant aux éléments participant à la qualité de vie, ou, au contraire, par l'absence de tout *gold standard* qui empêche une sélection parmi les outils existants. L'inscription d'une mesure de la qualité de vie dans la démarche d'évaluation économique recommande pourtant une ligne de partage dans cet ensemble hétérogène. A côté d'une estimation de la durée de survie, d'une description fine des incapacités et handicaps induits par la maladie, le recours au calcul économique impose d'agréger ces in-

formations afin de rendre comparables des options concurrentes et d'en éclairer le choix. Il est donc souhaitable, en économie, de pouvoir dissocier, dans une méthode de mesure des conséquences, ce qui relève de la description de l'état de santé, de ce qui relève de l'opinion, l'importance, la valeur que les individus attachent au fait d'être dans tel ou tel état de santé.

La réflexion sur la valeur d'un état de santé était, dans les premiers travaux économiques, relativement sommaire. A partir d'une mesure de l'efficacité exprimée en nombre d'années de vie gagnées, des coefficients aux valeurs souvent arbitraires ou établies par consensus d'experts médicaux, étaient utilisés pour pondérer ces années et rendre compte du fait que celles-ci n'étaient pas toutes associées aux mêmes effets morbides. Ainsi, dans le domaine de l'évaluation économique des stratégies de prévention de l'ostéoporose, Weinstein et Schiff estimaient, dans un article publié au début des années 1980 [1], qu'une année de vie, en cas de cancer de l'endomètre, était l'équivalent, en termes de qualité de vie pour la femme concernée, de 0,8 année en bonne santé, alors que Tosteson *et al.*, dans un article daté de 1990 [2], considéraient qu'une année dans les mêmes conditions morbides « ne valait » que 0,4 année. La conversion d'années avec maladie en années en bonne santé, aussi critiquables qu'aient été ses modalités, avait pour

but de rendre comparable des états de santé différents et d'exprimer les résultats de l'analyse économique en termes d'années de vie pondérées par la qualité ou QALY (voir *m/s* 2000, n° 4, p. ??).

Par la suite, dans le cadre de l'évaluation économique des actions de soins, deux approches vont se développer pour intégrer la valeur d'un état de santé à la mesure des conséquences. La première est de révéler directement, pour chaque étude d'évaluation économique, l'utilité (au sens économique) des états de santé particuliers identifiés dans les études cliniques auxquelles s'adosse la démarche économique. La seconde approche, qui tend aujourd'hui à prendre le pas sur la première, mobilise l'emploi d'index d'états de santé génériques pondérés par les préférences. Les états de santé sont dits génériques car ils sont définis en toute indépendance d'un contexte clinique ou des particularités d'une population. Ils peuvent donc être utilisés dans tout groupe quelle que soit la maladie concernée, voire en population générale. Les index sont dits pondérés par les préférences car chaque état est associé à un poids obtenu, une fois pour toute, en interrogeant des personnes malades ou des personnes appartenant à la population générale avec des méthodes de révélation des préférences. Ce poids reflète donc la valeur de chaque état de la classification relative, le plus souvent, à l'état de bonne santé et au décès.

La mesure directe des préférences

Cette méthode de construction des coefficients de pondération des états de santé consiste à révéler, pour chaque situation décisionnelle particulière, les préférences pour des scénarios décrivant des états de santé considérés comme pertinents (à l'issue d'un essai clinique par exemple), en mobilisant généralement une ou plusieurs des méthodes suivantes: l'échelle visuelle analogique (graduée ou non), la méthode de l'équivalent probabiliste (ou *standard gamble*), l'arbitrage temporel (ou *Time Trade Off*). Sous sa forme la plus simple, l'échelle visuelle analogique (EVA) est une

ligne dont les extrémités sont clairement identifiées. Dans l'ensemble des états de santé qui sont à évaluer, la personne interrogée est invitée à isoler les deux états qu'elle considère comme étant le meilleur et le pire et à les utiliser comme bornes. Elle doit ensuite indiquer le désir ressenti pour les états restants en les plaçant l'un après l'autre sur le segment de droite défini par les deux états-bornes. Une variante de cette méthode (échelle catégorielle, *category scaling*) consiste à faire apparaître sur le segment de droite des intervalles égaux (souvent 10) et à demander à la personne dont on veut connaître les préférences de placer les états dans un des intervalles, en faisant l'hypothèse que le passage d'un intervalle à l'intervalle adjacent est associé à un changement identique de différences de désir. Des aides représentant, par exemple, un thermomètre peuvent être utilisées.

Avec la méthode de l'équivalent probabiliste, l'objectif est de déterminer, en cohérence avec la théorie de l'utilité espérée, le niveau de préférence que donne la personne interrogée à un état de santé *i* différent de deux états qui vont être les bornes entre lesquelles devra être comprise l'évaluation

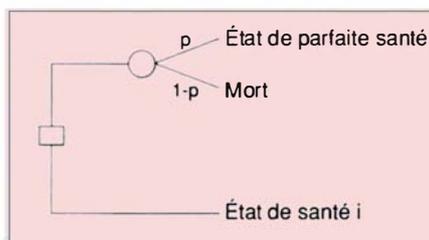


Figure 1. **Méthode de l'équivalent probabiliste (standard gamble).** Avec la méthode de l'équivalent probabiliste, l'objectif est de déterminer l'utilité que donne la personne interrogée à un état de santé *i* différent de l'état mort et de l'état de santé parfaite. Pour cela, elle doit comparer une situation certaine, le fait d'être dans l'état *i*, avec une situation probabiliste (ou loterie) où elle a une probabilité *p* d'être dans l'état de santé parfaite et une probabilité (1-*p*) d'être dans l'état mort. La valeur de *p* est modifiée jusqu'à ce que la personne interrogée déclare être indifférente entre les deux situations.

tion de l'état *i*. L'un des états placés aux bornes est jugé comme le pire possible et donc posé comme valant 0 (il s'agit souvent de la mort). L'autre état est l'état de santé le meilleur possible. Il est posé comme valant 1. Pour déterminer son niveau de préférence pour l'état *i*, la personne interrogée doit comparer une situation certaine: le fait d'être dans l'état *i* pendant *t* années (*t* est donné) avec une situation probabiliste ou loterie où elle a une probabilité *p* d'être dans l'état le meilleur possible et une probabilité (1-*p*) d'être dans l'état le pire (figure 1).

La valeur de *p* est modifiée jusqu'à ce que la personne déclare être indifférente entre la situation certaine et la situation probabiliste du fait des probabilités associées aux deux conséquences de celle-ci.

La valeur de *p*, *p*^{*}, pour laquelle il y a indifférence entre les situations certaine et probabiliste est la valeur de l'état *i*. En effet, lorsqu'il y a indifférence, si la personne effectue un choix rationnel au sens de la théorie de l'utilité espérée, alors le fait qu'elle soit indifférente entre les deux situations proposées signifie que celles-ci sont associées au même niveau d'utilité. En d'autres termes, l'utilité espérée de la loterie est égale à l'utilité de la situation certaine donc

$$p^* U(\text{meilleur état possible}) + (1-p^*)$$

$$U(\text{pire état possible}) = U(\text{état } i)$$

or $U(\text{meilleur état possible}) = 1$ et $U(\text{pire état possible}) = 0$

donc

$$p^* \times 1 + (1-p^*) \times 0 = U(\text{état } i)$$

$$U(\text{état } i) = p^*$$

Pour faciliter la détermination de *p*^{*}, il est fréquent d'utiliser un support permettant d'illustrer le choix proposé et de visualiser les différentes probabilités (figure 2).

Au cours d'un arbitrage temporel, l'objectif est de déterminer la valeur que donne la personne interrogée à un état de santé *i* différent du pire état possible (qui peut être la mort) et du meilleur. Comme pour la méthode précédente, les valeurs de ces deux états sont posées arbitrairement comme étant respectivement 0 et 1. La personne interrogée doit choisir entre l'option 1 ou l'option 2, celle qu'elle préfère (figure 3). L'option 1

est définie comme être dans un état de santé i pendant t années, tandis que l'option 2 est caractérisée par le fait de vivre dans l'état de santé le meilleur possible pendant x années (avec $x < t$). La valeur de x varie jusqu'à ce que la personne interrogée se déclare indifférente entre les deux options.

La valeur de l'état i est égale à x/t . En effet, toujours sous l'hypothèse d'un comportement rationnel de maximisation de l'utilité, lorsque la personne ne peut plus choisir entre vivre x^* dans l'état de santé le meilleur possible et t dans l'état i , alors

$$x^* \times U(\text{meilleur état possible}) = t \times U(\text{état } i)$$

$$\text{or } U(\text{meilleur état possible}) = 1$$

donc

$$U(\text{état } i) = x^*/t$$

Pour cette méthode également, la situation de choix proposée peut être représentée graphiquement et permettre, en particulier, de visualiser les différentes durées.

La mesure directe des préférences est lourde à mettre en œuvre. Elle ne peut être développée que de façon prospective et suppose que le nombre d'états soumis au jugement des personnes dont on veut connaître les préférences, soit limité. Enfin, la comparabilité des résultats obtenus à l'occasion d'études particulières est restreinte par la grande hétérogénéité des situations décisionnelles spécifiques analysées et des méthodes de description des états et de révélation des préférences retenues.

Les index d'états de santé pondérés par les préférences

Un index est composé de deux éléments: un système générique de classification des états de santé et une fonction de pondération dont le rôle est de permettre le classement des états et leur comparaison sur une échelle d'intervalle ou de ratio, tenant compte des préférences individuelles. Une échelle d'intervalle permet de savoir quelle est la distance entre deux objets, ce qui est une information plus riche que celle obtenue par les échelles ordinales qui indiquent simplement le rang de ces objets, mais elle ne permet pas de connaître le niveau de désir absolu associé à chaque objet. Avec une échelle de ratio, la distance entre un 0 qui n'est pas arbitraire et au moins un objet est connue. Il est donc possible de déterminer le degré de désir pour un objet. Seules les valeurs d'une échelle de ratio peuvent faire l'objet de calculs algébriques.

Le système de classification des états de santé

Un système de classification est un ensemble fini et exhaustif d'états de santé génériques ou spécifiques (à une population, à une pathologie...). Il s'appuie sur une représentation qui fait de la santé la résultante de plusieurs composantes ou dimensions: physique, psychologique... Pour chaque dimension sont définis, en nombre fini, des niveaux qui correspondent à des situations mutuellement exclusives. Ces niveaux peuvent être ou non ordonnés. L'état de santé d'un individu à un moment donné est donc obtenu en choisissant, pour chaque dimension, le niveau le plus pertinent dans l'ensemble des niveaux possibles. Le choix de l'état de santé qui, dans le système de classification, décrit le mieux une personne à un moment donné peut se faire, soit en demandant à l'individu lui-même (ou à un proche) de répondre à un questionnaire dans lequel il devra choisir, pour chaque question, une seule des réponses proposées (questionnaire fermé), soit en laissant un professionnel de santé (mé-

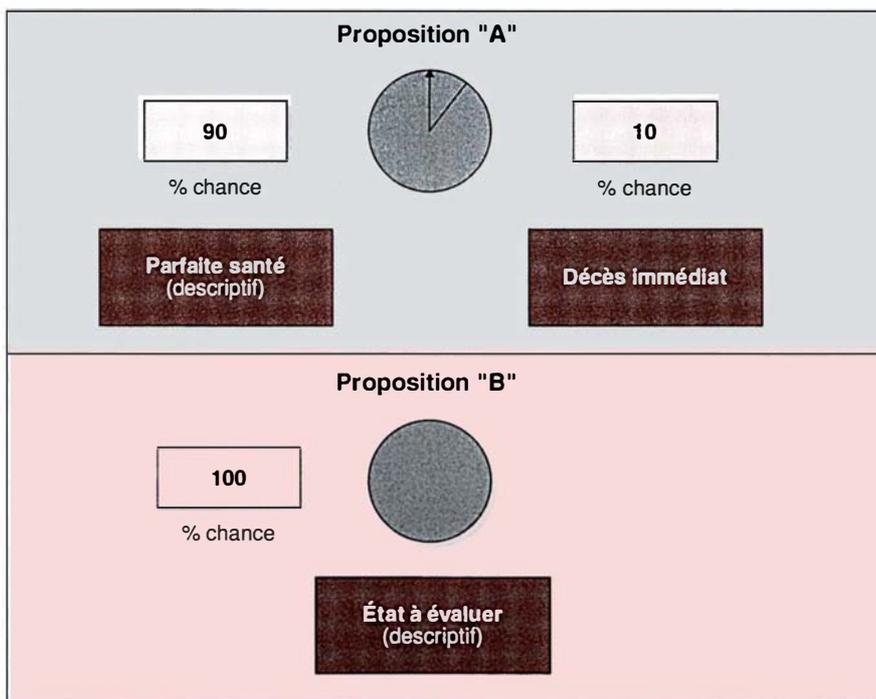


Figure 2. Exemple d'un support illustrant le choix proposé lors d'un équivalent probabiliste. Sur cet exemple, dans la proposition A, la probabilité associée au fait d'être en parfaite santé est 90%, celle d'être dans l'état le pire possible 10%. La proposition B illustre la situation certaine. La personne doit dire si elle préfère la proposition A ou la proposition B ou encore si elle n'a pas de préférence entre les deux. Tant que la personne n'est pas indifférente, entre les deux propositions, un nouveau jeu de probabilités est proposé. Par exemple, si A est préférée à B, cela signifie que la personne accepte la probabilité de 10% d'être dans l'état le pire possible, le couple (90, 10) devient (85, 15) afin de savoir si elle accepte maintenant ce nouveau risque plus élevé compte tenu de la description de l'état présenté dans la proposition B.

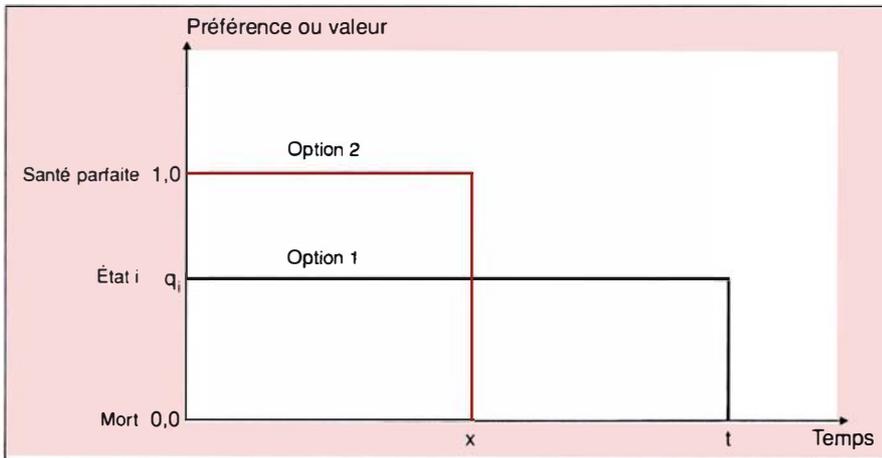


Figure 3. **Méthode de l'arbitrage temporel (Time Trade Off).** L'objectif de la méthode de l'arbitrage temporel est de déterminer la valeur que donne la personne interrogée à un état de santé i différent de l'état mort et de l'état de santé parfaite. Pour cela, elle doit choisir entre l'option 1 ou l'option 2, celle qu'elle préfère. L'option 1 est caractérisée par le fait d'être dans l'état de santé i pendant t années, tandis que l'option 2 est caractérisée par le fait d'être dans l'état de santé parfaite pendant x années (avec $x < t$). La valeur de x est modifiée jusqu'à ce que la personne interrogée se déclare indifférente entre les deux options.

decin, infirmière...) choisir directement, pour chaque dimension, le niveau pertinent. Un tel système permet de mettre en évidence la morbidité d'une population dans son expression multidimensionnelle et le degré de sévérité de cette morbidité. Il existe plusieurs systèmes de classification. Ils diffèrent, généralement, par le nombre et le type de dimensions retenues pour décrire la notion d'état de santé et par le nombre d'états qu'ils contiennent. Il ne peut être question, ici, de les comparer car chacun est issu de longues années de recherches dans des équipes de différentes sensibilités, poursuivant un objectif général commun, mais ayant souvent fait, dès leurs premiers travaux, des choix assez différents en termes de représentation de la notion d'état de santé et de méthodes de construction des pondérations. Par ailleurs, il n'existe pas d'instrument étalon, de *gold standard*, dans le domaine des index d'états de santé pondérés par les préférences. Trois index sont plus particulièrement connus et utilisés par les économistes: le *Health Utilities Index* (HUI) [3-12], l'*EuroQol* [13-16], le *Quality of*

Well-Being Scale ou index de bien-être [17-20].

Le système de classification du HUI est un ensemble fini et exhaustif d'états de santé génériques. La plupart des dimensions sont fonctionnelles: la vue, l'ouïe, la capacité à marcher... La première version du HUI (HUI Mark 1), publiée en 1982 [3], ne distinguait que 4 dimensions déclinables en 4 à 8 niveaux soit, au total, un système de 960 états de santé. Dix ans plus tard, le HUI2 reposait sur 7 dimensions ayant 3 à 5 niveaux possibles soit 24 000 états. Le HUI3, publié en 1995 au Canada [5, 6] et plus récemment en France [9-12], retient 8 dimensions (la vue, l'ouïe, l'élocution, la capacité à marcher, la dextérité, l'état psychologique, la mémoire et la capacité à réfléchir, la douleur) avec chacune 5 à 6 niveaux, soit un système de 972 000 états. Dans le système HUI3, un état est donc un vecteur de 8 éléments, le premier indique le niveau pertinent sur la première dimension (la vue) pour l'individu dont l'état de santé est décrit, le second élément indique le niveau sur la seconde dimension (l'ouïe), etc.

L'*EuroQol* est un instrument dont le développement a été, dès ses débuts, mené simultanément dans plusieurs pays d'Europe [14]. Une préoccupation centrale de ce groupe de chercheurs était, en effet, de pouvoir procéder aux comparaisons internationales de valeurs nationales obtenues sur des états de santé identiques. L'*EuroQol* s'est, tout d'abord, appuyé sur un système de classification à 6 dimensions: mobilité, autonomie de la personne, activité principale, relations sociales, douleur et humeur soit 216 états de santé. Plus récemment, une 2^e version de l'instrument a été publiée [13]. Ce nouveau système repose sur 5 dimensions (mobilité, autonomie de la personne, activités courantes, douleur/gêne, anxiété/dépression) ayant chacune 3 niveaux soit 243 états auxquels sont ajoutés l'état « mort » et l'état « inconscience ». Chacun des 243 états est donc un vecteur à 5 composantes.

Le développement du *Quality of Well Being* s'est inscrit, dès ses débuts, dans la perspective plus large de construction d'un modèle général d'évaluation des politiques de santé [17]. La classification d'états de santé de cet index générique, qui repose sur les travaux de Fanshel et Bush [21], isole d'une part trois domaines, la mobilité, l'activité physique et l'activité sociale et, d'autre part, un ensemble de symptômes/problèmes. Cet indicateur a été retenu par l'état américain de l'Oregon, lorsque celui-ci a souhaité élargir la population éligible au système *Medicaid* sur la base des résultats d'évaluations économiques intégrant simultanément les effets des programmes de santé sur la mortalité et sur la morbidité des populations.

La fonction de pondération par les préférences

Le système ou la fonction de pondération est le second élément constituant l'index d'états de santé pondéré par les préférences. L'existence de cette fonction et les spécificités de sa conception différencient les index d'états de santé pondérés par les préférences des autres méthodes de qualité de vie (figure 4).

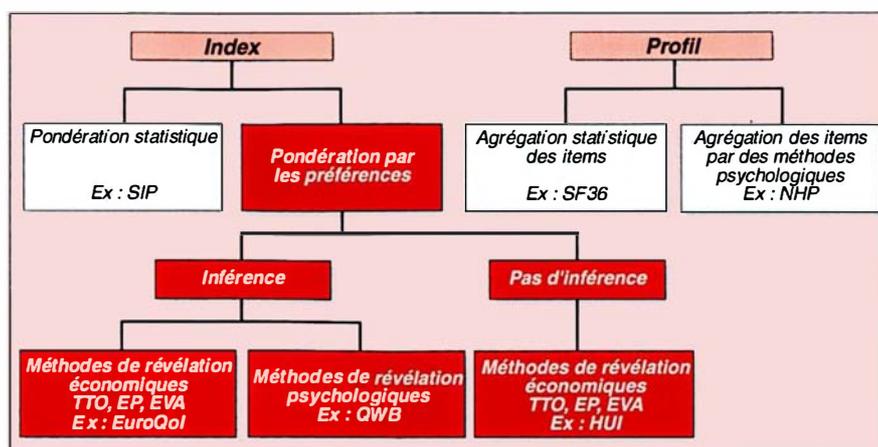


Figure 4. **Typologie des méthodes de mesure de la qualité de vie.** Tous les instruments de mesure de la qualité de vie partagent une conception multidimensionnelle de la notion de qualité de vie au sens où celle-ci est le résultat de plusieurs éléments, l'état physique du sujet, ses sensations, son état psychologique ses relations sociales et ses rapports avec son environnement familial, professionnel... Les instruments sont tous composés de questions ou items qui sont des descripteurs de cette qualité de vie. Ces items sont regroupés, généralement par des méthodes statistiques d'analyse de données ou des méthodes psychométriques, en plusieurs dimensions. La mesure ainsi obtenue prenant alors la forme d'un score pour chaque dimension, il s'agit donc d'un profil de qualité de vie. Si les scores de toutes les dimensions sont agrégés, la mesure devient unidimensionnelle, c'est un index. Il existe différentes méthodes qui permettent de passer d'un résultat multidimensionnel à un score unique. Les techniques de pondération peuvent être soit statistiques soit faire appel aux préférences de sujets interrogés spécifiquement dans cette perspective. Toutefois, pour rendre ces enquêtes plus faciles à réaliser, il est courant de ne recueillir qu'un nombre limité d'évaluations ou de jugements obtenus par arbitrage temporel (TTO), équivalent probabiliste (EP) ou échelle visuelle analogique (EVA) ou par des méthodes strictement psychométriques. Le système de pondération sera développé en utilisant les données ainsi recueillies pour estimer un modèle algébrique des préférences par des techniques statistiques de régression (inférence statistique). Une seconde procédure permet explicitement à la personne interrogée de diviser la tâche d'évaluation en plusieurs tâches plus simples. La théorie de l'utilité multiattribut de Keeney et Raiffa [22], prolongement de la théorie de l'utilité espérée lorsque les conséquences sont multidimensionnelles, permet alors de construire une fonction de pondération en supposant que chaque dimension est un attribut de la fonction d'utilité de la personne interrogée. Cette procédure explicitement décomposée permet que soient analysés les effets de chaque dimension. Il est donc possible de retracer la manière dont les différentes personnes ont combiné les différentes dimensions pour construire leur évaluation globale.

Premièrement, la fonction de ce type d'index est associée aux états de santé multidimensionnels eux-mêmes et non aux dimensions qui définissent l'état de santé comme c'est le cas pour les profils psychométriques. Deuxièmement, à la différence des index psychométriques, les méthodologies de construction de ces fonctions font explicitement référence à la

notion de préférence, que les méthodes de leur révélation soient issues de la psychologie ou de l'économie. Les principes fondamentaux de la pondération des index économiques sont d'une part, que les individus ont des préférences pour les états de santé, qu'ils peuvent les révéler et, d'autre part, que ces préférences doivent être prises en compte lorsqu'il s'agit de

comparer ou de choisir entre différents états. La stratégie de mesure est un élément-clé de la construction d'une fonction de pondération par les préférences. Elle comprend, premièrement, la structure générale adoptée pour poser les questions aux personnes dont on souhaite connaître les préférences (les dimensions retenues pour définir un état de santé seront-elles présentées une par une ou toutes ensemble...) et, deuxièmement, la méthode adoptée pour analyser les informations qui vont être ainsi recueillies (méthodes de régression, analyse de la variance, s'il s'agit de méthodes statistiques...). Deux grandes types de stratégie peuvent être distingués. Le premier, fréquemment nommé approche holistique, suppose que le répondant donne une valeur à tous les états de santé de la classification. L'autre type est l'approche dite décomposée. Les valeurs vont être obtenues pour tous les états sans que le répondant ne doive les donner pour chacun d'entre eux. La valeur d'un état de santé sera exprimée comme une fonction décomposée de ses dimensions. L'intérêt majeur est bien sûr de réduire la tâche qui incombe à chaque personne interrogée. Deux procédures d'application de cette approche décomposée sont envisageables. Une première procédure a pour objectif de développer un modèle algébrique des préférences du décideur à partir de jugements portant sur un sous-ensemble d'états: ce sont les modèles statistiquement inférés ou plus exactement reposant sur l'utilisation de techniques statistiques d'estimation (EuroQol, QWB). Une seconde procédure permet de diviser explicitement la tâche d'évaluation en plusieurs tâches plus simples dans le cadre de la théorie économique: ce sont les modèles explicitement décomposés (HUI). La théorie de l'utilité multi-attribut de Keeney et Raiffa [22], prolongement de la théorie de l'utilité espérée lorsque les conséquences sont multidimensionnelles, permet alors de construire une fonction de pondération en supposant que chaque dimension est un attribut de la fonction d'utilité de la personne interrogée. L'approche explicitement décomposée permet que soient analy-

sés les effets de chaque dimension. Il est donc possible de retracer la manière dont les personnes ont combiné les différentes dimensions pour construire leur évaluation globale.

Lorsque la stratégie de mesure est définie, les données sont recueillies à partir des méthodes de révélation, que sont l'échelle visuelle analogique, l'équivalent probabiliste et l'arbitrage temporel.

Des méthodes encore imparfaites

Les questions méthodologiques les plus importantes apparaissent surtout à l'étape de construction des coefficients ou des fonctions de pondération des états de santé.

D'une part, les résultats empiriques obtenus par les méthodes de révélation des préférences doivent avoir une validité de contenu et être reproductibles. Or, la plupart du temps, les états de santé présentés sont des états hypothétiques pour les personnes interrogées qui ne disposent pas toujours d'informations suffisantes pour se représenter véritablement les situations décrites. De nombreux travaux empiriques ont montré combien les préférences révélées pouvaient être influencées par le mode de présentation des états de santé. Le biais de « labellisation » a été observé selon le mode de description de situations pathologiques (le fait de nommer une maladie par exemple) [23, 24]. De même, le biais de « formatage » a été décrit à de nombreuses reprises (*framing effect*). Il rend compte, par exemple, de la variabilité des préférences révélées selon qu'une même information est présentée en termes positifs (probabilité de survie) ou négatifs (probabilité de décès) [23-26]. Enfin, plusieurs travaux ont évalué la sensibilité des résultats à la manière dont la méthode de révélation était mise en œuvre empiriquement sans pour l'instant permettre d'en tirer des recommandations méthodologiques [27-29].

A ces premiers facteurs de variabilité des résultats, il faut ajouter la comparabilité limitée des méthodes elles-mêmes. De nombreuses études ont comparé les trois méthodes élémentaires de révélation des préférences

que sont l'échelle visuelle analogique, l'équivalent probabiliste et l'arbitrage temporel, soit dans le domaine de la santé [30-32] soit dans d'autres domaines [33, 34]. En règle générale, elles ont observé que la préférence associée à un même état de santé était plus élevée lorsqu'elle était révélée par un équivalent probabiliste que par un arbitrage temporel et que la préférence la plus faible était obtenue avec l'échelle visuelle analogique. Or, l'échelle visuelle analogique comme l'arbitrage temporel ne permettent pas de prendre en compte l'attitude individuelle à l'égard du risque alors que, selon la théorie économique de l'utilité espérée, cette attitude est susceptible de modifier le jugement individuel sur un état de santé (*figure 4*). Il a, en effet, été montré que les préférences d'une personne pour la même conséquence étaient différentes selon que ces préférences étaient révélées en

univers certain, risqué (c'est-à-dire probabiliste) ou incertain. Depuis les années 1950, la conception dominante du risque en économie est celle de von Neumann et Morgenstern [35]. Elle implique qu'une personne confrontée au choix entre une loterie et une conséquence certaine révèle sa préférence entre les deux options et donc son attitude vis-à-vis du risque. Si elle préfère la conséquence certaine à la loterie ayant la même espérance de gain, cela signifie qu'elle préfère éviter les risques associés à la loterie. Son comportement traduit une aversion vis-à-vis du risque et l'importance de celle-ci s'exprime par la concavité plus ou moins forte de la fonction d'utilité de la personne. Si, au contraire, elle préfère la loterie, elle fait preuve d'une préférence pour le risque. Enfin, en cas d'indifférence, le comportement de la personne est dit neutre à l'égard du risque. La *figure 5* illustre ces trois

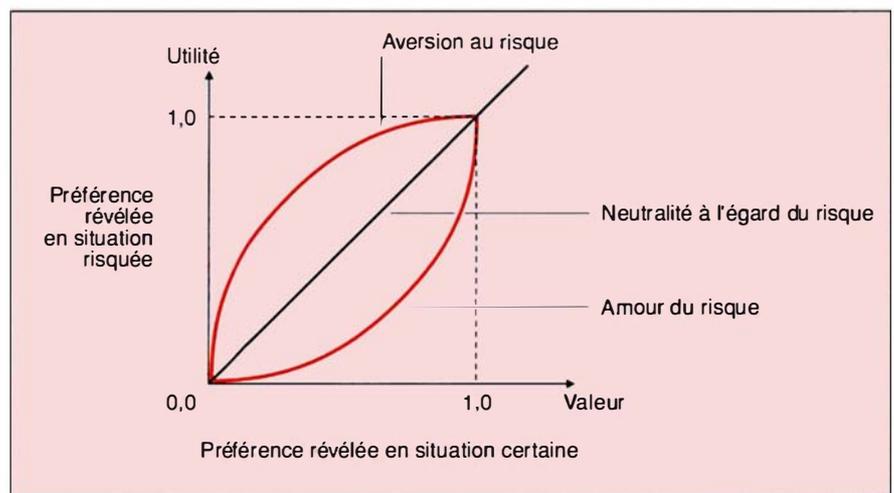


Figure 5. Influence de l'attitude à l'égard du risque sur les préférences selon qu'elles sont révélées en situation certaine (valeur) ou probabiliste (utilité). Les différences empiriques observées entre l'équivalent probabiliste, l'échelle visuelle analogique et l'arbitrage temporel révèlent, en particulier, une attitude différenciée à prendre en compte l'attitude individuelle à l'égard du risque. Selon von Neumann et Morgenstern [35], si une personne préfère une conséquence certaine à une situation probabiliste ou loterie ayant la même espérance de gain, cela signifie qu'elle préfère éviter les risques associés à la loterie. On parle alors d'aversion vis-à-vis du risque et l'importance de celle-ci se traduit par la concavité plus ou moins forte de la fonction d'utilité de la personne. Si, au contraire, la personne préfère la loterie, elle fait preuve d'une préférence pour le risque, sa fonction d'utilité est convexe. Enfin, en cas d'indifférence, le comportement de la personne est dit neutre à l'égard du risque.

situations. Les différences empiriques observées entre l'équivalent probabiliste, l'échelle visuelle analogique et l'arbitrage temporel révèlent donc une aptitude différente qui doit être prise en compte lors du choix de la méthode de révélation pour que les préférences obtenues soient cohérentes avec le cadre théorique de leur utilisation ultérieure. C'est la raison pour laquelle de nombreux économistes privilégient la méthode de l'équivalent probabiliste.

D'autre part, les réflexions des économistes ont rapidement porté sur le choix des personnes dont les valeurs devaient fonder la pondération des index économiques de mesure de la qualité de vie. Trois groupes sont, en effet, susceptibles d'être interrogés : les professionnels de santé (ou des experts), les malades directement concernés (ou les anciens malades) et la population générale. Le premier groupe a été surtout sollicité dans les premiers travaux empiriques mais, très vite, le fait qu'il ne reflète pas les valeurs des patients ou de la population générale a remis en cause la possibilité de s'appuyer sur les préférences de ce groupe bien informé pour construire une mesure des préférences pour des états de santé. Le choix entre les deux autres groupes reste discuté. Les personnes qui sont effectivement dans l'état de santé évalué ont tendance à le juger différemment de celles issues de la population générale [36, 37]. Leurs jugements sont, toutefois, vraisemblablement plus proches de ceux des futurs malades que ceux des personnes qui doivent imaginer être dans un état. Les personnes effectivement concernées sont également susceptibles, toutefois, de surestimer leur propre situation par rapport à d'autres groupes, biaisant ainsi la comparaison générale. Enfin, l'interrogation de telles personnes peut se révéler impossible (personnes atteintes de troubles mentaux, handicapés, enfants...). Le fait que les personnes interrogées ne connaissent ni les situations de maladie ni les traitements ne constitue pas un argument définitif pour rejeter l'intérêt de leurs valeurs. De nombreuses décisions individuelles sont prises à partir d'une in-

formation partielle ou imparfaite. L'usage qui doit être fait de ces index pondérés par les préférences oriente le choix de la population enquêtée. Lorsque les préférences doivent être utilisées pour allouer des ressources collectives, plusieurs groupes d'experts ont, ainsi, recommandé d'interroger la population « des citoyens informés bien que les patients puissent constituer une solution de rechange acceptable » [38]. L'objet de ce débat est, en réalité, de décider s'il est plus acceptable de fonder des valeurs collectives à partir de jugements de personnes imparfaitement informées ou de le faire à partir des jugements de groupes mieux informés sur certains états de santé mais qui peuvent avoir des intérêts spécifiques. Lorsqu'il s'agit, en revanche, de développer des outils d'aide à la décision individuelle, comme c'est le cas pour les travaux portant sur la décision médicale partagée, ce sont, plus particulièrement, les valeurs des patients qui sont au cœur des études de préférences ■

RÉFÉRENCES

1. Weinstein MC, Schiff I. Cost-effectiveness of hormone replacement therapy in the menopause. *Obstet Gynecol Surv* 1983; 38: 445-55.
2. Tosteson AN, Rosenthal DI, Melton LJ. Cost effectiveness of screening perimenopausal white women for osteoporosis: bone densitometry and hormone replacement therapy. *Ann Intern Med* 1990; 113: 594-603.
3. Torrance G, Boyle MH, Horwood SP. Application of multi-attribute theory to measure social preferences for health states. *Operations Res* 1982; 30: 1043-69.
4. Feeny D, Leiper L, Barr RD, et al. The comprehensive assessment of health status in survivors of childhood cancer: application to high-risk acute lymphoblastic leukaemia. *Br J Cancer* 1993; 67: 1047-52.
5. Feeny D, Torrance GW, Furlong W. Health utilities index. In: Spilker B, ed. *Quality of life and pharmacoeconomics in clinical trials*. Philadelphia: Lippincott-Raven Publishers, 1996: 239-242.
6. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems - health utilities index. *Pharmaco Economics* 1995; 7: 490-502.
7. Torrance GW, Feeny D, Furlong W, Barr RD, Zhang Y, Wang Q. Multiattribute utility

function for a comprehensive health status classification system. Health utilities index mark 2. *Med Care* 1996; 34: 702-22.

8. Boyle MH, Furlong W, Feeny D, Torrance GW, Hatcher J. Reliability of the health utilities index-mark III used in the 1991 cycle 6 Canadian general social survey health questionnaire. *Quality Life Res* 1995; 4: 249-57.

9. Le Galès C, Buron C, Costet N, Rosman S, Slama G. Développement d'un index d'états de santé pondéré par les préférences en population française. In: *L'état de la réforme*. Colloque du Collège des Économistes de la Santé, Paris (France), 3-4 février 2000, publication sur CD-Rom.

10. Le Galès C, Costet N, Gentet JC, et al. Cross-cultural adaptation of a health status classification in children with cancer. First results of the French adaptation of the HUI 2 and 3. *Int J Cancer* 1999 (suppl 12): 112-8.

11. Costet N, Le Galès C, Buron C, et al. French cross-cultural adaptation of the health utilities indexes mark 2 (HUI 2) and 3 (HUI 3) classification systems. *Quality Life Res* 1998; 7: 245-56.

12. Le Galès C, Costet N, Buron C et le Groupe HUI Méthodologie de l'adaptation transculturelle d'un index d'états de santé pondéré par les préférences. *J Econom Med* 1997; 15: 529-48.

13. Essink-Bot ML, Stouthard ME, Bonsel GJ. Generalizability of valuations on health states collected with the EuroQol questionnaire. *Health Economics* 1993; 2: 237-46.

14. EuroQol Group EuroQol. A new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199-208.

15. Dolan P. Modelling valuations for EuroQol health states. *Med Care* 1997; 35: 1095-108.

16. Dolan P, Gudex C, Kind P, Williams A. *A social tariff for EuroQol: results from a UK general population survey*. University of York (GB): Center for Health Economics, 1995: discussion paper 138.

17. Kaplan RM, Anderson JP. A general health policy model: update and applications. *Health Serv Res* 1988; 23: 203-35.

18. Kaplan R, Ernst JA. Do rating scales produce biased preference weights for a health index? *Med Care* 1983; XXI: 193-207.

19. Kaplan RM, Bush JW, Berry CC. Health status index: category ratings versus magnitude estimation for measuring levels of well being. *Med Care* 1979; 17: 501-25.

20. Kaplan RM, Bush JW, Berry CC. The reliability, stability and generalisability of a health status index. Social statistics section. *Am Stat Assoc Proc* 1978: 704-9.

21. Fanshel S, Bush JW. A health-status index and its application to health-services outcomes. *Operations Res* 1970: 1021-66.

RÉFÉRENCES

22. Keeney R, Raiffa H. *Decision with multiple objectives*. Cambridge, London: Cambridge University Press, 1993.
23. McNeil BJ, Pauker SG, Sox HC, et al. On the elicitation of preferences for alternative therapies. *N Engl J Med* 1982; 306: 1259-62.
24. Mazur DJ, Hickam M. Treatment preference of patients and physicians: influence of summary data when framing effects are controlled. *Med Decision Making* 1990; 10: 2-5.
25. Fischhoff B, Slovic P, Lichtenstein S. Knowing what you want: measuring labile values. In: Wallsten T, ed. *Cognitive processes in choice and decision behavior*. Hillsdale NJ (USA): Erlbaum, 1980: 117-41.
26. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 1979; 47: 263-91.
27. Lenert L, Cher D, Goldstein M, Garber A. Randomized trial for search procedures for computer-based utility elicitation. Stanford University Ca (USA): Stanford Center for Study of Patient Preferences, 1995.
28. Lenert L, Morss S, Goldstein M, Bergen M and Garber A. *Measurement of the validity of utility elicitation*. Stanford University CA (USA): Stanford Center for Study of Patient Preferences, 1996.
29. Van Wijck EEE, Bosch JL, Hunink M. Time-trade off values and standard gamble utilities assessed during telephone interviews versus face-to-face interviews. *Med Decision Making* 1998; 18: 400-5.
30. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. Describing health states: methodological issues in obtaining values for health states. *Medical Care* 1984; 22: 543-52.
31. Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. *Socio-Economic Plann Sci* 1976; 10: 128-36.
32. Bombardier C, Wolfson AD, Sinclair AJ, McGeer A. Comparison of three preference measurement methodologies in the evaluation of a functional index. In: Deber R, Thompson G, eds. *Choices in health care: decision making and evaluation of effectiveness*. Toronto, Canada: University of Toronto Press, 1982.
33. Hershey JC, Schoemaker PJH. Probability versus certainty equivalence methods in utility measurement: are they equivalent? *Management Sci* 1985; 31: 1213-31.
34. Machina MJ. The economic theory of individual behaviour toward risk. Stanford University Ca (USA): Institute for Mathematical Studies in the Social Sciences, 1983: technical report 433.
35. Von Neumann J, Morgenstern O. *Theory of games and economic behaviour*. Princeton NJ (USA): Princeton University Press, 1947 (2^e ed), 1944 (1^{re} ed).
36. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. *J Chron Dis* 1978; 31: 697-704.
37. Llewellyn-Thomas HA, Sutherland HJ, Thiel EC. Do patients' evaluations of a future health state change when they actually enter the state? *Med Care* 1993; 31: 1002-12.
38. Canadian Coordinating Office for Technology Assessment. *Guidelines for economic evaluation of pharmaceuticals*, 2^e ed. Ottawa (Canada), 1997, Traduction française, 1998.

Catherine Le Galès

CREGAS, Centre de recherche en économie et gestion appliquée à la santé, Inserm U. 537, Cnrs UPRESA 8052, Pavillon de la Force, 80, rue du Général-Leclerc, 94276 Le Kremlin-Bicêtre Cedex, France.

TIRÉS À PART

C. Le Galès.

Summary

Measurement techniques of quality of life in health economics

The aim of this article is to present and discuss methods used in economic evaluation of health care to measure quality of life. Broadly speaking, the quality adjustment is always based on a set of weights called preferences or utilities. In a specific decisional context, the simplest approach is to identify all the health states experienced by patients and to ask subjects to reveal directly preferences for each state. The three most widely used techniques are the rating scale and its variants, the standard gamble and the time trade-off. This direct simple approach is resource consuming and a very complex task, and the comparability of results collected from different studies is limited because of great variability of methods used for describing health states, selecting subjects and applying techniques of preference measurement. A recent alternative which is very attractive and is frequently used, is to by-pass the measurement task by using one of the pre-scored multi-attribute health status classification systems that exists. The classification is based upon a multidimensional conception of health. It allows us to describe an individual's or a population's health status at a given moment and over a period of time. The weighting function allows the different health states to be scored in relation to one another on a scale from 0 to 1. The three main available systems (Health Utilities Index, EuroQol and Quality of Well-Being index) are presented. Main methodological issues associated with these indexes are addressed in the context of their use.