



HAL
open science

Mesure et interprétation du changement d'utilité dérivée d'instruments de qualité de vie psychométriques

Kossar Hosseini

► To cite this version:

Kossar Hosseini. Mesure et interprétation du changement d'utilité dérivée d'instruments de qualité de vie psychométriques. Médecine humaine et pathologie. Université de Lorraine, 2013. Français. NNT : 2013LORR0277 . tel-01750622

HAL Id: tel-01750622

<https://hal.univ-lorraine.fr/tel-01750622v1>

Submitted on 14 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

Mention : « Science de la Vie et de la Santé »

par Kossar HOSSEINI

**Mesure et interprétation du changement d'utilité dérivée
d'instruments de qualité de vie psychométriques**

Le 13 décembre 2013

Membre du Jury :

Rapporteurs : Mme Isabelle DURAND-ZALESKI

M. François BOYER

**PU-PH, AP-HP, Unité de Recherche
Clinique en Économie de la Santé,
Paris**

**PU-PH, Université de Champagne-
Ardenne, EA 3797, Reims**

Examineurs : Mme Mariette MERCIER

M. Philippe TESSIER

Mme Anne-Christine RAT

M. Francis GUILLEMIN

**PR, Université de Franche-Comté,
Besançon**

**MCU, EA 4275 – SPHERE, Université
de Nantes**

**MCU-PH, Université de Lorraine,
Université Paris Descartes, APEMAC,
EA 4360, Nancy**

**PU-PH, Université de Lorraine,
Université Paris Descartes, APEMAC,
EA 4360, Nancy
Directeur de Thèse**

REMERCIEMENTS

Au terme de ce travail, je souhaite adresser mes sincères remerciements à toutes les personnes qui ont contribué à sa réalisation et ont permis par leur soutien et leurs conseils, de le mener à bien.

En premier lieu, je tiens à remercier mon directeur de thèse, professeur Francis Guillemin, pour la confiance qu'il m'a accordée en acceptant d'encadrer ce travail doctoral, pour ses multiples conseils, et surtout la patience qu'il a manifestée à mon égard. Je lui suis également reconnaissante pour le temps conséquent qu'il m'a accordé. J'aimerais également lui dire à quel point j'ai apprécié ses qualités pédagogiques et scientifiques, et j'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de ce travail.

Une thèse est un document jugé par des pairs. Toute ma reconnaissance va vers celles et ceux qui ont accepté d'évaluer ce travail.

J'adresse mes profonds remerciements aux professeurs Isabelle Durand-Zaleski et François Boyer qui m'ont fait l'honneur d'accepter d'être les rapporteurs de ce travail de thèse.

J'exprime également ma gratitude au professeur Mariette Mercier et au docteur Philippe Tessier pour l'intérêt qu'ils ont porté à ce travail et d'avoir accepté d'être examinateurs dans le cadre de ce jury de thèse. Je témoigne également toute ma reconnaissance au docteur Anne-Christine Rat pour son enthousiasme et sa sympathie, et d'avoir accepté d'examiner ce travail.

J'adresse mes sincères remerciements au professeur Jacques Hubert, investigateur principal de l'étude Dovirein, et aux membres du comité scientifique de la cohorte KHOALA, pour m'avoir accordé leur confiance et la possibilité de réaliser mes travaux de thèse.

Je remercie également le professeur Serge Briançon et l'ensemble du service d'épidémiologie du CHU de Nancy, et en particulier les personnes qui ont contribué aux projets de recherches ayant permis ce présent travail.

Je remercie mon laboratoire d'accueil l'EA 4360 et toute l'équipe de l'école de santé publique de Nancy.

Depuis mes premiers pas dans le service, j'ai eu la chance d'être entourée par des personnes d'une grande générosité, qui par leur présence et leur réconfort dans les moments de doute m'ont permis d'avancer. Mes sincères remerciements à mes collègues et surtout amies, Elodie, Christine, Ira, Laetitia, Nelly, Lucie et Marie-Line.

Je remercie Guillaume qui, malgré les circonstances, a su rester un soutien pour moi.

Mes remerciements vont aussi à ma famille.

A mes parents, pour votre amour inconditionnel, votre soutien sans faille et tout ce que vous avez pu m'apporter pour franchir les obstacles les plus difficiles. Vous m'avez donné le goût de la recherche scientifique. Sans vous je n'en serais pas là aujourd'hui.

A mes formidables frères Vahid, Mohammad, Abolfazl et Rouhallah. Vous êtes la plus belle sécurité dans ma vie.

A mes grands-parents, que je porte dans mon cœur.

A mes tantes Susan et Sepideh qui m'ont toujours encouragé.

A Claire, Anne-Laure et Alix pour votre chaleureuse présence.

A Timothée, pour sa patience et ses encouragements.

A difference is a difference only if it makes a difference

Darrell Huff

TABLE DES MATIERES

Résumé	3
Abstract.....	4
Liste des travaux liés à la thèse	5
Liste des publications	5
Liste des communications orale et affichée.....	6
1. Utilité indirecte : concept et mesure.....	8
1.1. Le concept d'utilité.....	8
1.2. Mesure des préférences en santé	9
1.2.1. Approche directe	9
1.2.2. Approche indirecte	12
1.3. La mesure d'utilité dans l'analyse économique en santé	15
1.4. La mesure du changement cliniquement significatif.....	16
1.4.1. Les méthodes de mesure du changement	17
1.5. Le changement d'utilité pertinent.....	19
1.6. Comorbidité et qualite de vie.....	20
1.7. Le concept de response-shift	21
2. Présentation des axes de recherche	23
2.1. Les enjeux de la mesure du changement d'utilité.....	23
2.2. Objectifs de la thèse.....	24

2.3. Travaux réalisés dans le cadre de la thèse	24
3. Premier travail	28
4. Mesure d'utilité et comorbidités.....	29
4.1. Deuxième travail.....	31
5. Mesure du seuil de changement de l'utilité en présence de comorbidité.....	32
5.1. Troisième travail.....	35
6. Le seuil de changement d'utilité chez le donneur vivant de rein.....	36
6.1. Quatrième travail	39
7. Discussion.....	41
8. Perspectives	47
References.....	49
Annexes	57

RESUME

Dans le domaine de la santé, la mesure d'utilité permet d'exprimer la valeur que peut donner un individu à un état de santé, comprise généralement entre 0 (la mort) et 1 (meilleur état de santé envisageable). L'utilité permet de valoriser simultanément le résultat clinique et son impact sur l'état de santé du patient, et d'exprimer ainsi la qualité de vie liée à la santé.

Nous nous sommes intéressés dans ce travail à l'approche indirecte d'estimation d'utilité, consistant à dériver une utilité à partir des échelles psychométriques de mesure de l'état de santé (questionnaires génériques). Cette approche permet de répéter la mesure de l'utilité dans le temps et ouvre des perspectives pour l'analyse du changement d'utilité.

Les objectifs étaient de définir un seuil de changement pour l'utilité dérivée des instruments de mesure psychométrique de la qualité de vie, et de prendre en compte dans l'interprétation du seuil la présence des comorbidités ou d'un phénomène de response-shift.

Afin de répondre à nos objectifs, nous avons étudié, d'une part, le changement d'utilité mesurée par le SF-6D dans une cohorte de patients atteints d'arthrose des membres inférieurs (cohorte KHOALA) et l'effet des comorbidités sur l'utilité et son changement. D'autre part, nous avons estimé l'utilité de donneur vivant de rein à partir des questionnaires EQ-5D et SF-6D avant et 3 mois après le prélèvement du rein. Un seuil de changement a été estimé et un phénomène de response-shift a été recherché.

Les résultats ont montré l'effet délétère des comorbidités sur la mesure de l'utilité chez les patients atteints d'arthrose, et mis en évidence l'importance de la sévérité fonctionnelle de l'arthrose dans l'expression de l'utilité et son changement. Chez les donneurs vivants de rein nous avons mis en évidence une diminution significative de l'utilité à 3 mois, avec une proportion de plus de 30% de donneurs ayant atteint le seuil de détérioration significative. L'estimation du seuil de changement diffère selon la population étudiée, l'instrument de mesure de l'utilité, et la méthodologie appliquée pour définir ce seuil.

En conclusion, l'évaluation du changement d'utilité doit tenir compte de l'existence des comorbidités ainsi que de l'adaptation du jugement de valeur que l'individu porte sur son état de santé lorsque celui-ci évolue. La complexité des méthodes psychométriques pour la mesure de l'utilité doit amener à être prudent sur leur prise en compte dans les modèles d'évaluation économique. L'identification d'un seuil pour l'interprétation du changement peut être un élément complémentaire utile dans la démarche d'analyse décisionnelle.

Mots clés : utilité indirecte, qualité de vie, SF-6D, EQ-5D, différence minimale importante

ABSTRACT

In health economic, the concept of utility was used to describe a state of health in patients. The unit value of health state utility ranges between 0, representing “death” state, and 1 representing a “perfect” health state. It is possible to recover through the measurement of utility both the clinical outcome and their impact on the health of the patient, and thus express the quality of life related to health. In this work, we used the indirect approach for estimating utility that is derived from a generic utility scales (quality of life questionnaires).

The objectives of this thesis were to define a threshold of change for the utility derived from psychometric instruments measuring quality of life, and to take into account for the change interpretation, the presence of comorbidities or a response-shift phenomenon.

We used data from a cohort of patients with osteoarthritis of lower limbs (KHOALA cohort) for estimate the impact of comorbidities in utility measured by the SF-6D its change over 3 years. We are also interested in the particular case of living kidney donors. The utility was estimated by the EQ-5D and SF-6D before and 3 months after kidney donation. A threshold of significant deterioration has been estimated and the presence of a response-shift phenomenon was controlled.

The results showed the deleterious effect of comorbidities on utility measurement for patients with lower limbs osteoarthritis, and highlighted the importance of the functional severity of osteoarthritis in the expression of utility and its change over time. Living kidney donors experienced a significant decrease in their utility score in the short term, and more than third of donors have reached the threshold of significant deterioration. The threshold of utility change varies depending on the population studied, the measurement instrument, and the methodology used to define the threshold.

In conclusion, the assessment of utility change must take into account the existence of other diseases and the adaptation of the individual on his health state variation. The complexity of measuring individual preferences by psychometric instruments should lead to be careful about introducing them in the economic evaluation. The use of a threshold for the interpretation of utility change can be an additional element in the decision-making process.

Key words: utility, health related quality of life, SF-6D, EQ-5D, minimal important difference

LISTE DES TRAVAUX LIES A LA THESE

LISTE DES PUBLICATIONS

1. Measuring change of utility for the individual: relevance and integration in cost-utility analysis

Hosseini Kossar, Gaujoux-Viala Cecile, Guillemin Francis

Soumis en révision

2. Impact of comorbidities on measuring indirect utility by the Medical Outcomes Study Short Form 6D in lower-limb osteoarthritis

Hosseini Kossar, Gaujoux-Viala Cecile, Coste Joel, Pouchot Jacques, Fautrel Bruno, Rat Anne-Christine, Guillemin Francis

Best Pract Res Clin Rheumatol 2012 Oct; 26(5):627-35.

3. Impact of comorbidities on utility changes in lower-limb osteoarthritis : KHOALA Study

Hosseini Kossar, Rat Anne-Christine, Gaujoux-Viala Cecile, Coste Joel, Pouchot Jacques, Fautrel Bruno, Guillemin Francis

En préparation

4. Decrease in health state utility among living kidney donors

Hosseini Kossar, Jacques Hubert, Marc Ladrière, Guillemin Francis

Soumis

LISTE DES COMMUNICATIONS ORALE ET AFFICHEE

1. Hosseini Kossar, Gaujoux-Viala Cecile, Coste Joel, Pouchot Jacques, Fautrel Bruno, Rat Anne-Christine, Guillemin Francis. Impact of comorbidities on measuring indirect utility by the Medical Outcomes Study Short Form 6D in lower-limb osteoarthritis.
 - Communications affichées: ISPOR Annual European Congress 2012, American College of Rheumatology Congress 2012, 6ème Conférence Francophone d'Épidémiologie Clinique 2012, 25^{ème} congrès français de rhumatologie

2. Hosseini Kossar, Jacques Hubert, Marc Ladrière, Guillemin Francis. Decrease in health state utility among living kidney donors.
 - Communication orale: 7ème Conférence Francophone d'Épidémiologie Clinique 2013.
 - Communication affichée : ISPOR Annual European Congress 2013.

3. Hosseini Kossar, Rat Anne-Christine, Gaujoux-Viala Cecile, Coste Joel, Pouchot Jacques, Fautrel Bruno, Guillemin Francis. Impact of comorbidities on utility changes in lower-limb osteoarthritis: KHOALA Study.
 - Communication affichée: ISPOR Annual European Congress 2013

1. UTILITE INDIRECTE : CONCEPT ET MESURE

1.1. LE CONCEPT D'UTILITE

Le concept d'utilité est utilisé par les économistes pour définir le degré de bien-être ou de satisfaction apporté à un individu par la consommation (obtention) d'un bien ou service. Cette notion d'utilité en économie est essentiellement liée à la prise de décision en situation de risque, et reflète le choix des individus selon un profil défini (averse au risque, averse aux pertes).

Dans le domaine de la santé, le concept d'utilité introduit par les économistes, s'inspire directement de la théorie contemporaine de l'utilité développée par von Neumann et Morgenstern en 1944 (1). Ce modèle, qui introduit la théorie des jeux et la prise de décision en situation d'incertitude, permet de décrire comment un individu (agent) prend une décision dans une situation où les conséquences d'une ou plusieurs actions sont incertaines.

L'utilité espérée décrit par von Neumann et Morgenstern et revu par Savage en 1954 (2), est basée sur cinq axiomes : la réflexivité, la complétude, la transitivité, la continuité, et l'indépendance. Les trois premiers axiomes permettent de s'assurer que l'ensemble de préférences associé à un agent forme un ordre. L'axiome de continuité est une condition technique qui permet de représenter cet ordre de préférences par une fonction d'utilité. Enfin, l'axiome d'indépendance est une propriété fondamentale qui garantit que l'on peut représenter les préférences de l'agent par une fonction d'utilité correspondant à la somme pondérée des utilités associées à chaque état.

Dans ce cadre théorique, l'agent est dit « rationnel » au sens économique si ses choix sont cohérents avec ses préférences, et s'il choisit l'action dont l'utilité espérée est la plus grande lorsque les choix sont effectués dans un contexte où leurs conséquences ne sont pas certaines.

L'axiome d'indépendance admet que les pondérations des utilités sont des probabilités et que les préférences sont linéaires par rapport aux probabilités, a été remis en cause par les paradoxes d'Allais (3) et d'Ellsberg (4). La mise en évidence de ces paradoxes a conduit au développement d'autres théories sur l'espérance d'utilité que nous n'introduirons pas dans ce présent travail car ne sont pas à ce jour appliquées dans le domaine de la santé.

1.2. MESURE DES PREFERENCES EN SANTE

Dans le domaine de la santé, la mesure d'utilité permet d'exprimer la préférence d'un patient confronté à un choix face à des stratégies diagnostiques et thérapeutiques différentes, et exprime plus particulièrement la valeur que peut donner un individu à un état de santé donné. Cette valeur est comprise généralement entre 0 représentant la mort, et 1 étant le meilleur état de santé envisageable. Il est possible de valoriser à travers l'utilité, à la fois le résultat clinique et leur impact sur l'état de santé général du patient, et ainsi exprimer la qualité de vie liée à l'état de santé (Health Related Quality of Life). Cette mesure des préférences de l'individu est une étape essentielle de l'évaluation économique en santé et permet par l'expression d'une unité commune (utilité) dans les analyses médico-économiques, la comparaison des échantillons de patients avec des traitements et pathologies différents. Elle peut donc faciliter la prise de décision et l'interprétation des résultats destinés au décideur final dans l'objectif d'une amélioration du bien-être social dans une perspective *welfariste*¹.

L'évaluation de l'utilité en santé est possible par deux approches : l'approche directe (*preference-based*) et l'approche indirecte (*health status*) en utilisant les instruments de mesure psychométriques.

1.2.1. APPROCHE DIRECTE

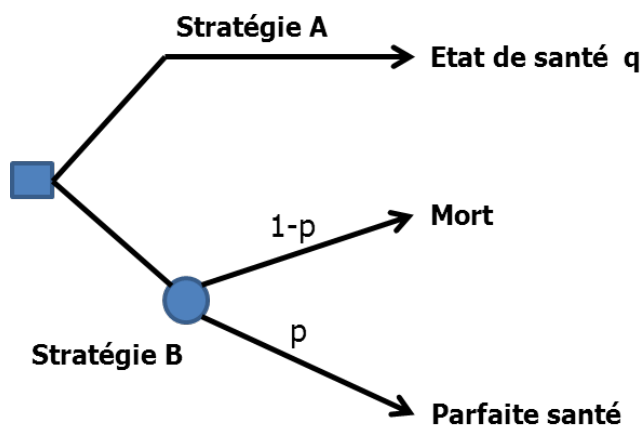
Les deux principales méthodes directes de mesure des préférences en santé sont la méthode des loteries (standard gamble) et la méthode du marchandage des temps (time-trade off). Elles sont respectivement basées sur le risque de décès pour la première, et la réduction de durée de vie pour la seconde, consenti(e) par un individu pour l'obtention d'une meilleure santé.

¹ *Le welfarisme en économie désigne la politique de l'Etat Providence. Dans ce cadre (et à l'inverse d'un état fondé sur le capitalisme) les pouvoirs publics interviennent afin de réguler les marchés ou assurer la protection sociale des individus contre la maladie, la vieillesse ou encore le chômage. L'objectif étant toujours d'améliorer le bien-être des individus, mesuré par l'utilité.*

1.2.1.1. METHODE DES LOTERIES OU STANDARD GAMBLE (SG)

Cette méthode est directement inspirée de la théorie des jeux telle que présentée par von Neumann et Morgenstern, et permet de mesurer l'utilité dans un contexte de risque et en y affectant une valeur cardinale (5-7). L'extension de la théorie des jeux aux champs de l'analyse décisionnelle en santé a été décrite par Torrance (8). Cette méthode consiste à proposer un jeu de loterie à un individu : d'une part un état de santé q (correspondant à un état dégradé mais préféré à la mort) pour le restant de sa vie, et d'autre part un état de santé parfait mais avec une probabilité $(1-p)$ de décès immédiat associée à celui-ci. Il s'agit donc pour l'individu de choisir entre une vie certaine dans un état de santé dégradé ou la possibilité d'un meilleur état de santé mais avec un risque de décès immédiat. La valorisation des préférences de l'individu consiste alors à faire varier la probabilité p jusqu'à ce que l'individu soit indifférent entre l'option risqué et l'état de santé (certain) dégradé.

FIGURE 1. JEU DE LOTERIES (9)



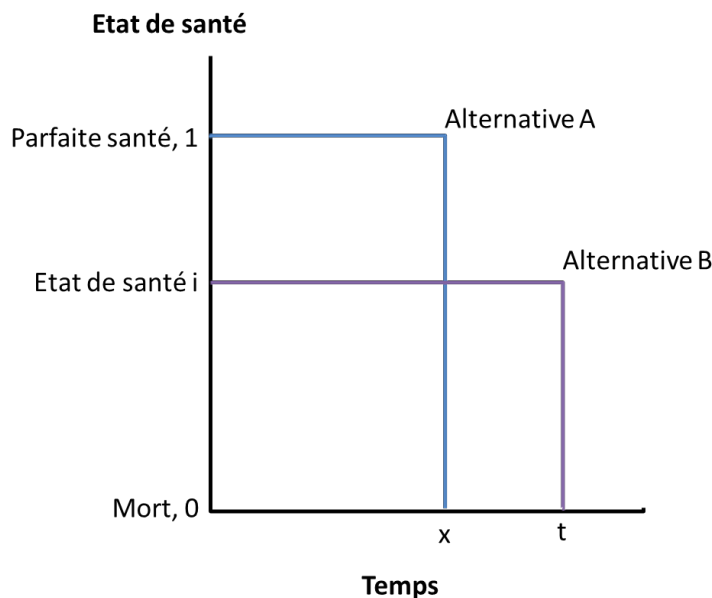
Cette méthode a été remise en question par la théorie des perspectives de Kahneman et Tversky (10) qui suppose une aversion pour les pertes et non pour le risque comme énoncé pour la fonction d'utilité de von Neumann et Morgenstern, indiquant ainsi que les individus ont tendances à déformer subjectivement les probabilités de gains et de pertes(11).

1.2.1.2. METHODE DU « MARCHANDAGE TEMPS » OU TIME TRADE-OFF (TTO)

La méthode développée par Torrance (12) est un peu moins compliquée et conceptuellement différente de la précédente. Elle n'intègre plus la notion d'incertitude dans le processus

décisionnel et permet de s'affranchir de la question de l'aversion au risque posée par la précédente méthode (13). Il s'agit cette fois de déterminer la préférence d'un individu entre deux résultats certains. Le schéma présenté à l'individu est composé d'une première alternative d'un état de santé pour une durée déterminée t , et d'une alternative concurrente offrant un meilleur état de santé (état optimal) mais pour une durée plus courte ($t-1$). Les deux alternatives sont suivies conventionnellement de la mort. Il s'agit alors de trouver, pour un individu, le point d'indifférence entre les deux alternatives en faisant varier la durée passée dans le meilleur état de santé. Au même titre que la méthode des loteries, le marchandage-temps a également des limites, dont la principale liée à une préférence pour le présent des individus.

FIGURE 2. METHODE DE MARCHANDAGE TEMPS (9)



Ces méthodes directes d'évaluation des préférences ont des limites engendrées par les hypothèses sur lesquelles elles reposent, et ont tendance à donner des valeurs d'utilité différentes pour un même échantillon (12;14-17). La valeur estimée de l'utilité par la méthode des loteries excède généralement celle obtenue par le marchandage temps (18-20). Il n'y a pas à ce jour de consensus sur la méthode à adopter pour l'évaluation des préférences individuelles. Cependant certains auteurs considèrent que le marchandage temps reflète mieux les préférences car exclut l'aversion au risque (15) et est plus facile à comprendre (16).

Ces méthodes directes nécessitent un entretien en face à face avec l'individu, et sont donc complexes à mettre en œuvre, en particulier si l'évaluation des préférences doit être répétée dans le temps. C'est pourquoi une approche indirecte de la mesure de l'utilité a été par la suite développée.

1.2.2. APPROCHE INDIRECTE

L'approche indirecte d'estimation d'utilité ou utilité multi-attributs consiste à dériver les préférences individuelles à partir des échelles génériques de mesure de qualité de vie liée à la santé (Health related quality of life), et est de plus en plus utilisée pour l'analyse médico-économique (21;22). Ces échelles génériques multidimensionnelles proposent une description de l'état de santé de l'individu à travers plusieurs dimensions (physique, psychique...). Cette description se fait sur l'état de santé actuel des individus, alors que la mesure de l'utilité directe permet d'estimer les préférences pour un état donné en projetant l'individu dans un avenir (résultat de santé) conditionnel ou incertain.

Cette méthode indirecte est moins complexe et plus aisée à mettre en œuvre, et donne la possibilité de mesurer une évolution intra-sujet en répétant la mesure dans le temps. Elle consiste à synthétiser dans un indicateur unique les différentes dimensions de la qualité de vie (mobilité, douleur, anxiété...) mesurée chez un individu, et par la suite pondérer cet indicateur par la valeur sociétale accordée à celui-ci. Cette démarche de pondération se fait à l'aide d'un algorithme préalablement défini : la pondération sociétale des états de santé est obtenue par mesure directe (jeu des loteries ou marchandage temps) auprès d'échantillons représentatifs de la population générale.

Les instruments de mesure de l'utilité indirecte les plus connues et utilisées sont les questionnaires multi-attributs EQ-5D de l'Euroqol (23), le SF-6D (24) et le HUI (Health Utility Index) (25). Nous ne développerons dans le cadre de ce présent travail les deux premiers instruments.

Ces auto-questionnaires permettent de déterminer les préférences des individus à un moment donné et/ou au cours du temps. Cependant, les pondérations sociétales sont obtenues indépendamment, à l'aide des méthodes directes d'évaluation de l'utilité, et dans ce sens les

limites discutées précédemment s'appliquent également aux résultats obtenus par la méthode indirecte.

1.2.2.1. L'EQ-5D DE L'EUROQOL

Il s'agit d'un questionnaire de mesure de qualité de vie avec 5 dimensions : mobilité, autonomie de la personne, activité courante, douleur/gêne, et anxiété/dépression. Chaque dimension est représentée par un item avec 3 niveaux de réponses possibles. Les réponses données à chaque dimension sont alors combinées afin d'obtenir l'état de santé de l'individu. Ce questionnaire permet de décrire 243 états de santé qui par la suite et à l'aide d'un algorithme sont pondérés par la valeur sociétale accordée à chaque état. L'algorithme dérivé de la population générale française (26), utilisant la méthode du marchandage temps, permet d'obtenir une valeur d'utilité entre -0.529 et 1, 1 étant un état de santé parfait, 0 la mort et les valeurs négatives traduisant un état que l'on qualifie pire que la mort.

1.2.2.2. LE SF-6D

Ce deuxième questionnaire est une version courte du questionnaire de qualité de vie SF-36 (27), comportant 36 items et 8 dimensions : activités physiques, limitations dues à l'état de santé, douleurs physiques, vie et relations avec les autres, santé psychique, limitations dues à l'état psychique, et la vitalité et santé perçue. Le SF-6D s'obtient à partir de 11 questions du SF-36 (28) et comporte 6 dimensions : fonction physique, limitation de rôle, fonction sociale, douleur, santé mentale et vitalité. Pour chaque dimension l'individu a le choix entre 4 à 6 niveaux de réponses, et dans ce sens le SF-6D permet de décrire 18000 états de santé différents. La fonction de pondération utilisée pour dériver l'utilité a été obtenue à partir d'une enquête par la méthode du jeu des loteries auprès d'un échantillon représentatif au Royaume-Uni. L'utilité obtenue par ce questionnaire varie de 0.29 à 1, 1 étant un état de santé parfait. Il n'existe pas à ce jour de pondération estimée à partir de la population française pour ce questionnaire.

Les méthodes directe et indirecte de mesure d'utilité ont des avantages et limites synthétisés dans le tableau 1. Cependant l'accessibilité de la méthode indirecte et sa capacité à pouvoir se répéter dans le temps a conduit à une plus grande utilisation de celle-ci dans le domaine de l'évaluation économique.

TABLEAU 1. AVANTAGES ET INCONVENIENTS DES DIFFERENTES MESURES D'UTILITE (29-31)

	Avantages	Inconvénients
Utilité directe	<ul style="list-style-type: none"> - Donne une valeur unique pour la mesure de la qualité de vie liée à la santé - Estimation directe des préférences. 	<ul style="list-style-type: none"> - Peut ne pas aussi sensible au changement qu'un instrument de mesure de l'état de santé - Difficile à mettre en œuvre - Nécessite un entretien en face à face - Le concept peut être difficile à comprendre pour certains individus - Peut dépendre de l'attitude de l'individu vis-à-vis de la valeur monétaire, le risque ou le temps - Les limites méthodologiques du SG* et TTO** conduisent à une divergence des résultats obtenus
Utilité indirecte	<ul style="list-style-type: none"> - Donne une valeur unique pour la mesure de la qualité de vie liée à la santé - Peut être considérée plus approprié par les décideurs car utilise les pondérations sociétales pour la valorisation de l'utilité - Ne nécessite pas d'entretien individuel en face à face - Facile à administrer. - Permet la comparaison entre différentes populations avec différentes pathologies 	<ul style="list-style-type: none"> - Pourrait ne pas être aussi sensible au changement que l'instrument de mesure de l'état de santé - Les pondérations sociétales sont basées sur les méthodes directes d'évaluation de l'utilité et conduisent en partie à une des résultats divergeant

* SG : Standard Gamble; **TTO : Time Trade Off

1.3. LA MESURE D'UTILITE DANS L'ANALYSE ECONOMIQUE EN SANTE

L'objectif principal de l'évaluation économique en santé est de maximiser le bien être des patients (gain individuel) en minimisant le coût social de la prise en charge de la maladie. Dans ce sens l'évaluation médico-économique devient un outil majeur de prise de décision pour les pouvoirs publics et permet d'évaluer les stratégies diagnostiques et thérapeutiques concurrentes.

L'analyse médico-économique peut intégrer les gains (efficacité) cliniques (exemple de la diminution de mortalité dans le cancer de poumon), ou les préférences individuelles (liées à la santé des patients atteints d'un cancer de poumon). Dans ce deuxième cas, on parle d'analyse coût-utilité, permettant de prendre en compte d'une part le coût des stratégies concurrentes et d'autre part les préférences des individus exprimées par l'utilité pour chaque stratégie. Le résultat est exprimé en termes de ratio coût-utilité déterminé par le rapport de différence des coûts et d'utilités mesurés pour chaque stratégie.

Les méthodes indirectes de mesure d'utilité évoquées précédemment, ont permis une plus large application de l'analyse coût-utilité dans un but décisionnel, et ont facilité l'intégration du point de vue des patients dans cette démarche. Ces méthodes permettent plus aisément de répéter la mesure des préférences lorsqu'il s'agit d'observer l'évolution de l'état de santé des individus dans le temps. Elles mènent ainsi à comparer le gain d'utilité apporté (variation entre les sujets) par chaque stratégie concurrente et le coût associé à celui-ci.

Cependant, cette possibilité de répéter les mesures de préférences dans le temps engage la discussion sur une évaluation du changement intra-individu. Ce point est rarement évoqué dans l'analyse économique alors que la mesure du changement intra-individu garde son importance dans le domaine clinique. On peut avec intérêt expliquer l'évolution de la qualité de vie dans un groupe de patients avec pour objectif leur prise en charge clinique, mais la littérature scientifique actuelle n'en fait plus mention dès lors qu'il s'agit d'une mesure des préférences individuelles. Or, la question du changement d'utilité peut amener à envisager la question de l'analyse coût-utilité différemment en observant, dans un premier temps et avant même de comparer les stratégies, la variabilité intra-individu des préférences, engendrée par

la mise en œuvre d'une nouvelle stratégie. Cette nouvelle perspective nous conduit à définir la notion de changement d'utilité et de son interprétation dans une optique décisionnelle.

1.4. LA MESURE DU CHANGEMENT CLINIQUEMENT SIGNIFICATIF

Le résultat d'un essai clinique est généralement reporté en termes de significativité statistique. Par exemple pour comparer deux techniques chirurgicales on peut évaluer la douleur par une échelle visuelle analogique (EVA) et montrer une différence statistiquement significative de ce score selon la technique de chirurgie utilisée. Cependant cette information sur la significativité statistique ne peut pas conduire directement à une interprétation sur la significativité clinique, et ne peut pas refléter une amélioration (ou détérioration) cliniquement significative pour un individu. Comme évoqué par Osoba et al. (32) un changement statistiquement significatif dépend de la taille de l'échantillon étudié et ne conduit pas systématiquement à un changement réellement perçu par l'individu.

Pour répondre à cette limite d'interprétation des résultats, le concept de changement cliniquement significatif, a introduit à la fin des années 80 la notion de différence minimale importante (MID) ou de différence minimale cliniquement importante (MCID) (33). Cette dernière permet une interprétation du résultat clinique perçu par le patient en plus d'une significativité statistique, par la détermination d'un seuil de changement du score. Cette différence cliniquement importante a été définie par les auteurs comme "la plus petite modification de score perçue comme bénéfique par les patients et qui en l'absence d'effets secondaires importants ou de surcoût excessif justifierait une modification de la prise en charge médicale"² (34).

Cet indicateur permet une interprétation du changement de score de qualité de vie liée à la santé pour un individu autre que l'interprétation statistique. Un changement inférieur au seuil défini est considéré comme étant cliniquement non significatif pour l'individu (35).

² *"The smallest difference in score in the domain of interest which patients would perceive as beneficial and which would mandate, in the absence of troublesome side-effects and excessive cost, a change in the patient's management."*

Plus récemment a été développé le concept d'un état symptomatique acceptable par le patient (PASS) défini comme « la valeur au-delà de laquelle les patients se considèrent bien »³ (36). Les concepts du MID/MCID ou PASS peuvent être considérés comme complémentaires; le premier permet au patient d'exprimer le changement du score comme le fait "d'aller mieux" et le second le fait "d'aller bien" (37).

1.4.1. LES METHODES DE MESURE DU CHANGEMENT

De nombreuses méthodes ont été développées pour estimer le seuil de changement cliniquement significatif (MID/MCID) (33;38). Ces méthodes peuvent être distinguées en deux groupes : la méthode basée sur une référence externe (anchor-based method), et la méthode basée sur la distribution de l'échantillon (distribution-based method) (39). Nous donnerons ici une brève présentation de ces méthodes.

1.4.1.1. LA METHODE BASEE SUR REFERENCE EXTERNE (ANCHOR-BASED)

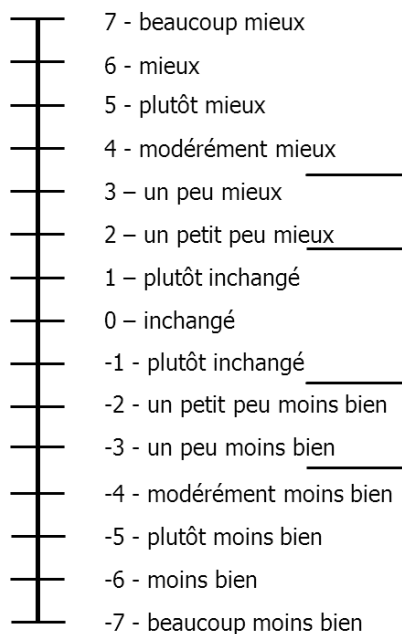
Cette approche, initialement définie par Jaeschke et al. (34) et revue par Juniper et al. (40), est basée sur un jugement extérieur de la pertinence clinique de la variation observée du score d'intérêt. Bien qu'il soit préférable dans ce domaine de donner plus de poids au jugement du patient et de l'interroger directement, d'autres références externes cliniquement pertinentes peuvent être utilisées, comme la tolérance au traitement ou la dépendance physique jugée par un médecin.

La méthode conventionnelle utilisée consiste à faire une évaluation globale du changement (référence externe) auprès des patients, à laquelle seront rattachées les variations du score de ces derniers.

Ainsi, on soumet à l'avis du patient la classification de son état de santé globale à l'aide d'une échelle allant de « beaucoup mieux » à « beaucoup moins bien » comme le présente la figure suivante.

³ *“The value beyond which patients consider themselves well”*

FIGURE 3. ECHELLE D'ÉVALUATION DE LA SANTÉ GLOBALE DU PATIENT



Le MCID se définit alors comme la moyenne des valeurs absolues du changement de score pour les patients estimant leur état de santé par les modalités -3, -2, 2, 3 de l'échelle ci-dessus (40).

1.4.1.2. METHODE BASEE SUR LA DISTRIBUTION DE L'ÉCHANTILLON

Cette seconde approche repose sur des paramètres statistiques de variabilité du score dans l'échantillon. Il existe plusieurs méthodes permettant d'évaluer le MID/MCID par cette approche qui ont été très bien détaillées par Crosby et al. (35) et définies en 3 types basées sur :

- La signification statistique de la variation ; la plus utilisée est la déviation standard (DS) : Norman et al. (41) ont défini $0,5DS$ du score initial comme une valeur robuste pour déterminer le MID/MCID.
- La variabilité au sein de l'échantillon ; l'Effect Size (ES) proposé par Cohen (42), est calculé à partir la moyenne de la variation des scores divisée par l'écart type du score initial. Le résultat indique alors la taille de l'effet traitement au niveau du groupe et s'interprète de la façon suivante : faible si ES de 0,20, moyen si de 0,50 et grand si ES à 0,80 ou plus.

- La précision de la mesure ; l'erreur de mesure standard (SEM) (43) est calculée à partir de l'écart type de la moyenne du score initial multiplié par la racine carrée de 1 moins la précision du test. Cette méthode prend en compte la possibilité qu'une partie du changement soit la conséquence d'une erreur de mesure aléatoire (indépendante de l'échantillon).

Par leurs différences méthodologiques, les méthodes d'estimation du seuil de changement aboutissent à des valeurs différentes. L'impression peut être donnée que l'on s'éloigne de l'objectif initial qui était de déterminer un seuil (unique) pour aider les cliniciens à interpréter le changement observé chez un patient. Etant donné que les deux approches ont des avantages et inconvénients propres à leur méthodologie, il n'est pas possible de déterminer la supériorité de l'une par rapport à l'autre. C'est pourquoi la recommandation actuelle (33;35) est d'utiliser dans la mesure du possible les deux approches pour établir un intervalle du seuil de changement.

1.5. LE CHANGEMENT D'UTILITE PERTINENT

La notion de changement cliniquement significatif est généralement appliquée à l'interprétation des mesures de changements de qualité de vie liée à la santé et très rarement à l'utilité dérivée de la qualité de vie. Il existe quelques travaux sur l'estimation d'un seuil de changement pour l'utilité dont le principal étant celui de Walters and Brazier (44) sur une différence minimale importante de l'EQ-5D et du SF-6D. Cependant les auteurs rappellent que dans le cadre d'une allocation des ressources, une analyse médico-économique s'impose, et qu'observer un changement de score seul ne suffit pas (45).

Cette mesure du changement d'utilité et d'estimation de seuil d'interprétation d'utilité doit tenir compte du fait que celle-ci est dérivée des instruments de mesure de la qualité de vie et qu'il peut donc être affecté par des éléments ayant un impact sur cette mesure comme la présence des comorbidités ou le phénomène de response-shift (RS). Ces deux concepts interviennent sur la mesure de l'utilité mais sont de natures différentes ; le premier est un phénomène de santé qui se conjugue à l'état actuel de l'individu, et le second est un phénomène interne à l'individu qui traduit son adaptation à son état de santé.

1.6. COMORBIDITE ET QUALITE DE VIE

La définition du terme de comorbidité est donnée initialement par Feinstein et al. (46) comme étant “toute entité diagnostique additionnelle qui apparait ou existe durant l’histoire clinique du patient”⁴.

Comme le souligne Groot et al. (47), la recherche d’une comorbidité dans l’histoire de la maladie a de plus en plus d’importance car elle peut potentiellement être un facteur de confusion dans le diagnostic de la maladie qui nous intéresse et influencer le pronostic du patient. De plus, comme mentionné par Redelmeier et al. (48) le traitement d’une maladie (prise en charge du patient) peut être influencé par la présence d’une ou plusieurs comorbidités.

Dans une revue systématique de l’effet des comorbidités sur la qualité de vie, Fortin et al.(49) ont montré une relation inverse entre ces deux concepts; la présence de comorbidité peut potentiellement affecter l’ensemble des domaines de la qualité de vie. Pour la majorité des études reprises dans cette revue, la qualité de vie était estimée par le questionnaire générique SF-36. Ainsi, si l’effet délétère des comorbidités sur la qualité de vie est avéré, on peut alors s’attendre à un effet similaire sur l’utilité dérivée de cette qualité de vie.

D’autre part, la plupart des études qui s’intéressent à la question sont transversales (50-52), alors que l’effet des comorbidités peut varier dans le temps et peut donc avoir un impact non négligeable sur la mesure du changement d’utilité dérivée de la qualité de vie.

Parmi les indices de mesure des comorbidités les plus couramment utilisés, on peut citer l’indice de Charlson et le CIRS (Cumulative Illness Rating Scale). L’indice de comorbidité de Charlson (53), est basé sur un critère de mortalité. Les 19 pathologies qui composent cet indice ont été déterminées à partir du risque relatif de décès respectif de chacune. Le CIRS (54) utilisé couramment pour la déterminer le poids des comorbidités prend en compte la sévérité des pathologies, allant de 0 pour un poids nul à 4 une extrême sévérité, qui peuvent toucher les 14 principaux domaines anatomique (cardiaque, vasculaire, respiratoire...).

L’indice de Groll (55), plus récemment développé, a cette particularité d’avoir été établi à partir du poids des comorbidités sur la dimension physique de la qualité de vie. Et prend ainsi en compte le retentissement physique des pathologies. Fortin et al. (56) ont montré que

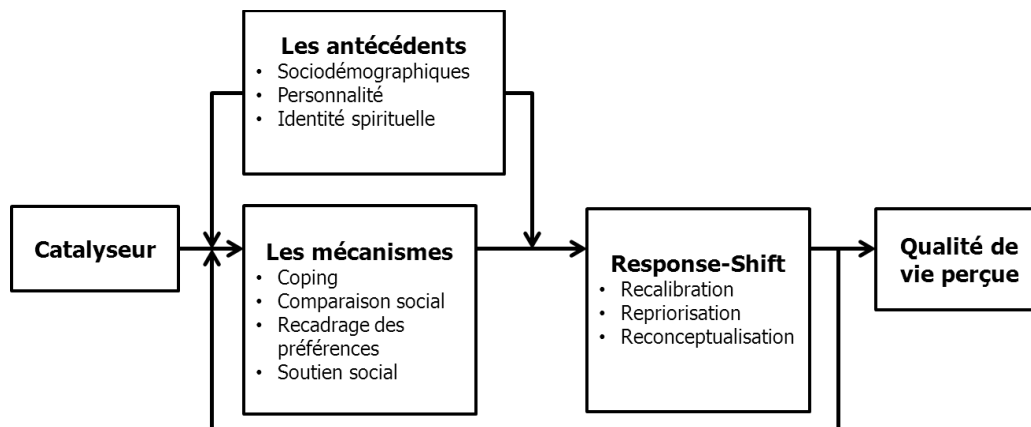
⁴ “Any distinct additional entity that has existed or may occur during the clinical course of patient who has the index disease under study”.

l'indice de Groll, au même titre que le CIRS, avait une meilleure corrélation avec toutes les dimensions de la qualité de vie mesurée par le SF-36 comparé à l'indice Charlson.

1.7. LE CONCEPT DE RESPONSE-SHIFT

Ce phénomène a été dans un premier temps observé chez des patients atteints de cancer (57). Par la suite, Sprangers and Schwartz (58) ont défini les 3 composantes du response-shift qui sont : recalibration, repriorisation et reconceptualisation des valeurs propres à l'individu, et qui vont le conduire à changer ses « références de mesure » suite à au cours de sa maladie. Le patient peut en effet revoir ses priorités ou sa propre échelle de valeurs au cours de sa maladie et sa conception des symptômes et exprimer un changement des critères de jugement, modifiant ainsi la valeur accordée à son état de santé (41). Il est donc important dans le cadre d'études longitudinales utilisant des mesures de préférences individuelles de tenir compte de ce phénomène afin d'évaluer le changement observé plus justement.

FIGURE 4. MODELE THEORIQUE D'INTERACTION DU PHENOMENE DE RESPONSE-SHIFT DANS LA MESURE DE LA QUALITE DE VIE (58).



Spranger et Schwartz donne une conceptualisation théorique de ce phénomène (figure 3). Ce modèle théorique a cinq composantes principales: (1) un catalyseur, (2) les antécédents, (3) les mécanismes, (4) le changement de réponse et (5) la qualité de vie perçue. Le « catalyseur » désigne tout changement de l'état de santé de l'individu en rapport ou pas avec la mise en place d'un traitement. Les « antécédents » font référence aux

caractéristiques fixes de la personne, et les « mécanismes » au processus comportemental ou cognitive d'adaptation au changement induit par le « catalyseur ». Ainsi un changement de l'état de santé des individus peut à travers un mécanisme d'ajustement refléter des caractéristiques et capacités individuelles qui conduisent à ce phénomène de response-shift qui est une composante de la mesure des changements de préférences exprimés par les individus.

Il existe deux approches méthodologiques pour estimer le phénomène de response-shift : une mesure directe par le biais de mesures pretest-posttest auprès de l'individu, et une mesure basée sur une méthodologie statistique de mesure pretest ou posttest.

La première approche est la méthode du pretest-posttest désignée également comme la méthode du then-test. Il s'agit de faire compléter dans les conditions prévues par l'étude un questionnaire avant *l'intervention* et un questionnaire de suivi après l'intervention, et en complément de ce dernier le patient devra compléter le même questionnaire mais en se posant dans la situation initiale. Il s'agit donc d'une mesure supplémentaire à la fin du suivi de la santé perçue initialement par l'individu. Etant donné que cette mesure then-test est effectuée à distance de l'intervention et au même moment que le suivi du patient, on peut alors considérer que ce dernier utilise pour celle-ci les mêmes jugements de valeurs de mesure qu'au suivi. La mesure du response-shift se fait alors par la différence observée entre le pretest et le then-test (59).

La deuxième approche est basée sur une analyse purement statistique des données par les méthodes suivantes : 1) le modèle à équation structurelle, 2) le modèle d'analyse du « biais d'item » ou DIF (Differential Item Functioning), et 3) la classification par arbre de régression. Ces méthodes varient selon la possibilité de prise en compte du point de vue de l'individu ou celle d'agréger les données au niveau du groupe et nécessitent par leur méthodologie un nombre non négligeable de sujets.

2. PRESENTATION DES AXES DE RECHERCHE

2.1. LES ENJEUX DE LA MESURE DU CHANGEMENT D'UTILITE

Comme évoqué précédemment, l'objectif principal de l'analyse médico-économique en santé est de contribuer à la prise de décision en discriminant les différentes stratégies concurrentes par leur coût et leur conséquence clinique. Dans ce cadre, l'analyse coût-utilité a été développée pour permettre d'intégrer le point de vue de l'individu, exprimé par la mesure de ses préférences pour le résultat de chaque stratégie. Si l'utilité a l'avantage d'exprimer les préférences de l'individu, il serait d'autant plus intéressant de pouvoir interpréter le changement d'utilité observé pour un groupe de patients sur la base d'un seuil de changement acceptable.

En ce qui concerne l'interprétation du résultat de l'analyse coût-utilité, un seuil d'acceptabilité peut être défini selon le PIB/habitant d'un pays et utilisé par la suite pour la prise de décision au niveau national. Cependant les pouvoirs publics, selon la politique en vigueur dans le pays, doivent prendre en compte d'autres éléments décisionnels comme le nombre de patients concernés par la stratégie et l'investissement total que cela représente. Il y a en effet, dans l'analyse économique un enjeu d'équité et la nécessité d'être rationnel sur les choix de financement, puisque l'euro dépensé pour un motif particulier ne le sera pas pour un autre.

On constate aujourd'hui que l'analyse médico-économique est cloisonnée sur un modèle unique d'utilité évaluée entre deux groupes permettant la confrontation des coûts et conséquences de ces groupes. Cette comparaison des groupes ne laisse pas de place à une observation du changement réel induit chez l'individu. Introduire la notion de seuil dans l'interprétation du changement d'utilité peut apporter une information complémentaire sur la pertinence de la stratégie évaluée mais aussi dans l'interprétation du ratio coût-utilité (notamment dans une situation de dilemme : la nouvelle stratégie est plus chère et plus efficace, ou la nouvelle stratégie est moins chère et moins efficace). Ce seuil d'interprétation des préférences individuelles peut aussi être considéré comme un élément préalable à l'analyse coût-utilité. Au même titre que le concept de différence minimale importante (MID) s'applique pour les scores cliniques, il permettra d'évaluer si un changement pertinent est observé au niveau individuel.

Dans le cadre de ce travail, nous nous intéressons à la mesure d'utilité indirecte dérivée des instruments de mesure de générique de qualité de vie liée à la santé. Etant donné que la mesure de l'utilité indirecte nécessite une pondération par les préférences sociétales, cela revient à intégrer dans la mesure de l'état de santé de l'individu le regard que porte la société sur ce dernier. Ainsi dans cette optique, estimer un seuil d'interprétation pour le changement d'utilité, revient alors à définir un seuil socialement significatif du changement d'utilité.

2.2. OBJECTIFS DE LA THESE

Les objectifs de ce projet de thèse sont :

1. De déterminer dans un premier temps un seuil de changement significatif pour une utilité dérivée des instruments psychométriques ;
2. De déterminer dans quelle mesure ce seuil de changement de l'utilité est sensible à la présence de comorbidités ;
3. D'identifier l'existence d'un phénomène de Response-Shift dans la mesure du changement de l'utilité.

L'hypothèse principale est que la mesure d'état de santé par des instruments psychométriques, permet de définir une utilité qui conduit à la détermination d'un seuil de changement d'utilité socialement significatif. L'hypothèse secondaire est que ce changement d'utilité est sensible à la présence de comorbidités ou d'un phénomène de Response-Shift.

2.3. TRAVAUX REALISES DANS LE CADRE DE LA THESE

Afin de répondre aux objectifs fixés, nous nous sommes intéressés à deux contextes indépendants : la pathologie chronique à travers l'arthrose et le don d'organe chez donneur vivant de rein. Les travaux actuellement en cours au sein de l'EA 4360 APEMAC avec le CIC-EC inserm du service d'épidémiologie et évaluation cliniques du CHU de Nancy nous ont permis de saisir cette opportunité de positionner notre question de recherche dans ces deux contextes qui ont un point commun : l'intérêt de la prise en compte des préférences de l'individu dans une analyse médico-économique qui sera conduite ultérieurement.

D'une part, l'arthrose est une affection chronique et fréquente de l'articulation (60), avec un coût économique important (61;62). Elle se développe généralement avec l'âge et se conjugue très souvent avec d'autres maladies (63;64). Au vu du poids économique que représente cette pathologie (65;66), il est donc intéressant de documenter le changement d'utilité que l'on peut observer chez ces patients en prenant en compte la présence des comorbidités.

D'autre part, le don de rein par donneur vivant est aujourd'hui de plus en plus encouragé par les pouvoirs publics, mais il existe peu d'étude sur les préférences individuelles associées à cet état. L'insuffisance rénale terminale (IRT) occupe une place particulière en Santé Publique en France compte tenu de l'augmentation de son incidence (110 personnes par million d'habitants/an), de sa prévalence (700 personnes par million d'habitants) et des coûts des traitements de suppléance. En France, le coût moyen des traitements par patient est estimé à 55 K€ par an pour la dialyse, à 70 K€ pour la transplantation la première année et 7,5 K€ par an ultérieurement. Plus de 1,5 milliard d'euros sont ainsi consacrés annuellement au traitement de l'IRT, soit environ 1,5 % du budget de l'Assurance Maladie. La greffe a donc un impact financier majeur, puisque ses coûts sont amortis en un peu plus d'un an. Cependant, sur le plan individuel, le point de vue du donneur est très peu documenté. La plupart des études de qualité de vie menées à ce jour sur le don de rein portent sur les receveurs. Etant donné que la société accepte et encourage ce don, il peut être pertinent d'estimer le changement des préférences individuelles induit par celui-ci pour compléter ce point de vue sociétal.

Les travaux suivants ont été menés dans le cadre de cette thèse, dans l'ordre d'apparition dans ce présent document.

- Un premier travail situant le cadre de la thèse a été effectué sur le positionnement théorique de la mesure du changement intra-individuel d'utilité dérivée des instruments psychométriques, de son intérêt et son application, ainsi que les phénomènes pouvant affecter la mesure de ce changement.
- Les deux travaux suivants permettent de répondre aux objectifs 1 et 2 de la thèse :

- Pour le second travail nous nous sommes intéressés au lien existant entre la présence des comorbidités et la mesure de l'utilité. Pour cela nous avons analysé les données à l'inclusion des patients avec une arthrose symptomatique des membres inférieurs inclus dans la cohorte KHOALA. L'utilité était estimée par le questionnaire générique SF-6D et les comorbidités évaluées par l'indice de Groll (FCI).
- Le troisième travail a porté sur une analyse longitudinale des données de la cohorte KHOALA. Nous avons dans un premier temps déterminé un seuil de changement d'utilité à partir des données à l'inclusion et à trois ans de suivi dans la cohorte. Puis nous avons exploré le lien entre la présence de comorbidité et un changement pour les patients, positif ou négatif, de l'utilité comparé au seuil prédéfini.
- Le dernier travail répond aux objectifs 1 et 3 de la thèse :
 - Dans le quatrième et dernier travail, nous avons estimé un seuil de changement d'utilité obtenue par deux instruments de mesure générique, l'EQ-5D et le SF-6D, chez des donneurs vivants de rein. L'étude du phénomène de Response-Shift a été intégrée dans ce travail.

3. PREMIER TRAVAIL

Measuring change of utility for the individual: relevance and integration in cost-utility analysis

Hosseini Kossar, Gaujoux-Viala Cecile, Guillemain Francis

Soumis en révision

Measuring change of utility for the individual: relevance and integration in cost-utility analysis

Authors: Hosseini Kossar^{1,2,3}, Gaujoux-Viala Cecile⁴, Guillemin Francis^{1,2,3}

Affiliation

¹ INSERM, CIC-EC CIE6, Nancy, France

² CHU Nancy, Clinical Epidemiology and Evaluation, Nancy, France

³ Université de Lorraine, University Paris Descartes, EA 4360 Apemac, Nancy, France

⁴ EA 2415, Montpellier I University; Department of Rheumatology, Nîmes University Hospital, Nîmes, France

Contact information for the corresponding authors:

Francis Guillemin

Centre d'épidémiologie clinique
Inserm CIC-EC CIE6
Service d'épidémiologie et évaluation
cliniques
Hôpitaux de Brabois
Allée du Morvan
54505 Vandoeuvre-lès-Nancy
tel. : +33 83 85 21 68
fax : +33 83 85 12 05
e-mail: francis.guillemin@chu-nancy.fr

Kossar Hosseini

Centre d'épidémiologie clinique
Inserm CIC-EC CIE6
Service d'épidémiologie et évaluation
cliniques
Hôpitaux de Brabois
Allée du Morvan
54505 Vandoeuvre-lès-Nancy
tel. : +33 83 85 93 02
fax : +33 83 85 12 05
e-mail: k.hosseini@chu-nancy.fr

Contact information for the author to receive reprint requests:

Francis Guillemin

Centre d'épidémiologie clinique
Inserm CIC-EC CIE6
Service d'épidémiologie et évaluation
cliniques
Hôpitaux de Brabois
Allée du Morvan
54505 Vandoeuvre-lès-Nancy

e-mail: francis.guillemin@chu-nancy.fr

Key words: utility, utility change, health-related quality of life, preference-based instrument, mediceconomic analysis, minimal clinically important difference

Word Count: 1477

Abstract

The concept of utility is used in medicoeconomic analysis to explain choices and preferences for health states by patients and society. Indirect or multi-attribute utility is estimated from generic health-related quality of life (HRQoL) instruments. In clinical trials or observational studies, HRQoL results can be expressed as the minimal important difference (MID) that provides information about the meaning of HRQoL change for patients. The MID could be also calculated to interpret change in utility: a change within individuals or difference in utility score may be compared by a threshold such as the MID before being introduced in an economic analysis. Using indirect utility implies considering conditions that can affect HRQoL measurement such as multi-morbidity or the response-shift phenomenon. In contrast with direct utility, indirect utility measures are relatively simple, so decision makers can use them to assess utility change within individuals. The determination of a threshold for interpreting such utility change should be considered and used for comparisons between several strategies.

Background

The concept of utility was developed in health economics to compare the benefit of one health care strategy or programme over another, i.e. running comparison between groups (programmes). Cost-utility analysis refers to the comparison of differential costs of strategies related to differential utility between those strategies. The recent development of indirect utility measurement allows considering within patient changes. Therefore, the question arises whether within individual utility change adds on comparison between several strategies for decision maker.

Using utility change involves defining a new framework for interpretation of cost-utility analysis. It may be an alternative and a complementary element for use by the policy maker. In this way determining a threshold for meaningful change in utility should be considered.

Introduction

The concept of utility in health was developed from the Von Neumann and Morgenstern (VNM) utility theory [1] to explain patient choices and preferences for health states. Torrance [2] expressed utility as the “desirability or preference that individuals exhibit for the condition”. The unit value of health state utility ranges between 0, representing “death,” and 1 representing “perfect” health. Some instruments allow for a negative value labelled as a state “worse than death”.

Utility methodology in health-related quality of life (HRQOL) assessment allows for comparing different treatment strategies for one or more diseases with a single unit of measurement and may be used in medicoeconomic evaluations [3] that usually report a variation within group. Therefore, one wonders whether within-individual utility change can add to the comparison of several strategies by decision makers. How can one consider whether a variation within individuals, instead of or beyond a difference between groups, of 0.2 unit of utility on 0 to 1 scale is significant for society? As well, how can we consider the real change in patients’ preferences for health state in decision analysis?

Direct and indirect utility

Health utility can be assessed as direct or indirect utility.

Direct utility is usually ascertained by face-to-face interview. The most common methods are the standard gamble (SG) and the time trade-off (TTO). The utility by SG was calculated from the risks (of death usually) that the patient would take to improve the health state [4]. The TTO method, developed by Torrance et al. [5], is a method that determines how many healthy days (months or years) one is willing to trade against a better health state. Several authors discussed the difference between these 2 methods and the utility score obtained. In a cross-sectional study of patients with irritable bowel syndrome, Puhan et al. [6] concluded that the SG reflected HRQoL better than did the TTO. Others argued against this method, highlighting that expected utility is not descriptively valid and does not describe individual preferences well (risk aversion) [7]. They considered that the TTO may be a better measurement because it is more consistent with individual preferences (better reflection of individual preferences), easier to understand [8] and less riskfull. The utility measurement is sensitive to the method chosen and therefore can affect the result of cost-utility analysis and bias the resource allocation decision [7]. In general, the SG overestimates the utility, and the TTO both under- or overestimates it [9].

Indirect or multi-attribute utility [10] measures are estimated from generic HRQoL instruments to which one applies the weight assigned by the general population for each

particular health state. The most popular instruments are the Health Utility Index (HUI) [11], the EuroQol (EQ-5D) [12] and the Medical Outcomes Survey Short Form 6D (SF-6D) [13]. It is easy to derive a utility score from generic instruments because they can be easily and quickly used as self-administered questionnaire[14].

Many studies reported differences in utility assessment between instruments [15,16], but to date, consensus is lacking on the preferred instrument. As for the direct method, the choice of instrument may affect the cost-utility analysis and then the resource allocation decision [17].

Estimation of magnitude of utility change

In clinical trials or observational studies, results are usually reported as statistically significant changes in intervention. These results are insufficient for informed decision making by practitioners. The minimal important difference (MID), is expressed as “the smallest difference in score in the domain of interest which patients would perceive as beneficial and which would mandate, in the absence of troublesome side-effects and excessive cost, a change in the patient’s management” [18]. This concept is specifically used to capture the real change in HRQoL score within individual patients. Differences in score smaller than the threshold are considered unimportant [19]. Osoba et al. [20] argued that clinically meaningful differences are especially important because small numerical differences in mean HRQoL scores might give statistically significant results with large sample sizes, but statistical significance is not equivalent to clinical significance. More recently, the concept of patient acceptable symptom (PASS) was proposed as “the value beyond which patients consider themselves well” [21]. These two concepts are complementary: the former is considered as the change in score a patient needs to “feel better” and the latter as the change in score the patient needs to reach an acceptable symptom state to “feel good” [22].

For interpreting utility change, the concept of MID could be introduced. If variability in score of a utility contributed by a new strategy is significant compared to the MID (regarded as important change) it could be evaluated even before comparing it to another strategy. Walters and Brazier [23] reported on a comparison of the MID for utility score estimated from the SF-6D and the EQ-5D generic questionnaires. They suggested, like Drummond [24], that for resource allocation decisions, results must be described as cost-effectiveness differences, and a change alone may not be meaningful. However, for medico-economic evaluation of strategies, we must know the real value added by a new strategy for within-individual change and its acceptability by the society. In this way, a change or difference in utility score must be

compared to a threshold such as the MID or PASS before being introduced in an economic analysis.

A literature review was conducted to examine the use of MID in utility change measures. The terms “minimally important difference,” “change,” “EQ-5D,” “Euroqol,” “SF-36,” “SF-6D,” “health utility index” and “utility” were searched. In total, 107 articles met the search criteria, and 33 were selected after an initial reading of abstracts. Finally, we chose 3 articles that discussed the notion of MID for utility change.

Walters and Brazier [24] first explored the relation between MID and utility by applying the MID to the SF-6D with data from 7 longitudinal studies. Using an anchor-based method, the authors found a mean MID of 0.033 for the SF-6D (range 0.010–0.048). They suggested that this result was meaningful for examining within-patient changes. In another review of 11 studies [23] comparing the MID for health-state utility measures by the EQ-5D and SF-6D, the authors found that the MID for the 2 measures were not equal: the mean MID for the EQ-5D was 0.074 (range -0.011–0.140) and for the SF-6D was 0.041 (range 0.011–0.097).

Pickard et al. [25], in a cross-sectional study of cancer patients, reported a mean MID of 0.09 for the EQ-5D utility score. The results of MID estimation depend on the method chosen; each method (anchor- or distribution-based) produced a different MID value [26].

Conditions likely affecting indirect utility

Using indirect utility implies a consideration of conditions such as multi-morbidity and response-shift (RS) phenomenon that can affect the measurement of HRQoL and subsequent utility change assessment.

Originally, Feinstein et al. [27] described comorbidity as “any distinct additional entity that has existed or may occur during the clinical course of patient who has the index disease under study”. In a systematic review of the effect of multi-morbidity on HRQoL, Fortin et al. [28] concluded an inverse relationship between these 2 conditions and showed that multimorbidity could affect all domains of HRQoL. Most studies considering the effect of multimorbidity on QoL were cross-sectional [28-30], but the effect may vary over time, and this type of study cannot capture the real effect on HRQoL nor predict the direction of change over time.

Another issue to consider when using utility derived from HRQoL is the RS phenomenon. As defined by Sprangers and Schwartz [31], the RS refers to change in an individual’s health that can alter the individual’s standards, values or conceptualization. RS can occur in any situation involving self-reported outcomes. It may bias the results of studies reporting HRQoL because people with chronic diseases often continue to report a high HRQoL over time, probably because of adaptation to their illness.

Conclusion

The use of indirect utility is increasingly widespread; problems induced by measurement issues specific to the characteristics of psychometric instruments should not be neglected. As well, an evaluation of utility change derived from HRQoL must consider other conditions that can affect HRQoL, such as multi-morbidity or RS

Defining a new framework of interpretation in cost-utility analysis may be an alternative and a complementary element for use by the policy maker. It is a standard to consider the difference in utility between strategies in economic analysis in health. A utility change is usually investigated in a longitudinal study for 2 or more measurement times for subjects in different groups, mostly with chronic diseases. In this context, determining a threshold for interpretation should be reconsidered. Issues to consider are an appropriate threshold for meaningful acceptable utility change within individuals and whether the process can simplify the analysis and ratio interpretation with use of QALYs in decision making. Interpretation of within group difference to comparing strategies may be conditioned to an existing meaningful within individual change in utility. Given the lack of accurate knowledge and guides for interpreting the use of indirect utility, defining an MID for interpreting cost-utility analysis seems a challenge.

Competing interest: The authors declare that they have no competing interests.

Acknowledgement: We thank Agnes Peignier, Document manager, for her contribution to the literature research.

Authors' contributions: KH, FG - Conceived and designed the work. KH, FG analyzed and interpreted the results of literature research. KH, FG wrote the manuscript. CG made important suggestions for editing the manuscript. All authors approved the final version.

Reference List

1. Neumann Jv, Morgenstern O: *Theory of Games and Economic Behavior*, sec ed. 1947 edn. Princeton University Press; 1944.
2. Torrance GW: **Utility approach to measuring health-related quality of life.** *J Chronic Dis* 1987, **40**: 593-603.
3. Beresniak A, Russell AS, Haraoui B, Bessette L, Bombardier C, Duru G: **Advantages and limitations of utility assessment methods in rheumatoid arthritis.** *J Rheumatol* 2007, **34**: 2193-2200.
4. Bakker C, van der LS: **Health related utility measurement: an introduction.** *J Rheumatol* 1995, **22**: 1197-1199.
5. Torrance GW, Thomas WH, Sackett DL: **A utility maximization model for evaluation of health care programs.** *Health Serv Res* 1972, **7**: 118-133.
6. Puhan MA, Schunemann HJ, Wong E, Griffith L, Guyatt GH: **The standard gamble showed better construct validity than the time trade-off.** *J Clin Epidemiol* 2007, **60**: 1029-1033.
7. Bleichrodt H: **A new explanation for the difference between time trade-off utilities and.** *Health Econ* 2002, **11**: 447-456.
8. Hornberger JC, Redelmeier DA, Petersen J: **Variability among methods to assess patients' well-being and consequent effect on a cost-effectiveness analysis.** *J Clin Epidemiol* 1992, **45**: 505-512.
9. Dolan P, Gudex C, Kind P, Williams A: **Valuing health states: a comparison of methods.** *J Health Econ* 1996, **15**: 209-231.
10. Petrou S, Hockley C: **An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population.** *Health Econ* 2005, **14**: 1169-1189.
11. Feeny D, Furlong W, Boyle M, Torrance GW: **Multi-attribute health status classification systems. Health Utilities Index.** *Pharmacoeconomics* 1995, **7**: 490-502.

12. **EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group.** *Health Policy* 1990, **16**: 199-208.
13. Ware JE, Jr., Sherbourne CD: **The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.** *Med Care* 1992, **30**: 473-483.
14. Kopec JA, Willison KD: **A comparative review of four preference-weighted measures of health-related quality of life.** *J Clin Epidemiol* 2003, **56**: 317-325.
15. Gaujoux-Viala C, Rat AC, Guillemin F, Flipo RM, Fardellone P, Bourgeois P *et al.*: **Comparison of the EQ-5D and the SF-6D Utility Measures in 813 Patients with Early Arthritis: Results from the ESPOIR Cohort.** *J Rheumatol* 2011, **38**: 1576-1584.
16. Marra CA, Esdaile JM, Guh D, Kopec JA, Brazier JE, Koehler BE *et al.*: **A comparison of four indirect methods of assessing utility values in rheumatoid arthritis.** *Med Care* 2004, **42**: 1125-1131.
17. Marra CA, Marion SA, Guh DP, Najafzadeh M, Wolfe F, Esdaile JM *et al.*: **Not all "quality-adjusted life years" are equal.** *J Clin Epidemiol* 2007, **60**: 616-624.
18. Jaeschke R, Singer J, Guyatt GH: **Measurement of health status. Ascertaining the minimal clinically important difference.** *Control Clin Trials* 1989, **10**: 407-415.
19. Crosby RD, Kolotkin RL, Williams GR: **Defining clinically meaningful change in health-related quality of life.** *J Clin Epidemiol* 2003, **56**: 395-407.
20. Osoba D, Rodrigues G, Myles J, Zee B, Pater J: **Interpreting the significance of changes in health-related quality-of-life scores.** *J Clin Oncol* 1998, **16**: 139-144.
21. Tubach F, Wells GA, Ravaud P, Dougados M: **Minimal clinically important difference, low disease activity state, and patient acceptable symptom state: methodological issues.** *J Rheumatol* 2005, **32**: 2025-2029.
22. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N *et al.*: **Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state.** *Ann Rheum Dis* 2005, **64**: 34-37.

23. Walters SJ, Brazier JE: **Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D.** *Qual Life Res* 2005, **14**: 1523-1532.
24. Walters SJ, Brazier JE: **What is the relationship between the minimally important difference and health.** *Health Qual Life Outcomes* 2003, **1**:4.: 4.
25. Pickard AS, Neary MP, Cella D: **Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer.** *Health Qual Life Outcomes* 2007, **5**: 70.
26. Copay AG, Subach BR, Glassman SD, Polly DW, Jr., Schuler TC: **Understanding the minimum clinically important difference: a review of concepts and methods.** *Spine J* 2007, **7**: 541-546.
27. Feinstein AR: **The pre-therapeutic classification of co-morbidity in chronic disease.** *Journal of Chronic Diseases* 1970, **23**: 455-468.
28. Fortin M, Lapointe L, Hudon C, Vanasse A, Ntetu AL, Maltais D: **Multimorbidity and quality of life in primary care: a systematic review.** *Health Qual Life Outcomes* 2004, **2**: 51.
29. Taylor AW, Price K, Gill TK, Adams R, Pilkington R, Carrangis N *et al.*: **Multimorbidity - not just an older person's issue. Results from an Australian biomedical study.** *BMC Public Health* 2010, **10**: 718.
30. Loza E, Abasolo L, Jover JA, Carmona L: **Burden of disease across chronic diseases: a health survey that measured prevalence, function, and quality of life.** *J Rheumatol* 2008, **35**: 159-165.
31. Sprangers MA, Schwartz CE: **The challenge of response shift for quality-of-life-based clinical oncology research.** *Ann Oncol* 1999, **10**: 747-749.

4. MESURE D'UTILITE ET COMORBIDITES

L'arthrose qui est une affection chronique et fréquente de l'articulation (5), et qui se développe généralement avec l'âge et se conjugue très souvent avec d'autres maladies (8-10).

Nous nous sommes donc intéressés pour ce deuxième travail à estimer l'effet des comorbidités sur la mesure de l'utilité indirecte à travers une analyse transversale chez des individus atteints d'une arthrose des membres inférieures.

Nous avons pour cette analyse utilisée les données à l'inclusion de la cohorte KHOALA (67) (Knee & Hip OsteoArthritis Long-term Assessment), constituée d'avril 2007 à mars 2009, représentative, multirégionale (Amiens, Brest, Nancy, Nice, Paris, Toulouse), de 878 sujets ayant une arthrose symptomatique de genou et/ou de hanche, tous recrutés en population générale. Les patients de la cohorte ont été recrutés à partir d'une étude de prévalence 3000 arthroses (68). Cette cohorte est donc composée de cas prévalent d'arthrose de hanche et/ou de genou, suivi chaque année.

L'utilité a été mesurée par le questionnaire SF-6D. La présence des comorbidités à l'inclusion a été évaluée par l'indice de Groll (Functional Comorbidity Index), composé d'une liste de 18 pathologies associée à une altération de la dimension physique de la qualité de vie simplifiée en 9 pathologies dans le cadre de cette étude. A partir de cet indice, le nombre de comorbidités pour chaque patient a été calculé. Etant donné que tous les patients de la cohorte avaient une arthrose des membres inférieurs, cet item a été retiré de l'indice pour cette étude.

Et enfin nous avons pris en compte dans le cadre de cette analyse la sévérité symptomatique de la maladie estimée par le score de WOMAC, allant de 0 à 100 (plus grande limitation). Cet index se définit par un score global et 3 dimensions distinctes: limitation fonctionnelle, douleur et raideur.

Pour répondre à l'objectif de l'étude, nous avons mené deux analyses par régression linéaire multiple. Une première incluant la somme de comorbidités par individus allant de 0 à 9, et la seconde en prenant comme autant de variables indépendantes les 9 catégories de comorbidités. Cette différenciation des deux analyses nous a permis de nous affranchir du

problème de multicollinéarité qui subsistait si les deux variables étaient incluses dans le même modèle. D'autre part et afin de simplifier le modèle nous avons utilisé uniquement le score fonctionnel de WOMAC, qui s'est avéré le plus corrélé au score d'utilité.

Au préalable, une analyse bivariée a été menée avec les autres covariables, et celle significative au seuil de $p < 0.10$ ont été inclus dans l'analyse multivariée.

Cette première analyse bivariée a montré qu'il n'y avait pas d'effet de l'âge ou de la localisation de l'arthrose (hanche, genou) sur l'estimation du score d'utilité. Le nombre de comorbidité ainsi que toutes les comorbidités, excepté l'ostéoporose, avaient un effet délétère significatif sur le score d'utilité.

Les deux analyses multivariées ont mis en évidence l'importance du degré de sévérité symptomatique de la maladie. Le score de fonction WOMAC expliquait la plus grande part de variance du score d'utilité, rendant ainsi l'effet des comorbidités marginal.

Ainsi, cette analyse transversale a montré un effet délétère des comorbidités discutable comparé à la pathologie de l'arthrose qui visiblement amène les patients à focaliser plus leur attention sur celle-ci au moment de leur participation à l'étude. Bien que les patients atteints d'arthrose aient une plus grande prévalence de comorbidité, il faut retenir comme élément principal de leur prise en charge le retentissement fonctionnel de la pathologie elle-même.

4.1. DEUXIEME TRAVAIL

Impact of comorbidities on measuring indirect utility by the Medical Outcomes Study Short Form 6D in lower-limb osteoarthritis

Hosseini Kossar, Gaujoux-Viala Cecile, Coste Joel, Pouchot Jacques, Fautrel Bruno, Rat Anne-Christine, Guillemin Francis

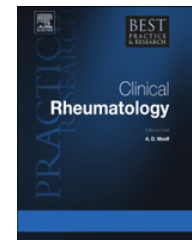
Best Pract Res Clin Rheumatol 2012 Oct; 26(5):627-35



ELSEVIER

Contents lists available at [SciVerse ScienceDirect](#)

Best Practice & Research Clinical Rheumatology

journal homepage: www.elsevierhealth.com/berh

5

Impact of co-morbidities on measuring indirect utility by the Medical Outcomes Study Short Form 6D in lower-limb osteoarthritis

Kossar Hosseini^a, Cécile Gaujoux-Viala^{a,b}, Joel Coste^a, Jacques Pouchot^a, Bruno Fautrel^b, Anne-Christine Rat^a, Francis Guillemin^{a,*}

^a Université de Lorraine, Paris Descartes University, APEMAC, EA 4360, F-54 000, France

^b Paris 6 – Pierre et Marie Curie University, Department of Rheumatology, Pitié-Salpêtrière Hospital, Paris, France

Keywords:

Co-morbidity

Utility

SF-6D

Health related quality of life

Osteoarthritis

Introduction: Co-morbidities can influence generic measurement of health indirect utility. We investigated their impact to assess indirect utility with the Medical Outcomes Study Short Form 6D (SF-6D) in patients with osteoarthritis (OA).

Methods: In patients with hip and knee OA from the Knee and Hip Osteo-Arthritis Long-term Assessment (KHOALA) study, co-morbidities were assessed by the Functional Co-morbidity Index. Multivariate linear regressions were used to determine predictors of utility score.

Results: For the 878 patients included, the mean (standard deviation (SD)) utility score for 808 patients was 0.66 (11; range 0.32–1.00) and mean number of co-morbidities 2.05 (1.58). Number of co-morbidities (beta = -0.30; $p = 0.002$), psychiatric disease (beta = -0.043; $p < 0.0001$) and degenerative disc disease (beta = -0.014; $p = 0.018$) were predictors of low utility score. The WOMAC functional score had a higher significant effect (beta = -0.003; $p < 0.0001$) and explained a higher percentage of the model variance.

Discussion: Compared to greater negative effect of functional severity of OA, co-morbidities have a negative but relatively marginal impact on indirect utility score. This suggests that, clinically, considering the functional severity of OA remains a first priority.

© 2012 Elsevier Ltd. All rights reserved.

* Corresponding author. Tel.: +33 0 383 852165.

E-mail address: francis.guillemin@chu-nancy.fr (F. Guillemin).

Introduction

The concept of utility in health was developed by economists to explain individual choices and preferences for health states [1]. The health-state utility score ranges from 0, for 'death', to 1, 'perfect' health. Utility methodology in assessing health-related quality of life (HRQoL) allows for comparing different treatment strategies for one or more diseases by a single unit of measurement and may be used in medico-economic evaluations for allocating resources [2].

There are two approaches for estimating utility in health: direct or indirect. Direct measurement usually involves face-to-face interview methods such as the standard gamble (SG) [3] and the time trade-off methods (TTO) [4]. The utility by SG is calculated from the risks (of death usually) that the patient would take to improve the health state [5]. The TTO method, developed by Torrance et al. [4], is a method that determines how many healthy days (months or years) one is willing to trade against a better health state. Indirect measurement, also called multi-attribute utility [6], involves estimation by generic HRQoL instruments with weighting for the general population for each health state. The complexity of the direct method and its implementation have led to the increasing use of indirect methods because they can be easily and quickly applied. The most popular instruments are the Health Utility Index (HUI) [7], the EuroQol (EQ-5D) [8] and the Medical Outcomes Survey Short Form 6D (SF-6D) [9]. Evaluation of the health state differs in two ways. First, the method used to determine the weight assigned in the general population to health state varies between instruments: health state values for the EQ-5D are derived with the TTO method while values for the SF-6D and HUI are obtained from the SG. Second, the number of dimensions and levels (variation in descriptive system) are lowest in the EQ-5D, compared to the SF-6D [10], and highest in the HUI. The SF-6D, obtained from the SF-36, is likely most popular.

Co-morbidities refer to chronic co-occurring disorders in the same individual. The concept was described by Feinstein et al. [11] as "any distinct additional entity that has existed or may occur during the clinical course of patient who has the index disease under study". Co-morbidities are inversely and negatively correlated with HRQoL [12]. As indirect utility measurement involves HRQoL, it should consider co-morbidities, which can affect the measurement and subsequent utility assessment. The treatment of a related disorder can be influenced by a chronic condition and therefore disturb the management of disease [13].

Osteoarthritis (OA) is a chronic and frequent degenerative joint disease [14] influencing QoL and incurs important economic burden [15,16]. As it develops with ageing, patients with OA have a risk of developing co-morbidities [17–19]. QoL and utility are associated linearly in OA as in other illnesses – with decreased QoL, utility decreases and vice versa – and as for QoL, the range of utility values varies by co-morbidity [20]. Therefore, QoL with OA can be influenced by co-morbidities, which significantly affects utility measurement by a multi-attribute instrument. Whether the relation of utility and co-morbidity is of the same magnitude as with QoL and co-morbidity is of interest.

Multi-attribute utility involves quality-adjusted life years resulting from utility weighting of life expectancy [21] and is generally used in cost-utility analysis to compare the differential costs of strategies related to differential utility among those strategies. In a literature review conducted using the terms 'comorbidity' and 'multimorbidity' with 'utility', we found only one study that analysed the impact of co-morbidities on utility scores for patients with OA [22]. Studies have focussed on the QoL of patients with OA without taking into account the impact of multimorbidities or co-morbidities in utility measurement [15,23–27]. The possible influence of co-morbidities on the condition of individuals and the cost of illness may affect utility measurement and cost-utility analysis.

We aimed to determine in a cross-sectional study the effect of co-morbidities in utility assessment by the Medical Outcomes Study SF-6D of OA patients.

Methods

Data collection

We used data for patients from the Knee and Hip Osteo-Arthritis Long-term Assessment (KHOALA) cohort study [28] including 878 patients recruited from the 3000 OA survey undertaken in France from

2007 to 2009 [29]. The 3000 OA study was a two-phase population-based survey conducted in six regions of France. A random sample of households was investigated with the next-birthday method in each household for the detection of cases by telephone-administered questionnaire, and further clinical examination of positive cases allowed confirmation of diagnosis. Data were for patients of both sexes; 40–75 years old; with uni or bilateral, symptomatic hip and/or knee OA; and with clinical diagnosis confirmed and fulfilling American College of Rheumatology criteria for knee [30] and hip [31] OA. Patients are followed annually with a clinical visit or by mailed self-reporting questionnaire. We used data collected at baseline. Sociodemographic characteristics (age, sex, education level, occupation and marital status) were collected, and co-morbidity was assessed by the Functional Comorbidity Index (FCI) [32], which was specifically developed to assess the impact of co-morbidities on the physical dimension of QoL. From this index, we calculated the number of co-morbidities for each patient. To simplify the analysis, we grouped the 18 conditions of this index into nine categories (Table 1). Given that all patients in this cohort have OA, we removed the item OA from the index. Limitation in activities and pain was measured by the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [33], with high scores reflecting increased limitation in activity. Functional, pain and stiffness WOMAC scores range from 0 to 100.

Utility was assessed by the SF-6D from the MOS 36-item Short Form Health Survey (SF-36) [9]. The six dimensions are physical functioning, role limitation, social functioning, pain, mental health and vitality, with 4 or 6 levels for each dimension. A scoring algorithm was developed with the SG method from a representative sample of the population in the United Kingdom for scores ranging from 0.29 to 1.00.

Statistical analysis

The sociodemographic and clinical characteristics of patients and SF-6D scores are described with means (SD) and numbers (percentages). We determined the impact of co-morbidities on utility values by bivariate analysis. Simple linear regression model was used for quantitative variables and analysis of variance for qualitative variables. Variables with $p \leq 0.10$ on bivariate analysis were included in multivariate linear stepwise regression models. Because of multicollinearity between the number of co-morbidity expressed by the FCI index and the co-morbidity by categories, two separate linear regression models were fitted. The *R*-square of each regression model was used to compare the models to determine the adequacy between the two models. A $p < 0.05$ was considered statistically significant. Analyses involved use of Statistical Analysis Software (SAS) 9.1 (SAS Inst, Cary, NC, USA).

Table 1

The 18 Functional Comorbidity Index items grouped into 9 items.

	Functional Comorbidity Index	Grouping
1	Arthritis (rheumatoid and osteoarthritis)	Not retained
2	Osteoporosis	Osteoporosis
3	Asthma	Respiratory conditions
4	COPD, ARDS, or emphysema	
5	Angina	Cardiac disease
6	Congestive heart failure (or heart disease)	
7	Heart attack (myocardial infarction)	
10	Peripheral vascular disease	
8	Neurological disease (such as multiple sclerosis or Parkinson's disease)	Neurologic disease
9	Stroke or transient ischaemic attack	
11	Diabetes type I and II	Diabetes and obesity
18	Obesity or body mass index > 30	
12	Upper gastrointestinal disease (ulcer, hernia, reflux)	Upper gastrointestinal disease
13	Depression	Psychiatric disease
14	Anxiety or panic disorders	
15	Visual impairment (such as cataracts, glaucoma, macular degeneration)	Visual and hearing impairment
16	Hearing impairment (very hard of hearing, even with hearing aids)	
17	Degenerative disc disease	Degenerative disc disease

COPD, chronic obstructive pulmonary disease; ARDS, acquired respiratory distress syndrome.

Results

The sociodemographic and clinical characteristics of the 878 patients are in Table 2. The mean utility score from the SF-6D for 808 patients was 0.66 (± 0.11 ; range 0.32–1.00). The mean number of co-morbidities was 2.05 (± 1.58) (excluding OA); the most common were degenerative disc disease (52.5%), diabetes and obesity (41.5%), upper gastrointestinal disease (22.2%) and psychiatric disease (22%).

On bivariate analysis, women had lower mean utility scores than men (0.64 vs. 0.69, $p < 0.0001$). The utility score was negatively affected by number of co-morbidities ($\beta = -0.02$, $p < 0.0001$) and

Table 2
Characteristics of patients with osteoarthritis included in the Knee and Hip Osteo-Arthritis Long-term Assessment (KHOALA) cohort by affected joint ($n = 878$).

	Osteoarthritis joint					
	Hip <i>n</i> = 222		Knee <i>n</i> = 607		Hip and knee <i>n</i> = 49	
Age, mean (SD)	61.31	(8.76)	62.07	(8.47)	64.67	(7.99)
Sex						
Male	74	(33.3)	178	(29.3)	17	(34.7)
Female	148	(66.7)	429	(70.7)	32	(65.3)
Education						
Primary	45	(20.3)	144	(23.9)	11	(22.4)
Secondary	110	(49.6)	322	(53.0)	28	(57.2)
University	67	(30.2)	139	(23.1)	10	(20.4)
Marital status						
Married or living together	161	(72.6)	404	(67.0)	36	(73.4)
Occupation						
Farmer and worker	39	(17.7)	95	(15.8)	8	(16.3)
Manager	21	(9.5)	31	(5.2)	1	(2.0)
Executive	34	(15.5)	86	(14.3)	7	(14.3)
Employee and craftman	117	(53.2)	333	(55.5)	28	(57.2)
Not working	9	(4.1)	55	(9.2)	5	(10.2)
Body mass index						
18.5–24.9	81	(37.3)	112	(18.6)	11	(23.4)
25–29.9	92	(42.4)	217	(36.0)	17	(36.2)
>30	44	(20.3)	274	(45.4)	19	(40.4)
No. of comorbidities, mean (SD)	2.00	(1.52)	2.07	(1.62)	2.06	(1.39)
Functional Comorbidity Index						
Osteoporosis	15	(6.8)	42	(6.9)	5	(10.2)
Degenerative disc disease	142	(64.0)	289	(47.6)	30	(61.2)
Respiratory conditions	16	(7.2)	63	(10.4)	6	(12.2)
Cardiologic disease	33	(14.9)	100	(16.5)	6	(12.2)
Neurologic disease	9	(4.1)	20	(3.3)	3	(6.1)
Hearing and visual impairment	42	(18.9)	84	(13.8)	9	(18.4)
Psychiatric disease	47	(21.2)	139	(22.9)	7	(14.3)
Diabetes and obesity	51	(23.0)	293	(48.3)	20	(40.8)
Upper gastrointestinal disease	60	(27.0)	126	(20.8)	9	(18.4)
SF-6D utility score (<i>N</i>)	206		557		45	
Mean (SD)	0.66	(0.11)	0.66	(0.11)	0.63	(0.11)
min–median–max	0.37–0.64–1.00		0.39–0.66–0.96		0.32–0.61–0.86	
WOMAC score [0–100]						
Function (<i>N</i>)	213		584		47	
Mean (SD)	30.02	(18.43)	30.26	(20.05)	36.22	(20.33)
min–median–max	0–28.13–77.94		0–29.41–95.59		1.47–38.24–82.35	
Pain (<i>N</i>)	214		588		47	
Mean (SD)	31.65	(18.88)	32.80	(19.33)	37.77	(19.33)
min–median–max	0–30.00–95.00		0–30.00 90.00		5.00–40.00–85.00	
Stiffness (<i>N</i>)	214		587		47	
Mean (SD)	42.58	(20.90)	40.91	(22.57)	43.35	(21.31)
min–median–max	0–50.00–100		0–37.50–100		12.50–50.00–87.50	

Data are no. (%) unless indicated.

the WOMAC function score (beta = -0.003, $p < 0.0001$). The location of OA was not related to utility score ($p = 0.154$). The utility score was higher for patients with a higher (university > 3 years) than lower level of education ($p < 0.0001$) and was higher for patients with an executive occupation than other work or not working ($p < 0.0001$).

Presence of all co-morbidity categories, except for osteoporosis, was associated with a low utility score ($p < 0.05$). The mean utility score was lower for patients with psychiatric disease (mean = 0.59, $p < 0.0001$), neurologic disease (mean = 0.62, $p = 0.038$) and pneumologic disease (mean = 0.62, $p = 0.01$).

In both multivariate regression models (Table 3), profession significantly predicted utility score ($p = 0.001$): utility scores were higher for patients with a manager or executive occupation than those

Table 3

Factors affecting the Medical Outcomes Study Short Form 6D (SF-6D) utility score for patients with hip or knee osteoarthritis by 2 multivariate linear regression models.

	Bivariate analysis		Model 1 ^a		Model 2 ^b	
	Mean/ r^d	p value	Multivariate analysis with no. of comorbidities		Multivariate analysis with comorbidities by categories	
			$n = 790, R^2 = 0.42^c$		$n = 790, R^2 = 0.44^c$	
			Beta	p value	Beta	p value
<i>Intercept of model</i>			0.759	<0.0001	0.811	<0.0001
Age	-0.001	0.108				
Sex		<0.0001				
Male	0.695					
Female	0.642					
Married or living together		0.255				
Yes	0.662					
No	0.652					
Education		<0.0001				
Primary	0.633					
Secondary	0.656					
University	0.693					
Profession		<0.0001		0.001		0.001
Farmer and worker	0.657		0.026	0.047	0.024	0.067
Manager	0.668		0.038	0.019	0.035	0.027
Executive	0.706		0.046	0.001	0.050	0.000
Employee and craftman	0.652		0.017	0.140	0.020	0.084
Not working	0.618		0.000		0.000	
Location of osteoarthritis		0.154				
Hip	0.655					
Knee	0.662					
Hip and knee	0.630					
WOMAC Function Score [0–100]	-0.003	<0.0001	-0.003	<0.0001	-0.003	<0.0001
No. of comorbidities [0–9]	-0.02	<0.0001	-0.030	0.002		
<i>Functional Comorbidity Index</i>						
Osteoporosis		0.105				
Yes	0.635					
No	0.660					
Upper gastrointestinal disease		0.006				
Yes	0.639					
No	0.665					
Degenerative disc disease		<0.0001				0.018
Yes	0.636				-0.014	
No	0.684				0.000	
Respiratory conditions		0.011				
Yes	0.628					
No	0.662					
Cardiologic disease		0.026				
Yes	0.639					
No	0.662					

(continued on next page)

Table 3 (continued)

	Bivariate analysis		Model 1 ^a		Model 2 ^b	
	Mean/ <i>r</i> ^d	<i>p</i> value	Multivariate analysis with no. of comorbidities		Multivariate analysis with comorbidities by categories	
			<i>n</i> = 790, <i>R</i> ² = 0.42 ^c		<i>n</i> = 790, <i>R</i> ² = 0.44 ^c	
			Beta	<i>p</i> value	Beta	<i>p</i> value
Neurologic disease		0.038				
Yes	0.618					
No	0.660					
Hearing and visual impairment		0.006				
Yes	0.634					
No	0.663					
Psychiatric disease		<0.0001				<0.0001
Yes	0.595				−0.043	
No	0.676				0.000	
Diabetes and obesity		<0.0001				
Yes	0.632					
No	0.678					

^a Model 1: no. of comorbidities measured by the Functional Comorbidity Index; variables with $p = 0.10$ on bivariate analysis entered stepwise into the model.

^b Model 2: comorbidities by categories established by the Functional Comorbidity Index; variables with $p = 0.10$ on bivariate analysis entered stepwise into the model.

^c *R*-square of linear regression model.

^d Mean utility score for qualitative variables and coefficient of simple linear regression for quantitative variables.

not working. Worsened function (increased WOMAC function score) significantly reduced the utility score (beta = -0.003 , $p < 0.0001$).

In the first multivariate model, with number of co-morbidities, for each additional co-morbidity (range 0–9), the mean utility score decreased 0.03 point (beta = -0.03 , $p < 0.0001$).

In the second model, including co-morbidities by categories, age and sex did not predict utility score. Among co-morbidities, only psychiatric disease (beta = -0.043 , $p < 0.0001$) and degenerative disc disease (beta = -0.014 , $p = 0.018$) predicted low utility score.

In the first multivariate model, the number of co-morbidities explained 2% of the variance in utility score (partial *R*-square = 0.02) and psychiatric and degenerative disc diseases explained 2% (partial *R*-square = 0.025) and 0.7% (partial *R*-square = 0.007), respectively, of the variance in utility score, whereas the WOMAC function score explained 38% of the variance in both models (partial *R*-square = 0.38).

Discussion

The effect on utility assessment of number and type of co-morbidities for patients with hip or knee OA was examined. We performed two regression models to examine the impact of co-morbidities on utility scores. OA patients in the KHOALA study had two co-morbidities on average (range 0–9), and the utility score decreased (0.01 points) with additional co-morbidities. Utility estimated by the HRQoL generic instrument SF-6D revealed the negative impact of psychiatric and degenerative disc diseases. The effect of co-morbidities on utility score was independent of the OA-affected joint. The bivariate analysis results confirmed our hypothesis of the impact on indirect utility of conditions that can affect measurement of HRQoL. However, both multivariate regression models revealed that co-morbidities explained a small proportion of the variance in utility score (mean 2%). The WOMAC function score, which reflects the impact on physical functioning in OA, explained most of the variance.

Psychological diseases had more of an impact than other co-morbidities on utility score measured by the SF-6D. Depression and psychological disorders adversely affect pain and other physical dimensions of patients with OA [34,35], and can be expressed in utility measures.

Only one other study investigated the effect of co-morbidities on utility scores in patients with OA [22] with the generic 15-D multi-attribute utility scale [36]. The study found no differences by type of

co-morbidity, as reported by a medical doctor, among nine groups of co-morbidities. Utility scores were worse for patients with than without co-morbidities. Another study involving the utility index from the European Quality of Life instrument (EQ-5D) and a sum of possible co-occurring conditions in patients with musculoskeletal conditions (classified into four diagnostic groups) found a negative effect of co-morbidities on utility score, but the authors also focussed on related QoL estimated from the SF-36 [37].

To explain the weight of various disease conditions, we differentiated the co-morbidities by the FCI. Given the importance of physical function for patients with OA, we used the FCI index to determine co-morbidity in our study, since it may be the most appropriate way to measure co-morbidities that affect physical function [38]. Other indices such as the Charlson co-morbidity index [39] focus on diagnoses that predict mortality. The primary objective of these instruments is to determine the risk of complications and death, and their use in predicting mortality is more suitable for studies with a survival outcome.

van Dijk et al. [27] have highlighted the negative association of number of co-morbidities and limitation in activities (estimated by the WOMAC physical function and SF-36 physical dimension) with use of the Cumulative Illness Rating Scale (CIRS) [40] in older adults with hip or knee OA. One study, involving the SF-36 and the Charlson co-morbidity index, found a negative significant effect of co-morbidities on the HRQoL in patients undergoing surgery of the shoulder [41].

We established two independent multivariate models, including in the first the number of co-morbidities, and in the second, co-morbidities by categories, to understand how each model can affect utility measurement. The models were similar in explaining the variance in utility score. We found a negative effect with the model containing number of co-morbidities, but with the model by categories of co-morbidities, only two co-morbidities had a significant negative impact. Therefore, we conclude that in a clinical perspective the presence of an additional co-morbidity must be taken into account for better clinical management of OA.

We used data from the first French national population-based cohort of symptomatic patients with hip or knee OA. Patients were recruited independent of any medical consultation, which allowed for a representative sample of cases of OA from the general population. Patients may be more focussed on the current disease and motivated to obtain care, and therefore the utility score may be more influenced by OA than other disorders. Investigating the utility of OA if attention was focussed on another disease would be interesting. In our sample, the impact of disease measured by the WOMAC score in the utility was more relevant than co-morbidity. The degree of functional deterioration reflected the weight of the present illness for patients compared to other co-occurring diseases. Given the aim of our study, the relationship between WOMAC and co-morbidity was not explored. However, studies have shown that other musculoskeletal pain [42] or number of co-morbidities [43] may also affect physical function and thus the WOMAC score.

Patients with OA have a high prevalence of co-morbidities [17–45], which can be predictors of use of health services and medical costs [18,44,46].

In summary, we showed the negative impact of co-morbidities on utility scores by use of a multi-attribute instrument in a cross-sectional study of OA patients. The effect may be similar by number of co-morbidities or by categories. However, this effect seems marginal compared to the functional impact of OA measured by the WOMAC score. This suggests that, clinically, considering the functional severity of OA remains a first priority.

Summary

Data from a representative sample of OA patients were used, in a cross-sectional study, to understand how the co-morbidities can impact the multi-attribute or indirect utility. The number and the categories of co-morbidities were defined by the FCI index, and included in two separate linear regression models. The severity of the OA disease reflected by the WOMAC function score was strongly associated with the utility measurement while a few co-morbidities showed only marginal association. This suggests that OA patients were more focussed on the present illness compared to other co-occurring diseases such as psychiatric or degenerative disc disease. Although patients with OA have a high prevalence of co-morbidities, our results imply that a first clinical priority remains to consider the functional severity of OA.

Practice points

- Patients with symptomatic OA have on average 2.05 (± 1.58) co-morbidities.
- The number of co-morbidities on one hand, and psychiatric and degenerative disc diseases on the other hand, have an independent impact on the indirect measurement.
- The impact of the functional effect of OA was more relevant than co-morbidities.

Research agenda

- Effects of co-morbidities must be further explored in a longitudinal study.
- The impact of OA, when occurring as a co-morbidity, on indirect utility must be also investigated when patients' attention is focussed on another disease.

Funding

Funding for the study was obtained from public sources (INSERM, ANR 2010-17186, CHU de Nancy, Conseil Régional de Lorraine, Société Française de Rhumatologie) and unrestricted grants from pharmaceutical companies (Expanscience, Genevrier, Grünenthal, Merck & Co, Inc, Pfizer, Pierre Fabre Médicaments, Sanofi-Aventis France). Opinions expressed in the present article are those of the authors and do not necessarily reflect those of the sponsors.

The study sponsors did not take part in study design, collection, analysis and interpretation of data; writing of the report or the decision to submit the article for publication.

References

- [1] Torrance GW. Utility approach to measuring health-related quality of life. *Journal of Chronic Diseases* 1987;40(6):593–603.
- *[2] Beresniak A, Russell AS, Haraoui B, Bessette L, Bombardier C, Duru G. Advantages and limitations of utility assessment methods in rheumatoid arthritis. *Journal of Rheumatology* 2007 Nov;34(11):2193–200.
- [3] Gafni A. The standard gamble method: what is being measured and how it is interpreted. *Health Services Research* 1994 Jun;29(2):207–24.
- [4] Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. *Health Services Research* 1972;7(2):118–33.
- [5] Bakker C, van der LS. Health related utility measurement: an introduction. *Journal of Rheumatology* 1995 Jun;22(6):1197–9.
- [6] Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health Economics* 2005 Nov;14(11):1169–89.
- [7] Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. Health utilities index. *Pharmacoeconomics* 1995 Jun;7(6):490–502.
- [8] EuroQol – a new facility for the measurement of health-related quality of life. The EuroQol group. *Health Policy* 1990 Dec; 16(3):199–208.
- [9] Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 1992 Jun;30(6):473–83.
- [10] Bryan S, Dolan P. Discrete choice experiments in health economics. For better or for worse? *The European Journal of Health Economics* 2004 Oct;5(3):199–202.
- [11] Feinstein AR. The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Diseases* 1970 Dec; 23(7):455–68.
- *[12] Fortin M, Lapointe L, Hudon C, Vanasse A, Ntetu AL, Maltais D. Multimorbidity and quality of life in primary care: a systematic review. *Health and Quality of Life Outcomes* 2004;2:51.
- *[13] Redelmeier DA, Tan SH, Booth GL. The treatment of unrelated disorders in patients with chronic medical diseases. *The New England Journal of Medicine* 1998 May 21;338(21):1516–20.
- [14] Martel-Pelletier J, Boileau C, Pelletier JP, Roughley PJ. Cartilage in normal and osteoarthritis conditions. *Best Practice & Research Clinical Rheumatology* 2008 Apr;22(2):351–84.
- *[15] Bitton R. The economic burden of osteoarthritis. *American Journal of Managed Care* 2009 Sep;15(8 Suppl.):S230–5.

- [16] Le TK, Montejano LB, Cao Z, Zhao Y, Ang D. Health care costs in US patients with and without a diagnosis of osteoarthritis. *Journal of Pain Research* 2012;2012(5):23–30.
- [17] Caporali R, Cimmino MA, Sarzi-Puttini P, Scarpa R, Parazzini F, Zaninelli A, et al. Comorbid conditions in the AMICA study patients: effects on the quality of life and drug prescriptions by general practitioners and specialists. *Seminars in Arthritis and Rheumatism* 2005 Aug;35(1 Suppl. 1):31–7.
- [18] Gabriel SE, Crowson CS, O'Fallon WM. Comorbidity in arthritis. *Journal of Rheumatology* 1999 Nov;26(11):2475–9.
- [19] Kadam UT, Jordan K, Croft PR. Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consultants in England and Wales. *Annals of the Rheumatic Diseases* 2004 Apr;63(4):408–14.
- [20] Ruchlin HS, Insinga RP. A review of health-utility data for osteoarthritis: implications for clinical trial-based evaluation. *Pharmacoeconomics* 2008;26(11):925–35.
- [21] Guillemin F. The value of utility: assumptions underlying preferences and quality adjusted life years. *Journal of Rheumatology* 1999 Sep;26(9):1861–3.
- [22] Tuominen U, Blom M, Hirvonen J, Seitsalo S, Lehto M, Paavolainen P, et al. The effect of co-morbidities on health-related quality of life in patients placed on the waiting list for total joint replacement. *Health and Quality of Life Outcomes* 2007 Mar 15;5:16.
- *[23] Aktekin LA, Eser F, Malhan S, Oksuz E, Keskin D, Bodur H. A comparison of four different HRQoL generic questionnaires in five different patient groups. *Rheumatology International* 2009 Nov;30(1):63–7.
- [24] Blanchard C, Feeny D, Mahon JL, Bourne R, Rorabeck C, Stitt L, et al. Is the health utilities index valid in total hip arthroplasty patients? *Quality of Life Research* 2004 Mar;13(2):339–48.
- [25] Ethgen O, Vanparijs P, Delhalle S, Rosant S, Bruyere O, Reginster JY. Social support and health-related quality of life in hip and knee osteoarthritis. *Quality of Life Research* 2004 Mar;13(2):321–30.
- [26] Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. *Aging Clinical and Experimental Research* 2005 Aug;17(4):255–63.
- [27] van Dijk GM, Veenhof C, Schellevis F, Hulsmans H, Bakker JP, Arwert H, et al. Comorbidity, limitations in activities and pain in patients with osteoarthritis of the hip or knee. *BMC Musculoskeletal Disorders* 2008 Jun 26;9:95.
- *[28] Guillemin F, Rat AC, Roux CH, Fautrel B, Mazieres B, Chevalier X, et al. The KHOALA cohort of knee and hip osteoarthritis in France. *Joint Bone Spine* 2012 Jun 20 [Epub ahead of print].
- [29] Guillemin F, Rat AC, Mazieres B, Pouchot J, Fautrel B, Euller-Ziegler L, et al. Prevalence of symptomatic hip and knee osteoarthritis: a two-phase population-based survey. *Osteoarthritis Cartilage* 2011 Nov;19(11):1314–22.
- [30] Altman RD. Osteoarthritis. Aggravating factors and therapeutic measures. *Postgraduate Medical Journal* 1986 Aug;80(2):150–63.
- [31] Altman RD. Criteria for classification of clinical osteoarthritis. *Journal of Rheumatology. Supplement* 1991 Feb;27:10–2.
- *[32] Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. *Journal of Clinical Epidemiology* 2005 Jun;58(6):595–602.
- [33] Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *Journal of Rheumatology* 1988 Dec;15(12):1833–40.
- [34] Rosemann T, Gensichen J, Sauer N, Laux G, Szecsenyi J. The impact of concomitant depression on quality of life and health service utilisation in patients with osteoarthritis. *Rheumatology International* 2007 Jul;27(9):859–63.
- [35] Rosemann T, Laux G, Szecsenyi J, Wensing M, Grol R. Pain and osteoarthritis in primary care: factors associated with pain perception in a sample of 1,021 patients. *Pain Medicine* 2008 Oct;9(7):903–10.
- [36] Sintonen H. An approach to measuring and valuing health states. *Social Science & Medicine – Medical Economics* 1981 Jun;15(2):55–65.
- [37] Salaffi F, De AR, Stancati A, Grassi W. Health-related quality of life in multiple musculoskeletal conditions: a cross-sectional population based epidemiological study. II. The MAPPING study. *Clinical and Experimental Rheumatology* 2005 Nov;23(6):829–39.
- *[38] Groll DL, Heyland DK, Caesar M, Wright JG. Assessment of long-term physical function in acute respiratory distress syndrome (ARDS) patients: comparison of the Charlson comorbidity index and the functional comorbidity index. *American Journal of Physical Medicine & Rehabilitation* 2006 Jul;85(7):574–81.
- [39] Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases* 1987;40(5):373–83.
- [40] Conwell Y, Forbes NT, Cox C, Caine ED. Validation of a measure of physical illness burden at autopsy: the cumulative illness rating scale. *Journal of the American Geriatrics Society* 1993 Jan;41(1):38–41.
- [41] Wylie JD, Bershady B, Iannotti JP. The effect of medical comorbidity on self-reported shoulder-specific health related quality of life in patients with shoulder disease. *Journal of Shoulder and Elbow Surgery* 2010 Sep;19(6):823–8.
- *[42] Suri P, Morgenroth DC, Kwok CK, Bean JF, Kalichman L, Hunter DJ. Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. *Arthritis Care & Research (Hoboken)* 2010 Dec;62:1715–23.
- [43] Juhakoski R, Tenhonen S, Anttonen T, Kauppinen T, Arokoski JP. Factors affecting self-reported pain and physical function in patients with hip osteoarthritis. *Archives of Physical Medicine and Rehabilitation* 2008 Jun;89(6):1066–73.
- *[44] Gore M, Tai KS, Sadosky A, Leslie D, Stacey BR. Clinical comorbidities, treatment patterns, and direct medical costs of patients with osteoarthritis in usual care: a retrospective claims database analysis. *Journal of Medical Economics* 2011;14(4):497–507.
- [45] Schellevis FG, van d V J, van de Lisdonk E, van Eijk JT, van WC. Comorbidity of chronic diseases in general practice. *Journal of Clinical Epidemiology* 1993 May;46(5):469–73.
- [46] Harris T, Kovar MG, Suzman R, Kleinman JC, Feldman JJ. Longitudinal study of physical ability in the oldest-old. *American Journal of Public Health* 1989 Jun;79(6):698–702.

5. MESURE DU SEUIL DE CHANGEMENT DE L'UTILITE EN PRESENCE DE COMORBIDITE

Nous avons vu dans le précédent travail et à travers une analyse transversale, l'effet délétère des comorbidités sur la mesure d'utilité chez des patients atteints d'une pathologie chronique.

Les objectifs de ce troisième travail étaient de définir un seuil de changement pour l'utilité et de déterminer l'effet des comorbidités sur l'atteinte de ce seuil de changement par les patients.

Cette nouvelle analyse s'est basée sur les données longitudinales de la cohorte KHOALA (en page 29) composée de patients avec une arthrose de hanche et/ou genou. Les individus inclus dans cette cohorte sont suivis annuellement et nous disposons des données sur 3 ans de suivi.

Les 878 sujets inclus dans la cohorte ont été interrogés chaque année sur leur santé perçue à travers le questionnaire générique SF-36 et une mesure algofonctionnelle de la maladie par le WOMAC, et les comorbidités ont été évaluées par l'indice de Groll (FCI) à l'inclusion.

Le tableau ci-dessous présente le score d'utilité obtenue par le SF-6D dérivé du SF-36 pour chaque année de suivi dans la cohorte.

	Visite de suivi			
	Inclusion (n=808)	1 ^{ère} année (n=713)	2 ^{ème} année (n=641)	3 ^{ème} année (n=697)
Score d'utilité par SF-6D				
Moyenne	0,659 (0,110)	0,664 (0,111)	0,664 (0,111)	0,665 (0,112)
(écart-type)				
min - médiane - max	0,319 - 0,647 - 1	0,319 - 0,648 - 0,958	0,334 - 0,667 - 0,958	0,319 - 0,664 - 0,958

Il n'y avait pas de différence statistiquement significative de la moyenne de l'utilité entre les 4 temps de mesure ($p=0,667$). Pour l'analyse longitudinale les données à l'inclusion et à la 3^{ème} année de suivi ont été retenues. Au total, 650 patients ont répondu au questionnaire à ces deux temps de mesure. Un test t sur échantillon apparié a été réalisé et n'a pas montré de

différence statistiquement significative ($p= 0,340$) de la moyenne des scores d'utilité entre l'inclusion et le suivi à 3 ans.

La moyenne d'âge des 650 patients était de 61.6 ans (± 8.3), et ils présentaient en moyenne 1.96 (± 1.52) comorbidités à l'inclusion. Le score moyen de fonction WOMAC était de 29.54 (± 19.41) à l'inclusion et 30.32 (± 19.55) à trois ans.

En l'absence d'une mesure de changement de santé globale comme référence externe précise et structurée dans la cohorte, nous avons utilisé uniquement la méthode basée sur la distribution de l'échantillon pour établir le seuil de changement.

Le choix a été fait d'établir ce seuil par la méthode de l'erreur de mesure standard (SEM) car celle-ci a l'avantage d'être plus précise ou indépendante de la taille de l'échantillon. Le calcul s'est fait selon la formule suivante (mettre réf):

$$SEM = SD_{t1} \times \sqrt{1 - ICC_{t1,t2}}$$

- SD_{t1} : Ecart-type de la moyenne du score d'utilité à l'inclusion
- $ICC_{t1,t2}$: Coefficient de corrélation intraclasse entre le score d'utilité à l'inclusion et à 3 ans de suivi

Le seuil de changement (MID) ainsi obtenu pour l'utilité estimée par le SF-6D était de 0,071. A partir de ce seuil, les 650 patients ont été classés selon la variation de leur score d'utilité en : 1) changement négatif, 2) changement positif, et 3) inchangé à 3 ans. Sur notre échantillon, la majorité des patients (53%) étaient inchangés, 156 (24%) patients avaient un changement positif, et 147 (23%) un changement négatif du score d'utilité.

Afin d'établir l'effet des comorbidités sur l'atteinte du seuil de changement (positif ou négatif) par les patients, 2 modèles par régression logistique multinomiale ont été menés : un modèle avec les 9 catégories de comorbidités et le second avec le nombre de comorbidité (allant de 0 à 9). Dans les deux modèles ont été inclus l'ensemble des variables socio-démographiques, ainsi que la différence des moyennes du score de fonction de WOMAC (une augmentation de la différence des moyennes traduit une détérioration de l'état fonctionnel) entre les deux temps de mesure.

Dans le premier modèle seule la présence d'une pathologie pneumologique avait un effet significatif sur un changement négatif de l'utilité (OR=2.34; $p=0.01$). Les résultats obtenus

ont montré un effet du score de WOMAC sur un changement positif et négatif d'utilité. Dans les deux modèles, une détérioration fonctionnelle due à l'arthrose augmente significativement le risque d'avoir un changement négatif et de façon corollaire elle diminue la probabilité d'un changement positif.

Cette analyse longitudinale a permis de définir un seuil de changement pour l'interprétation de l'utilité, et a mis en évidence l'importance de la sévérité fonctionnelle de l'arthrose dans l'expression de l'utilité et son changement comparé aux comorbidités. Il serait intéressant de compléter celle-ci en prenant en compte l'existence d'un phénomène de response-shift au sein de notre cohorte.

5.1. TROISIEME TRAVAIL

Impact of comorbidities on utility changes in lower-limb osteoarthritis : KHOALA Study

Hosseini Kossar, Rat Anne-Christine, Gaujoux-Viala Cecile, Coste Joel, Pouchot Jacques, Fautrel Bruno, Guillemin Francis

En préparation

Impact of comorbidities on utility changes in lower-limb osteoarthritis : results from the KHOALA Study

Authors: Hosseini Kossar¹, Rat Anne-Christine¹, Gaujoux-Viala Cecile^{2,3}, Coste Joel¹, Pouchot Jacques¹, Fautrel Bruno⁴, Guillemin Francis¹

¹ Université de Lorraine, Paris Descartes University, APEMAC, EA 4360, F- 54 000, France

² Montpellier I University, Nîmes University Hospital, Rheumatology, France

³ EA 2415, Montpellier I University, Nîmes University Hospital, Rheumatology Department, France

⁴ Paris 6 – Pierre et Marie Curie University; Department of Rheumatology, Pitié-Salpêtrière Hospital, Paris, France

Authorship: KH and FG conceived and designed the study. ACR, FG, JC and BF collected data. KH and FG conducted the analysis and interpreted data. KH and FG drafted the manuscript. ACR, CGV, JC, JP, BF helped to revise critically the manuscript. All authors read and approved the final manuscript.

Corresponding author

Kossar Hosseini

Centre d'Epidémiologie Clinique

Inserm CIC-EC

Service d'Epidémiologie et Evaluation Cliniques

CHU de Nancy

Hopitaux de Brabois

Allée du Morvan

54 505 Vandoeuvre-lès-Nancy

France

Tel.: +33(0)3 83 85 93 02 / Fax : +33(0)3 83 85 12 05 Email : k.hosseini@chu-nancy.fr

Short title : Utility changes in lower-limb osteoarthritis

Abstract

Objectives: 1) to determine the minimal important change (MID) of indirect utility score for patients with osteoarthritis (OA) over 3-year follow up; and 2) to estimate the impact of comorbidity on patients likelihood to reach the MID.

Methods: 878 patients with symptomatic knee or/and hip OA of the KHOALA cohort study were included . Comorbidity assessed by the Functional Comorbidity Index (FCI) at baseline. Limitation in activities and pain was assessed by the WOMAC function score (0-100, 0 the better), and utility score using the SF-6D at baseline and 3-year. The MID of utility score was determined using the standard error of measurement (SEM). Accordingly, patients were classified as negative/positive change or unchanged. Two separate multinomial logistic models were fitted to determine predictors of a positive or negative change over the MID.

Results: 650 patients completed the questionnaire at both times. The mean (SD) utility score was 0.664 (± 0.110) at baseline and 0.667 (± 0.110) at 3-year. Patients had on average 1.96 (± 1.52) comorbidities. The MID of SF-6D utility score was 0.071, and 147 patients had a negative change and 156 a positive change. In both multinomial models, patients with a decrease of the WOMAC function score had an increased utility ($p < .0001$).

Conclusions: According to the MID, about half of patients had a positive or a negative change in their utility score. Compared to functional severity of OA, comorbidities have a relatively marginal impact on indirect utility score. This suggests that clinically, considering the functional severity of OA remains a first priority.

Word count: 250

Key words: utility, minimal important difference, distribution-based, comorbidity, osteoarthritis, SF-6D, health related quality of life, cohort study, long-term follow-up, functional severity

Introduction

Utility methodology in assessing health-related quality of life (HRQoL) allows for comparing different treatment strategies for one or more diseases by a single unit of measurement and is generally used in cost-utilities analysis to compare the differential costs of strategies related to differential utility among those strategies [1]. The main objective of cost-utility analysis is to relate cost changes to utility changes across treatment strategies and to define the incremental cost-utility ratio, further used in decision analysis.

The interpretation of utility changes can be anchored on the concept of minimal important difference (MID). This concept was developed by Jaeschke et al. [2] as “the smallest difference in score in the domain of interest which patients would perceive as beneficial and which would mandate, in the absence of troublesome side-effects and excessive cost, a change in the patient’s management”. Such MID helps calibrate the magnitude of observed changes. It could be useful to interpret changes of utility score over time induced by a strategy (within group), and to compare changes between strategies (between groups). The most common methods for estimating the MID of HRQoL instruments are distribution-based and anchor-based [3, 4].

Osteoarthritis (OA) is a frequent chronic degenerative joint disease [5] influencing HRQoL and incurring important economic burden [6, 7]. Because it develops with ageing, patients with OA have a risk of developing comorbidities [8-10] and a high prevalence of comorbidities [11-13], which can be predictors of use of health services and medical costs [9, 12, 14]. In a previous cross-sectional study [15], we showed the negative impact of comorbidities on utility scores by use of a multi-attribute instrument in a cross-sectional study of OA patients.

The purpose of this study was twofold: 1) to determine the MID of indirect utility score for patients with OA over 3-years follow-up; 2) to estimate the impact of comorbidity on patients likelihood to reach the MID.

Methods

Data collection

We used data for patients from the Knee and Hip Osteo-Arthritis Long-term Assessment (KHOALA) cohort study [16] including 878 patients recruited from the 3,000 OA population-based prevalence survey undertaken in France from 2007 to 2009 [17]. Inclusion criteria were patients of both sexes; 40 to 75 years old; with uni- or bilateral, symptomatic hip and/or knee OA; and with clinical diagnosis confirmed and fulfilling American College of Rheumatology

criteria for knee [18] and hip [19] OA. Patients are followed annually with a clinical visit or by mailed self-reporting questionnaire. Sociodemographic characteristics (age, sex, education level, occupation, marital status) were collected at baseline. Comorbidity was assessed at baseline by the Functional Comorbidity Index (FCI) [20], which was specifically developed to assess the impact of comorbidities on the physical dimension of QoL. Given that all patients in this cohort have OA, we removed the item OA from the index and grouped the 18 conditions of this index into 9 categories for analysis (Table 1). From this index, we calculated the number of comorbidities for each patient.

We used data collected at baseline and 3-year for limitation in activities and physical function measured by the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [21]. Functional WOMAC score (17 items, 5 point Likert scale for each item) range from 0 to 100, with high scores reflecting increased limitation in activity.

Utility outcome

Utility was assessed by the SF-6D [22], from the MOS 36-item Short Form Health Survey (SF-36) [23] at baseline and 3-year. The SF-6D was developed from 11 questions out of the SF-36. It is a multidimensional health classification system assessing the 6 domains of physical functioning, role limitation, social functioning, pain, mental health, and vitality, with 4–6 levels for each domain. An SF-6D health state is defined by 1 level from each domain, which results in a total of 18,000 possible health states. The SF-6D scoring algorithm was developed using the standard gamble method from a sample of 249 SF-6D health states from a representative sample of the UK population. Utility scores generated by the SF-6D range from 0.29 to 1.00, with 1.00 representing full health and 0.29 representing the worst possible health state defined by the SF-6D (i.e., all domains being at the worst level).

MID estimation

The MID for the SF-6D utility score was estimated using the standard error of measurement (SEM) [24] in the entire sample by the standard deviation of the instrument at baseline multiplied by the square root of one minus the intraclass correlation coefficient (ICC) between utility score measured at baseline (t1) 3-year (t2) :

$$SEM = \sigma_{t1} * \sqrt{1 - ICC_{(t1,t2)}}$$

Patient with a negative or positive change equal or greater than the MID were considered to have a meaningful change.

Statistical analysis

The sociodemographic and clinical characteristics of patients were described with means (SD) and numbers (percentages). For the SF-6D utility score the mean (SD) and the distribution of the score at both time were computed.

For the MID assessment, the ICC (95% confidence interval) was estimated between baseline and 3-year utility scores from an ANOVA, and further used for SEM calculation.

According to the MID patients were classified as negative change, positive change or unchanged. Two multinomial logistic models were fitted to determine predictors of a positive or a negative change over the MID threshold. Because of multicollinearity between the number of comorbidity recorded by the FCI index and by the comorbidity by categories, two separate models were fitted. The first one included the number of comorbidity and the second comorbidity by categories expressed by the FCI index. Both models included sociodemographic characteristics and the mean difference of Womac function score between baseline and 3-year (a negative difference indicates a better function at 3-year) .

A $p < 0.05$ was considered statistically significant. Analyses were performed with SAS 9.3 (SAS Inst, Cary, NC, USA).

Results

For the 878 patients included in the cohort study, 650 responded to the SF-6D questionnaire at baseline and 3-year and were retained for analysis. The mean (SD) age was 61.6 (8.3), 67.5% were women and 68.9% living in couple at baseline. The mean (SD) Womac function score was 29.5 (19.4) at baseline and 30.3 (19.5) at 3-year. Patient had on average 1.96 (± 1.52) comorbidities at baseline. The most common comorbidities were degenerative disc disease (53.7%) and diabetes and obesity (39.5%).

The mean (SD) utility score using the SF-6D at baseline was 0.664 (0.110)(range 0.370–1.00), and 0.667 (0.110) range (0.319–0.958) at 3-year. The distribution of utility score was near normal for both measurement time.

MID estimation

For a ICC (95% CI) of 0.583 (0.531;0.632) the MID estimate by the SEM method was 0.071.

According to the MID more than half of patients (53.4%) was considered having no change in their utility score at 3-year, 156 (24%) experienced a positive change and 147 (22.6%) patients were rated as negative change, i.e. reporting a deterioration. Patients with a negative change had a worse functional score at 3-year (33.30 ± 19.85) and more comorbidities (2.03 ± 1.42) compared to unchanged patients (Table 2). In the first multinomial model (Table 3), including comorbidity by category, compared to unchanged patients, the presence of a pneumologic comorbidity increased the probability for patients to experienced a negative utility change (OR=2.34; p=0.01).

In the second model including the number of comorbidities (Table 4), there was no significant effect of the number of comorbidities on a positive or negative change in the utility score. In both models patients living in couple were less likely to have a positive change, and those with a deterioration of the WOMAC function had an impairment in utility score (OR=1.02; p=0.003) compared to unchanged patients.

Discussion

We determined the MID of the SF-6D from a sample of 650 patients with hip or knee OA. According to the MID less than half of patients had a positive or negative change in utility at 3-year follow-up. Our results showed a marginal effect of comorbidities on utility change.

Our study has several strengths. First, we have response rate of 74% of utility assessment at 3 years follow-up. Second, we used data from the first French national population-based cohort of symptomatic patients with hip or knee OA. Patients were recruited independently of any medical consultation, on a population-based screening process. This provided a representative sample of cases of OA from the general population and gives representative data regarding the distribution of disease severity, and distribution of comorbidities associated with OA. This allows for generalizing the results. Compared to other studies reporting a minimal change for utility score, we had a large sample of patients with long term follow-up. Subjects were followed by a medical examination in each of the 6 investigating centres at baseline and 3-year follow-up in “real life”.

There were also some limitations. First, while authors recommended to use both anchor- and distribution-based method to calculate a range of MID [25], we used only a distribution-based method to assess the MID. Although there is no consensus on which method is the best, authors reported limitations for both MID estimation methods. The major criticism of anchor-based methods is that it relies on the use of retrospective rating scale which introduces a

potential response-shift and recall bias due to the extend time between assessments [26]. Jeaschke et al. [2] used a length of time between the two measurements of 2 weeks and Juniper et al. [27] of 4 weeks, which is much smaller than the 3-year follow-up of our cohort study. According to our data, applying an anchor-based method seemed less appropriate to determine the MID of utility score than the distribution-based method. However, distribution-based methods were considered to provide no direct information about the minimal change and the clinical relevance of change by expressing the observed change in a stanrdized metric [25].

Second, we expressed the MID using data between the inclusion in cohort and 3-year follow up. Although results showed no significant changes in utility over time on average in our cohort , the MID estimate may vary during follow-up and may not necessarily be equal over consecutive shorter periods. Further study is needed to investigate how the MID varies over time for the same sample of patients.

Third, we showed a marginal effect of comorbidities on a positive or negative change in patients reaching the MID. Given the importance of physical function for patients with OA, we used the FCI index to determine comorbidity in our study, since it may be the most appropriate way to measure comorbidities that affect perceived physical functioning [35]. The FCI index allows to determine the presence of 18 categories of comorbidity without giving them a weight depending on the severity as with the Cumulative Illness Rating Scale (CIRS) [36]. It appears that a simple determination of comorbidities is insufficient for determining effect on utility change for patients with OA and it would be interesting to see if taking into account the severity of comorbidities affects the utility changes.

This study is the first that reports MID for SF-6D utility scores using SEM methods. According to classical test theory, Wyrwich et al. [24] considered the SEM as sample independent and better for determining a meaningful change as compared with other distribution-based methods for MID estimation such as the effect size. The SEM is then considered to provide an equivalent meaning across different patient samples, allowing use of this threshold in further studies.

We reported a MID of 0.07 for the SF-6D scores when other studies reported a MID ranging from 0.01 to 0.10 using anchor or distribution-based method or both [28-32]. Considering a sample of 59 knee OA patients with 6 months follow-up, Walters and Brazier [31] used both methods and reported a MID of SF-6D scores range from 0.03 to 0.49.

Since the magnitude of a meaningful change score is not the same for improved or deteriorated patients [33, 34], we have attempted to show how the changes in score may affected each group separately, compared to unchanged patients. Results showed that the change in utility score was dependent on functional impact of OA on the individual. These longitudinal findings completes previous cross-sectionnal results from a previous sectional study [15] conducted on baseline data of the KHOALA cohort which showed a greater effect of OA compared to other co-occurring diseases.

The WOMAC is a disease-specific questionnaire developed for patients with hip or knee OA. It comprises a multidimensional scale measuring 3 dimensions: pain, stiffness and physical function. We retained for our analysis only the functional score which is the most correlated with the utility score. Despite the fact that studies have shown that other musculoskeletal pain [37] or number of comorbidities [38] may also affect physical function and thus the WOMAC score, considering the aim of our study, the relationship between WOMAC and comorbidity was not explored.

In conclusion, we reported a MID as threshold to determine a meaningful improvement or deterioration in utility change over time in individuals with OA. Results suggest that in a long-term follow-up patients were more focused in the present illness compared to other co-occurring diseases. Although patients with OA have a high prevalence of comorbidities [11, 13, 39], which can be predictors of use of health services and medical costs [14, 40], our results imply that clinically a first priority remains to consider the functional severity of OA.

Acknowledgements

The authors thank the Clinical Investigation Center (Clinical Epidemiology) for the coordination and monitoring, and all investigators who recruited and followed subjects in the cohort.

Conflict of interest statement

Authors declare no conflicts of interest for the submitted paper.

Funding sources: Funding for the study was obtained from public sources (INSERM, ANR, CHU de Nancy, Conseil Régional de Lorraine, Société Française de Rhumatologie) and unrestricted grants from pharmaceutical companies (Expanscience, Genevrier, Grünenthal, Merck & Co, Inc, Pfizer, Pierre Fabre Médicaments, Sanofi-Aventis France). Opinions expressed in the present article are those of the authors and do not necessarily reflect those of the sponsors.

The study sponsors did not take part in study design, collection, analysis and interpretation of data; writing of the report or the decision to submit the article for publication.

Table 1. The 18 Functional Comorbidity Index items grouped into 9 items

	Functional Comorbidity Index	Grouping
1	Arthritis (rheumatoid and osteoarthritis)	Not retained
2	Osteoporosis	Osteoporosis
3	Asthma	Pneumologic disease
4	COPD, ARDS, or emphysema	
5	Angina	Cardiac disease
6	Congestive hearth failure (or heart disease)	
7	Heart attack (myocardial infarction)	
10	Peripheral vascular disease	
8	Neurological disease (such as multiple sclerosis or Parkinson's disease)	Neurologic disease
9	Stroke or transient ischemic attack	
11	Diabetes types I and II	Diabetes and obesity
18	Obesity or body mass index >30	
12	Upper gastrointestinal disease (ulcer, hernia, reflux)	Upper gastrointestinal disease
13	Depression	Psychiatric disease
14	Anxiety or panic disorders	
15	Visual impairment (such as cataracts, glaucoma, macular degeneration)	Visual and hearing impairment
16	Hearing impairment (very hard of hearing, even with hearing aids)	
17	Degenerative disc disease	Degenerative disc disease

COPD, chronic obstructive pulmonary disease; ARDS, acquired respiratory distress syndrome

Table 2. Characteristics of patients with osteoarthritis included in this study, by change in their utility between baseline and 3-year according to the MID (n=650).

	Change in utility score		
	Unchanged N=347	Positive change N=156	Negative change N=147
Age, mean (SD)	61.4 (8.6)	61.2 (7.9)	62.2 (7.9)
Sex			
Male	113 (32.6)	50 (32.1)	48 (32.7)
Female	234 (67.4)	106 (67.9)	99 (67.3)
Married or living together	251 (72.3)	101 (64.7)	96 (65.3)
Education			
Primary	76 (21.9)	33 (21.2)	36 (24.5)
Secondary	176 (50.7)	79 (50.6)	73 (49.7)
University	95 (27.4)	42 (26.9)	37 (25.2)
Occupation			
Farmer and worker	53 (15.3)	22 (14.1)	32 (21.8)
Manager	22 (6.3)	6 (3.8)	9 (6.1)
Executive	46 (13.3)	25 (16.0)	28 (19.0)
Employee and craftman	204 (58.8)	88 (56.4)	61 (41.5)
Not working	20 (5.8)	15 (9.6)	14 (9.5)
SF-6D utility score at baseline (N)	347	156	147
mean (SD)	0.662 (0.102)	0.601 (0.092)	0.733 (0.104)
min - median - max	0.403 - 0.652 - 0.887	0.370 - 0.595 - 0.810	0.403 - 0.751 - 1
SF-6D utility score at 3-year (N)	347	156	147
mean (SD)	0.661 (0.105)	0.735 (0.099)	0.610 (0.096)
min - median - max	0.398 - 0.648 - 0.887	0.517 - 0.751 - 0.958	0.319 - 0.611 - 0.845
WOMAC function score at baseline [0-100] (N)	346	155	144
mean (SD)	29.13 (19.32)	32.57 (20.30)	27.30 (18.33)
min - median - max	0.00 - 26.47 - 95.59	0.00 - 32.35 - 83.82	0.00 - 24.26 - 86.76
WOMAC function score at 3-year [0-100] (N)	345	155	147
mean (SD)	30.85 (19.83)	26.31 (18.02)	33.30 (19.85)
min - median - max	0.00 - 29.41 - 83.93	0.00 - 25.00 - 76.47	0.00 - 33.82 - 94.12
No. of comorbidities [0-9], mean (SD)	1.97 (1.58)	1.87 (1.47)	2.03 (1.42)
Functional comorbidity index			
Osteoporosis	26 (7.5)	4 (2.6)	9 (6.1)
Pneumologic disease	23 (6.6)	12 (7.7)	20 (13.6)
Cardiologic disease	48 (13.8)	16 (10.3)	22 (15.0)

	Change in utility score		
	Unchanged N=347	Positive change N=156	Negative change N=147
Neurologic disease	12 (3.5)	9 (5.8)	4 (2.7)
Diabetes and obesity	137 (39.5)	64 (41.0)	56 (38.1)
Upper gastrointestinal disease	83 (23.9)	32 (20.5)	35 (23.8)
Psychiatric disease	76 (21.9)	32 (20.5)	27 (18.4)
Hearing and visual impairment	50 (14.4)	19 (12.2)	24 (16.3)
Degenerative disc disease	183 (52.7)	88 (56.4)	78 (53.1)

Data are no. (%) unless indicated.

Table 3. Factors affecting a positive or negative change of utility score compared to unchanged patients, for patients with hip or knee osteoarthritis, using multinomial regression model (Model 1)

Model 1 Including comorbidities by category				
	Positive change		Negative change	
	Odds Ratio	95% CI	Odds Ratio	95%CI
Age	1.01	0.99; 1.04	1.01	0.97; 1.04
Female	0.92	0.58; 1.47	1.06	0.66; 1.69
Married or living together	0.56*	0.36; 0.89	0.64	0.41; 1.02
Education (<i>Ref=Primary</i>)				
Secondary	1.30	0.74; 2.29	1.02	0.59; 1.75
University	1.10	0.55; 2.22	0.75	0.37; 1.53
Occupation (<i>Ref=Not working</i>)				
Farmer and worker	0.58	0.23; 1.46	0.79	0.33; 1.92
Manager	0.34	0.10; 1.14	0.52	0.17; 1.56
Executive	0.81	0.31; 2.11	0.95	0.36; 2.48
Employee and craftman	0.53	0.25; 1.16	0.41*	0.19; 0.91
WOMAC function score mean difference [0-100]	0.95**	0.93; 0.96	1.02*	1.01; 1.04
Functional comorbidity index				
Osteoporosis	0.23*	0.07; 0.73	0.70	0.29; 1.66
Pneumologic disease	1.50	0.69; 3.29	2.35*	1.19; 4.63
Cardiologic disease	0.64	0.32; 1.25	1.03	0.57; 1.87
Neurologic disease	1.83	0.69; 4.87	0.95	0.29; 3.14
Diabetes and obesity	1.20	0.77; 1.85	0.92	0.60; 1.43
Upper gastrointestinal disease	0.81	0.48; 1.35	0.81	0.49; 1.34
Psychiatric disease	0.70	0.41; 1.21	0.65	0.37; 1.13
Hearing and visual impairment	0.71	0.37; 1.37	0.99	0.55; 1.79
Degenerative disc disease	1.16	0.76; 1.77	1.16	0.76; 1.77
n=634				
AIC	1249.01			
2 Log likelihood	1169.01			
Nagelkerke R ²	0.187			

Table 4. Factors affecting a positive or negative change of utility score compared to unchanged patients, for patients with hip or knee osteoarthritis, using multinomial regression model (Model 2)

Model 2 Including the number of comorbidity				
	Positive change		Negative change	
	Odds Ratio	95% CI	Odds Ratio	95%CI
Age	0.99	0.97; 1.03	1.01	0.98; 1.03
Female	0.85	0.54; 1.34	1.01	0.64; 1.57
Married or living together	0.59*	0.37; 0.92	0.68	0.43; 1.06
Education (<i>Ref=Primary</i>)				
Secondary	1.16	0.67; 2.01	1.01	0.60; 1.72
University	0.97	0.49; 1.92	0.76	0.38; 1.53
Occupation (<i>Ref= Not working</i>)				
Farmer and worker	0.57	0.23; 1.41	0.80	0.34; 1.91
Manager	0.36	0.11; 1.17	0.57	0.19; 1.68
Executive	0.81	0.32; 2.07	0.98	0.38; 2.53
Employee and craftman	0.55	0.26; 1.19	0.42*	0.19; 0.92
WOMAC function score mean difference [0-100]	0.95**	0.94; 0.97	1.03*	1.01; 1.04
No. of comorbidities [0-9]	0.93	0.81; 1.07	0.99	0.87; 1.14
n=634				
AIC	1238.93			
2 Log likelihood	1190.93			
Nagelkerke R ²	0.153			

References

1. Beresniak A, Russell AS, Haraoui B, Bessette L, Bombardier C, Duru G Advantages and limitations of utility assessment methods in rheumatoid arthritis. *J Rheumatol* 34; 2193-200.
2. Jaeschke R, Singer J, Guyatt GH Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 10; 407-15.
3. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 77; 371-83.
4. Norman GR, Sridhar FG, Guyatt GH, Walter SD Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care* 39; 1039-47.
5. Martel-Pelletier J, Boileau C, Pelletier JP, Roughley PJ Cartilage in normal and osteoarthritis conditions. *Best Pract Res Clin Rheumatol* 22; 351-84.
6. Bitton R The economic burden of osteoarthritis. *Am J Manag Care* 15; S230-S235.
7. Le TK, Montejano LB, Cao Z, Zhao Y, Ang D Health care costs in US patients with and without a diagnosis of osteoarthritis. *J Pain Res* 2012;5; 23-30.
8. Caporali R, Cimmino MA, Sarzi-Puttini P *et al.* Comorbid conditions in the AMICA study patients: effects on the quality of life and drug prescriptions by general practitioners and specialists. *Semin Arthritis Rheum* 35; 31-7.
9. Gabriel SE, Crowson CS, O'Fallon WM Comorbidity in arthritis. *J Rheumatol* 26; 2475-9.
10. Kadam UT, Jordan K, Croft PR Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consultants in England and Wales. *Ann Rheum Dis* 63; 408-14.
11. Caporali R, Cimmino MA, Sarzi-Puttini P *et al.* Comorbid conditions in the AMICA study patients: effects on the quality of life and drug prescriptions by general practitioners and specialists. *Semin Arthritis Rheum* 35; 31-7.
12. Gore M, Tai KS, Sadosky A, Leslie D, Stacey BR Clinical comorbidities, treatment patterns, and direct medical costs of patients with osteoarthritis in usual care: a retrospective claims database analysis. *J Med Econ* 2011;14; 497-507.
13. Schellevis FG, van d, V, van de Lisdonk E, van Eijk JT, van WC Comorbidity of chronic diseases in general practice. *J Clin Epidemiol* 46; 469-73.
14. Harris T, Kovar MG, Suzman R, Kleinman JC, Feldman JJ Longitudinal study of physical ability in the oldest-old. *Am J Public Health* 79; 698-702.
15. Hosseini K, Gaujoux-Viala C, Coste J *et al.* Impact of co-morbidities on measuring indirect utility by the Medical Outcomes Study Short Form 6D in lower-limb osteoarthritis. *Best Pract Res Clin Rheumatol* 26; 627-35.

16. Guillemin F, Rat AC, Roux CH *et al.* The KHOALA cohort of knee and hip osteoarthritis in France. *Joint Bone Spine* 79; 597-603.
17. Guillemin F, Rat AC, Mazieres B *et al.* Prevalence of symptomatic hip and knee osteoarthritis: a two-phase population-based survey. *Osteoarthritis Cartilage* 19; 1314-22.
18. Altman RD Osteoarthritis. Aggravating factors and therapeutic measures. *Postgrad Med* 80; 150-63.
19. Altman RD Criteria for classification of clinical osteoarthritis. *J Rheumatol Suppl* 27; 10-2.
20. Groll DL, To T, Bombardier C, Wright JG The development of a comorbidity index with physical function as the outcome. *J Clin Epidemiol* 58; 595-602.
21. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 15; 1833-40.
22. Brazier J, Roberts J, Deverill M The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 21; 271-92.
23. Ware JE, Jr., Sherbourne CD The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 30; 473-83.
24. Wyrwich KW, Tierney WM, Wolinsky FD Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 52; 861-73.
25. Revicki D, Hays RD, Cella D, Sloan J Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 61; 102-9.
26. Crosby RD, Kolotkin RL, Williams GR Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 56; 395-407.
27. Juniper EF, Guyatt GH, Willan A, Griffith LE Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 47; 81-7.
28. Gandhi PK, Ried LD, Bibbey A, Huang IC SF-6D utility index as measure of minimally important difference in health status change. *J Am Pharm Assoc (2003)* 52; 34-42.
29. Kwakkenbos L, Fransen J, Vonk MC *et al.* A comparison of the measurement properties and estimation of minimal important differences of the EQ-5D and SF-6D utility measures in patients with systemic sclerosis. *Clin Exp Rheumatol* 31; 50-6.
30. Lee BB, King MT, Simpson JM *et al.* Validity, responsiveness, and minimal important difference for the SF-6D health utility scale in a spinal cord injured population. *Value Health* 11; 680-8.

31. Walters SJ, Brazier JE What is the relationship between the minimally important difference and health. *Health Qual Life Outcomes* 1:4.; 4.
32. Walters SJ, Brazier JE Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 14; 1523-32.
33. Hays RD, Woolley JM The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? *Pharmacoeconomics* 18; 419-23.
34. Stratford PW, Riddle DL Assessing sensitivity to change: choosing the appropriate change coefficient. *Health Qual Life Outcomes* 3; 23.
35. Groll DL, Heyland DK, Caeser M, Wright JG Assessment of long-term physical function in acute respiratory distress syndrome (ARDS) patients: comparison of the Charlson Comorbidity Index and the Functional Comorbidity Index. *Am J Phys Med Rehabil* 2006 Jul;85; 574-81.
36. Linn BS, Linn MW, Gurel L Cumulative illness rating scale. *J Am Geriatr Soc* 16; 622-6.
37. Suri P, Morgenroth DC, Kwok CK, Bean JF, Kalichman L, Hunter DJ Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative. *Arthritis Care Res (Hoboken)* 2010 Dec;62; 1715-23.
38. Juhakoski R, Tenhonen S, Anttonen T, Kauppinen T, Arokoski JP Factors affecting self-reported pain and physical function in patients with hip osteoarthritis. *Arch Phys Med Rehabil* 2008 Jun;89; 1066-73.
39. Gore M, Tai KS, Sadosky A, Leslie D, Stacey BR Clinical comorbidities, treatment patterns, and direct medical costs of patients with osteoarthritis in usual care: a retrospective claims database analysis. *J Med Econ* 14; 497-507.
40. Gabriel SE, Crowson CS, O'Fallon WM Comorbidity in arthritis. *J Rheumatol* 26; 2475-9.

6. LE SEUIL DE CHANGEMENT D'UTILITE CHEZ LE DONNEUR VIVANT DE REIN

Pour ce dernier travail nous nous sommes intéressé à déterminer un seuil de changement d'utilité chez le donneur vivant de rein.

En ce qui concerne le don de rein, il existe de nombreuses études sur la qualité de vie des receveurs. Les auteurs se penchent moins sur le cas des donneurs, surtout à court terme. Bien évidemment on a initialement montré que les donneurs avaient à moyen et long terme (plus d'un an après l'intervention) une meilleure qualité de vie que la population générale (47;69). Et les bénéfices connus du don par donneur vivant (une meilleure survie du greffon, moins d'attente...) amènent les pouvoirs publics à promouvoir ce don. Or si on considère que le donneur est un individu en bon état de santé avant le don de son rein, on peut s'attendre qu'à court terme (moins d'un an) et suite à l'opération il puisse vivre une détérioration provisoire de son état de santé. D'une autre façon, cela revient à accepter par la société et le donneur lui-même, une diminution de son état de santé général avec une espérance de gains futures (indirects pour le donneur). Il serait donc intéressant de définir un seuil d'acceptabilité de la détérioration de l'état de santé du donneur.

Ainsi l'objectif de ce travail était de décrire la variation d'utilité chez les donneurs vivants de rein à court terme et de définir un seuil de détérioration minimale importante (DÉMI) chez ces donneurs.

Pour cela nous avons utilisé les données préliminaires de l'étude DOVIREIN qui est menée sur 22 centres hospitaliers en France et a pour objectif de réaliser l'évaluation coût-efficacité de différentes techniques de prélèvement sur la qualité de vie des donneurs vivants de rein. Tous les couples donneur / receveur, pris en charge dans les différents centres participants à l'étude et satisfaisant les critères d'inclusion, ont été invités à participer à l'étude.

Dans le cadre de ce présent travail, nous avons utilisé les données des couples donneur/receveur inclus entre juin 2010 et novembre 2012.

Les données sociodémographiques et cliniques ont été recueillies à l'inclusion avant l'intervention chirurgicale. La santé perçue des donneurs a été évaluée par les questionnaires EQ-5D de l'Euroqol et le SF-6D à la veille de l'intervention (J-1) et à 3 mois après le don (J90).

Pour estimer le seuil DÉMI de l'utilité nous avons utilisé les deux approches basées sur une référence externe et la distribution de l'échantillon.

Pour la première, nous avons soumis au donneur à J90 une question de santé globale formulée de la façon suivante :

Par rapport à il y a environ 3 mois, avant le don de votre rein, comment qualifiez-vous votre état de santé en ce moment ?

- *Beaucoup mieux*
- *Mieux*
- *Un peu mieux*
- *Inchangé*
- *Un peu moins bon*
- *Moins bon*
- *Beaucoup moins bon*

Pour établir la DÉMI par cette méthode, nous avons pris comme seuil la différence des moyennes du score d'utilité pour les donneurs estimant leur état « un peu moins bon ».

Pour la seconde approche, nous avons utilisé l'erreur de mesure standard (SEM).

Et enfin, nous avons exploré la présence d'un éventuel phénomène de response-shift chez les donneurs vivants en intégrant pour le SF-6D un questionnaire then-test à J90.

Les résultats de notre étude ont montré une diminution significative de l'utilité à 3 mois après l'intervention. La détérioration de l'utilité exprimée par les donneurs n'est pas affectée par un phénomène de response-shift en ce qui concerne le SF-6D.

On constate à travers la question de santé globale (J90), que les donneurs ont tendance à estimer leur état de santé inchangé ou meilleur alors qu'on observe en moyenne une diminution de l'utilité.

Les méthodes appliquées pour le calcul du seuil ont permis d'établir une DÉMI entre -0,113 et -0,076 pour l'EQ-5D, et entre -0,128 et -0,073 pour le SF-6D. A partir de ces seuils nous avons pu établir que 18% à 43% des donneurs avaient une détérioration significative.

Enfin, nous avons voulu savoir si le fait d'atteindre ou pas ce seuil DÉMI par les donneurs était influencé par l'état du receveur et en particulier les complications post chirurgie.

Donneurs avec seuil de changement DéMI atteint EQ-5D	Complication chez le receveur				
	NON (N=60)		OUI (N=151)		
DéMI par la méthode de référence externe					
OUI	16	(26.7%)	23	(15.2%)	p* = 0.054
NON	44	(73.3%)	128	(84.8%)	
DéMI par la méthode SEM					
OUI	26	(43.3%)	47	(31.1%)	p*= 0.093
NON	34	(56.7%)	104	(68.9%)	

*test de chi-2

Donneurs avec seuil de changement DéMI atteint – SF-6D	Complication chez le receveur				
	NON (n=58)		OUI (n=134)		
DéMI par la méthode de référence externe					
OUI	13	(22.4%)	46	(34.3%)	p*= 0.100
NON	45	(77.6%)	88	(65.7%)	
DéMI par la méthode SEM					
OUI	20	(34.5%)	63	(47.0%)	p*= 0.108
NON	38	(65.5%)	71	(53.0%)	

*test de chi-2

Comme le montre les deux tableaux ci-dessus, pour les deux instruments de mesure d'utilité et quelle que soit la méthode d'estimation du DéMI, le fait d'avoir un changement significatif pour le donneur est indépendant des complications survenues chez son receveur.

Les résultats ont montré une diminution significative de l'utilité à court terme, avec une proportion non négligeable de donneurs au seuil DéMI. Il serait intéressant de compléter ce travail par une analyse des facteurs, liés au donneur ou son receveur, influençant une détérioration significative.

6.1. QUATRIÈME TRAVAIL

Decrease in health state utility among living kidney donors

Hosseini Kossar, Jacques Hubert, Marc Ladrière, Guillemin Francis

Soumis

Decrease in health state utility among living kidney donors

Authors: Hosseini Kossar ^{a,b,c}, Jacques Hubert^d, Marc Ladrière^e, Guillemin Francis ^{a,b,c}

Affiliation

^a INSERM, CIC-EC CIE6, Nancy, France

^b CHU Nancy, Clinical Epidemiology and Evaluation, Nancy, France

^c Université de Lorraine, University Paris Descartes, EA 4360 Apemac, Nancy, France

^d CHU Nancy, Department of Urology, Nancy, France

^e CHU Nancy, Department of Nephrology, Nancy, France

Corresponding author

Kossar Hosseini

Centre d'Epidémiologie Clinique

Inserm CIC-EC

Service d'Epidémiologie et Evaluation Cliniques

CHU de Nancy

Hopitaux de Brabois, Allée du Morvan

54 505 Vandoeuvre-lès-Nancy

France

Tel.: +33(0)3 83 85 93 02

Fax : +33(0)3 83 85 12 05

Email : k.hosseini@chu-nancy.fr

Key words: Kidney donors, health related quality of life, utility, minimal important difference

Word count of text: 2474 (2700 max)

Word count of Abstract: 245 (250max)

Abstract

Background: Changes in the utility of donors in the post-transplantation period have been little investigated. For kidney donors, we aimed to determine the change in utility from before to 3 months after donation and 2) To determine the minimal important decrease (MIDe) in health utility.

Methods: We used data from a prospective multicenter observational study measuring the health-related quality of life of kidney donors before and 3 months after transplantation. Utility scores were estimated by the SF-6D and EQ-5D. Two methods were used to estimate the MIDe: the anchor-based method by the Global Rating of Change (GRoC) for donors rating their global health “somewhat worse” 3 months after donation and the distribution-based method by the standard error of measurement (SEM).

Results: 211 and 192 donors completed the EQ-5D and SF-6D, respectively. Utility score was significantly decreased at 3 months ($p < .0001$). Mean (SD) utility score before transplantation was 0.932 (± 0.091) and 0.834 (± 0.085) with the EQ-5D and SF-6D, respectively; mean scores were 0.882 (± 0.151) and 0.757 (± 0.118), respectively, 3 months after donation. Half of donors (53.9%, $n=123$) rated their global health “about the same” and 15.4% ($n=35$) “somewhat worse” at follow-up. MIDe was -0.113 and -0.128 for the EQ-5D and SF-6D, respectively, by the GRoC and -0.076 and -0.073, respectively, by SEM method.

Conclusion: Health utility scores for kidney donors were significantly lower at 3 months with both utility measurements. About one third of donors experience a significant decrease in health utility according to the MIDe.

Introduction

The concept of utility in health was developed by economists to explain individual choices and preferences for health states (1). Torrance (réf 2 p1) expressed utility as “desirability or preference that individuals exhibit for the condition”. The unit value of health state utility ranges between 0, representing “death” state, and 1 representing “perfect health” state. Some instruments allow a negative value corresponding to state perceived as “worse than death”. Indirect multi-attribute utility (2) measures are estimated from generic instruments to which one applies the weight assigned by the general population for each particular health state. Utility allows for comparing different treatment strategies for one or more diseases by a single unit of measurement and may be used in medicoeconomic evaluations (3). The main objective of cost-utility analysis is to relate cost changes to utility changes across treatment strategies and define the incremental cost-utility ratio, further used in decision making. This analysis can provide better information to decision makers by taking into account patient and society views in the analysis.

The interpretation of health utility changes can be based on the concept of Minimal Important Difference (MID). This concept was developed by Jaeschke et al. (4) as “the smallest difference in score in the domain of interest which patients would perceive as beneficial and which would mandate, in the absence of troublesome side-effects and excessive cost, a change in the patient’s management”. Such MID helps calibrate the magnitude of observed changes. It could be useful to interpret changes in utility scores over time induced by a strategy (within group) or compare changes between strategies (between groups). The most common methods for estimating the MID of health-related quality of life (HRQoL) instruments are distribution-based and anchor-based (5;6).

In living kidney donors, generic HRQoL instruments have been used to estimate the short- and long-term effects of kidney transplantation. Studies with long-term follow-up after transplantation have shown that donors have, on average, the same or better HRQoL as healthy people (7-9) or the general population (10;11). Most donors had a positive experience or experienced no changes in HRQoL (12), but some experience a decrease in quality of life, especially regarding psychological functioning (8;13;14). To our knowledge, changes in the HRQoL of donors in the post-transplantation period have been little investigated. Since donors are considered healthy before donation, we may expect impaired HRQoL and associated health utility measure in the short-term. Therefore, the minimal important decrease of utility is useful to determine where such decrease is meaningful.

Repeated measurement of HRQoL over time and subsequent utility change can be affected by a response-shift (RS) phenomenon, defined by Sprangers and Schwartz (15) as a change in an individual's health that can alter the individual's standards, values or conceptualization. RS can occur in any situation involving self-reported outcomes. In living donors, it may bias the results because donors may continue to report a high HRQoL after the intervention, in part because of adaptation to their condition, which affects the magnitude of observed utility change. In this way, an RS phenomenon may affect the change in utility.

The purpose of this study was to determine 1) the change in utility for kidney donors between before and 3 months after donation and 2) to determine the Minimal Important Decrease (MIDe) in utility for donors.

Methods

Kidney donors

We used data from donors included in the DOVIREIN study from June 2010 to November 2012. The DOVIREIN study is an observational multicenter prospective study measuring quality of life of living kidney donors according to surgical techniques used for kidney sampling. The study started in June 2010 and was approved by the institutional review board (CCTIRS notification OMG/JD/10.033; CNIL authorization 910068, ClinicalTrials.gov no. 10.039). From 22 public hospitals covering the whole French territory, all donor–receiver couples were invited to participate and were included after they gave their written informed consent before transplantation surgery.

Demographics and clinical data were collected before the transplantation. The HRQoL of donors, measured by generic instruments (the EQ-5D of Euroqol (16) and the Medical Outcomes Study 36-item Short Form [SF-36] (17)), was collected the day before the transplantation and at 3-month follow-up.

Utility measurement

The SF-6D (18) was developed from 11 questions of the SF-36. It is a multidimensional health classification system assessing the 6 domains of physical functioning, role limitation, social functioning, pain, mental health, and vitality, with 4 to 6 levels for each domain. An SF-6D health state is defined by 1 level from each domain, which results in a total of 18,000 possible health states. The SF-6D scoring algorithm was developed with the standard gamble method (19) from a sample of 249 SF-6D health states from a representative sample of the

UK population. Utility scores generated by the SF-6D range from 0.29 to 1.00, with 1.00 representing full health and 0.29 representing the worst possible health state defined by the SF-6D (i.e., all domains being at the worst level) (19).

The EQ-5D consists of 5 items or dimensions (mobility, self-care, usual activity, pain/discomfort, and anxiety/depression) with 3 answer categories: no problem (1), some problems (2), or severe problems (3), for 243 possible health states described as vectors ranging from 11111 (no problems in any domain) to 33333 (severe problems in all 5 domains). The EQ-5D can be reported in terms of a five-digit profile indicating the level on each dimension or a preference-based single number. The latter is obtained by applying algorithms that link the five-digit health state description with average values obtained for the general population in France (20) using the time trade-off method. This algorithm yields a range from -0.50 to 1, with 1.00 representing “full health”, 0 “death” and negative value a health state labelled “worse than death” .

MIDe estimation

We derived the minimal important decrease (MIDe) for each instrument using two methods: anchor-based and distribution-based methods (21).

For the anchor-based method, we used the Global Rating of Change (GRoC) scale with 7 levels, asking donors at the end of follow-up how they rated their general health as compared with 3 months ago: “much better” (1), “better” (2), “somewhat better” (3), “about the same” (4), “somewhat worse” (5), “worse” (6) or “much worse” (7). Given the context of organ donation from living kidney donors, we assumed that a response of “somewhat worse” represents a meaningful change for the donor. The MIDe of utility was the mean change score between baseline and follow-up for donors rating their health “somewhat worse” (4;22).

For the distribution-based method, MIDe was defined as one standard error of measurement (SEM) (23), which is estimated in the entire sample by the standard deviation of the instrument at baseline (σ_x) multiplied by the square root of 1 minus the intraclass correlation coefficient ($ICC_{1,2}$) between the utility scores measured before transplantation (t1) and 3

months after (t2): $SEM = \sigma_x * \sqrt{1 - ICC_{(t1,t2)}}$

For the interpretation of MIDE in the context of living kidney donors, data are presented as negative values. Only donors with a negative change equal or greater than the MIDE were considered to have a meaningful decrease in utility score.

Response-Shift consideration

The most commonly used approach for detecting change in internal standards or “recalibration response shift” (15) is the retrospective pre-test–post-test design (24), also known as the “then-test”. At the end of follow-up, the SF-36 then-test was administered after donors completed the SF-36 post-test questionnaire. Donors were asked to think back to how they felt at the start of the study (before donation) and answer retrospectively. A significant difference between the then-test and pre-test (SF-6D) mean utility score was considered an RS effect.

Statistical analysis

The sociodemographic characteristics of donors and utility scores by the SF-6D and EQ-5D were described with means (SD) and numbers (percentages). The mean utility scores at baseline and 3 months as well as at baseline and the then-test were compared by paired *t* test. Because global measures of change are expected to be correlated with the 3-month and not the initial state(6) (25), we calculated Spearman rank-order (ρ) correlations for the GROC scale with baseline and 3-month utility scores to ensure the validity of the scale.

The mean (SD) difference in utility score was described for each level of the GROC scale by each instrument.

For the anchor-based method to estimate the MIDE, we used the mean (95% confidence interval) change in utility for donors who rated their health as “somewhat worse” at 3 months. For the distribution-based method, we calculated the standard deviation of the mean utility at baseline and the intraclass correlation coefficient (95% confidence interval) estimated between baseline and 3-month utility scores by ANOVA for SEM calculation.

A $p < 0.05$ was considered statistically significant. Analyses involved SAS 9.3 (SAS Inst, Cary, NC, USA).

Results

For the 228 donors included, the mean (SD) age was 49.6 (10.7), and 59.2% were women (Table 1). In total, 211 and 192 donors completed the EQ-5D and SF-6D, respectively, at 3-month follow up, and 141 completed the SF-6D then-test questionnaire (Figure 1).

Utility measurement

The mean (SD) utility score before transplantation was 0.932 (0.091; range 0.487–1.00) with the EQ-5D, and 0.834 (0.085; range 0.579–1.00) with the SF-6D (Table 2).

For the 197 donors responding to the GROC scale, 123 (53.9%) considered their global health “about the same”, and 35 (15.4%) “somewhat worse” 3 months after donation (Table 3). Nine (3.9%) donors rated it “much better” and only 2 (0.9%) “much worse”.

We found a significant negative correlation between the GROC and utility scores with the SF-6D ($\rho = -0.295$; $p < .0001$) and EQ-5D ($\rho = -0.198$; $p = 0.0064$) at follow-up: donors with a worse health state at 3 months had decreased utility score. No significant correlation was observed with baseline scores.

Change in utility and response-shift

Table 2 shows the mean utility score for donors at each time. At all times, utility scores were higher by the EQ-5D than the SF-6D. Mean utility score decreased between before and after the donation by each instrument ($p < .0001$).

The mean utility score using the then-test for the SF-6D was 0.831 (± 0.094), with no significant difference in mean score utility between before donation and the then-test ($p = 0.50$).

MIDe estimation

Using the anchor-based method, the MIDe of a worsening change in health state was -0.113 and -0.128 with the EQ-5D and SF-6D, respectively (Table 4). With the EQ-5D, 18.5% of donors reached the MIDe threshold and 30.7% with the SF-6D.

The MIDe was lower by distribution-based method: -0.076 and -0.073 with the EQ-5D and SF-6D, respectively; 34.6% and 43.2% of donors reached the MID threshold with the EQ-5D and SF-6D, respectively.

Discussion

Results showed a significant decrease in utility scores for kidney donors after three months donation, and a related MIDe between instruments when using the same method. As expected (25), we found a significant correlation between the GROC and utility scores at follow-up, which suggests a decrease in utility score when donors report worse general health and no significant correlation between the GROC and utility scores at baseline.

Strengths

We used data from a representative sample of French kidney donors. Our study is the first to estimate a threshold of MIDE for living kidney donors using 2 generic and popular instruments. We found no significant RS by the SF-6D then-test, which suggested no change in an individual's standards, values or conceptualization after kidney donation.

Our goal was to define a threshold for the clinician to identify donors with a meaningful difference in utility score. Thus, we chose the short-term estimation for utility and MIDE, when other studies focused on the long-term (>1 year) estimation. The short-term estimation better reflects the incidence of surgical and medical complications in donors. From a literature review, Parikh et al. (26) concluded that living donors returned to their baseline QoL within 6 months.

Given that consensus is lacking on determining the MID (6), we chose to use the anchor-based method with a patient-based indicator and the SEM as the only distribution-based method.

Limitations

We controlled the RS phenomenon only for the SF-6D questionnaire. Compared to the EQ-5D, the SF-6D has more dimensions and more levels (variation in descriptive system). Hence, the utility reported by the EQ-5D then-test may be different from that measured by the SF-6D and lead to a different result for the RS.

Our goal was to define an MIDE for donors and not compare MIDE estimates using different methods, so we used only the SEM as a distribution-based method.

Results vary between utility instruments for 2 reasons. First, when using the anchor-based method, MIDE was estimated only for donors who rated their health "somewhat worse", whereas the distribution-based method included all donors. According to classical test theory, Wyrwich et al. (27) considered the SEM as sample-independent and better for determining a meaningful change as compared with other distribution-based methods for MID estimation such as the effect size. For the interpretation, MIDE was expressed as a negative value with both methods. A negative value corresponds to a decrease in utility change beyond which donor health status may be considered meaningfully deteriorated. The absolute value of the MIDE by the SEM method corresponds to the minimal difference reflecting a meaningful change (positive or negative) for donors. By the anchor-based approach, the MIDE with both

utility instruments was larger than results found by Walters and Brazier (28) using data from 8 longitudinal studies in 11 patient groups, which suggests that the context differs for living donors as compared with other patients. Second, the method used to estimate utility score differed between both HRQoL instruments: health state estimates with the EQ-5D were derived by the time trade-off method and with the SF-6D by the standard gamble method. The standard gamble in general overestimates the utility, whereas the time trade-off method can under- or overestimate the utility (29).

In the short term, donors overestimated their global health. A global health state “about the same”, “somewhat better” or “better” was associated with a negative utility score change with both HRQoL instruments. When the MIDE was applied to our sample, a high proportion of donors reached the threshold. The decrease in utility may be explained by donors not having entirely recovered physically at 3 months after transplantation, but they probably already experienced enhanced self-esteem, which led them to feel better.

In perspective, the MIDE can be expressed as a threshold to identify donors with a significant decrease in health state in terms of clinical management and surgical technique. Further work is necessary to explain the factors leading donors to reach the threshold (such as clinical condition of the recipient) and apply the MIDE in medico-economics study to determine how it can discriminate different strategies of kidney removal.

In conclusion, we report a threshold for a significant decrease in health utility score in the short term for kidney donors that can be used for donor management after transplantation. Further investigation is necessary to examine whether the utility score and the MIDE vary beyond 3 months

Acknowledgements

The study was coordinated by INSERM CIC-EC at Nancy University hospital and involved in centers in France. We thank all the investigators and research nurses for data collection and verification and Juliette Oudot for study organisation. We also thank the QV-DVR study especially for collecting data on quality of life.

Funding

This project was funded by the Ministry of Health (DHOS STIC 2009). Opinions expressed in the present article are those of the authors and do not necessarily reflect those of the sponsors. The study sponsors did not take part in the study design, collection, analysis and interpretation of data, writing of the report or decision to submit the article for publication

Table 1. Sociodemographic characteristics of kidney donors (n=228)

Variables	N (%)*
Age (years), mean (SD)	49.6 (10.7)
Sex	
Male	93 (40.8)
Female	135 (59.2)
Married or living together	182 (79.8)
With children	189 (82.9)
Family degree relation with the recipient	
Father	27 (11.8)
Mother	32 (14.0)
Brother	31 (13.6)
Sister	45 (19.7)
Partner	69 (30.3)
Son	7 (3.1)
Daughter	1 (0.4)
Other	16 (7.0)
Occupation	
Farmer	7 (3.1)
Manager	16 (7.0)
Executive	40 (17.5)
Craftman	17 (7.5)
Employee	83 (36.4)
Worker	21 (9.2)
Not working	2 (0.9)
Student	6 (2.6)
Other	35 (15.4)

*Data are number (%) unless indicated.

Table 2. Mean utility score with the EQ-5D and SF-6D health-related quality of life (HRQoL)

instruments at each measurement time

Utility measurement	Before donation	3 months after donation	Then-test 3 months after donation
EQ-5D			
No.	222	212	
Mean (SD)	0.932 (0.091)	0.882 (0.151)	
Median (min–max)	1.000 (0.487–1.000)	0.910 (0.127–1.000)	
p ¹	--	<0.0001	
SF-6D			
No.	211	204	150
Mean (SD)	0.834 (0.085)	0.757 (0.118)	0.831 (0.094)
Median (min–max)	0.845 (0.579–1.000)	0.784 (0.471–1.000)	0.845 (0.568–1.000)
P value ¹	--	<0.0001	0.50

¹ Paired *t* test, compared to before donation

Table 3. Number of donors and magnitude of change in health utility score with each HRQoL instrument by category of Global Rating of Change scale.

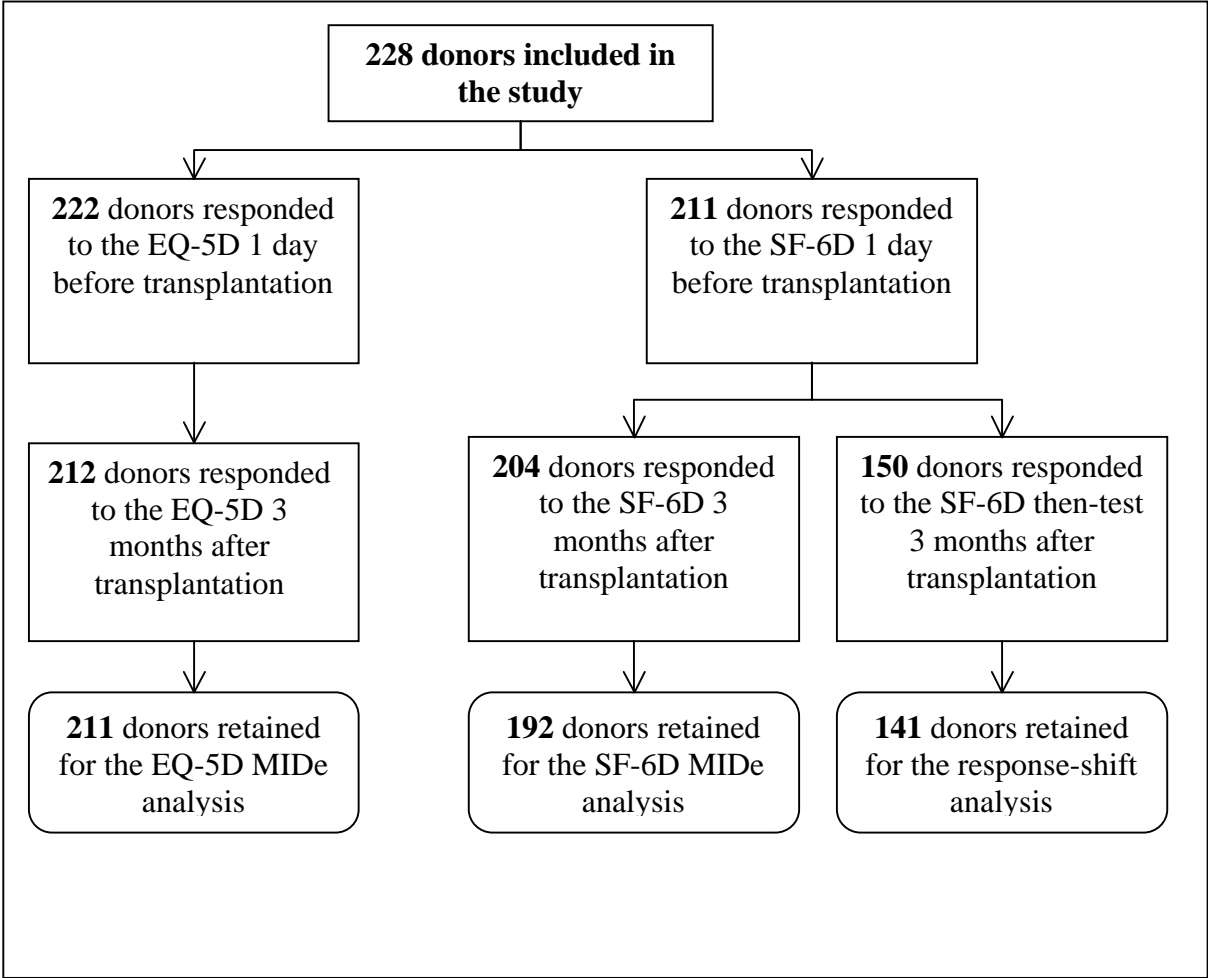
Global Rating of Change scale (n=197)							
Instrument	Much better	Better	Somewhat better	About the same	Somewhat worse	Worse	Much worse
EQ-5D							
No.	8	13	8	120	31	6	2
Mean (SD)	0.040 (0.301)	-0.046 (0.145)	-0.072 (0.144)	-0.025 (0.146)	-0.113 (0.149)	-0.056 (0.100)	-0.421 (0.088)
SF-6D							
N	8	12	6	111	28	5	2
Mean (SD)	0.041 (0.125)	-0.075 (0.131)	0.007 (0.146)	-0.059 (0.125)	-0.128 (0.116)	-0.142 (0.146)	-0.347 (0.055)

Table 4. MIDE assessed for each HRQoL instrument by the anchor-based and distribution-based methods.

Instrument	MIDE by anchor-based method ¹			MIDE distribution-based method ²			
	No.	Mean change (95% CI) ³	SD	No.	SD at baseline	ICC ⁴ (95% CI)	SEM
EQ-5D	31	-0.113 (-0.167 ; -0.058)	0.149	211	0.091	0.308 (0.184 ; 0.423)	-0.076
SF-6D	28	-0.128 (-0.173 ; -0.083)	0.116	192	0.085	0.267 (0.135 ; 0.390)	-0.073

¹Anchor-based method with the Global Rating of Change scale (mean score change); ²Distribution-based method with the standard error of measurement (SEM); ³95% CI, 95% confidence interval; ⁴ICC, intraclass correlation coefficient

Figure 1. Donors with completed questionnaires at baseline and follow-up



Reference List

- (1) Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987;40(6):593-603.
- (2) Petrou S, Hockley C. An investigation into the empirical validity of the EQ-5D and SF-6D based on hypothetical preferences in a general population. *Health Econ* 2005 Nov;14(11):1169-89.
- (3) Beresniak A, Russell AS, Haraoui B, Bessette L, Bombardier C, Duru G. Advantages and limitations of utility assessment methods in rheumatoid arthritis. *J Rheumatol* 2007 Nov;34(11):2193-200.
- (4) Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989 Dec;10(4):407-15.
- (5) Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002 Apr;77(4):371-83.
- (6) Norman GR, Sridhar FG, Guyatt GH, Walter SD. Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care* 2001 Oct;39(10):1039-47.
- (7) Clemens KK, Thiessen-Philbrook H, Parikh CR, Yang RC, Karley ML, Boudville N, et al. Psychosocial health of living kidney donors: a systematic review. *Am J Transplant* 2006 Dec;6(12):2965-77.
- (8) Johnson EM, Anderson JK, Jacobs C, Suh G, Humar A, Suhr BD, et al. Long-term follow-up of living kidney donors: quality of life after donation. *Transplantation* 1999 Mar 15;67(5):717-21.
- (9) Ku JH. Health-related quality of life of living kidney donors: review of the short form 36-health questionnaire survey. *Transpl Int* 2005 Dec;18(12):1309-17.
- (10) de Groot IB, Stiggelbout AM, van der Boog PJ, Baranski AG, Marang-van de Mheen PJ. Reduced quality of life in living kidney donors: association with fatigue, societal participation and pre-donation variables. *Transpl Int* 2012 Sep;25(9):967-75.
- (11) Maglakelidze N, Pantsulaia T, Managadze L, Chkhotua A. Assessment of health-related quality of life in living kidney donors. *Transplant Proc* 2011 Jan;43(1):373-5.
- (12) Garcia MF, Andrade LG, Carvalho MF. Living kidney donors--a prospective study of quality of life before and after kidney donation. *Clin Transplant* 2013 Jan;27(1):9-14.

- (13) Reimer J, Rensing A, Haasen C, Philipp T, Pietruck F, Franke GH. The impact of living-related kidney transplantation on the donor's life. *Transplantation* 2006 May 15;81(9):1268-73.
- (14) Smith GC, Trauer T, Kerr PG, Chadban SJ. Prospective psychosocial monitoring of living kidney donors using the Short Form-36 health survey: results at 12 months. *Transplantation* 2004 Nov 15;78(9):1384-9.
- (15) Sprangers MA, Schwartz CE. The challenge of response shift for quality-of-life-based clinical oncology research. *Ann Oncol* 1999 Jul;10(7):747-9.
- (16) EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy* 1990 Dec;16(3):199-208.
- (17) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992 Jun;30(6):473-83.
- (18) 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 Nov;51(11):1115-28.
- (19) Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002 Mar;21(2):271-92.
- (20) Chevalier J, de Pouvourville G. Valuing EQ-5D using time trade-off in France. *Eur J Health Econ* 2013 Feb;14(1):57-66.
- (21) Copay AG, Subach BR, Glassman SD, Polly DW, Jr., Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007 Sep;7(5):541-6.
- (22) Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 1994 Jan;47(1):81-7.
- (23) Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care* 1999 May;37(5):469-78.
- (24) Sprangers MA, Van Dam FS, Broersen J, Lodder L, Wever L, Visser MR, et al. Revealing response shift in longitudinal research on fatigue--the use of the thentest approach. *Acta Oncol* 1999;38(6):709-18.
- (25) Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008 Feb;61(2):102-9.

- (26) Parikh ND, Ladner D, Abecassis M, Butt Z. Quality of life for donors after living donor liver transplantation: a review of the literature. *Liver Transpl* 2010 Dec;16(12):1352-8.
- (27) Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999 Sep;52(9):861-73.
- (28) Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 2005 Aug;14(6):1523-32.
- (29) Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. *J Health Econ* 1996 Apr;15(2):209-31.

7. DISCUSSION

Ce travail avait pour objectif d'estimer un seuil de changement pour l'utilité dérivée des instruments psychométriques et de déterminer comment ce changement peut être influencé par la présence de comorbidités ou un phénomène de response-shift.

Nous avons estimé un seuil de changement pour l'utilité mesurée par les instruments génériques de qualité de vie dans deux groupes de sujet; les donneurs vivants de rein et les personnes avec une arthrose des membres inférieurs. Dans le premier groupe l'utilité a été estimée par les questionnaires EQ-5D et SF-6D, et dans le second uniquement par le SF-6D. La chronologie de la recherche par rapport aux études en cours au sein du laboratoire de recherche nous a permis de prendre en compte la présence des comorbidités dans l'échantillon de patients avec une arthrose, et de rechercher la présence du response-shift seulement chez les donneurs vivants de rein.

Si les comorbidités peuvent avoir un effet sur la valorisation de l'utilité à un temps donné, celui-ci peut devenir complètement marginal ou même inexistant lorsqu'on observe le changement sur le long terme. A travers notre cohorte d'individus avec une arthrose, nous avons pu observer la prédominance de la pathologie principale sur l'amplitude du changement par rapport aux autres comorbidités. L'évaluation des comorbidités a été effectuée à partir d'un index basé sur le retentissement des maladies sur la dimension physique de la qualité de vie, et n'intègre donc pas la notion de pondération de la gravité de la maladie. Il est important de préciser qu'au-delà des résultats obtenus dans nos travaux, il serait intéressant d'étudier l'impact de la gravité des autres maladies présentes chez un même individu dans l'interprétation du changement d'utilité.

L'estimation du seuil de changement dépend de la méthode appliquée (70 ;71). Comme nous avons pu le voir dans le cas du donneur vivant, le seuil estimé peut être très différent selon la méthode utilisée. Il n'existe pas à ce jour de standard pour la mesure de ce seuil, et il est recommandé d'appliquer les deux méthodes basées sur une référence externe et la distribution de l'échantillon afin de donner un intervalle pour l'interprétation de ce changement (72). Cependant certains auteurs font bien la distinction selon la méthode d'estimation du seuil de changement utilisée, de ce qui peut être traduit comme changement minimal ou bien changement significatif. On considère qu'un seuil défini à partir d'une

référence externe basée sur le point de vue des patients permet d'établir un changement significatif au sens clinique, alors que les mesures basées sur la distribution de l'échantillon renvoient plutôt à la notion de changement minimal détectable (72). Une des principales critiques que l'on fait à la méthode basée sur la distribution est de ne pas tenir compte du point de vue du patient sur l'évolution de sa santé.

Cependant, la méthode de référence externe est aussi critiquable, notamment du fait de l'utilisation d'une question de santé globale (GROc). Celle-ci permet au patient d'évaluer aujourd'hui son état de santé du passé. Ainsi cette mesure ne peut pas être exempte d'un biais de mémoire et ou d'un phénomène de response-shift (38). Ces deux points sont à prendre en considération dès lors que l'appréciation de la variation se fait sur une longue période de temps. La mémorisation d'un état de santé antérieur peut être influencée par des événements récents ou frappants, mais aussi par l'état de la personne au moment de répondre à la question. La présence du phénomène de response-shift pourrait se traduire par une re-conceptualisation et re-priorisation des valeurs pour le patient qui l'amèneront à changer sa vision de sa maladie. Les deux travaux initiaux sur cette méthode ont utilisés un délai de 2 semaines (34) ou 4 semaines (40) entre la mesure initiale et l'évaluation du changement. Bien que la détermination du délai soit importante, il faut également tenir compte des événements marquants qui peuvent se produire dans le temps pour un individu (changement de traitement, chirurgie, aggravation marquante de la maladie).

Une autre critique que l'on peut faire à cette méthode est la précision de l'échelle utilisée (35;73). La validité de cette échelle n'est pas démontrée et la précision du résultat obtenu dépend du nombre de modalités de réponse mais aussi de l'expression de celles-ci (35). Cette méthode ne tient donc pas compte de la variabilité de la mesure intra individu.

Nous avons appliqué cette méthode de référence externe pour estimer le changement chez les donneurs vivants de rein sur une période de 3 mois. Nous avons pu constater l'absence du phénomène de response-shift sur cet échantillon. Cependant, une détérioration de l'utilité a été observée parmi les donneurs reportant leur état de santé inchangé à 3 mois, et les individus avaient tendance à surévaluer leur utilité. Dans le contexte particulier du don du rein qui s'apparente pour l'individu à un acte altruiste, ce décalage entre le point de vue du donneur et la variation observée pourrait traduire surtout la manifestation d'une plus grande estime de soi suite au don. Pour les deux instruments de mesure utilisés, le seuil estimé par

cette méthode était bien plus élevée que celui obtenu par la méthode basée sur la distribution de l'échantillon.

Pour l'application de la méthode basée sur la distribution nous avons fait le choix d'utiliser l'erreur standard de mesure (SEM) (74) qui prend en compte la possibilité qu'une partie du changement soit due à une erreur de mesure aléatoire (75). Elle est donc une expression de la précision de la mesure tout étant indépendante des caractéristiques de l'échantillon (76) et doit rester relativement semblable entre différents échantillons. En appliquant cette méthode dans nos deux échantillons, nous avons estimé un seuil de changement similaire pour l'utilité mesurée par le SF-6D.

Au-delà de la discordance observée pour le seuil estimé selon la méthode, on peut aussi dans le cadre de l'étude sur le donneur vivant revenir sur celle existante entre les deux instruments de mesure d'utilité. Sur un même échantillon de patients et en appliquant la même méthode d'estimation de seuil nous avons obtenu des résultats sensiblement différents pour le SF-6D et l'EQ-5D. On peut identifier clairement deux raisons à cette discordance, qui découle d'une différence d'échelle de mesure et de la valorisation des pondérations pour les deux instruments (77). D'une part ces deux instruments sont différents dans leur système descriptif de la santé. Le SF-6D donne plus de possibilités de réponses à l'individu et pourrait ainsi mieux capter une petite variation de l'état de santé. La valorisation des pondérations sociétales des états de santé diffère entre les deux instruments : le marchandage temps (TTO) a été utilisé pour l'EQ-5D et le jeu de loteries (SG) pour le SF-6D. Ces méthodes de valorisation ont une conception théorique distincte qui se traduit par une différence de l'utilité estimée pour un même échantillon d'individu (15;30;78-80). Cette différence se répercute sur la valorisation des états de santé, mais aussi sur la prise de décision dans le cadre d'analyses médico-économiques (81-83).

A l'inverse des autres études (84-86), nous avons observé dans notre échantillon de donneurs vivants en moyenne une utilité estimée par l'EQ-5D plus élevée que celle par le SF-6D. Cependant, l'amplitude du changement observé était en moyenne plus élevée par le SF-6D, rappelant ainsi que cet instrument peut être plus sensible à une variation.

La notion de changement que nous avons appliqué à la mesure d'utilité vient du domaine clinique et de la nécessité d'établir un seuil pertinent du changement en complément d'une significativité statistique. L'interprétation du changement dans le domaine clinique s'adresse bien évidemment en premier lieu au clinicien et pour la prise en charge de la maladie. Cette démarche permet de revenir vers l'individu et de se défaire de la complexité de l'analyse statistique qui impose le point de vue du groupe. Il existe un seuil d'interprétation pour le changement de qualité de vie mesurée à partir des questionnaires génériques (comme le SF-36). Ce seuil peut s'appliquer au score obtenu pour chaque dimension de la qualité. Cependant du fait de la modification métrique et d'application des pondérations sociétales, ce seuil n'est pas transposable à l'utilité dérivée de la qualité de vie.

La mesure d'utilité permet d'exprimer les préférences des individus pour un état de santé donné dans un contexte propre à l'individu. L'estimation du seuil de changement de l'utilité va donc dépendre de ce contexte et sera différente selon la maladie et sa gravité dans un échantillon de patients (87). On peut ainsi dire qu'il existe des seuils de changement pour une utilité estimée selon la population concernée et le travail de Walters et Brazier (44) sur l'estimation d'un seuil d'utilité sur différents échantillons de patients résume bien ce point (MID estimé pour 11 échantillons de patients varie de 0.011 à 0.097 pour le SF-6D et de -0.011 à 0.140 pour l'EQ-5D).

Comme évoqué, l'utilisation des instruments de mesure psychométrique de la qualité de vie permet de dériver une utilité reflétant les préférences de l'individu pour un état de santé. Ce passage de la mesure de qualité de vie à l'expression des préférences de l'individu se fait par une étape de pondération par les préférences sociétales. De ce fait, l'utilité indirecte n'est pas le reflet des préférences de l'individu seul, et intègre en partie le point de la société. On peut ainsi évoquer pour le changement d'utilité un seuil socialement significatif. L'utilité estimée a globalement un impératif comme objectif : permettre de comparer des stratégies diagnostiques et thérapeutiques dans un but décisionnel selon la valeur accordée par la société à celles-ci. Dans ce cadre, cette notion de seuil socialement significatif prend son sens, car permet de différencier les individus avec une amélioration ou une détérioration.

La définition et l'interprétation du seuil doivent s'adapter au contexte. Et nous avons pu voir que l'expression du seuil varie lorsqu'on est sur un échantillon de donneurs vivants de rein. La particularité du contexte nous a amené à déterminer un seuil uniquement pour la détérioration de l'utilité ; les individus étant en bonne santé avant le don, l'intérêt principal

était de voir la proportion de détériorations. Dans ce contexte particulier, le changement d'utilité ne s'interprète plus comme socialement significatif mais comme socialement acceptable : au sens sociétal, on accepte et on tolère une détérioration provisoire de l'utilité pour les donateurs en espérant un gain futur pour la société.

Comme nous l'avons évoqué, la notion d'utilité en santé a son application dans l'analyse médico-économique en rapprochant les coûts des stratégies et leurs efficacités exprimées par les préférences des individus. L'approche coût-utilité, qui tend à devenir la méthode de référence des analyses médico-économiques, a conduit à l'intégration d'un indicateur unique qui est l'année de vie ajustée sur la qualité (QALY). Ce concept permet de pondérer la quantité de vie (durée de vie restante) par un coefficient rendant compte de sa qualité (88;89). En théorie la détermination du QALY permet de comparer des actions de santé quelle que soit la pathologie concernée. Comme le souligne Weinstein et al. (90), le QALY rend compte d'un état de santé et ne donne pas d'indication sur le changement (variation) de celui-ci. Ce concept repose sur des hypothèses qui imposent que la relation entre la quantité et la qualité de vie soit constante au cours du temps et que l'utilité associée à un état de santé ne varie pas avec le temps. Cependant selon le contexte, l'horizon temporel étudié, et la survenue d'évènements associés à une maladie, la perception des individus sur leur état de santé peut être modifiée. On peut ainsi dire que la complexité de la mesure des préférences individuelles doit amener à être prudent sur la prise en compte de celles-ci dans les modèles d'évaluation économique. Et l'utilisation d'un seuil pour l'interprétation du changement peut être un élément complémentaire dans la démarche d'analyse décisionnelle.

Walters et Brazier (44) considèrent qu'un seuil de changement d'utilité défini pour une période d'un an peut être directement considéré comme un seuil d'interprétation du QALY. Bien évidemment dans une analyse médico-économique l'utilisation de ce seuil seul n'a pas d'intérêt et les résultats doivent intégrer la notion de coût.

Conclusion

Les instruments de mesure psychométrique de la qualité de vie permettent de dériver un score d'utilité, et facilitent ainsi l'appréciation des préférences individuelles. Cette approche indirecte de l'utilité est moins complexe à comprendre par les individus que l'approche directe, et peut être aisément répétée au cours du suivi des patients. La répétition de la

mesure des préférences conduit à une analyse de la variation de l'utilité pour les individus au cours du temps et à l'interprétation du changement observé. Dans ce contexte, la détermination d'un seuil s'avère nécessaire pour l'interprétation du changement d'utilité. La mesure de l'utilité indirecte implique une pondération des préférences individuelles par la valeur sociétale accordée à chaque état de santé. De ce fait la détermination du seuil de changement d'utilité indirect amène à exprimer un seuil de changement socialement significatif. Cependant, la démarche appliquée à cette fin utilise plusieurs méthodes et nous contraint à considérer un intervalle pour le changement et non pas un seuil unique. L'évaluation du changement d'utilité chez un patient doit tenir compte de l'existence des pathologies autres ainsi que l'adaptation du jugement de valeur de l'individu sur l'évolution de son état de santé.

8. PERSPECTIVES

L'enjeu principal de la détermination d'un seuil de changement pour l'utilité est son application dans l'analyse médico-économique.

Une première perspective à venir de ce travail sera d'utiliser le seuil de changement dans une analyse médico-économique comparant les différentes techniques de chirurgie de prélèvement de rein chez le donneur vivant. La comparaison inter-groupe nécessaire pour ce type d'analyse identifiera la quantité de changement observé selon la technique de chirurgie. En complément, le seuil d'utilité permettra de rendre compte de la proportion de donneurs avec une détérioration significative au sens sociétal dans chaque groupe, et apportera ainsi une information supplémentaire pour le décideur final pour l'allocation des ressources. Cette information aura d'autant plus d'importance si les différentes techniques de chirurgie pour le prélèvement de rein s'avèrent équivalentes sur l'efficacité clinique pour le donneur.

A partir de ce même échantillon de donneurs vivants nous pourrions analyser l'effet des facteurs sociodémographique et cliniques (du donneur et de son receveur) sur l'atteinte du seuil de changement et déterminer le profil des donneurs qui vivent une détérioration de l'utilité.

On peut également poser la question de la période pour la détermination du seuil. Comment varie ce seuil si la mesure est effectuée à court ou long terme ? Y a-t-il une modification du seuil chez les mêmes individus au cours du temps ? Et dans quelle mesure la survenue d'un événement marquant durant la maladie peut-elle modifier ce seuil ? Ces points pourront être considérés à travers les données de suivi de la cohorte KHOALA dont nous disposons maintenant.

Et enfin il serait intéressant de considérer le phénomène de response-shift dans un suivi de long terme et son impact sur la mesure du changement d'utilité. Est-ce que ce phénomène a le même effet tout au long de l'histoire de la maladie ? Ou peut-il s'amplifier ou au contraire s'atténuer avec l'évolution de la maladie et dans quelle mesure ?

Les perspectives énoncées ici doivent prendre en compte un point important de nos travaux : le recours aux instruments psychométriques pour la mesure de l'utilité. Bien que l'estimation de l'utilité indirecte soit admise pour l'application à l'analyse médico-économique, il ne faut pas oublier que le standard reste la mesure directe de l'utilité qui

répond le mieux aux exigences de la théorie économique pour la modélisation des préférences. Nous avons fait le choix d'étudier l'utilité indirecte, car comparée à la méthode directe elle permet de répéter plus aisément la mesure des préférences au cours du temps. Mais il serait pertinent de comparer le changement observé par les deux méthodes entre deux temps de mesure pour deux stratégies comparées (différence inter-stratégies des différences intra-groupes), et de mettre en rapport l'amplitude du changement observé par la différence d'utilité indirecte à celui mesuré par une méthode directe (différence inter-stratégies des valeurs dans chaque groupe).

*The difficulty lies, not in the new ideas,
but in escaping from the old ones.*

John Maynard Keynes

REFERENCES

- (1) Neumann Jv, Morgenstern O. *Theory of Games and Economic Behavior*. sec ed. 1947 ed. Princeton University Press; 1944.
- (2) Savage LJ. *The Foundations of Statistics*. Dover Publications; 1972.
- (3) Allais M. Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica* 1953 Oct;21(4):503-46.
- (4) Ellsberg D. Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics* 1961;75(4):643-69.
- (5) van Osch SM, Stiggelbout AM. The construction of standard gamble utilities. *Health Econ* 2008 Jan;17(1):31-40.
- (6) Treadwell JR, Lenert LA. Health values and prospect theory. *Med Decis Making* 1999 Jul;19(3):344-52.
- (7) Bakker C, van der LS. Health related utility measurement: an introduction. *J Rheumatol* 1995 Jun;22(6):1197-9.
- (8) Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. *Health Serv Res* 1972;7(2):118-33.
- (9) Drummond F, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press; 2005.
- (10) Kahneman D, Tversky A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 1979 Mar;47(2):263-92.
- (11) Abellan-Perpinan JM, Bleichrodt H, Pinto-Prades JL. The predictive validity of prospect theory versus expected utility in health utility measurement. *J Health Econ* 2009 Dec;28(6):1039-47.
- (12) Torrance GW. Toward a utility theory foundation for health status index models. *Health Serv Res* 1976;11(4):349-69.
- (13) Dolan P, Gudex C. Time preference, duration and health state valuations. *Health Econ* 1995 Jul;4(4):289-99.

- (14) Bleichrodt H, Johannesson M. Standard gamble, time trade-off and rating scale: experimental results on the ranking properties of QALYs. *J Health Econ* 1997 Apr;16(2):155-75.
- (15) Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health Econ* 2002 Jul;11(5):447-56.
- (16) Hornberger JC, Redelmeier DA, Petersen J. Variability among methods to assess patients' well-being and consequent effect on a cost-effectiveness analysis. *J Clin Epidemiol* 1992 May;45(5):505-12.
- (17) Lenert LA, Treadwell JR, Schwartz CE. Associations between health status and utilities implications for policy. *Med Care* 1999 May;37(5):479-89.
- (18) Krabbe PF, Essink-Bot ML, Bonsel GJ. The comparability and reliability of five health-state valuation methods. *Soc Sci Med* 1997 Dec;45(11):1641-52.
- (19) Read JL, Quinn RJ, Berwick DM, Fineberg HV, Weinstein MC. Preferences for health outcomes. Comparison of assessment methods. *Med Decis Making* 1984;4(3):315-29.
- (20) Stavem K. Quality of life in epilepsy: comparison of four preference measures. *Epilepsy Res* 1998 Feb;29(3):201-9.
- (21) Brauer CA, Rosen AB, Greenberg D, Neumann PJ. Trends in the measurement of health utilities in published cost-utility analyses. *Value Health* 2006 Jul;9(4):213-8.
- (22) Brazier J, Deverill M, Green C. A review of the use of health status measures in economic evaluation. *J Health Serv Res Policy* 1999 Jul;4(3):174-84.
- (23) EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. *Health Policy* 1990 Dec;16(3):199-208.
- (24) 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 Nov;51(11):1115-28.
- (25) Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems. Health Utilities Index. *Pharmacoeconomics* 1995 Jun;7(6):490-502.

- (26) Chevalier J, de Pouvourville G. Valuing EQ-5D using time trade-off in France. *Eur J Health Econ* 2013 Feb;14(1):57-66.
- (27) Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992 Jun;30(6):473-83.
- (28) Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002 Mar;21(2):271-92.
- (29) Amjadi SS, Maranian PM, Paulus HE, Kaplan RM, Ranganath VK, Furst DE, et al. Validating and assessing the sensitivity of the Health Assessment Questionnaire-Disability Index-derived Short Form-6D in patients with early aggressive rheumatoid arthritis. *J Rheumatol* 2009 Jun;36(6):1150-7.
- (30) Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. *J Health Econ* 1996 Apr;15(2):209-31.
- (31) Khanna D, Tsevat J. Health-related quality of life--an introduction. *Am J Manag Care* 2007 Dec;13 Suppl 9:S218-23.:S218-S223.
- (32) Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 1998 Jan;16(1):139-44.
- (33) Beaton DE, van ED, Smith P, van d, V, Cullen K, Kennedy CA, et al. Minimal change is sensitive, less specific to recovery: a diagnostic testing approach to interpretability. *J Clin Epidemiol* 2011 May;64(5):487-96.
- (34) Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989 Dec;10(4):407-15.
- (35) Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003 May;56(5):395-407.
- (36) Tubach F, Wells GA, Ravaud P, Dougados M. Minimal clinically important difference, low disease activity state, and patient acceptable symptom state: methodological issues. *J Rheumatol* 2005 Oct;32(10):2025-9.
- (37) Tubach F, Dougados M, Falissard B, Baron G, Logeart I, Ravaud P. Feeling good rather than feeling better matters more to patients. *Arthritis Rheum* 2006 Aug 15;55(4):526-30.

- (38) Copay AG, Subach BR, Glassman SD, Polly DW, Jr., Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007 Sep;7(5):541-6.
- (39) Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002 Apr;77(4):371-83.
- (40) Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 1994 Jan;47(1):81-7.
- (41) Norman G. Hi! How are you? Response shift, implicit theories and differing epistemologies. *Qual Life Res* 2003 May;12(3):239-49.
- (42) Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. L. Erlbaum Associates; 1988.
- (43) Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999 Sep;52(9):861-73.
- (44) Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 2005 Aug;14(6):1523-32.
- (45) Walters SJ, Brazier JE. What is the relationship between the minimally important difference and health. *Health Qual Life Outcomes* 2003 Apr 11;1:4.:4.
- (46) Feinstein AR. The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Diseases* 1970 Dec;23(7):455-68.
- (47) de Groot IB, Stiggelbout AM, van der Boog PJ, Baranski AG, Marang-van de Mheen PJ. Reduced quality of life in living kidney donors: association with fatigue, societal participation and pre-donation variables. *Transpl Int* 2012 Sep;25(9):967-75.
- (48) Redelmeier DA, Tan SH, Booth GL. The treatment of unrelated disorders in patients with chronic medical diseases. *N Engl J Med* 1998 May 21;338(21):1516-20.

- (49) Fortin M, Lapointe L, Hudon C, Vanasse A, Ntetu AL, Maltais D. Multimorbidity and quality of life in primary care: a systematic review. *Health Qual Life Outcomes* 2004;2:51.
- (50) Fortin PR, Clarke AE, Joseph L, Liang MH, Tanzer M, Ferland D, et al. Outcomes of total hip and knee replacement: preoperative functional status predicts outcomes at six months after surgery. *Arthritis Rheum* 1999 Aug;42(8):1722-8.
- (51) Lozano LM, Segur JM, Macule F, Nunez M, Torner P, Castillo F, et al. Intramedullary versus extramedullary tibial cutting guide in severely obese patients undergoing total knee replacement: a randomized study of 70 patients with body mass index >35 kg/m². *Obes Surg* 2008 Dec;18(12):1599-604.
- (52) Taylor AW, Price K, Gill TK, Adams R, Pilkington R, Carrangis N, et al. Multimorbidity - not just an older person's issue. Results from an Australian biomedical study. *BMC Public Health* 2010;10:718.
- (53) Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40(5):373-83.
- (54) Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. *J Am Geriatr Soc* 1968 May;16(5):622-6.
- (55) Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. *J Clin Epidemiol* 2005 Jun;58(6):595-602.
- (56) Fortin M, Hudon C, Dubois MF, Almirall J, Lapointe L, Soubhi H. Comparative assessment of three different indices of multimorbidity for studies on health-related quality of life. *Health Qual Life Outcomes* 2005;3:74.
- (57) Breetvelt IS, Van Dam FS. Underreporting by cancer patients: the case of response-shift. *Soc Sci Med* 1991;32(9):981-7.
- (58) Sprangers MA, Schwartz CE. The challenge of response shift for quality-of-life-based clinical oncology research. *Ann Oncol* 1999 Jul;10(7):747-9.
- (59) Sprangers MA, Van Dam FS, Broersen J, Lodder L, Wever L, Visser MR, et al. Revealing response shift in longitudinal research on fatigue--the use of the thentest approach. *Acta Oncol* 1999;38(6):709-18.

- (60) Martel-Pelletier J, Boileau C, Pelletier JP, Roughley PJ. Cartilage in normal and osteoarthritis conditions. *Best Pract Res Clin Rheumatol* 2008 Apr;22(2):351-84.
- (61) Bitton R. The economic burden of osteoarthritis. *Am J Manag Care* 2009 Sep;15(8 Suppl):S230-S235.
- (62) Le TK, Montejano LB, Cao Z, Zhao Y, Ang D. Health care costs in US patients with and without a diagnosis of osteoarthritis. *J Pain Res* 2012;2012;5:23-30.
- (63) Caporali R, Cimmino MA, Sarzi-Puttini P, Scarpa R, Parazzini F, Zaninelli A, et al. Comorbid conditions in the AMICA study patients: effects on the quality of life and drug prescriptions by general practitioners and specialists. *Semin Arthritis Rheum* 2005 Aug;35(1 Suppl 1):31-7.
- (64) Gabriel SE, Crowson CS, O'Fallon WM. Comorbidity in arthritis. *J Rheumatol* 1999 Nov;26(11):2475-9.
- (65) Gore M, Tai KS, Sadosky A, Leslie D, Stacey BR. Clinical comorbidities, treatment patterns, and direct medical costs of patients with osteoarthritis in usual care: a retrospective claims database analysis. *J Med Econ* 2011;2011;14(4):497-507.
- (66) Harris T, Kovar MG, Suzman R, Kleinman JC, Feldman JJ. Longitudinal study of physical ability in the oldest-old. *Am J Public Health* 1989 Jun;79(6):698-702.
- (67) Guillemin F, Rat AC, Mazieres B, Pouchot J, Fautrel B, Euller-Ziegler L, et al. Prevalence of symptomatic hip and knee osteoarthritis: a two-phase population-based survey. *Osteoarthritis Cartilage* 2011 Nov;19(11):1314-22.
- (68) Guillemin F, Rat AC, Roux CH, Fautrel B, Mazieres B, Chevalier X, et al. The KHOALA cohort of knee and hip osteoarthritis in France. *Joint Bone Spine* 2012 Dec;79(6):597-603.
- (69) Maglakelidze N, Pantsulaia T, Managadze L, Chkhotua A. Assessment of health-related quality of life in living kidney donors. *Transplant Proc* 2011 Jan;43(1):373-5.
- (70) Luo N, Johnson J, Coons SJ. Using instrument-defined health state transitions to estimate minimally important differences for four preference-based health-related quality of life instruments. *Med Care* 2010 Apr;48(4):365-71.

- (71) Hays RD, Woolley JM. The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? *Pharmacoeconomics* 2000 Nov;18(5):419-23.
- (72) Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008 Feb;61(2):102-9.
- (73) Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther* 2009;17(3):163-70.
- (74) Harvill LM. Standard Error of Measurement. *Educational Measurement: Issues and Practice* 1991 Jun 1;10(2):33-41.
- (75) Wyrwich KW, Wolinsky FD. Identifying meaningful intra-individual change standards for health-related quality of life measures. *J Eval Clin Pract* 2000 Feb;6(1):39-49.
- (76) Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: is there a connection? *J Biopharm Stat* 2004 Feb;14(1):97-110.
- (77) Bryan S, Longworth L. Measuring health-related utility: why the disparity between EQ-5D and SF-6D? *Eur J Health Econ* 2005 Sep;6(3):253-60.
- (78) Dolan P. A clarificatory note on 'Valuing health states: a comparison of methods'. *J Health Econ* 1997 Oct;16(5):617.
- (79) Dolan P. Valuing health-related quality of life. Issues and controversies. *Pharmacoeconomics* 1999 Feb;15(2):119-27.
- (80) Puhan MA, Schunemann HJ, Wong E, Griffith L, Guyatt GH. The standard gamble showed better construct validity than the time trade-off. *J Clin Epidemiol* 2007 Oct;60(10):1029-33.
- (81) Joore M, Brunenberg D, Nelemans P, Wouters E, Kuijpers P, Honig A, et al. The impact of differences in EQ-5D and SF-6D utility scores on the acceptability of cost-utility ratios: results across five trial-based cost-utility studies. *Value Health* 2010 Mar;13(2):222-9.

- (82) Marra CA, Marion SA, Guh DP, Najafzadeh M, Wolfe F, Esdaile JM, et al. Not all "quality-adjusted life years" are equal. *J Clin Epidemiol* 2007 Jun;60(6):616-24.
- (83) McDonough CM, Tosteson AN. Measuring preferences for cost-utility analysis: how choice of method may influence decision-making. *Pharmacoeconomics* 2007;25(2):93-106.
- (84) Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ* 2004 Sep;13(9):873-84.
- (85) Grieve R, Grishchenko M, Cairns J. SF-6D versus EQ-5D: reasons for differences in utility scores and impact on. *Eur J Health Econ* 2009 Feb;10(1):15-23.
- (86) Kopec JA, Willison KD. A comparative review of four preference-weighted measures of health-related quality of life. *J Clin Epidemiol* 2003 Apr;56(4):317-25.
- (87) Crosby RD, Kolotkin RL, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. *J Clin Epidemiol* 2004 Nov;57(11):1153-60.
- (88) Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ* 1986 Mar;5(1):1-30.
- (89) Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med* 1977 Mar 31;296(13):716-21.
- (90) Weinstein MC, Torrance G, McGuire A. QALYs: the basics. *Value Health* 2009 Mar;12 Suppl 1:S5-S9.

ANNEXES

Annexe 1 : Questionnaire SF-36

COMMENT REpondre : Les questions qui suivent portent sur votre santé, telle que vous la ressentez. Ces informations nous permettront de mieux savoir comment vous vous sentez dans votre vie de tous les jours.

Veillez répondre à toutes les questions en cochant la case correspondant à la réponse choisie, comme il est indiqué. Si vous ne savez pas très bien comment répondre, choisissez la réponse la plus proche de votre situation.

1. Dans l'ensemble, pensez-vous que votre santé est :

*cochez la case qui correspond
à votre choix*

- Excellente **q**
- Très bonne **q**
- Bonne **q**
- Médiocre **q**
- Mauvaise **q**

2. **Par rapport à l'année dernière à la même époque**, comment trouvez-vous votre état de santé **en ce moment** ?

*cochez la case qui correspond
à votre choix*

- Bien meilleur que l'an dernier **q**
- Plutôt meilleur **q**
- A peu près pareil **q**
- Plutôt moins bon **q**
- Beaucoup moins bon **q**

3. Voici une liste d'activités que vous pouvez avoir à faire dans votre vie de tous les jours. Pour chacune d'entre elles indiquez si **vous êtes limité(e) en raison de votre état de santé actuel**.

cochez la case qui correspond à votre choix, une par ligne

Liste d'activités	oui, beaucoup limité(e)	oui, un peu limité(e)	non, pas du tout limité(e)
a. Efforts physiques importants tels que courir, soulever un objet lourd, faire du sport	0	0	0
b. Efforts physiques modérés tels que déplacer une table, passer l'aspirateur, jouer aux boules	0	0	0
c. Soulever et porter les courses	0	0	0
d. Monter plusieurs étages par l'escalier	0	0	0
e. Monter un étage par l'escalier	0	0	0
f. Se pencher en avant, se mettre à genoux, s'accroupir	0	0	0
g. Marcher plus d'un km à pied	0	0	0
h. Marcher plusieurs centaines de mètres	0	0	0
i. Marcher une centaine de mètres	0	0	0
j. Prendre un bain, une douche ou s'habiller	0	0	0

4. Au cours de ces **4 dernières semaines**, et en raison de votre **état physique**,

cochez la case qui correspond à votre choix, une par ligne

	OUI	NON
a. Avez-vous réduit le temps passé à votre travail ou à vos activités habituelles	0	0
b. Avez-vous accompli moins de choses que vous auriez souhaité	0	0
c. Avez-vous dû arrêter de faire certaines choses	0	0
d. Avez-vous eu des difficultés à faire votre travail ou toute autre activité (par exemple, cela vous a demandé un effort supplémentaire)	0	0

5. Au cours de ces **4 dernières semaines**, et en raison de votre **état émotionnel** (comme vous sentir triste, nerveux(se) ou déprimé(e)),

cochez la case qui correspond à votre choix, une par ligne

	OUI	NON
a. Avez-vous réduit le temps passé à votre travail ou à vos activités habituelles	<input type="radio"/>	<input type="radio"/>
b. Avez-vous accompli moins de choses que vous auriez souhaité	<input type="radio"/>	<input type="radio"/>
c. Avez-vous eu des difficultés à faire ce que vous aviez à faire avec autant de soin et d'attention que d'habitude	<input type="radio"/>	<input type="radio"/>

6. Au cours de ces **4 dernières semaines** dans quelle mesure votre état de santé, physique ou émotionnel, vous a t-il gêné(e) dans votre vie sociale et vos relations avec les autres, votre famille, vos amis, vos connaissances ?

cochez la case qui correspond à votre choix

- Pas du tout
- Un petit peu
- Moyennement.....
- Beaucoup.....
- Enormément

7. Au cours de ces **4 dernières semaines**, quelle a été l'intensité de vos **douleurs physiques** ?

cochez la case qui correspond à votre choix

- Nulle
- Très faible.....
- Faible
- Moyenne.....
- Grande
- Très grande.....

8. Au cours de ces **4 dernières semaines**, dans quelle mesure vos douleurs physiques vous ont-elles limité(e) dans votre travail ou vos activités domestiques ?

*cochez la case qui correspond
à votre choix*

- Pas du tout..... **0**
- Un petit peu..... **0**
- Moyennement **0**
- Beaucoup..... **0**
- Enormément..... **0**

9. Les questions qui suivent portent sur comment vous vous êtes senti(e) **au cours de ces 4 dernières semaines**. Pour chaque question, veuillez indiquer la réponse qui vous semble la plus appropriée. **Au cours de ces 4 dernières semaines**, y a-t-il eu des moments où :

*cochez la case qui correspond à votre choix,
une par ligne*

	en permanence	très souvent	souvent	quelquefois	rarement	jamais
a. Vous vous êtes senti(e) dynamique ?	0	0	0	0	0	0
b. Vous vous êtes senti(e) très nerveux (se) ?	0	0	0	0	0	0
c. Vous vous êtes senti(e) si découragé(e) que rien ne pouvait vous remonter le moral ?	0	0	0	0	0	0
d. Vous vous êtes senti(e) calme et détendu(e) ?	0	0	0	0	0	0
e. Vous vous êtes senti(e) débordant(e) d'énergie ?	0	0	0	0	0	0
f. Vous vous êtes senti(e) triste et abattu(e) ?	0	0	0	0	0	0
g. Vous vous êtes senti(e) épuisé(e) ?	0	0	0	0	0	0
h. Vous vous êtes senti(e) heureux (se) ?	0	0	0	0	0	0
i. Vous vous êtes senti(e) fatigué(e) ?	0	0	0	0	0	0

10. Au cours de ces **4 dernières semaines** y a-t-il eu des moments où votre **état de santé, physique ou émotionnel**, vous a gêné(e) dans votre vie sociale et vos relations avec les autres, votre famille, vos amis, vos connaissances ?

Cochez la case qui correspond à votre choix

- En permanence **0**
- Une bonne partie du temps **0**
- De temps en temps **0**
- Rarement **0**
- Jamais **0**

11. Indiquez, pour **chacune** des phrases suivantes, dans quelle mesure elles sont vraies ou fausses dans votre cas :

cochez la case qui correspond à votre choix, une par ligne

	totalem vraie	plutôt vraie	je ne sais pas	plutôt fausse	totalem fausse
a. Je tombe malade plus facilement que les autres	0	0	0	0	0
b. Je me porte aussi bien que n'importe qui	0	0	0	0	0
c. Je m'attends à ce que ma santé se dégrade	0	0	0	0	0
d. Je suis en excellente santé	0	0	0	0	0

Annexe 2 : Questionnaire EQ-5D

Veillez indiquer, pour chacune des rubriques suivantes, l'affirmation qui décrit le mieux votre état de santé aujourd'hui, en cochant la case appropriée (X)

Mobilité

- Je n'ai aucun problème pour me déplacer à pied
- J'ai des problèmes pour me déplacer à pied
- Je suis obligé(e) de rester alité(e)

Autonomie de la personne

- Je n'ai aucun problème pour prendre soin de moi
- J'ai des problèmes pour me laver ou m'habiller tout(e) seul(e)
- Je suis incapable de me laver ou de m'habiller toute(e) seul(e)

Activités courantes (*exemples: travail, études, travaux domestiques, activités familiales ou loisirs*)

- Je n'ai aucun problème pour accomplir mes activités courantes
- J'ai des problèmes pour accomplir mes activités courantes
- Je suis incapable d'accomplir mes activités courantes

Douleurs/gêne

- Je n'ai ni douleurs ni gêne
- J'ai des douleurs ou gêne(s) modérée(s)
- J'ai des douleurs ou gêne(s) extrême(s)

Anxiété/Dépression

- Je ne suis ni anxieux(se) ni déprimé(e)
- Je suis modérément anxieux(se) ou déprimé(e)
- Je suis extrêmement anxieux(se) ou déprimé(e)

Par rapport au niveau général de ma santé au cours des 12 derniers mois, mon état de santé, Aujourd'hui, est :

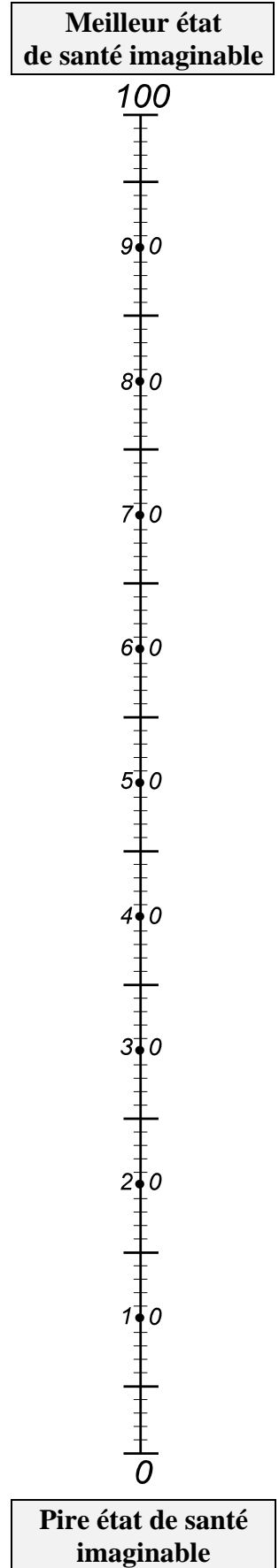
- Meilleur
- A peu près le même
- Pire

COCHEZ UNE
CASE

Pour vous aidez à indiquer dans quelle mesure tel ou tel état de santé est bon ou mauvais, nous avons tracé une échelle graduée (comme celle d'un thermomètre) sur laquelle 100 correspond au meilleur état de santé que vous puissiez imaginer et 0 au pire état de santé que vous puissiez imaginer.

Nous aimerions que vous indiquiez sur cette échelle où vous situez votre état de santé aujourd'hui. Pour cela, veuillez tracer une ligne allant de l'encadré ci-dessous à l'endroit qui, sur l'échelle, correspond à votre état de santé actuel.

Votre état de santé aujourd'hui



Annexe 3 : Questionnaire then-test SF-36

COMMENT REpondre : Les questions qui suivent portent sur votre santé, telle que vous **la ressentiez avant le don de votre rein**. Ces informations nous permettront de mieux savoir comment vous vous sentez dans votre vie de tous les jours.

Veillez répondre à toutes les questions en cochant la case correspondant à la réponse choisie, comme il est indiqué. Si vous ne savez pas très bien comment répondre, choisissez la réponse la plus proche de votre situation.

2. Dans l'ensemble, pensez-vous que votre santé **était** :

*cochez la case qui correspond
à votre choix*

- Excellente **q**
- Très bonne **q**
- Bonne **q**
- Médiocre **q**
- Mauvaise **q**

3. **Par rapport à l'année dernière à la même époque**, comment trouviez-vous votre état de santé **avant le don de votre rein** ?

*cochez la case qui correspond
à votre choix*

- Bien meilleur que l'an dernier **q**
- Plutôt meilleur **q**
- A peu près pareil **q**
- Plutôt moins bon **q**
- Beaucoup moins bon **q**

4. Voici une liste d'activités que vous pouvez avoir à faire dans votre vie de tous les jours. Pour chacune d'entre elles indiquez si **vous étiez limité(e) en raison de votre état de santé avant le don de votre rein.**

cochez la case qui correspond à votre choix, une par ligne

Liste d'activités	oui, beaucoup limité(e)	oui, un peu limité(e)	non, pas du tout limité(e)
a. Efforts physiques importants tels que courir, soulever un objet lourd, faire du sport	0	0	0
b. Efforts physiques modérés tels que déplacer une table, passer l'aspirateur, jouer aux boules	0	0	0
c. Soulever et porter les courses	0	0	0
d. Monter plusieurs étages par l'escalier	0	0	0
e. Monter un étage par l'escalier	0	0	0
f. Se pencher en avant, se mettre à genoux, s'accroupir	0	0	0
g. Marcher plus d'un km à pied	0	0	0
h. Marcher plusieurs centaines de mètres	0	0	0
i. Marcher une centaine de mètres	0	0	0
j. Prendre un bain, une douche ou s'habiller	0	0	0

5. Au cours des **4 dernières semaines précédant le don de votre rein**, et en raison de votre **état physique**,

cochez la case qui correspond à votre choix, une par ligne

	OUI	NON
a. Avez-vous réduit le temps passé à votre travail ou à vos activités habituelles	0	0
b. Avez-vous accompli moins de choses que vous auriez souhaité	0	0
c. Avez-vous dû arrêter de faire certaines choses	0	0
d. Avez-vous eu des difficultés à faire votre travail ou toute autre activité (par exemple, cela vous a demandé un effort supplémentaire)	0	0

6. Au cours des **4 dernières semaines précédant le don de votre rein**, et en raison de votre **état émotionnel** (comme vous sentir triste, nerveux(se) ou déprimé(e)),

cochez la case qui correspond à votre choix, une par ligne

	OUI	NON
a. Aviez-vous réduit le temps passé à votre travail ou à vos activités habituelles	<input type="radio"/>	<input type="radio"/>
b. Aviez-vous accompli moins de choses que vous auriez souhaité	<input type="radio"/>	<input type="radio"/>
c. Aviez-vous eu des difficultés à faire ce que vous aviez à faire avec autant de soin et d'attention que d'habitude	<input type="radio"/>	<input type="radio"/>

7. Au cours des **4 dernières semaines avant le don de votre rein** dans quelle mesure votre état de santé, physique ou émotionnel, vous a-t-il gêné(e) dans votre vie sociale et vos relations avec les autres, votre famille, vos amis, vos connaissances ?

cochez la case qui correspond à votre choix

- Pas du tout
- Un petit peu
- Moyennement.....
- Beaucoup.....
- Enormément

8. Au cours des **4 dernières semaines précédant le don de votre rein**, quelle a été l'intensité de vos **douleurs physiques** ?

cochez la case qui correspond à votre choix

- Nulle
- Très faible.....
- Faible.....
- Moyenne.....
- Grande
- Très grande.....

9. Au cours des **4 dernières semaines précédant le don de votre rein**, dans quelle mesure vos douleurs physiques vous ont-elles limité(e) dans votre travail ou vos activités domestiques ?

cochez la case qui correspond à votre choix

- Pas du tout..... **0**
- Un petit peu..... **0**
- Moyennement **0**
- Beaucoup..... **0**
- Enormément..... **0**

10. Les questions qui suivent portent sur comment vous vous êtes senti(e) **au cours des 4 dernières semaines précédant le don de votre rein**. Pour chaque question, veuillez indiquer la réponse qui vous semble la plus appropriée. **Au cours de ces 4 dernières semaines**, y a-t-il eu des moments où :

cochez la case qui correspond à votre choix, une par ligne

	en permanenc	très souvent	souvent	quelquefoi s	raremen t	jamais
a. Vous vous êtes senti(e) dynamique ?	0	0	0	0	0	0
b. Vous vous êtes senti(e) très nerveux (se) ?	0	0	0	0	0	0
c. Vous vous êtes senti(e) si découragé(e) que rien ne pouvait vous remonter le moral ?	0	0	0	0	0	0
d. Vous vous êtes senti(e) calme et détendu(e) ?	0	0	0	0	0	0
e. Vous vous êtes senti(e) débordant(e) d'énergie ?	0	0	0	0	0	0
f. Vous vous êtes senti(e) triste et abattu(e) ?	0	0	0	0	0	0
g. Vous vous êtes senti(e) épuisé(e) ?	0	0	0	0	0	0
h. Vous vous êtes senti(e) heureux (se) ?	0	0	0	0	0	0
i. Vous vous êtes senti(e) fatigué(e) ?	0	0	0	0	0	0

11. Au cours des **4 dernières semaines précédant le don de votre rein**, y a-t-il eu des moments où votre **état de santé, physique ou émotionnel**, vous a gêné(e) dans votre vie sociale et vos relations avec les autres, votre famille, vos amis, vos connaissances ?

Cochez la case qui correspond à votre choix

- En permanence **0**
- Une bonne partie du temps **0**
- De temps en temps **0**
- Rarement **0**
- Jamais **0**

12. Indiquez, pour **chacune** des phrases suivantes, dans quelle mesure elles sont vraies ou fausses dans votre cas **avant le don de votre rein** :

cochez la case qui correspond à votre choix, une par ligne

	totalem	plutôt vraie	je ne sais pas	plutôt fausse	totalem
a. Je tombais malade plus facilement que les autres	0	0	0	0	0
b. Je me portais aussi bien que n'importe qui	0	0	0	0	0
c. Je m'attendais à ce que ma santé se dégrade	0	0	0	0	0
d. J'étais en excellente santé	0	0	0	0	0

Résumé

Dans le domaine de la santé, la mesure d'utilité permet d'exprimer la valeur que peut donner un individu à un état de santé, comprise généralement entre 0 (la mort) et 1 (meilleur état de santé envisageable). L'utilité permet de valoriser simultanément le résultat clinique et son impact sur l'état de santé du patient, et d'exprimer ainsi la qualité de vie liée à la santé. Nous nous sommes intéressés à l'approche indirecte d'estimation d'utilité dérivée à partir des échelles psychométriques de mesure de l'état de santé. Celle-ci permet de répéter la mesure dans le temps et ouvre les perspectives d'analyses à l'évaluation du changement d'utilité au cours du temps. Les objectifs étaient de définir un seuil de changement pour l'utilité dérivée des instruments de mesure psychométrique de la qualité de vie, et de prendre en compte dans l'interprétation du seuil la présence des comorbidités ou d'un phénomène de response-shift. Afin de répondre à nos objectifs, nous avons étudié, d'une part, le changement d'utilité mesurée par le SF-6D dans une cohorte de patients atteints d'arthrose des membres inférieurs (cohorte KHOALA) et l'effet des comorbidités sur l'utilité et son changement. D'autre part, nous avons estimé l'utilité de donneur vivant de rein à partir des questionnaires EQ-5D et SF-6D avant et 3 mois après le prélèvement du rein. Un seuil de changement a été estimé et un phénomène de response-shift a été recherché. Les résultats ont montré l'effet délétère des comorbidités sur la mesure d'utilité, et mis en évidence l'importance de la sévérité fonctionnelle de l'arthrose dans l'expression de l'utilité et son changement. Chez les donneurs vivants nous avons mis en évidence une diminution significative de l'utilité à court terme, avec une proportion de plus de 30% de donneurs ayant atteint le seuil de changement. L'estimation du seuil diffère selon la population étudiée, l'instrument de mesure de l'utilité, et la méthodologie appliquée pour définir ce seuil. En conclusion, l'évaluation du changement d'utilité doit tenir compte de l'existence des comorbidités ainsi que de l'adaptation du jugement de valeur que l'individu porte sur son état de santé lorsque celui-ci évolue. La complexité des méthodes psychométriques pour la mesure de l'utilité doit amener à être prudent sur leur prise en compte dans les modèles d'évaluation économique. L'identification d'un seuil pour l'interprétation du changement peut être un élément complémentaire utile dans la démarche d'analyse décisionnelle.

Mots clés : utilité indirecte, qualité de vie, SF-6D, EQ-5D, différence minimale importante

Abstract

In health economic, the concept of utility was used to describe a state of health in patients. The unit value of health state utility ranges between 0, representing "death" state, and 1 representing a "perfect" health state. In this work, we used the indirect approach for estimating utility that is derived from a generic utility scales (quality of life questionnaires). The objectives of this thesis were to define a threshold of change for the utility derived from psychometric instruments measuring quality of life, and to take into account for the change interpretation, the presence of comorbidities or a response-shift phenomenon. We used data from a cohort of patients with osteoarthritis of lower limbs (KHOALA cohort) for estimate the impact of comorbidities in utility measured by the SF-6D its change over 3 years. We are also interested in the particular case of living kidney donors. The utility was estimated by the EQ-5D and SF-6D before and 3 months after kidney donation. A threshold of significant deterioration has been estimated and the presence of a response-shift phenomenon was controlled. The results showed the deleterious effect of comorbidities on utility measurement for patients with lower limbs osteoarthritis, and highlighted the importance of the functional severity of osteoarthritis in the expression of utility and its change over time. Living kidney donors experienced a significant decrease in their utility score in the short term, and more than third of donors have reached the threshold of significant deterioration. The threshold of utility change varies depending on the population studied, the measurement instrument, and the methodology used to define the threshold. In conclusion, the assessment of utility change must take into account the existence of other diseases and the adaptation of the individual on his health state variation. The complexity of measuring individual preferences by psychometric instruments should lead to be careful about introducing them in the economic evaluation. The use of a threshold for the interpretation of utility change can be an additional element in the decision-making process.

Key words: utility, health related quality of life, SF-6D, EQ-5D, minimal important difference