



**Actualisation partielle du Guide
méthodologique pour l'évaluation économique
des stratégies de santé**

Olivia Clément
sous la Direction de
Thomas Barnay et Claude Le Pen

Table des matières

1. Synthèse : points de convergence et éléments de débats.....	4
1.1 Comparateur	4
1.2 Coûts.....	5
1.3 Incertitude	6
1.4 Évaluation contingente.....	6
1.5 Méthode des choix discrets.....	7
1.6 Transférabilité	8
2. Introduction	9
3. Méthodologie.....	12
3.1 Guidelines internationaux	12
3.2 Bases de données	12
3.3 Etude des bibliographies	12
3.4 Recherches bibliographiques	13
4. Résultats par item	15
4.1 Comparateur	15
4.1.1 Prise en compte du cadre de l'étude	15
4.1.2 Comparateurs fréquemment recommandés	16
4.1.3 Description du comparateur et type de comparateur	17
4.1.4 Prise en compte du futur.....	18
4.1.5 Comparaison indirecte et modèles	19
4.2 Prise en compte des coûts.....	21
4.2.1 Perspective	21
4.2.2 Présentation et estimation des coûts.....	21
4.2.3 Typologie : Coûts directs et coûts indirects	24
4.2.4 Prise en compte de différents coûts	25
4.2.5 Long terme vs court terme	30
4.3 Incertitude	31
4.3.1 Incertitude liée aux paramètres et au modèle	31
4.3.2 Méthodes appréhendées dans la littérature scientifique	34
4.4 Evaluation contingente.....	39
4.4.1 Historique et points importants	40
4.4.2 Construction du scénario hypothétique.....	45
4.4.3 Choix de la méthode de révélation des préférences.....	48

4.4.4	Réalisation de l'enquête, traitement des données et validité du questionnaire.....	50
4.4.5	Existence de biais	54
4.5	Méthode des choix discrets.....	62
4.5.1	Origines et bases théoriques	62
4.5.2	Mise en place de la méthode des choix discrets.....	63
4.5.3	Biais	71
4.5.4	Validité.....	73
4.6	Transférabilité	75
4.6.1	Éléments empiriques, de définition et guidelines.....	75
4.6.2	Facteurs affectant la transférabilité et éléments à prendre en compte lors de l'étude de la transférabilité	81
4.6.3	Aperçu des méthodes existantes ayant trait à la généralisabilité	86
4.6.4	Essais cliniques internationaux.....	98
5.	Annexes	103
	Annexe 1 – Tableau récapitulatif des recommandations des comparateurs	103
	Annexe 2 – Niveau de variation attendu dans les études médico-économiques.....	103
	Annexe 3 – Facteurs à prendre en compte lorsque l'on souhaite évaluer la qualité d'une méthode des choix discrets.....	104
	Annexe 4 – Facteurs de transférabilité.....	105
	Annexe 5 – L'arbre de décision dans le modèle de Welte.....	106
	Annexe 6 – Tableau comparatif : Nice, IQWiG et France	107
	Bibliographie.....	109
1.	Guidelines internationaux	109
2.	Comparateur	109
3.	Prise en compte des coûts.....	109
4.	Incertitude	111
5.	Méthode des Choix Discrets.....	112
6.	Evaluation contingente.....	114
7.	Transférabilité	117

1. Synthèse : points de convergence et éléments de débats

Cette synthèse vise à mettre en lumière les éléments de convergence et ceux qui restent en débat dans la littérature économique autour des points suivants : comparateur, coûts, incertitude, évaluation contingente, méthode des choix discrets et transférabilité.

1.1 Comparateur

Les différents guidelines s'accordent sur l'importance du choix du comparateur car il influence les résultats de l'étude et leur pertinence pour les décideurs. Aucun consensus ne se dégage cependant certains comparateurs sont plus fréquemment cités :

- La technologie de santé existante *versus* la technologie alternative faisant l'objet de l'étude (guide polonais, belge, canadien, néo-zélandais et autrichien) ;
- La technologie la plus utilisée (guides anglais, suédois, néo-zélandais, canadien, polonais et hollandais) ;
- La technologie recommandée dans les guidelines de pratiques cliniques s'ils existent pour la pathologie concernée (guidelines anglais, autrichien, hollandais et belge dans certains cas).

D'autres comparateurs tels que les technologies les plus efficaces, les moins coûteuses ou les plus coûts-efficaces sont parfois mentionnées. On note que ces technologies sont toujours associées à l'un des comparateurs cités précédemment. Dans tous les cas, les guidelines estiment que le choix du comparateur doit être justifié et présenté de façon précise (nom, classe thérapeutique, forme posologique, voie d'administration, durée du traitement...).

Un élément qui ne fait pas nécessairement l'objet de débat mais pour lequel aucun consensus ne semble exister à l'heure actuelle est la prise en compte du futur. Certains pays (Canada, Belgique et Nouvelle-Zélande) recommandent de tenir compte du futur (possible développement de génériques si le brevet expire prochainement par exemple ou évolution des pratiques cliniques). Cependant, la prise en compte du futur complique le choix du comparateur et doit s'accompagner d'analyses de sensibilité.

1.2 Coûts

Dans les guidelines et la littérature, lors de la détermination des coûts, l'importance de la perspective est soulignée. Par ailleurs, on retrouve l'idée qu'il est important de présenter les coûts en quantités et en valeurs monétaires (sans toutefois donner d'information précise sur la devise à utiliser). Enfin, les guidelines recommandent des analyses de sensibilité lors du traitement des coûts car ils ne sont pas toujours faciles à valoriser. Les analyses de sensibilité limitent l'incertitude.

La prise en compte des coûts est l'objet de nombreux points de désaccord. Tout d'abord au niveau de leur estimation. Ce sujet est peu traité dans la littérature et les guidelines ne s'accordent pas sur une méthode en particulier : analyse marginale, coûts moyens, coûts d'opportunité... Toutefois, tous indiquent qu'il est important de justifier et de préciser le type de coûts utilisés.

Un autre point abordé différemment en fonction des guidelines est la typologie relative aux coûts. Certains guides font la distinction entre coûts directs et indirects et médicaux et non médicaux sans toutefois les définir clairement, partager une définition commune, attribuer les mêmes coûts à la même typologie. Peut-être serait-il judicieux de ne plus utiliser cette typologie comme c'est le cas au Royaume-Uni et en Nouvelle-Zélande.

La prise en compte du temps perdu peut également poser problème. Il est possible de considérer que le temps perdu est pris sur le temps de loisir des patients et des familles. Dans ce cas, il n'y a pas de coûts d'opportunité pour la société. Toutefois cette hypothèse a ses limites. Par ailleurs, prendre en compte ou non le temps perdu dépend de la perspective adoptée. Il est possible de le présenter séparément comme un élément supplémentaire d'aide à la décision.

Deux autres types de coûts doivent être abordés avec attention : les pertes de production et de productivité et les coûts futurs. Les pertes de productivité et de production peuvent être évaluées au moyen de la méthode du capital humain ou de la méthode des coûts de friction. Toutes deux ont leurs limites, cependant la méthode des coûts de friction requiert un travail empirique important et aboutit à des estimations inférieures. Les coûts futurs font l'objet d'un important débat dans la littérature.

La prise en compte des coûts, en raison de l'impact direct sur les résultats des évaluations médico-économiques, doit être traitée avec attention.

1.3 Incertitude

L'ensemble des guidelines s'entendent sur l'impact de l'incertitude et de ses causes sur les résultats. L'incertitude a deux sources : elle peut être liée aux paramètres et aux choix méthodologiques. Donner une définition plus précise des sources de l'incertitude semble judicieux (peu de guidelines précise ce qu'il en est réellement).

On observe que les guides et la littérature ne proposent pas de traiter l'incertitude de la même façon. Tous s'accordent sur l'importance des analyses de sensibilité sans toutefois recommander le même type d'analyses (analyses déterministes univariées, multivariées, probabilistes...). Par ailleurs, les analyses de sensibilité peuvent être réalisées en complément d'analyses de seuil ou d'études des cas extrêmes. Enfin, la littérature scientifique propose d'autres méthodes pas toujours abordées (ou alors simplement mentionnées) dans les guidelines : plan coût-efficacité, la méthode de la « boîte », le bootstrap, le théorème de Fieller, la courbe d'acceptabilité.

L'incertitude est un vaste sujet influençant les résultats et lié à de nombreux éléments des évaluations médico-économiques (prise en compte des coûts, transférabilité, choix méthodologiques...).

1.4 Évaluation contingente

L'évaluation contingente peut être utilisée pour obtenir des consentements à payer (CAP) ou des consentements à recevoir (CAR). La littérature scientifique, en se basant sur la théorie des perspectives de Kaneman et sur « l'aversion à la perte », montre que le CAP est plus utilisé que le CAR. L'étude des évaluations contingentes menées dans les différents secteurs confirme la préférence pour le CAP.

Lorsque l'on étudie l'évaluation contingente, les auteurs s'accordent sur l'importance à accorder à la définition du scénario car l'acceptation de ce dernier par les répondants est une étape primordiale à la réussite de l'enquête. Il est donc nécessaire de réaliser un arbitrage entre le niveau d'information donné, la durée du questionnaire et les capacités cognitives des individus. Une bonne définition du scénario est importante car elle permet de limiter le biais hypothétique ainsi que les possibles effets de structure. Par ailleurs, les différentes études soulignent l'importance d'informer les individus de l'existence de substituts et de leur contrainte budgétaire. Une telle précaution permet de limiter l'effet de certains biais tels que les biais d'inclusion. Enfin, les études convergent sur le fait qu'il est conseillé d'inclure des questions socio-économiques. Les réponses obtenues pourront par la suite être utilisées pour expliquer les différentes valeurs des CAP.

Les points de divergences concernent davantage son utilisation dans le domaine de la santé et la mise en place de la méthode. Les points débattus sont la méthode de révélation des préférences

(enchères, carte de paiement, référendum, question ouverte) et le mode d'interrogation (interview en face-à-face, téléphone, mail...). Ces éléments peuvent jouer un rôle sur la validité de la méthode et la recevabilité des résultats. La représentativité de l'échantillon dépend également fortement de la méthode utilisée. Aucun consensus n'a été trouvé dans la littérature. Par ailleurs, le choix du mode d'interrogation peut avoir un impact significatif sur les coûts et le temps nécessaire à la réalisation du questionnaire.

Enfin, il est nécessaire de s'interroger sur la recevabilité d'une étude en fonction du traitement économétrique. Le traitement des valeurs nulles ou extrêmes peut poser problème car il peut signifier une réponse protestataire de la part des répondants et biaiser les résultats.

1.5 Méthode des choix discrets

Comme dans le cas de l'EC, on s'accorde sur le fait que le développement du questionnaire et du scénario sont des étapes primordiales à la réussite de la MCD. Il est nécessaire que le contexte soit crédible pour être accepté par les répondants. Les auteurs s'accordent également sur le fait que les différents attributs doivent être pertinents au regard de la problématique de décision, avoir un intérêt pour les répondants et être compréhensibles. Il est possible de recourir à des experts ou à des consommateurs pour identifier les attributs. Le choix de mauvais attributs faussera les résultats. S'ils n'intéressent pas les répondants, leur variation n'influencera pas leur choix. Quant aux niveaux des attributs, ils doivent être plausibles, recevables et impliquer un arbitrage de la part du répondant. Il est possible que le répondant ne réalise pas les arbitrages nécessaires s'il perçoit les écarts comme étant insignifiants ou trop importants. On conseille également d'attribuer le même nombre d'attributs pour éviter l'effet-attribut.

L'étape du design expérimental ne fait pas l'objet d'un consensus. La littérature propose différentes méthodes telles que la méthode factorielle totale ou la méthode factorielle partielle. Le design expérimental influence les fonctions d'utilité et suppose de tenir compte des capacités cognitives des répondants. Si aucune méthode n'est conseillée en particulier, certaines règles doivent être respectées pour assurer la qualité de l'étude (équilibre des niveaux, orthogonalité, croisement minimum, équilibre des utilités).

Comme dans le cas de l'EC, aucun accord n'a été trouvé concernant la façon d'administrer le questionnaire. Les deux méthodes appartenant aux méthodes d'éllicitations des préférences révélées, elles ont en commun les biais propres à cette famille.

On note l'existence d'un débat concernant l'analyse des données. La plupart des études comparent l'impact relatif des attributs en étudiant la taille et la significativité des paramètres estimés. Lancsar *et al.* (2007) indiquent que ces paramètres ne sont pas directement comparables. Ils

présentent cinq méthodes permettant de comparer l'impact relatif des attributs. Chacune des méthodes présentant des avantages différents, le choix dépendra des objectifs de l'étude.

Enfin, comme dans le cas de l'EC, le traitement des réponses nulles peut poser problème. Un désaccord existe concernant le traitement économétrique de telles réponses. Faut-il les tronquer ou utiliser une autre méthode (telle que celle proposée par Dziegielewska et Mendelsohn) ?

1.6 Transférabilité

La littérature dans son ensemble reconnaît que le caractère transférable des études est un enjeu important. En effet, s'il pouvait être reconnu, il permettrait des gains financiers conséquents. Les auteurs s'accordent pour reconnaître la variabilité des données. Il est également connu que les données économiques varient davantage que les données liées aux gains de santé.

Par ailleurs, différents facteurs affectent la transférabilité des études : les caractéristiques des patients, de la pathologie, des professionnels de santé, du système de santé et les critères méthodologiques utilisés lors de la réalisation d'une étude. D'autres éléments tels que la perspective, l'horizon temporel, le choix du comparateur peuvent limiter la transférabilité des études.

Suite à l'étude de la littérature, différentes recommandations permettant d'accroître la transférabilité ont été observées. Réaliser des études transparentes est le premier élément impactant la transférabilité. En effet, il est conseillé de fournir un maximum de renseignements concernant les données (les présenter précisément, indiquer les sources, les prix, les quantités, les méthodes de calcul utilisées...). On préconise également de donner des renseignements concernant le choix du comparateur et la perspective adoptée. Enfin, présenter des analyses de sensibilité peut aider les utilisateurs potentiels de l'étude. Lorsqu'il n'est pas possible de faire figurer ces informations dans une étude, on recommande d'indiquer au lecteur comment il peut se les procurer.

Les points de divergence concernant la transférabilité ont trait aux méthodes utilisées pour évaluer le caractère transférable des études. Quatre méthodes ont été identifiées. La littérature scientifique ne s'accorde pas sur l'utilisation d'un modèle en particulier (même si le modèle de Welte est le plus fréquemment cité).

Au regard de la littérature, il n'existe pas d'accord quant à la transférabilité des études même s'il semble reconnu que les résultats d'une étude ne peuvent pas être directement appliqué à un autre contexte. Cependant, les auteurs peuvent décider d'utiliser certaines données ou d'apporter des modifications aux études existantes pour tenter de les adapter.

2. Introduction

En France, les premières recommandations en matière d'évaluation économique dans le domaine de la santé ont été émises en 1997¹. En 2003, le Collège des Economistes de la Santé a pris l'initiative de réviser ces recommandations. Cette décision a été motivée par les changements de contexte, de niveau d'exigence scientifique et du mode d'élaboration des évaluations économiques. Le contexte, institutionnel notamment, a permis de mettre en place une commission tripartite (industrie, pouvoirs publics et experts économistes) qui a identifié les principaux problèmes méthodologiques et a élaboré des réponses consensuelles. Le niveau d'exigence scientifique requis correspondait à celui des guidelines internationaux publiés alors.

La Haute Autorité de Santé a été créée par la loi du 13 août 2004 relative à l'assurance maladie. Cette loi s'inscrit dans le cadre de la réforme de l'ensemble de notre système de santé. Le statut de la HAS est celui d'une autorité publique indépendante à caractère scientifique. La loi confie cinq missions à la HAS dont l'évaluation médico-économique et en santé publique (dans le cadre de l'aide aux décisions de remboursement des produits et services médicaux aux pouvoirs publics).

Depuis sa création, La Haute Autorité de Santé développe ses compétences économiques. En témoigne la création de la CEESP en 2008 dont l'objectif est de contribuer à ce que la dimension d'efficience ou de coût d'opportunité soit prise en compte à la fois dans la décision publique et dans les décisions des professionnels. La CEESP offre une expertise médico-économique qui devrait être conjuguée à l'expertise médicale existante dans les commissions spécialisées que sont la Commission de la Transparence, la Commission d'Évaluation des Produits et Prestation et la Commission d'Évaluation des Actes Professionnels. Dans ce cadre, la HAS a chargé le CES de réaliser une actualisation partielle des guidelines de 2003. Ce travail consiste en une revue de la littérature portant principalement sur la période 2003-2010 et ciblant les six items suivants :

- Le comparateur ;
- La prise en compte des coûts ;
- L'incertitude ;
- La méthode d'évaluation contingente ;
- La méthode des choix discrets ;
- Le caractère transférable des études.

Ce rapport, qui n'a aucune vocation normative, propose un recensement exhaustif de la littérature scientifique et des différents guidelines internationaux rédigés depuis 2003. L'objectif est

¹ Publiées dans la Lettre du Collège des Economistes de la Santé d'avril 1997 et reprises dans le Journal d'Economie Médicale, 1998, **16**(4-5) :353-372.

de confronter les recommandations émises en 2003 pour chacun des six items avec les éléments les plus récents de la littérature (articles, guidelines publiés depuis 2003,...) afin d'évaluer leur pertinence en 2010. Ce travail est mené indépendamment pour chaque item.

Le travail présenté consistant en une revue de la littérature, une méthodologie a dû être définie. Nous avons tout d'abord identifié les guidelines internationaux rédigés depuis 2003 à partir du site de l'Ispor ou de sites spécialisés tels que ceux du NICE ou de l'IQWiG. Dans un second temps, nous avons effectué des recherches à partir de mots clés sur différentes bases de données (Medline, Web of Science, Science Direct et JSTOR) pour la période 2003-2010. Une lecture sélective des abstracts puis des articles a été faite afin d'identifier les articles pertinents d'un point de vue méthodologique. Enfin, nous avons étudiés les bibliographies des articles pour identifier d'autres articles (notamment ceux rédigés avant 2003 et étant toujours des articles de références). Un total de deux cent dix huit articles et de dix guidelines ont été utilisés pour la rédaction de ce rapport.

Il ressort de l'examen de la littérature et de l'étude des guidelines que les items sont inégalement traités. Une importante part traite par exemple de l'évaluation contingente dont de nombreux développements ont été observés dans les domaines tels que l'environnement et les transports. Cette littérature abondante est largement documentée dans le présent rapport ce qui le différencie sensiblement des autres guidelines qui font très peu référence à la méthode d'évaluation contingente et à la méthode des choix discrets. La méthode d'évaluation contingente avait été mentionnée à titre illustratif dans le guide français de 2003 ; elle ne permettait pas alors d'être mobilisée comme instrument d'aide à la décision publique. Le développement depuis de cette méthode dans le domaine de l'évaluation économique des interventions de santé justifie qu'une partie y soit consacrée. La méthode des choix discrets proche de l'évaluation contingente s'inscrit dans l'élicitation des préférences révélées et constitue à ce titre un piste de réflexion que nous avons suivie.

En revanche, la question du comparateur est peu traitée dans la littérature scientifique. C'est pourquoi les informations portant sur cet item proviennent des guidelines internationaux. Les autres items sont tout à fait classiques. La prise en compte des coûts, compte tenu de son impact financier et de la prise de décision qui en découle, est à cet égard fondamentale. La discussion n'échappe pas au choix déterminant de la perspective adoptée (le Royaume-Uni par exemple adopte toujours la position du payeur public alors que la France en 2003 recommandait plutôt une perspective sociétale tout en laissant une certaine flexibilité dans le choix final). Le thème de l'incertitude renvoie aussi à une littérature abondante, il est abordé de façon incomplète dans ce rapport, un groupe de travail à la HAS ayant spécifiquement investi cette question.

Le caractère plus ou moins transférables des études est un enjeu financier important (gains en temps, monétaires). Cependant, il nécessite que soient collectées de nombreuses données (données médicales, démographiques, de coûts, horizon temporel). La mise à disposition puis la prise en compte de ces données ont pu évoluer depuis 2003 influençant le caractère transférable des études.

Le rapport consistant en une revue de la littérature, nous allons tout d'abord présenter la méthodologie utilisée pour les recherches bibliographiques. Puis, les résultats obtenus à l'étude de la littérature scientifique et des guidelines internationaux seront présentés par item.

3. Méthodologie

3.1 Guidelines internationaux

Le travail présenté consiste en une revue de la littérature. Différentes sources ont été utilisées. Nous avons tout d'abord identifié les guidelines internationaux rédigés depuis 2003 à partir du site de l'Ispor. Dix guidelines ont été utilisés : les guides français, suédois et polonais (rédigés en 2003), les guides hollandais, autrichien et canadien (2006), le guide néo-zélandais rédigé en 2007 et les guides anglais et belge (2008) et allemand (2009).

3.2 Bases de données

Dans un second temps, nous avons effectué des recherches par mots-clés sur quatre bases de données (Medline, Web of Science, Science Direct et JSTOR). La sélection de ces bases de données nous a permis d'identifier des articles ne se référant pas uniquement au secteur de la santé. En effet, il nous a semblé pertinent d'étudier ce qui était fait dans d'autres domaines, notamment dans le cas de l'évaluation contingente et de la méthode des choix discrets ; ces deux méthodes ayant été utilisées dans des domaines tels que l'environnement et les transports avant d'être appliquée au secteur de la santé.

Les recherches ont été réalisées par item, sur la période 2003-2010 et seuls les articles rédigés en français et en anglais ont été retenus. La recherche a été effectuée en quatre étapes. L'identification par mots-clés, la suppression des doublons, une lecture sélective des abstracts et une lecture sélective des articles. Différents critères de sélection ont été retenus. Les articles s'appliquant à une pathologie particulière ont été rejetés au profit des articles liés à la méthodologie et/ou à la revue de la littérature. Nous avons également cherchés à identifier les articles apportant de nouvelles informations ou étant des articles de référence.

3.3 Etude des bibliographies

Afin d'identifier des articles de références antérieurs à 2003 ou des articles non identifiés à partir des recherches par mots-clés, nous avons étudié les bibliographies des guidelines et des articles retenus.

3.4 Recherches bibliographiques

Item	Comparateur		Coûts		Incertitude		Méthode des choix discrets		Evaluation contingente		Transférabilité	
Source	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
Guidelines internationaux	10	10	10	10	10	9	10	0	10	1	10	3
Medline	35	7	17	15	26	32	40	57	51	74	37	18
Web of Science	6		6		15		107		41		60	
Science Direct	2		10		3		28		66		12	
JSTOR	0		-		-		-		-		4	
Autres	0		12		18		31		40		7	
Total	53	17	54	27	72	50	216	57	208	75	130	28

Tableau 1 : Récapitulatif des recherches bibliographiques.

P1 : Nombre d'articles sélectionnés par mots clés et après suppressions des doublons.

P2 : Nombre d'articles après lecture sélective.

Lorsque l'on étudie les données relatives aux recherches bibliographiques on observe une disparité entre les différents items. La littérature scientifique traitant de l'évaluation contingente et de la méthode des choix discrets est riche (nous avons retenus 75 et 57 articles respectivement). Ce phénomène s'explique par le fait que la littérature ne provient pas uniquement du secteur de la santé. Les méthodes d'élicitation des préférences ont tout d'abord été utilisées dans le domaine de l'environnement et des transports. C'est également pour cette raison qu'une part conséquente des articles identifiés appartient à la catégorie « autres ». Cette catégorie fait référence aux articles trouvés suite à l'étude des bibliographies des articles. Dans le cas de l'évaluation contingente et de la méthode des choix discrets, les articles provenant des domaines autres que la santé sont souvent antérieurs à 2003 (du fait de l'utilisation de ces méthodes depuis plusieurs années déjà).

Dans le cas de l'incertitude, on note que c'est l'item (après l'EC et la MCD) pour lequel nous avons identifié le plus d'articles. L'intérêt des auteurs pour ce sujet témoigne du fait que de nombreux points sont encore débattus. Au contraire, peu d'articles pertinents ont été identifiés dans le cas du comparateur et de la prise en compte des coûts. Pour ces deux items, le rapport s'appuie principalement sur les guidelines internationaux. Tous ne s'accordent pas nécessairement sur la méthode à adopter mais les différences observées relèvent davantage de choix méthodologiques que de réels points de débat ou de points de désaccord. Le cas de la transférabilité est encore différent. Peu de guidelines traitent de ce thème (belge, français et canadien) et la littérature scientifique n'est pas encore développée même si l'on note que les auteurs conseillent fréquemment d'approfondir les recherches.

4. Résultats par item

4.1 Comparateur

La première étape d'une évaluation médico-économique consiste à en définir le cadre : hypothèses, population cible, perspective de l'étude, choix du comparateur, méthodes utilisées²... Afin que l'évaluation puisse être utilisée lors de la prise de décision, il est important que ce cadre soit le plus réaliste possible³. Il faut donc choisir des comparateurs appropriés⁴ car ce choix influencera le rapport coût-efficacité de l'intervention et la pertinence de l'étude pour les décideurs⁵. Les guidelines, s'ils ne développent pas le thème du comparateur (à l'exception des guides canadien et belge), donnent tout de même leurs recommandations.

Le guide polonais souligne que le choix du comparateur influence l'étude. C'est pourquoi, il propose différentes alternatives et laisse le chercheur relativement libre de son choix (du moment qu'il est justifié). On observe également que les guides français et allemand ne recommandent aucun comparateur en particulier. Par ailleurs, le sujet est peu traité dans la littérature scientifique.

4.1.1 Prise en compte du cadre de l'étude

On note qu'il est important que le comparateur soit adapté à la question étudiée⁶, aux indications thérapeutiques, à la population cible de l'étude et au but de l'intervention⁷. Les guides français et canadien recommandent de lister l'ensemble des solutions possibles, acceptables et pertinentes comme comparateurs potentiels. A partir de cette liste et en tenant compte du contexte, une sélection des comparateurs retenus pourra être réalisée. Il peut être utile de faire appel à des experts lors du processus de sélection du comparateur⁸. Le choix du comparateur peut être compliqué lorsqu'il existe une série de solutions de rechange approuvées pour une même indication ou si la pratique clinique varie au sein du pays. Lors de la sélection des comparateurs, il faut réaliser

² NICE. (2008). "Guide to the methods of technology appraisal."

³ Graf von der Schulenburg, J.-M. et C. Hoffmann (2000). "Review of European guidelines for economic evaluation of medical technologies and pharmaceuticals." *HEPAC*, 1: 2-8.

⁴ Walter, E. ; Zehetmayr, S. (2006). "Guidelines on Health Economic Evaluation." Institut für Pharmaökonomische Forschung. Page 5.

⁵ Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." 3 : 34.

⁶ Walter, E. ; Zehetmayr, S. (2006). "Guidelines on Health Economic Evaluation." Institut für Pharmaökonomische Forschung. Page 5.

⁷ Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." 3 : 34.

⁸ Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." 3 : 34.

un arbitrage entre la rigueur scientifique, la disponibilité des données, les contraintes de temps et la possibilité d'analyser un nombre important de comparateurs⁹.

4.1.2 Comparateurs fréquemment recommandés

A. Technologie de santé susceptible d'être remplacée

L'un des comparateur le plus fréquemment recommandé est la technologie de santé la plus susceptible d'être remplacée par la technologie faisant l'objet de l'étude si elle devait être adoptée (guides polonais, canadien, belge, néo-zélandais et autrichien s'il n'est pas possible de déterminer le traitement standard recommandé par les guidelines). Afin d'identifier ce traitement, il est possible de réaliser une étude du marché, des sondages, une analyse des bases de données...¹⁰ Dans le cas où la nouvelle technologie est un médicament ajouté au traitement existant, le comparateur est alors le traitement utilisé dans la pratique avant l'ajout du nouveau médicament¹¹.

B. Technologie la plus utilisée

Plusieurs guides recommandent également de choisir la technologie la plus utilisée comme comparateur (guides anglais, suédois, néo-zélandais, canadien et hollandais). Selon Cox *et al.* (2006), ce choix est le plus répandu. En Pologne, dans le cas où la technologie la plus susceptible d'être remplacée n'est pas la plus utilisée, il est recommandé de prendre les deux traitements comme comparateurs. En Angleterre, choisir la technologie la plus utilisée comme comparateur peut souvent mener à prendre en compte plusieurs comparateurs pertinents car les pratiques cliniques peuvent varier au sein du NHS.

C. Prendre en compte les guidelines de pratiques cliniques

Dans certains pays, des guidelines concernant les pratiques cliniques existent. Il est alors possible de choisir les soins de santé recommandés dans le guideline comme comparateur. Le NICE ainsi que les guidelines autrichien et hollandais recommandent de suivre leurs guidelines. En Belgique, on recommande de faire appel aux guidelines s'il n'est pas possible d'identifier la technologie la plus susceptible d'être remplacée. Dans le cas où des guidelines internationaux existent, il faut faire attention car les recommandations ne sont pas toujours pertinentes pour tous les pays¹². Au Canada,

⁹ Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." 3 : 34.

¹⁰ Cleemput, I. ; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. 28 : 20

¹¹ Cleemput, I. ; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. 28 : 20

¹² Cleemput, I. ; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. 28 : 20.

il est noté que les techniques les plus utilisées peuvent ne pas toujours refléter les soins recommandés ou qui sont les plus efficaces sur le plan clinique. Dans ce cas, les soins recommandés doivent être compris dans l'analyse (si on estime cette option possible et pertinente) en plus des soins habituels.

D. Autres comparateurs

Les comparateurs choisis peuvent également être les technologies les plus efficaces¹³, les moins coûteuses ou les plus coûts-efficace. On note toutefois que lorsque ces comparateurs sont mentionnés, ils sont toujours associés à l'une des technologies précédemment citées¹⁴. Dans tous les cas, il est important que le chercheur justifie le choix du comparateur¹⁵.

4.1.3 Description du comparateur et type de comparateur

A. Type de comparateur

La technologie étudiée doit être comparée à des traitements dont l'efficacité a été prouvée auparavant¹⁶. Les comparateurs peuvent être des traitements médicaux ou non. Il est possible de choisir l'option « ne rien faire » comme comparateur (si cela est indiqué au regard des pratiques cliniques)¹⁷. Aux États-Unis, la *Food and Drug Administration* utilise davantage les essais cliniques comparant la nouvelle technologie à un placebo¹⁸. En Pologne, cette option est acceptée si elle est l'unique alternative pertinente proposée aux patients. Dans le cas où « aucun traitement » est l'option la plus répandue dans la pratique mais qu'une autre alternative médicale existe, les deux options peuvent être prises en compte¹⁹. On note que l'option « ne rien faire » n'est pas nécessaire associée à des coûts et à des effets nuls²⁰. Dans le guide français, il est souligné que des essais cliniques sont réalisés en assimilant l'option « ne rien faire » à des placebos. Cette approximation

¹³ Walter, E. ; Zehetmayr, S. (2006). "Guidelines on Health Economic Evaluation." Institut für Pharmaökonomische Forschung. Page 5.

¹⁴ Orlewska, E. et P. Mierzejewski. (2003). "Polish guidelines for conducting pharmacoeconomic evaluations".

¹⁵ Cleemput, I. ; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. **28** : 20.

¹⁶ Cleemput, I. ; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. **28** : 20.

¹⁷ Cox, H. L., K. B. Laupland, et al. (2006). "Economic evaluation in critical care medicine." *Journal Of Critical Care* **21**(2): 117-124.

¹⁸ Sullivan *et al.* (2009) "Health Technology Assessment in Health-Care Decisions in the United States." *Value in Health* **12**(S2) : S39-S44.

¹⁹ Orlewska, E. et P. Mierzejewski. (2003). "Polish guidelines for conducting pharmacoeconomic evaluations".

²⁰ Cleemput, I. ; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. **28** : 20

peut conduire à sous-estimer la différence d'efficacité étant donné la présence d'un effet placebo²¹. On note que Draborg *et al.* (2005) assimilent le choix d'un placebo à « aucun comparateur ».

Lors du choix du comparateur, il est utile de prendre en compte la solution la moins coûteuse disponible pour la même indication. Le guide canadien recommande d'inclure le coût du médicament et tous les frais d'administration qui y sont liés. Il est le seul guide à mentionner les frais d'administration.

Dans le cas où des stratégies différentes peuvent être adoptées en fonction de sous-groupes, le guideline belge recommande de les indiquer, les effets pouvant différer d'un groupe à l'autre (en fonction de l'âge des patients par exemple). Au Canada, on recommande de distinguer les différentes stratégies de traitement en fonction des situations où l'intervention représente un élément supplémentaire de la stratégie, une séquence différente dans le traitement ou une solution distincte qui pourrait remplacer un autre élément du traitement. Chaque stratégie doit être analysée séparément.

B. Description du comparateur

Lorsque l'on a choisi le(s) comparateur(s), il est recommandé de le(s) décrire de façon précise : indiquer le nom (générique ou de marque^{22,23}), la classe thérapeutique, la forme posologique, la voie d'administration, la posologie quotidienne, la durée du traitement, le coût quotidien, les coûts de la durée habituelle du traitement (pour la technologie étudiée comme pour les comparateurs^{24,25})... Il est également recommandé de préciser si les schémas posologiques utilisés dans la pratique clinique diffèrent de ceux utilisés lors des essais cliniques²⁶.

4.1.4 Prise en compte du futur

Lorsque l'on réalise une étude, il peut être conseillé de tenir compte du futur. En effet, des éléments tels que les prix des médicaments (notamment dans le cas du développement de médicaments génériques) ou les pratiques cliniques peuvent évoluer dans le temps. Le choix du comparateur peut alors être compliqué. Par ailleurs, les comparateurs pris en compte peuvent être

²¹ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé."

²² NICE. (2008). "Guide to the methods of technology appraisal."

²³ Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." 3 : 34.

²⁴ Walter, E. ; Zehetmayr, S. (2006). "Guidelines on Health Economic Evaluation." Institut für Pharmaökonomische Forschung. Page 5.

²⁵ Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." 3 : 34.

²⁶ Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." 3 : 34.

obsolètes^{27,28}, ce qui peut encourager la prise en compte du futur. Les guidelines canadien, belge et néo-zélandais recommandent de tenir compte de ces possibles changements lors du choix du comparateur. Par exemple, le guide canadien recommande de prévoir des comparateurs futurs car l'absence de tels comparateurs peut entraîner une mauvaise estimation du ratio coût-efficacité incrémental. Il est possible d'inclure l'arrivée prévue d'un générique qui fera concurrence au produit étudié lorsque l'on sait que le brevet arrive prochainement à expiration. Lorsque l'on tient compte du futur, le guide belge recommande l'utilisation de comparaisons indirectes ou de modèles. La prise en compte du futur est source d'une incertitude additionnelle. L'utilisation d'analyses de sensibilité peut dans ce cas se révéler pertinente²⁹.

4.1.5 Comparaison indirecte et modèles

Le guide hollandais indique que lorsque l'on réalise une comparaison, la meilleure méthode reste un essai clinique randomisé « à l'aveugle ». Cette idée est soutenue par Edwards *et al.* (2009). Les essais cliniques randomisés sont établis de façon à minimiser les biais³⁰. Toutefois, il est nécessaire que ces essais soient menés dans les règles afin de détecter des possibles différences cliniques entre les interventions étudiées. Lorsque les essais cliniques sont de faible qualité, les résultats peuvent être erronés. Il est peu probable, selon Edwards *et al.* (2009), que toutes les nouvelles technologies fassent l'objet d'essais cliniques randomisés. Malgré l'absence de données, des recommandations doivent être faites ; les comparaisons indirectes peuvent alors être une solution.

Cinq méthodes peuvent être utilisées afin de réaliser des comparaisons indirectes :

- Comparaison des intervalles de confiance (95%) ;
- Réaliser un test statistique à partir des données estimées ;
- Réaliser une comparaison indirecte à partir d'un comparateur commun ;
- Comparaison à partir d'un mix de traitement (*Mixed treatment comparisons*) réalisé au moyen de différents essais cliniques ;
- Comparaison des « *point estimates* ».

²⁷ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé."

²⁸ EUCOMED, (2001). "Health Technology Assessment for Medical Devices in Europe – What has to be considered?"

²⁹ Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." **3** : 34.

³⁰ Song *et al.* (2003). "Validity of indirect comparisons for estimating efficacy of competing interventions: empirical evidence from published meta-analyses." *BMJ* **326**: 472-6.

Les comparaisons indirectes réalisées au moyen d'un comparateur semblent être la méthode la plus répandue. Par ailleurs, c'est la seule approche à avoir été validée de façon empirique³¹. Elles sont mentionnées dans les guides belge, français et canadien sous certaines conditions. On note qu'il est recommandé de les éviter si cela est possible.

Au Canada, s'il n'existe aucun essai clinique comparant directement l'intervention et les comparateurs retenus, il est possible d'employer des comparaisons indirectes en appliquant les techniques appropriées et en expliquant et en justifiant les méthodes utilisées pour synthétiser les résultats de ces comparaisons. Dans les deux guidelines, on recommande d'indiquer les limites de la comparaison indirecte et d'indiquer les effets que l'emploi de cette méthode peut avoir sur l'interprétation des résultats et il est demandé d'expliquer pourquoi l'on choisit une comparaison indirecte. Le guideline français accepte les comparaisons indirectes dans la mesure où aucune étude originale ne permet de confrontation directe. Cette solution n'est acceptable que provisoirement³².

³¹ Edwards, S. J. *et al.* (2009). "Indirect comparisons of treatments based on systematic reviews of randomised controlled trials." *The International Journal of Clinical Practice* **63**(6): 841-54.

³² Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé."

4.2 Prise en compte des coûts

4.2.1 Perspective

Les coûts pris en compte lors d'une analyse sont fortement liés au point de vue adopté qu'il est nécessaire d'explicitier clairement. Selon le cadre méthodologique (parcouru encore par de nombreux débats) et le point de vue adopté, le périmètre des coûts diffère dans les guidelines.

Le guide français ainsi que le guide autrichien³³ recommandent, sans que ce soit une obligation, de se placer dans une perspective sociétale.

Les guides suédois, écossais, hollandais et polonais ont été développés dans le cadre d'une perspective sociétale uniquement alors que les guides anglais, et canadien ont été élaborés en vue d'une aide à la décision publique. Le guide canadien adopte le point de vue de l'Assurance Maladie, le guide anglais recommande d'adopter celui du *National Health Service* et du *Personal Social Services* (NHS et PSS). Le guide élaboré par la Nouvelle-Zélande doit aider le financeur du traitement ou de la nouvelle technologie : le point de vue du « payeur » y est donc adopté. Le guide allemand prend en compte le point de vue des assurés du SHI. Pour finir, les études belges tiennent compte du point de vue du payeur (gouvernement ou patient) de soins.

Malgré la diversité des perspectives adoptées dans les études, tous les guidelines s'accordent sur l'importance de la prise en compte de celle-ci lorsque l'on détermine les coûts devant être considérés au cours de l'étude.

4.2.2 Présentation et estimation des coûts

A. Quantités

Pour des questions de transparence³⁴, la plupart des guides (anglais, canadien, autrichien, allemand, suédois, polonais, écossais, hollandais et néo-zélandais) recommande une double présentation des coûts en quantité et en valeur monétaire. Seuls les guides belge et français ne font pas référence aux quantités.

³³ Walter, E. ; Zehetmayr, S. (2006). "Guidelines on Health Economic Evaluation." Institut für Pharmaökonomische Forschung. Page 4.

³⁴ Walter, E. ; Zehetmayr, S. (2006). "Guidelines on Health Economic Evaluation." Institut für Pharmaökonomische Forschung. Page 5.

B. Présentation des coûts

Les guides belge³⁵ et canadien³⁶ recommandent que les coûts soient présentés dans la valeur courante de la monnaie, il est donc nécessaire d'utiliser les prix déflatés (au Canada, par exemple on utilise l'IPC).

C. Coût d'opportunité

Les guides autrichien et belge insistent sur la notion de coût d'opportunité, ce qui n'est pas le cas du guide français (hormis pour les soins de ville : « Si on se situe dans une perspective sociétale, il est toutefois légitime de regarder si les tarifs de remboursement constituent des estimations satisfaisantes des coûts d'opportunité correspondants³⁷. »). En effet, l'utilisation des coûts d'opportunité permet d'évaluer qu'elle aurait pu être la meilleure utilisation alternative des ressources utilisées.

Ces deux guides utilisent la même méthode pour évaluer les coûts d'opportunité : en l'absence de meilleures alternatives et pour des raisons d'uniformité entre les analyses, on utilise les prix de marché même s'ils ne reflètent pas réellement les coûts d'opportunité. S'ils ne sont pas connus, on peut toujours utiliser les prix moyens ou les montants remboursés. Lorsque les coûts ne sont pas disponibles, le guide autrichien est le seul à recommander aux auteurs des études de les estimer eux-mêmes à partir de données précédemment publiées en Autriche ou de valeurs moyennes.

Le guide anglais demande que les coûts financiers payés par le NHS/PSS soient intégrés au modèle, même s'ils ne reflètent pas les coûts d'opportunité.

D. Bases de données, sources utilisées

Le guide français détaille de façon assez précise les différentes bases de données pouvant être employées (bases de données issues du PMSI, Système d'Information de l'Assurance Maladie, tarifs de remboursement...) ainsi que leurs limites. Peu de guides indiquent les sources à la disposition des auteurs. Le guide néo-zélandais renvoie les auteurs vers quatre bases (PHARMAC, *the New Zealand Health Information Service*, Ministère de la santé et les District Health Boards), l'anglais recommande les bases de données développées par le NHS tout en étant le seul à préciser leurs limites. Les guides belge et suédois donnent des indications pour certains coûts uniquement (les prix des médicaments sont disponibles auprès des pharmacies suédoises - *Apoteket's Sales Prices* - et les coûts des séjours en milieu hospitalier sur le site de la Cellule Technique pour la gestion des données en Belgique).

³⁵ Cleemput, I. ; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. **28** : 18.

³⁶ Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." **3** : 29.

³⁷ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." Page 24.

Quant à la Nouvelle-Zélande, en raison de possibles différences au niveau des systèmes de santé et des pratiques, elle déconseille explicitement l'utilisation de données obtenues dans d'autres pays³⁸. La Belgique interdit également la conversion de coûts provenant de la littérature ou d'études réalisées à l'étranger³⁹ : les valeurs monétaires indiquées doivent refléter les coûts/prix belges et non pas des prix étrangers convertis en euro.

E. Analyse de sensibilité

Tous les guides confirment que les coûts peuvent être difficiles à évaluer et qu'une analyse de sensibilité est donc nécessaire. Les guides français et suédois recommandent une analyse de sensibilité dans leur partie traitant de l'incertitude^{40,41}. Le guide anglais la recommande lorsque différentes sources sont disponibles. En revanche, le guide hollandais en préconise uniquement pour les coûts ayant l'impact le plus important sur les résultats.

Le guide canadien préconise l'utilisation d'analyses de sensibilité pour une tout autre raison : elles doivent aider le payeur public à prendre ses décisions. Dans le cas où le point de vue adopté relève de plusieurs provinces, le scénario de référence sera élaboré à partir des coûts moyens canadiens (s'ils sont connus). Les coûts les plus élevés et les coûts les plus bas (supportés par les provinces) sont utilisés dans les analyses de sensibilité. Une analyse de sensibilité est également demandée lorsque les coûts sont partagés entre le payeur public et les patients afin de prendre en compte les possibles variations dans la proportion des coûts à la charge du payeur public. Le décideur peut alors prendre sa décision en se référant à « la proportion la plus pertinente pour les prévisions de remboursement sous sa responsabilité. »⁴².

F. Estimation des coûts

Les guides allemand et autrichien indiquent que la réalisation d'analyse marginale est souhaitable pour estimer le coût d'une unité supplémentaire. L'utilisation des seuls coûts moyens est tolérée lorsque les coûts marginaux ne sont pas disponibles^{43,44}. Dans tous les cas, une indication explicite du type de coût utilisé est indispensable.

³⁸ Pharmaceutical Management Agency. (2007). "Prescription for Pharmacoeconomic Analysis." Page 45.

³⁹ Cleemput, I. ; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. **28** : 20.

⁴⁰ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." Page 44.

⁴¹ Pharmaceutical Benefits Board. (2003). "General guidelines for economic evaluations from the Pharmaceutical Benefits Board." **3** : 2.

⁴² Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." **3** : 34.

⁴³ Graf von der Schulenburg, J.M. Greiner W. et al. (2008). "German recommendations on health economic evaluation: third and updated version of the Hanover consensus." *Value in Health*. **11**(4): 542.

Au Canada, il n'existe pas de consensus concernant la meilleure méthode de valorisation des coûts. Toutefois, la valeur marchande si elle est disponible doit être utilisée.

4.2.3 Typologie : Coûts directs et coûts indirects

Lors de la prise en compte des coûts, la distinction entre coûts directs et indirects est généralement faite (hormis pour la Nouvelle-Zélande et l'Angleterre). On note cependant que la typologie n'est pas toujours claire. Peut-être serait-il plus simple de ne pas l'utiliser comme c'est le cas en Angleterre et en Nouvelle-Zélande.

En France, les coûts directs couvrent toutes les ressources utilisées en raison d'une action de soins de santé et qui lui sont directement attribuables. Ils sont séparés en coûts directs médicaux et en coûts directs non médicaux. Les coûts directs médicaux comprennent tous les coûts médicaux résultants du traitement ou de la maladie.

On observe que d'autres pays (les guides hollandais, autrichien, polonais et allemand) font la distinction entre coûts directs médicaux et non-médicaux sans toutefois les détailler, les définir ou donner des informations sur leur évaluation.

Le guide allemand indique que la charge supportée par l'entourage des patients (coûts directs non médicaux), si elle peut être évaluée, pourra être présentée en annexe.

Lors de l'estimation des coûts directs médicaux, le guide belge recommande de prendre en compte les prix des médicaments génériques, même s'ils sont peu utilisés en Belgique⁴⁵. Lorsque le même médicament est vendu sous des noms et à des prix différents, on considère le prix le plus bas car il permet d'obtenir un meilleur ratio coût-bénéfice. Les coûts directs non médicaux sont exclus des évaluations belges.

Le Guide français détaille ces coûts directs et indique les sources à utiliser pour leur estimation. Les coûts directs médicaux sont les coûts liés :

- à l'hospitalisation (estimation à partir des bases de données issues du PMSI ou au moyen d'enquêtes *ad hoc*) ;
- au transport (obtenus grâce au Système d'Information de l'Assurance Maladie). Il faut distinguer transport primaire (domicile-hôpital) et secondaire (entre établissements) ;

⁴⁴ Walter, E. ; Zehetmayr, S. (2006). "Guidelines on Health Economic Evaluation." Institut für Pharmaökonomische Forschung. Page 5.

⁴⁵ Cleemput, I. ; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. 28 : 20.

- aux soins de ville (tarifs conventionnés utilisés comme base de valorisation des coûts en médecine de ville plus parts non remboursées et les dépassements d'honoraires lorsque la perspective est sociétale)⁴⁶.

Dans les études françaises, les coûts directs sont souvent assimilés aux coûts directs médicaux uniquement⁴⁷. Les coûts directs non médicaux étant assimilés à tort à des coûts indirects.

4.2.4 Prise en compte de différents coûts

A. Temps perdu et débours monétaires

Deux types de coûts peuvent être pris en compte : le temps « perdu » et les débours monétaires. Comme exemples pour le temps « perdu », le temps nécessaire pour accéder aux services de santé, le temps de visite de la famille, le temps de garde des enfants... La mesure de la perte de temps est réalisée à partir d'une enquête ad hoc. Lorsque la maladie est de courte durée, on pose l'hypothèse que ces coûts ne vont pas modifier les résultats (il faudra alors justifier cette hypothèse dans l'analyse). Plus la maladie est longue, plus il est pertinent de prendre en compte ces coûts.

Le bilan complet des coûts doit porter une attention toute particulière à ce type de coûts lorsque la technologie étudiée entraîne un transfert de charges (du milieu hospitalier à la famille par exemple). Des coûts médicaux moins lourds peuvent se traduire par un transfert de charge important sur les familles.

Le temps perdu est en général pris sur le temps de loisir. La décision de recourir à une aide extérieure dépend du revenu familial et de la valorisation du temps hors travail. Il est alors indiqué que ce temps perdu n'a pas de coût d'opportunité pour la société. Il est ainsi comptabilisé à part et représente un argument supplémentaire dans la comparaison de deux actions de soins. Ce raisonnement comporte toutefois des limites. Si une importante partie de la population exprime le besoin d'une aide permanente, la population aidante moins nombreuse ne serait plus disponible pour les activités domestiques ou pour consommer⁴⁸. Cette situation pourrait générer des effets négatifs sur la croissance économique. C'est l'importance des coûts attendus qui doit guider le choix d'une valorisation monétaire. Pour valoriser ce temps, il est possible de le faire au coût horaire correspondant au coût d'une ressource équivalente sur le marché. Dans une approche coût-bénéfice, le temps perdu est pris en charge par la propension à payer des patients (willingness to pay).

⁴⁶ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." Page 24.

⁴⁷ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." Page 22.

⁴⁸ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." page 25.

Les débours monétaires sont les coûts pour réaménager un domicile en raison d'un handicap. Dans ce cas, la valorisation des biens est faite au prix de marché et via des dépenses identifiées grâce à une enquête ad hoc auprès des familles.

Bien que le guide suédois n'identifie pas les coûts directs et les coûts indirects il précise néanmoins que tous les coûts associés au traitement ou à la maladie doivent être identifiés, quantifiés et évalués.

B. Les pertes de production et de productivité

Des patients actifs peuvent voir leur productivité diminuer en raison d'un traitement et d'une pathologie. Cette baisse de productivité impacte également la production de richesse.

Le Guide français est le seul à conseiller l'utilisation de trois méthodes :

- i) la méthode dite de capital humain
- ii) la méthode des « *friction costs* »
- iii) et une troisième méthode dérivée de la théorie du bien-être (réalisée au moyen des études coûts par QALYS)⁴⁹.

B.1. Méthode du capital humain

L'impact de la maladie et de sa prise en charge est valorisé par la mesure du nombre d'heures de travail perdues multipliées par le salaire brut augmenté des cotisations sociales employeur (Drummond et al, 2005). Cette évaluation des pertes de production ne concerne que le secteur salarié, dans lequel la rémunération du facteur travail peut-être mesurée de façon isolée (la question du travail indépendant n'y est pas traitée) (Sultan-Taiëb et al, 2009). L'impact de la maladie est alors le même quelque soit le secteur d'activité. Cette méthode est critiquée car elle part de l'hypothèse que nous sommes en situation de plein emploi et que la perte d'une journée de travail a un « impact proportionnel et mécanique sur la production⁵⁰ ».

B.2. Méthode des « friction costs »

En réalité, les employeurs ajustent la gestion de leurs ressources humaines lors de la survenue d'arrêts maladie. La méthode des « *friction costs* », en étudiant de façon empirique ces ajustements, pallie le manque de réalisme de la méthode du capital humain. L'absentéisme d'un individu ne se traduit pas nécessairement par une baisse de la production : lorsque son absence est de courte durée, un ajustement de la productivité des autres travailleurs ou le rattrapage de son travail par l'employé à son retour peut empêcher une baisse de la productivité ou de la production. Il peut toutefois exister une période d'ajustement au cours de laquelle l'employé ne sera pas remplacé, c'est le temps qu'il faut à l'employeur pour procéder aux ajustements nécessaires pour combler l'absence

⁴⁹ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." page 26.

⁵⁰ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." page 26.

d'un travailleur. La méthode des coûts de friction évalue ces laps de temps et les baisses de productivité et de production qui en découlent.

B.3. Approche dans le cadre d'études coûts par QALYS,

Ici les « inconvénients liés aux arrêts de travail sont évalués en les intégrant dans l'évaluation qui est faite par les patients de leur état de santé suite à un traitement donné. Les coûts indirects sont donc inclus au dénominateur du ratio coût-résultat, et n'apparaissent pas de façon explicite⁵¹ ».

B.4. Méthode recommandée

En France, le choix de la méthode est libre, toutefois, la méthode des coûts de friction requiert un travail empirique conséquent avant d'être appliquée.

Parmi les autres guides, aucun consensus n'existe quant au choix de la méthode la plus pertinente. Le guide canadien recommande les deux méthodes indifféremment alors que les guides suédois, polonais et autrichien se fondent sur la méthode du capital humain. Toutefois, en Autriche, dans le cas d'une absence prolongée ou d'une mort prématurée, la situation réelle du marché du travail peut-être prise en compte (l'employeur peut remplacer le poste laissé vacant). Dans ce cas, c'est la méthode des coûts de friction qui est préconisée et son utilisation doit-être justifiée.

Les Pays-Bas emploient la méthode des coûts de friction car elle donne une meilleure approximation des coûts supportés par la société. Notons tout de même que la méthode des *friction costs* fournit des estimations inférieures à celles obtenues à partir de la méthode du capital humain. Selon Oliva et al. (2005), les pertes de productions associées au cancer du sein en Espagne en 2003 s'élèvent à environ 256,8 millions d'euros selon la méthode du capital humain et à 11,6 millions d'euros selon la méthode des coûts de friction lorsqu'on applique un taux d'actualisation annuel de 3%.

C. Les coûts futurs

Lorsqu'une intervention a un impact sur la survie des patients, ils vont consommer des ressources durant ces années de vie gagnées. Le débat, toujours d'actualité, porte surtout sur les ressources médicales (liées à la technologie étudiée ou non). Faut-il tenir compte de ces coûts futurs dans les analyses médico-économiques ? Il n'y a pas de controverse au sujet des coûts médicaux futurs liés à l'intervention (Garber et Phelps, 1997; Meltzer, 1997; Brouwer et al., 2001 ; Nyman, 2004 ; Van Baal et al.; 2007 ; Lee, 2008).⁵² Le débat porte sur les futurs coûts médicaux non liés à l'intervention.

⁵¹ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." page 44.

⁵² Feenstra, T. L. ; van Baal, P. H. M. ; Gandjour, A. ; Brouwer, W. B. F. (2008). "Future costs in economic evaluation. A comment on Lee." *Journal of Health Economics*. 27: 1645-1649.

Selon le guide français, une hypothèse peut-être faite dans le cas d'un système d'assurance maladie universelle basé sur une solidarité intergénérationnelle fondé sur le concept de neutralité actuarielle id-est la somme actualisée des cotisations tout au long de la vie est égale à la somme actualisée des dépenses de santé.⁵³ Les dépenses futures de santé des survivants seront prises en compte uniquement si cette hypothèse n'est pas vérifiée.

Dans les guides, la prise en compte des coûts futurs est traitée de différentes façons, sans toutefois que des justifications soient systématiquement apportées. Le guide suédois recommande que les coûts futurs soient pris en compte lorsque le traitement affecte la survie des patients. Les guides hollandais et polonais prennent en compte les coûts futurs liés au traitement alors que ceux qui ne découlent pas du traitement ou de la pathologie sont exclus de l'analyse. Quant au guide belge, il recommande de ne pas tenir compte de ces coûts tandis que dans les guides allemand, autrichien et anglais le sujet n'est pas abordé. Le guide néo-zélandais, contrairement aux autres guides, précise les modalités d'étude de ces coûts. Il indique qu'en théorie, les coûts futurs doivent être pris en compte dans les analyses. Cependant, il est compliqué de les estimer du fait du manque de données disponibles. Recueillir ces données requiert un temps conséquent ; les coûts futurs sont en définitive absents des analyses.

Si peu d'indications concernant les coûts futurs sont données dans les guides, la littérature traitant de ce sujet est abondante. La prise en compte des futurs coûts médicaux non liés à la technologie peut biaiser le ratio coût-utilité en rendant la technologie étudiée moins attirante. Russel (1986) considère qu'il n'est pas raisonnable de pénaliser un programme de soins de santé en incorporant les dépenses de soins de santé réalisées par des survivants. Au contraire, Weinstein et al. (1980) avancent que puisque ces coûts existent, ils ne doivent pas être ignorés.

Nyman (2004) montre qu'il faut être cohérent dans la décision d'inclure ces coûts. Les coûts doivent être pris en compte au numérateur si l'utilité liée à la consommation des mêmes biens est prise en compte au dénominateur du ratio coût-utilité. Selon cet auteur, il faut inclure les coûts directement liés à la mesure de l'utilité et les coûts liés à l'intervention (même s'ils ne sont pas à l'origine de gains d'utilité). Les coûts, générant des gains d'utilité non pris en compte dans une analyse coût-bénéfice, seront exclus. Richardson et Olsen (2006) pensent que Nyman aboutit aux bonnes conclusions mais pour de mauvaises raisons. Ils suggèrent que les préférences individuelles ne reflètent pas nécessairement les préférences de la société et que c'est le bien-être social qui doit être pris en compte et non pas les utilités individuelles. La décision de prendre en compte les coûts futurs dépend de la fonction de bien-être social. Enfin, on remarque que le choix de prendre en

⁵³ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." Page 28.

compte ou non les coûts médicaux futurs dépend fortement de la modélisation et de la contrainte budgétaire utilisée (Lee, 2008). En effet, Lee montre qu'être cohérent (ou robuste) n'est pas suffisant et que les différents résultats obtenus dépendent du type de contraintes budgétaires utilisées. A ce stade, aucune réponse claire n'a pour le moment été donnée et les auteurs continuent de débattre de ce sujet.

D. Les coûts intangibles

Les coûts intangibles (coûts traduisant « la perte de bien-être du patient et de son entourage »⁵⁴) sont peu mentionnés. Le guide français ne les prend d'ailleurs pas en compte car ils sont compris dans les études QALYS ou coûts-bénéfice. Le guide polonais souligne les difficultés liées à leur estimation.

E. Non prise en compte des coûts « indirects »

Trois pays recommandent de ne pas prendre en compte les coûts indirects. Le guide anglais les exclut car ils ne sont pas à la charge du NHS ou du PSS (c'est leur point de vue qui est adopté dans l'analyse). En Belgique, il en va de même, mais si l'auteur les considère comme réellement important, il peut les présenter séparément à l'aide de la méthode du capital humain.

En Nouvelle-Zélande, ces coûts ne sont pas pris en compte pour différentes raisons. Tout d'abord, ces coûts sont difficiles à estimer. Les hypothèses sont parfois peu réalistes (plein emploi dans la méthode du capital humain) et peuvent rendre les résultats non valides. Le point de vue adopté n'étant pas une perspective sociétale, inclure des coûts supportés par la société peut fausser les résultats de l'analyse et in fine la prise de décision. Le temps nécessaire à l'évaluation de ces coûts étant important, les coûts d'opportunité sont conséquents. Enfin, un argument éthique est avancé : inclure les coûts liés aux pertes de productivité peut biaiser les résultats vis-à-vis de ceux qui ne font pas parti de la force de travail⁵⁵ (enfants, retraités...). Toutefois, même si ces coûts ne sont pas évalués, la Nouvelle-Zélande, tout comme la France, recommande qu'ils soient incorporés dans les études QALYS.

F. Cas de la Nouvelle-Zélande

Le guide néo-zélandais ne prend pas en compte les mêmes coûts et surtout, il ne fait pas de distinction claire entre coûts directs et indirects. Les coûts pris en compte sont traités selon qu'ils sont des coûts pharmaceutiques, des coûts directs médicaux à la charge des patients ou des coûts hospitaliers. Tout comme le guide anglais, ce guide ne prend en compte que les coûts financés par les fonds publics (NHS et PSS en Angleterre). Lorsque les coûts sont à la charge des patients, seuls ceux

⁵⁴ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé." Page 44.

⁵⁵ "incorporating differential earning levels will results in valuing one group of individuals more than another...they tend to bias against those who are not in the labour force... (Prescription of Pharmacoeconomic Analysis, 2007).

remboursés ou partiellement subventionnés par le gouvernement sont pris en compte. Ainsi, le guide néo-zélandais exclut les coûts indirects et les coûts directs non médicaux car ils ne sont pas à la charge du gouvernement.

4.2.5 Long terme vs court terme

Alors que les guides semblent supposer que les coûts sont stables dans le temps, certains articles traitant des coûts dans l'analyse médico-économique se réfèrent à leur flexibilité et aux différents éléments pouvant en être la cause.

Dans le guide de la Nouvelle-Zélande, la flexibilité des coûts fait référence à la perte de brevets et au développement de génériques. Si un brevet expire dans moins de dix ans, on recommande de tenir compte de la date à laquelle des médicaments génériques pourront être mis en vente sur le marché et de la baisse des prix qui en découlera. Si le brevet expire dans plus de dix ans, il faut donner une date approximative à laquelle les génériques seront développés et un prix possible. Smith (2003)⁵⁶ indique que pour certaines technologies, des royalties peuvent être payées lorsqu'un brevet existe : plus la technologie est répandue, plus son coût devrait diminuer.

Les coûts administratifs sont fixes à court terme mais peuvent varier dans le temps. Il est donc conseillé, dans le cas d'une analyse de long-terme, d'estimer les variations de coûts administratifs⁵⁷. Les coûts des interventions peuvent baisser dans le futur lorsque les praticiens deviennent plus entraînés et plus efficaces (effets d'apprentissage). Finalement, des économies d'échelles peuvent exister. Afin de tenir compte des variations possibles des coûts dans le temps, Smith recommande des analyses de sensibilité.

⁵⁶ Smith, M. W. and Barnett, P. G. (2003). "Direct measurement of health care costs." Medical Care Search and Reviews. 60(3): 74S-91S.

⁵⁷ Smith, M. W. and Barnett, P. G. (2003). "Direct measurement of health care costs." Medical Care Search and Reviews. 60(3): 74S-91S.

4.3 Incertitude

L'incertitude ne peut pas être dissociée des évaluations médico-économiques. Elle renvoie à la méthode d'évaluation utilisée relevant de choix subjectifs et aux paramètres du modèle. Les auteurs des analyses doivent également gérer l'incertitude liée aux projections cependant cette question est traitée la plupart du temps dans un chapitre intitulé *actualisation*. L'ensemble des guides recommande d'analyser les différentes incertitudes, l'objectif étant de tester la robustesse des résultats, de savoir dans quelles mesures ils peuvent être affectés afin de proposer des alternatives dans la prise de décision.

4.3.1 Incertitude liée aux paramètres et au modèle

A. Origines de l'incertitude

L'incertitude a deux origines : les choix méthodologiques et les paramètres. Dans le cas des choix méthodologiques, l'incertitude est liée à la méthode analytique, aux hypothèses, à la structure du modèle ou aux sources de données. L'incertitude liée aux paramètres renvoie à l'incertitude probabiliste entourant la vraie valeur d'un paramètre du modèle. Les données provenant de différentes sources peuvent être incomplètes et affectées par l'incertitude. Des hypothèses sont faites au sujet des valeurs de certains paramètres. Seul le guide belge⁵⁸ fait référence à une troisième source d'incertitude : lorsqu'on dispose de données sur des populations particulières ou liées à un contexte donné, la généralisation des résultats de l'analyse s'accompagne d'une incertitude élevée et de potentiels biais de mesure.

D'après le guide français⁵⁹, les origines de l'incertitude peuvent être combinées. Toutefois, on observe que l'impact d'un paramètre sur les résultats et son niveau d'incertitude dépendent de la façon dont le paramètre est intégré au modèle et de sa relation avec les autres paramètres du modèle⁶⁰. Lorsque l'on étudie l'incertitude des paramètres, les résultats sont valables pour le modèle étudié uniquement. L'incertitude du modèle doit donc être évaluée séparément. En effet, on note que plusieurs guides indiquent comment traiter séparément les deux sources d'incertitude (guides canadien, belge et néo-zélandais).

B. Incertitude liée aux paramètres

Afin d'étudier l'incertitude liée aux paramètres, la méthode suivante est recommandée (aux Pays-Bas, en Nouvelle-Zélande et au Canada) :

⁵⁸ Cleemput I.; Crott R.; Huybrechts, M. ; Van Wilder P. and Ramaekers, D. (2006). "Pharmaco-economic guidelines" KCE reports.28: 20.

⁵⁹ Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé". 44.

⁶⁰ Ishak J. ; Gestios D. et Caro J. (2009). "When does value of information analysis add value?" ISPOR Europe.

- Choix des paramètres : les paramètres les plus incertains et/ou ayant le plus d'impact sur les résultats (guide néo-zélandais). L'omission d'un paramètre dans l'analyse doit être justifiée (d'après les résultats d'une analyse préliminaire par exemple – le guide canadien –). Si les paramètres ont différentes composantes, le guide belge recommande de les analyser séparément.
- Identification et justification de la variabilité retenue pour les paramètres. Ces plages peuvent être définies au moyen de la littérature scientifique existante, des opinions d'experts (guides néo-zélandais, polonais, autrichien et anglais) ou en utilisant un intervalle de confiance précis autour de la moyenne dans le cas des données stochastiques (Drummond et al, 1997).
- Estimation des nouveaux résultats en référence à la situation initiale.
- Interprétation des résultats. Le guide hollandais rappelle qu'il est important d'adopter le point de vue de l'utilisateur. Pour l'interprétation, l'utilisation de l'élasticité est recommandée en Nouvelle-Zélande. Elle indique dans quelle proportion les résultats de l'analyse changent lorsque les paramètres varient d'une unité. L'utilisation de l'élasticité permet un jugement objectif. Plus elle est élevée, plus les résultats sont sensibles à la variation d'un paramètre.

La méthode recommandée pour obtenir les nouveaux résultats varient d'un guide à l'autre. Les analyses de sensibilité sont fréquemment utilisées (notamment *a minima* des analyses de sensibilité déterministes). Les analyses de sensibilités univariées et multivariées peuvent être associées à des analyses de seuil (guide hollandais) et à une étude des cas extrêmes (cas le plus favorable et le moins favorable) dans le cas des guides hollandais, néo-zélandais et canadien. Le guide polonais recommande l'utilisation des analyses de sensibilité déterministes uniquement lorsque les paramètres ne sont pas déterminés par calcul probabiliste. On remarque toutefois que les analyses de sensibilité déterministes ne permettent pas d'explorer simultanément un grand nombre de variables. En outre, certaines valeurs sont plus probables que d'autres. Cette limite, ainsi que les possibles corrélations entre variables ne sont pas prises en compte dans les analyses. Enfin, les résultats des analyses ne sont valables que pour une combinaison donnée de valeurs des paramètres constants⁶¹. Pour ces raisons, d'autres guides recommandent davantage l'utilisation d'analyses de sensibilité probabilistes. En Allemagne, on suggère l'utilisation de la méthode de Monte-Carlo (en complément des analyses de sensibilité déterministes). Ces analyses permettent une analyse plus

⁶¹ Riou França L. et Launois R. (2005). "Prise en compte de l'incertitude dans l'évaluation des technologies de santé. Simulation de Monte Carlo, Familles de distribution de probabilité, calibrage des distributions". REES France.

approfondie de l'incertitude et de meilleures estimations des résultats (guides anglais, canadien et néo-zélandais). En effet, elles prennent en compte les interactions des différents paramètres et fournissent des résultats quantitatifs pouvant être interprétés. Elles sont plus susceptibles de donner lieu à une estimation non biaisée des coûts et des effets moyens, ainsi qu'à des décisions optimales compte tenu de la non-linéarité des liens.

On observe que les guides canadien et écossais conseillent d'utiliser la simulation de Monte-Carlo afin de déterminer les intervalles de confiance et d'estimer l'impact sur les résultats de la variation simultanée des variables au sein des plages plausibles préalablement définies. D'après le guide écossais, les analyses de sensibilité probabilistes sont préférées aux analyses de sensibilité déterministes car elles permettent l'utilisation de la courbe d'acceptabilité (outil utilisé afin de présenter l'incertitude autour des ratios coût-efficacité). Enfin, les guides français et belge sont les seuls à aborder certaines méthodes utilisées pour étudier l'incertitude. Dans le guide français, plusieurs méthodes sont recommandées afin d'estimer l'intervalle de confiance du ratio coût-résultats lorsque l'on fait face à une incertitude d'échantillonnage :

- les approches paramétriques : la méthode de la boîte, des séries de Taylors, la méthode de l'ellipse de confiance, la méthode de Fieller (dans le guide belge également) ;
- les approches non paramétriques : le bootstrap naïf (tout comme dans le guide belge)

Ces méthodes produisent des intervalles très différents. L'examen de leurs propriétés statistiques suggère qu'elles ne peuvent être utilisées que sous certaines hypothèses spécifiques.

Comme le guide écossais, le guide français souligne qu'il est souvent intéressant d'établir la courbe d'acceptabilité, c'est-à-dire la courbe représentant la probabilité que la stratégie étudiée soit coût-efficace en fonction de différentes valeurs-plafond du ratio coût-résultat. Quant au guide belge, il recommande également l'utilisation du ratio coût-efficacité incrémental, du plan coût-efficacité et de la méthode du bénéfice net.

C. Incertitude liée au modèle

L'incertitude liée au modèle peut être plus forte que l'incertitude liée aux paramètres⁶². Toutefois, le sujet est moins abordé dans les guidelines. Le guide canadien, tout comme les guides français, néo-zélandais et belge, recommande que l'incertitude soit étudiée au moyen d'analyses de sensibilité déterministes (analyse univariée, multivariée, de seuil, de cas extrêmes) ; les analyses de sensibilité probabilistes étant utilisées pour l'étude de l'incertitude des paramètres uniquement. Il est possible de traiter cette incertitude en identifiant les hypothèses et les paramètres incertains pour lesquels une analyse de sensibilité est nécessaire. On choisit alors une hypothèse vraisemblable

⁶² Ishak J. ; Gestios D. et Caro J. (2009). "When does value of information analysis add value?" ISPOR Europe.

de variation de ces facteurs. Les analyses de sensibilité sont utiles car elles aident à estimer l'impact des différentes hypothèses sur les résultats. La réalisation d'analyses univariées successives sur l'ensemble des paramètres, à condition qu'ils soient bien indépendants les uns des autres, peut être suffisante pour rendre compte de la robustesse des résultats. Il est possible de réaliser une analyse de sensibilité multivariée sous forme de scénario afin d'atténuer les difficultés de présentation et d'interprétation des résultats qui peuvent être rencontrées. L'utilisation de scénarios extrêmes est également envisageable, ainsi qu'une analyse de sensibilité en termes de seuil (détermination des valeurs critiques). Enfin, les techniques de simulation de Monte-Carlo d'ordre 2 sont utilisées afin de déterminer la distribution des résultats qui dépend de la loi suivie par les paramètres. Les résultats sont obtenus sous forme probabiliste. Un intervalle de confiance peut alors être déterminé. De plus, la simulation de Monte-Carlo permet d'estimer la probabilité de dominance stricte d'une stratégie par rapport à une autre, ce qui constitue un outil d'aide à la décision.

Les guides recommandent l'utilisation de certaines méthodes sans vraiment justifier leur choix. A l'exception du guide néo-zélandais et du guide canadien, les guides n'indiquent pas tous les avantages et les inconvénients de ces différentes méthodes. La littérature scientifique apporte des éléments sur ces points.

4.3.2 Méthodes appréhendées dans la littérature scientifique

A. Ratio coût-efficacité incrémental

Lorsque l'on compare deux technologies de santé, le ratio coût-efficacité incrémental (RCEI) est fréquemment utilisé. Il s'agit du rapport entre la différence de coûts moyens (ΔC) et la différence d'effets moyens (ΔE) des deux groupes (traité et contrôle) : $RCEI = \Delta C / \Delta E$. Ce ratio est utilisé notamment dans deux types de calculs : le plan coût-efficacité et les intervalles de confiance (méthode de la boîte et courbe d'acceptabilité).

B. Plan coût-efficacité⁶³

L'objectif est de montrer la différence de coûts et de résultats entre le groupe contrôle et le groupe traité. L'axe des abscisses représente la différence d'effets moyens (ΔE , E = effets) et l'axe des ordonnées représente la différence de coûts moyens (ΔC). Les coordonnées d'un point sont (ΔE , ΔC). Le plan est divisé en quatre quadrants.

⁶³ Briggs, A. H. (2004). "Statistical approaches to handling uncertainty in health economic evaluation." *European Journal of Gastroenterology and Hepatology*. **16**: 551-561.

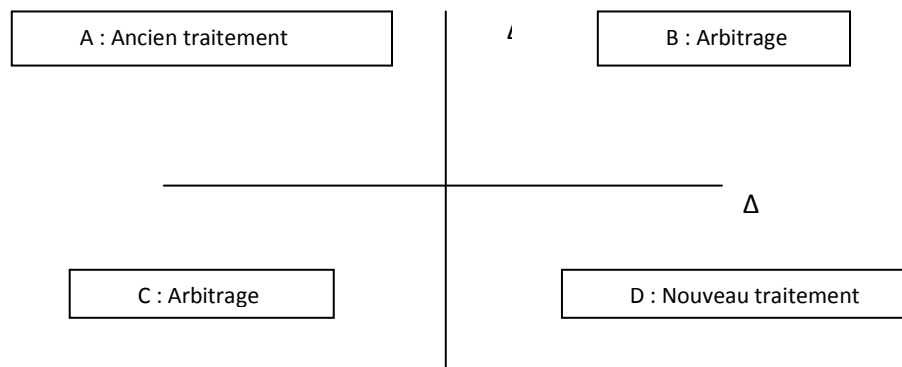


Figure 1 : Le plan coût-efficacité

S'il est possible de placer un point dans le quadrant sans incertitude, alors différentes situations existent. Lorsque le point appartient au quadrant D, le nouveau traitement est préféré (coût-efficace). Au contraire, si le point appartient au quadrant A, la stratégie de soins de santé du groupe contrôle est dominante. Dans ces deux situations, le choix se porte sur la stratégie dominante.

La réalité n'est pas toujours aussi claire. Les traitements sont souvent plus efficaces, mais également plus coûteux. Dans ce cas, la décision va dépendre de la volonté à payer : les bénéfices additionnels valent-ils les coûts supplémentaires ? Le traitement est adopté lorsque le ratio coût-efficacité incrémental est inférieur au « ratio plafond » (représentant la volonté à payer pour une unité d'efficacité supplémentaire) du payeur.

C. Méthode dite de la « boîte »^{64,65}

En réalité, nous ne connaissons que les estimations des différences de coûts et d'effets. Afin d'utiliser une méthode basée sur l'hypothèse de normalité, on accepte les hypothèses suivantes :

- Les coûts et les résultats sont distribués selon une loi normale.
- La taille de l'échantillon est supérieure à 30.
- Les différences de coûts et de résultats sont indépendantes.

Il est alors possible de calculer les intervalles de confiance (au moyen de méthodes statistiques standards) et de les placer dans le plan. Ils peuvent être obtenus à partir de méthodes paramétriques ou non paramétriques.

⁶⁴ O'Brien, B. J. et Briggs, A. H. (2002). "Analysis of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods". *Statistical Methods in Medical Research*. **11**: 455-468.

⁶⁵ Briggs, A. H. (2004). "Statistical approaches to handling uncertainty in health economic evaluation." *European Journal of Gastroenterology and Hepatology*. **16**: 551-561.

C.1. Le Bootstrap

Grâce au développement d'outils de plus en plus performants, la méthode du Bootstrap s'est peu à peu imposée (Briggs, 2004). Elle a pour but d'améliorer la robustesse d'un modèle prédictif et de renseigner sur la stabilité du modèle⁶⁶.

Le bootstrap procède de la façon suivante (O'Brien et Briggs, 2002) :

- Un nombre important d'échantillons bootstrap identiques à l'échantillon étudié sont artificiellement générés.
- Les calculs sont répétés pour tous les échantillons.
- Le résultat permet d'obtenir un vecteur de bootstrap qui estime de façon empirique la distribution de l'échantillon.

La procédure est réalisée pour le groupe traité et pour le groupe contrôle. On estime les coûts moyens et les résultats. En utilisant les moyennes obtenues, on peut calculer les différences de coûts et de résultats entre les deux groupes et le ratio coût-efficacité incrémental. Les résultats peuvent être placés dans le plan coût-efficacité et l'intervalle de confiance peut-être défini à l'aide des résultats du bootstrap (Carayanni, 2009).

C.2. Méthode paramétrique : le théorème de Fieller

Le théorème de Fieller, qui est basé sur une hypothèse stipulant que les différences de coûts et de résultats suivent une distribution normale, peut être utilisé afin d'estimer les intervalles de confiance et les densités jointes de la suite de variables aléatoires⁶⁷. Siani et de Peretti (2006) montrent que la méthode de Fieller a l'avantage, contrairement à d'autres méthodes (méthode du bootstrap par exemple) de ne souffrir d'aucune instabilité et d'être robuste. En effet, selon Siani et Moatti (2003), la meilleure méthode à utiliser est la méthode de Fieller tronquée.

C.3. Détermination finale de la « boîte »

En plaçant les intervalles de confiance dans le plan coût-efficacité, une boîte est obtenue (limites horizontales et verticales représentant respectivement les intervalles de confiance de la différence de résultats et de la différence de coûts). Le principal avantage de cette méthode est sa simplicité. Deux problèmes majeurs sont les hypothèses de normalité et d'indépendance.

⁶⁶ Dans cette optique, elle fournit des indications statistiques telles que la dispersion, la distribution, l'intervalle de confiance afin de connaître la précision des estimations réalisées.

⁶⁷ O'Brien, B. J. et Briggs, A. H. (2002). "Analysis of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods". *Statistical Methods in Medical Research*. **11**: 455-468.

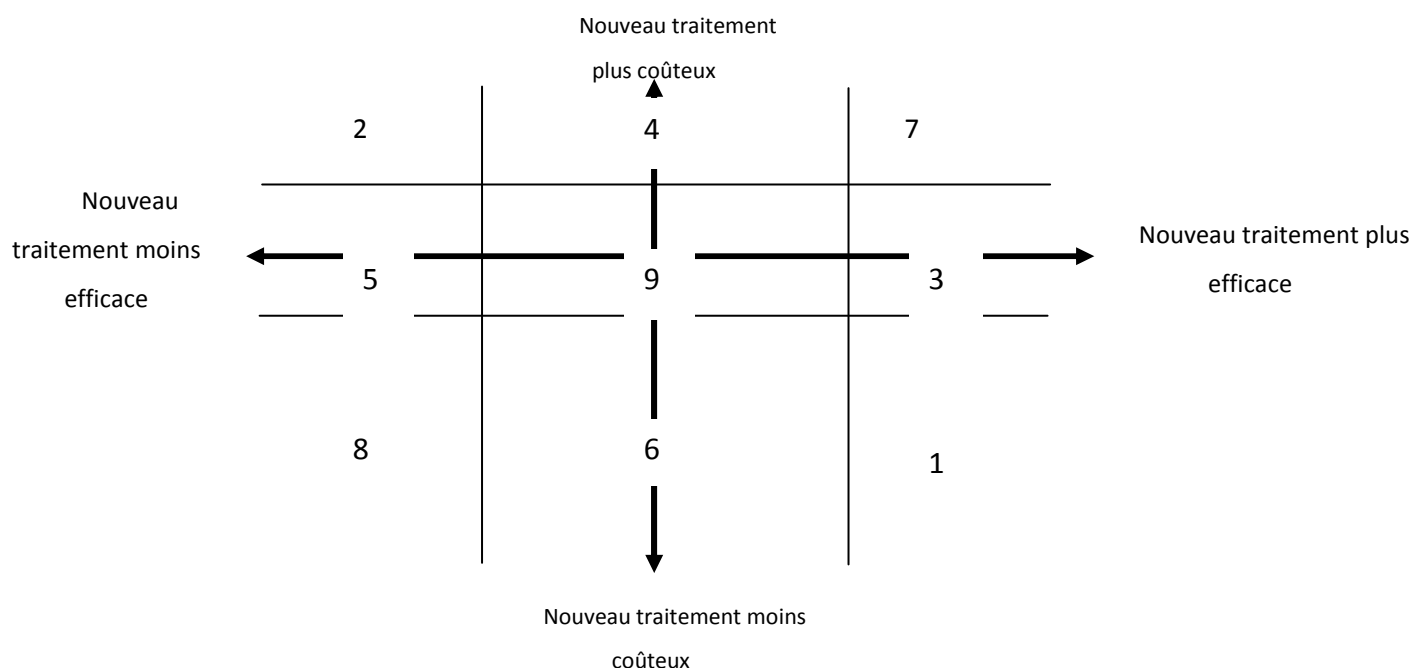


Figure 2 : La « boîte » - Source : Briggs et O'Brien (2002)

L'observation de la boîte nous permet d'identifier neuf situations. Dans les situations 1 et 2, nous avons vu qu'une des deux interventions était plus efficace et moins coûteuse. Il n'y a aucune ambiguïté lors de la prise de décision. Dans les cas 7 et 8, un des traitements est plus efficace mais également plus coûteux.

Dans les cas 3 à 6, la différence de coûts (3 et 5) ou la différence de résultats (4 et 6) n'est pas statistiquement significative (Briggs, 2004).

On observe que l'utilisation des ratios coût-efficacité incrémental n'est pas optimale dans les neuf situations. En effet, des ratios de même signe mais appartenant à des quadrants opposés ne sont pas comparables. Un ratio négatif appartenant au quadrant A (préfère le traitement actuel) est qualitativement différent d'un ratio négatif appartenant au quadrant D (préférence pour le nouveau traitement). On peut parler du problème de la décision « miroir » (Siani et Moatti, 2003). Le plan coût-efficacité, le ratio coût-efficacité et la méthode de la « boîte » ayant des limites, l'utilisation d'autres méthodes est nécessaire.

D. La courbe d'acceptabilité

Une nouvelle approche a été développée à partir de la règle régissant la prise de décision afin de dépasser les limites des autres méthodes. Si un ratio coût-efficacité est inférieur à un certain plafond, R_c , reflétant la volonté à payer des investisseurs (montant maximum qu'ils acceptent de payer pour

une unité d'efficacité supplémentaire), alors la nouvelle technologie doit être développée. Ainsi, dans le plan coût-efficacité, il est possible de tracer une courbe dont la pente est le ratio coût-efficacité optimal. Les résultats obtenus par la méthode du bootstrap qui sont situés sous et à droite de cette courbe, justifient le développement de la nouvelle technologie. Cette méthode permet d'estimer la proportion des résultats en faveur du nouveau traitement.

Même si R_c n'est pas connu, il est possible de faire varier sa valeur afin d'estimer la proportion de ratios coût efficacité en faveur de la nouvelle intervention. Cette présentation des résultats est appelée courbe d'acceptabilité. Même si la courbe d'acceptabilité a été développée afin de comparer deux interventions, elle peut également être utilisée dans le cas de décisions concernant de multiples technologies (Fenwick E. et al, 2001). Cette approche permet de quantifier sans ambiguïté la probabilité que l'intervention soit coût-efficace en fonction de divers plafonds fixés par le décisionnaire. Toutefois, là encore, il est nécessaire de définir la distribution de paramètres inconnus.

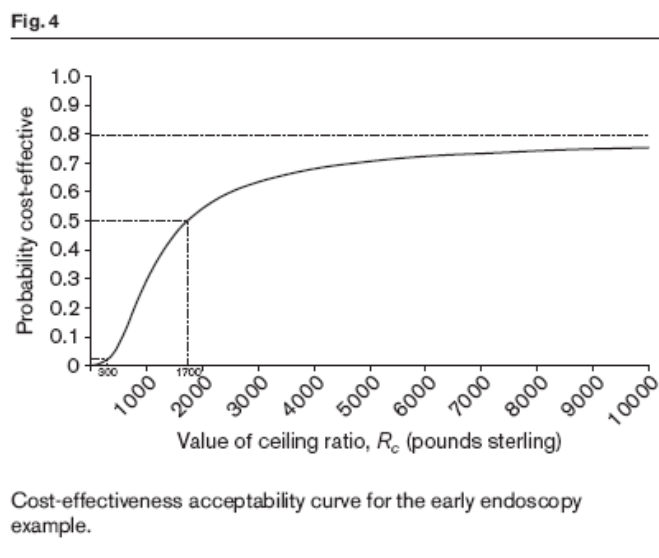


Figure 3 : Briggs A.H.(2004), Statistical approaches to handling uncertainty in health economic evaluation.

E. Bénéfice net

D'autres auteurs (Claxton et Posnett, 1996 ; Stinnet et Mulahy, 1998 ; Tambour et Zethraeus, 1998) ont redéfini le ratio coût-efficacité afin de contourner certains problèmes (ratios de même signes par exemples). La règle stipulant qu'une nouvelle technologie doit-être adoptée si le ratio coût-efficacité incrémental est inférieur à un ratio plafond peut-être modifiée et exprimée sous la forme de bénéfices nets (Net Monetary Benefit ou NMB et Net Health Benefit ou NHB).

$$NMB = R_c \Delta E - \Delta C > 0$$

$$NHB = \Delta E - (\Delta C / R_c) > 0$$

Avec R_c le ratio coût-efficacité plafond et ΔE et ΔC les différences d'efficacité et de coûts entre le groupe traité et le groupe contrôle.

L'utilisation des bénéfices nets facilitent les analyses statistiques⁶⁸. En effet, le bénéfice net peut-être exprimé comme étant une combinaison linéaire de la moyenne des bénéfices nets dans chaque groupe. Il est alors possible d'ajouter d'autres variables indépendantes et d'étudier l'interaction des différents éléments avec le traitement afin de voir si les effets de ce dernier sont homogènes parmi les différents sous-groupes. En outre, il est maintenant aisé de calculer la variance des estimateurs et les estimateurs des bénéfices nets suivent une distribution normale. L'obtention de la courbe d'acceptabilité est également facilitée car elle peut-être dérivée des p-value des bénéfices nets positifs.

4.4 Evaluation contingente

Un problème qui se pose en évaluation médico-économique est l'évaluation des biens non marchands⁶⁹ (environnement ou santé par exemple). Par définition, on ne dispose pas de valeurs monétaires pour les biens non marchands et aucun marché ne permet de révéler les préférences des individus. La méthode d'évaluation contingente a pour objectif de connaître les préférences d'une population déterminée⁷⁰. La révélation des préférences peut être obtenue sur un marché contingent c'est-à-dire un marché fictif, construit grâce aux réponses à un questionnaire permettant d'identifier les consentements à payer/à recevoir (CAP ou CAR) des individus. Il n'est pas possible d'observer les comportements des individus ; on se base sur leurs réponses uniquement. L'information recueillie est donc assez limitée. Dans ces conditions, il est difficile d'étudier les biais de comportement. Les critiques faites à l'encontre de l'évaluation contingente se basent principalement sur l'existence de ces biais.

En économie appliquée, ce sont les comportements réels des agents qui sont observés. Des hypothèses permettent ensuite de déduire une mesure des paramètres pertinents⁷¹. Lorsqu'on utilise la méthode d'évaluation contingente, ce sont les intentions des individus qui sont maintenant observées. On fait alors l'hypothèse qu'elles sont un bon indicateur des comportements.

⁶⁸ Zethraeus N. et al., "Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies." *Pharmacoeconomics*, **21**(1): 39-48.

⁶⁹ Selon l'INSEE, on considère qu'une unité rend des services non marchands lorsqu'elle les fournit gratuitement ou à des prix qui ne sont pas économiquement significatifs (éducation, santé, action sociale, administration). Les biens et services marchands sont destinés normalement à être vendus sur le marché à un prix calculé pour couvrir leur coût de production.

⁷⁰ Bonnieux, F. (1998). "Principes, mises en œuvre et limites de l'évaluation contingente". *Économie publique*.

1: 47-90.

⁷¹ *Ibid.*

L'évaluation contingente a tout d'abord été développée dans les secteurs des transports (Jones-Lee, 1976) et de l'environnement, notamment aux États-Unis, dans les années 1980 (cf. Évaluation contingente et le NOAA Panel). Une importante partie de la littérature étudiée est liée au domaine de l'environnement.

4.4.1 Historique et points importants

A. Méthodes d'élicitation des préférences

La méthode d'évaluation contingente appartient à la famille des méthodes de l'élicitations des préférences déclarées. Elle s'oppose aux méthodes étudiant les préférences déclarées. Les préférences déclarées sont basées sur les comportements observés des individus en situation réelle. Dans ce cas, il est difficile de mesurer des éléments intangibles et l'on ne peut pas prédire les réponses pour de nouvelles alternatives. Dans les cas des préférences révélées, on fait face à une situation hypothétique et il est possible de prendre en compte des éléments intangibles et d'éliciter les préférences pour de nouvelles alternatives. Différentes méthodes de préférences révélées existent parmi lesquelles la méthode des choix discrets (qui permet également d'évaluer des consentements à payer), l'analyse conjointe ou bien la méthode hédonique.

B. Evaluation contingente et le NOAA Panel

En 1980, la loi américaine a été modifiée et le principe de « pollueur payeur » est appliqué. Étant donné le montant des sommes en jeu, les méthodes d'évaluation devaient être codifiées rigoureusement. Un panel d'experts présidé par Kenneth Arrow et Robert Solow a été réuni : le National Oceanic and Atmospheric Administration Panel ou le NOAA Panel. Il en a résulté des recommandations précises sur la forme que doit prendre une évaluation contingente pour être jugée valide par les tribunaux. Ces recommandations constituent, encore aujourd'hui, un point de repère essentiel dans la pratique de l'évaluation contingente. Suite à la parution des recommandations, les débats et critiques, encore d'actualité, se sont concentrés sur cette méthode. La controverse porte principalement sur le fait de savoir si les réponses fournies sont valables ou si les défauts de la méthode nous empêchent de les prendre en compte.

C. Evaluation contingente et choix rationnels

Avant 1993, des résultats empiriques ont montré une incohérence (« inconsistance ») des résultats avec les hypothèses liées aux choix rationnels. Des choix rationnels impliquent une certaine cohérence : si le prix d'un panier de bien diminue, toutes choses étant égales par ailleurs, alors les quantités, l'utilité voire les préférences de l'individu devraient être modifiées. Tant qu'un individu n'a

pas atteint son niveau de satiété, une plus grande consommation devrait augmenter son utilité et se traduire par un consentement à payer supérieur.

Il semblerait que les choix des individus ne soient pas toujours rationnels⁷². Les détracteurs de la méthode d'évaluation contingente utilisent cet argument. Deux éléments justifient de vouloir des résultats rationnels :

- Nous ne savons pas comment raisonner face à des valeurs qui ne répondent pas aux hypothèses de rationalité.
- Il est difficile de trouver des alternatives nous permettant de vérifier les valeurs obtenues. Les estimations obtenues grâce à la méthode d'évaluation contingente sont supposées être représentatives de la réalité. Si les choix réalisés par les personnes interrogées ne sont pas rationnels, il ne sera pas possible de justifier de la concordance entre préférences hypothétiques et préférences révélées.

D. Evaluation contingente et secteur de la santé

La méthode d'évaluation contingente consiste à interroger des individus afin qu'ils révèlent le montant maximum qu'ils sont prêts à payer pour une politique de santé (un bien, une technologie de santé...) ou le montant minimum qu'ils seraient prêts à accepter pour y renoncer. Différents types d'évaluations peuvent être menés : il est possible d'explicitier un comparateur ou pas. La plupart des évaluations contingentes évaluent un programme uniquement sans comparateur explicite alors que d'autres évaluations comparent des substituts ou les bénéfices retirés de différentes ressources. On note également que dans le domaine de la santé, différentes catégories de répondants existent : les patients, la population générale ou un échantillon.

Il existe trois types d'arguments supportant l'idée qu'il faille utiliser la méthode d'évaluation contingente dans le secteur de la santé⁷³ :

- Le concept de « Consentement à payer » est théoriquement correct car tiré de l'économie du bien-être.
- Le CAP permet une estimation des bénéfices plus compréhensible que celle obtenue avec les QALYs. Trois types de bénéfices sont mentionnés par Olsen et Smith (2001). Le premier élément permettant de mesurer les bénéfices est caractérisé par les résultats de santé mesurés au moyen des éléments clés que sont l'état de santé, la durée et la probabilité. Le deuxième élément va au-delà des résultats de santé et renvoie à l'utilité liée aux

⁷² Arrow, K. *et al.* (1993). "Report of the NOAA Panel on Contingent Valuation".

⁷³ Olsen, J.A. et Smith, R. (2001). "Theory versus practice: A review of willingness to pay in health and health care." *Health Economics* 10(1):39-52.

caractéristiques du processus de soins. Le dernier élément fait référence aux externalités. Ainsi, différentes valeurs sont évaluées dans le domaine de la santé.

- Si l'on souhaite utiliser les analyses afin d'obtenir une meilleure allocation des ressources, il est nécessaire de prendre en compte les coûts d'opportunité. Lors de l'élaboration du scénario, il est possible d'informer les répondants de cet objectif et de lui donner des renseignements sur le contexte (quels soins de santé existants seront réduits, verront leur financement diminuer...).

E. Evaluation contingente et altruisme

Les choix des individus révélés lors d'une évaluation contingente dépendent de différents éléments selon le point de vue adopté.

- Perspective néo-classique

Les préférences exprimées par les individus peuvent être des préférences individuelles et les choix exprimés dépendent uniquement des attributs personnels. Arrow (1963) et Becker (1993) adoptent un point de vue quelque peu différent : tout élément perçu comme approprié par le consommateur est pris en compte dans sa fonction d'utilité et lors de sa prise de décision (et non plus uniquement ses attributs personnels). Toutefois, c'est toujours l'utilité individuelle qui entre en jeu lors de la prise de décision individuelle. Ainsi, si l'altruisme est un élément influençant les choix des individus, c'est seulement parce que l'idée de faire une bonne action accroît leur utilité⁷⁴. Le modèle néoclassique perçoit la rationalité des individus comme étant individuelle uniquement (ils maximisent leur utilité personnelle).

- Perspective institutionnaliste

Premièrement, les institutions jouent un rôle primordial dans le choix des individus (l'individu ne peut pas agir de façon rationnelle sans tenir compte des institutions sociales⁷⁵). Deuxièmement, lorsqu'un individu fait ses choix, il prend en compte le fait qu'ils peuvent influencer les opportunités des autres. Un aspect social et un aspect normatif entrent donc en jeu lors de la prise de décision. Les préférences sont alors perçues comme étant dépendantes du contexte dans lequel évolue l'individu. En effet, l'individu a intégré des normes, des valeurs et s'est développé en fonction des institutions qui l'entourent. Ces éléments sont intégrés lors de sa prise de décision.

Lorsque le CAP reflète des composantes altruistes, la sommation des valeurs individuelles peut alors entraîner un double comptage et biaiser les valeurs collectives associées aux programmes

⁷⁴ Vatn, A. (2009). "An institutional analysis of methods for environmental appraisal." *Ecological Economics*, 68(8-9): 2207-2215.

⁷⁵ Etzioni, A. (1988). "The moral dimension: toward a new economics". *The Free Press*, New York.

(Shiell et Rush, 2003). Par exemple, Arana et Leon (2002) trouvent un CAP de 14% plus élevé pour un programme de prévention réduisant le risque de contracter une grippe chez l'ensemble de la population par rapport à une vaccination qui supprimerait ce risque pour soi-même uniquement. Chanel, Luchine *et al.* (2004) suggèrent d'intégrer explicitement le facteur « altruisme » car celui-ci peut occuper une place essentielle dans le cas d'une évaluation contingente de la mise en œuvre publique d'un programme de prévention. Ne pas expliciter ce facteur peut mener à un double comptage et à une surévaluation du programme de santé étudié⁷⁶.

F. Consentement à payer et consentement à recevoir

Lors d'une évaluation contingente, deux estimations peuvent être réalisées⁷⁷ :

- le consentement à payer (ou CAP). Il représente la somme maximum que les individus sont prêts à payer pour acheter un bien. Le but de l'évaluation contingente est de révéler cette valeur. Pour ce faire, on propose à l'enquêté une situation d'échange marchand dans laquelle il va révéler le prix qu'il est prêt à payer pour acquérir un bien.
- Le consentement à recevoir (ou CAR). Ici, on demande à l'enquêté le prix à partir duquel il est prêt à vendre (renoncer) à un bien donné.

La théorie économique suppose que CAP et CAR s'égalisent à un niveau où l'individu est indifférent entre posséder l'objet ou l'équivalent monétaire.

Pourtant, lorsqu'on observe la littérature existante, on note qu'il n'est pas neutre d'utiliser l'un ou l'autre des modes d'interrogation (Brookshire et Coursey, 1987; Plott et Zeiler, 2005; Sayman et Öncüler, 2005). Flachaire et Hollard (2005) recensent une cinquantaine d'expériences ayant pour but de tester la divergence entre CAP et CAR. La divergence est un phénomène général (présente dans 90% des expériences analysées). Lorsque le ratio CAR/CAP est calculé, on constate que la divergence est souvent de grande ampleur. En effet, interroger un individu sur le montant maximum qu'il est prêt à payer ou le montant minimal auquel il faudrait l'indemniser ne conduit pas au même résultat essentiellement pour deux raisons :

- L'existence de substituts : en l'absence de substitut raisonnable, les agents peuvent exiger un prix minimal (CAR) très supérieur à ce qu'ils seraient prêts à payer pour acquérir ce bien. Schogren *et al.* (1994) montrent que lorsqu'il s'agit de la santé, bien pour lequel il n'existe par de substitut, un agent peut ne pas vouloir céder un bien qu'il ne pourra pas remplacer. Le CAR est alors supérieur au CAP.

⁷⁶ Chanel, O. *et al.* (2004). "Les consentements à payer pour des programmes de prévention sanitaire incluent-ils de l'altruisme?" *Revue Économique*, 55(5): 923-946.

⁷⁷ Flachaire, E. and Hollard, G. (2005). "Une approche comportementale de l'évaluation contingente".

- Les contraintes de revenu : le CAP ne peut excéder le budget d'un individu alors que la CAR n'est pas limitée à la hausse. En effet, une critique traditionnelle du CAR porte sur l'absence de contrainte budgétaire. Les personnes interrogées peuvent alors donner des réponses irréalistes.

Ces deux effets n'expliquent qu'en partie la divergence. La théorie du consommateur standard ne permet pas d'expliquer l'ensemble des comportements observés⁷⁸. La théorie des perspectives développées par Daniel Kahneman en 1979 permet d'expliquer en partie pourquoi le CAP est plus utilisé que le CAR. Rose et Masiero (2010) montrent que l'«aversion à la perte» influence les résultats. Les résultats de leur étude montrent que les estimations obtenues avec les CAP sont plus fiables que celles obtenues au moyen du CAR. Les paramètres estimés témoignent de la présence d'«aversion à la perte». Les CAR sont donc surévalués relativement aux CAP.

Un choix doit donc être réalisé lorsque l'on utilise la méthode d'évaluation contingente : souhaitons-nous évaluer le consentement à payer ou le consentement à recevoir ? Dans l'ensemble de la littérature étudiée, le consentement à payer est préféré.

Lorsque l'on réalise une évaluation contingente dans le secteur de la santé, Drummond *et al.* (1997) distinguent différents CAP en fonction des catégories de bénéfices résultants du programme de santé étudié :

- les bénéfices intangibles qui correspondent à l'amélioration de la santé de chaque utilisateur du programme ;
- les coûts des soins futurs évités ;
- et les gains de production liés à l'amélioration de l'état de santé.

Lorsque l'on évalue uniquement les bénéfices intangibles liés aux changements de l'état de santé de chaque individu, la diminution des coûts de santé futurs et les gains de production étant évalués en utilisant les prix de marché, on calcule le CAP « restreint ». Le CAP « global » est estimé lorsque l'on cherche à connaître la valeur qu'un consommateur individuel attribuerait à un programme de santé spécifique, dans un monde où les marchés de tous les biens seraient privés. Le consentement à payer dépend alors des bénéfices de santé, mais également des économies réalisées sur les dépenses futures et des effets sur le revenu de la diminution de l'absence au travail.

Une évaluation contingente peut donc servir à évaluer l'ensemble des bénéfices d'un programme, et non pas seulement les bénéfices de santé. Il est donc recommandé d'être prudent lors de la réalisation d'une évaluation contingente et d'indiquer aux personnes interrogées les

⁷⁸ Ibid.

bénéfices qu'elles doivent prendre en compte (afin d'éviter un double comptage)⁷⁹. Par exemple, il y aura un double comptage si la personne interrogée intègre les gains de revenu (lié à un moindre taux d'absentéisme au travail) alors que l'auteur réalise une étude au moyen de la méthode du capital humain évaluant les gains de productivité liés au programme.

G. Impact de la langue

La compréhension du scénario et du marché sont des éléments primordiaux pour la réussite d'une évaluation contingente. Il est conseillé de garder un vocabulaire relativement simple. En effet, l'emploi d'un vocabulaire trop compliqué pourrait entraîner un important taux de non réponse de la part d'individus ne comprenant pas le scénario. Dans le cas où ces individus répondraient tout de même au questionnaire, il est probable que leurs réponses ne correspondent pas à leur réel consentement à payer. Les résultats de l'évaluation contingente seraient donc faussés.

Des groupes témoins sont souvent utilisés afin de tester le questionnaire et de le préciser si nécessaire. Cette étape est réalisée pour accroître la crédibilité du scénario et diminuer les possibilités de rejet.

Le choix du vocabulaire doit être fait avec attention car un même mot peut avoir différentes interprétations. Il est important de faciliter la compréhension du scénario et d'éviter toute ambiguïté.

Un problème culturel peut également se poser. Une étude réalisée en français en France n'aura pas nécessairement la même signification pour des francophones d'autres nationalités. Enfin, il faut être attentif à la traduction d'une évaluation contingente car la probabilité d'obtenir différentes interprétations est alors accrue.

Lorsque l'on a décidé si l'on estimait CAP ou CAR et que l'on s'est interrogé sur la rationalité des individus et l'impact de la langue, il est nécessaire d'élaborer le scénario. Il joue un rôle essentiel dans la réussite de l'évaluation contingente.

4.4.2 Construction du scénario hypothétique

A. Le scénario

Une des étapes de l'évaluation contingente est la construction du scénario de référence. Les individus sont interrogés directement par le biais d'enquêtes. On cherche à évaluer, à l'aide de questions appropriées, combien les individus sont prêts à payer pour une modification donnée. Il est demandé aux individus de se placer dans un marché fictif et d'imaginer la réalisation d'une politique ;

⁷⁹ Drummond, M. F. *et al.* (1997). "Méthode d'évaluation économique des programmes de santé". Edition Economica pour la traduction française. 2^{ème} édition.

la situation est donc hypothétique. Les réponses obtenues sont des intentions. La réussite de l'étude dépend de la bonne compréhension du scénario. Il est donc nécessaire de décrire un marché hypothétique le plus vraisemblable possible. L'acceptation du scénario est une étape primordiale. Des individus peuvent indiquer un CAP nul uniquement parce qu'ils refusent le scénario hypothétique ou qu'ils ont des doutes sur la faisabilité du programme. Le scénario fournit une description du bien à valoriser et du fonctionnement du marché contingent.

Lors de l'élaboration du scénario, il faut s'interroger sur le niveau d'information à fournir aux répondants ainsi que sur le véhicule de paiement qui va être utilisé pour financer le projet. Il est également nécessaire d'informer les répondants de l'existence de substituts et de l'importance de prendre en compte leur contrainte budgétaire lors de leur estimation.

B. L'impact de l'information

De nombreuses études empiriques réalisées dans les domaines de la santé et de l'environnement (Bergstrom *et al.*, 1985; Munro *et al.*, 2001; Lee *et al.*, 1998) témoignent du fait que les CAP varient en fonction du degré d'information fourni aux répondants⁸⁰ (même s'il n'existe pas de consensus sur l'ampleur du phénomène ni sur l'effet que peuvent avoir des informations additionnelles).

Lorsque le questionnaire est développé, des groupes témoins sont utilisés afin de déterminer le niveau d'information optimal (Macmillan *et al.*, 2006; Ahlheim *et al.*, 2007). Il faut également réaliser un arbitrage entre la durée du questionnaire et les informations fournies aux répondants⁸¹. Malgré ces précautions, certains individus ne prennent pas toujours en compte les informations qui leur sont nécessaires afin de décider de leur attitude à l'égard du projet⁸². Frör (2008) étudie l'irrationalité des individus à partir de psychologie cognitive. Il se base sur l'hypothèse que, dans des situations réelles, les individus font face à des contraintes (informations limitées, contrainte de temps). Ils doivent donc utiliser des règles de décisions heuristiques, simplifiées et s'appuyant sur leurs expériences passées. Ainsi, les répondants ne traitent pas l'information de la même manière et les résultats peuvent varier.

C. Véhicule de paiement

Le véhicule de paiement est le mode de prélèvement du montant déclaré par les individus. Les réponses des individus sont conditionnées par ce véhicule de paiement⁸³. Il est important que les individus aient conscience du fait que leurs réponses peuvent avoir des conséquences et qu'ils

⁸⁰ Lamiraud, K. *et al.* (2009). "The impact of information on patient preferences in different delivery patterns: A contingent valuation study of prescription versus OTC drugs." *Health Policy*, **93**(2-3): 102-110.

⁸¹ Frör, O. (2008). "Bounded rationality in contingent valuation: Empirical evidence using cognitive psychology." *Ecological Economics*, **68**(1-2): 570-581.

⁸² *Ibid.*

⁸³ Champ, P. A. *et al.* (2002). "Contingent valuation and incentives". *Land Economics*, **78**(4): 591-604.

auront effectivement à payer le montant déclaré si le projet est implanté (Carson et Groves, 2007). En outre, si le véhicule de paiement proposé est peu crédible (les individus pensent que l'argent ne servira pas à financer la politique envisagée ou que la politique devrait être financée par d'autres acteurs...) alors ils peuvent le rejeter et les réponses aux questionnaires ne seront plus exploitables (non réponses ou CAP qui ne représente pas la réalité). Afin d'obtenir des estimations valables, il faut se référer à un mode de paiement réaliste⁸⁴. L'idéal est de proposer un véhicule de paiement déjà utilisé dans la réalité (possible si les individus ont l'habitude de payer). Dans le cas d'un programme visant l'assainissement des eaux, la référence à la facture d'eau comme véhicule de paiement s'impose et est bien acceptée par les individus interrogés⁸⁵. Au contraire, dans une analyse de valorisation d'une forêt périurbaine, la perception d'un droit d'entrée a été rejetée, les personnes interrogées estimant qu'une forêt est ouverte à tous⁸⁶. Afin d'identifier le rejet lié au véhicule de paiement, Arrow *et al.* (1993) recommandent de faire suivre les questions fermées d'une question ouverte demandant à l'individu pourquoi il a choisi telle réponse (Pourquoi avez-vous voté oui/non ? Le programme ne vaut pas la peine d'être financé ; Je ne sais pas ; Quelqu'un d'autre devrait financer ce programme...).

D. Contrainte budgétaire et substituts

Il est important de rappeler l'existence d'une contrainte budgétaire et de possibles substituts⁸⁷. Un individu peut indiquer un CAP de 100 euros, cependant, lorsqu'on l'interroge sur la consommation à laquelle il serait prêt à renoncer pour financer le programme, il risque de revoir son estimation à la baisse⁸⁸. En outre, si l'individu interrogé n'est pas conscient de l'existence de substituts, son estimation sera erronée.

E. Description du bien

L'individu devant être en mesure de comprendre la nature du bien qu'il évalue, la description du bien doit comprendre les quantités disponibles, la qualité du bien, les conditions de productions et les mesures prises pour augmenter la quantité/qualité du bien⁸⁹. La description du bien requiert une attention particulière, notamment si les individus ne sont pas familiers avec le bien étudié.

Il est important de souligner que les réponses des individus sont conditionnées par l'information donnée et peuvent être sensibles à des points tels que la formulation des questions, leur ordre et la

⁸⁴ Mitchell, R. C. and Carson, R. T. (1989). "Using surveys to value public goods: the contingent valuation method". Washington, DC: Resources for the Future.

⁸⁵ Bonnieux, F. (1998). "Principes, mises en œuvre et limites de l'évaluation contingente". Economie publique.

1: 47-90.

⁸⁶ *Ibid.*

⁸⁷ Luchini, S. (2002). "De la singularité de la méthode d'évaluation contingente". Economie et Statistique. **357:** 141-152.

⁸⁸ Arrow, K. *et al.* (1993). "Report of the NOAA Panel on Contingent Valuation".

⁸⁹ Luchini, S. (2002). "De la singularité de la méthode d'évaluation contingente". Economie et Statistique. **357:** 141-152.

présentation de l'information⁹⁰. Ce phénomène est connu sous le nom d'effet de structure (Luchini, 2002). La description orale peut-être combinée à des documents tels que des photos, dessins, cartes, schémas... Toutes ces techniques ont pour but de faire comprendre à la personne interrogée le fonctionnement du marché et la nature du bien pour lequel elle va révéler la nature de ses préférences. Mathews, Freeman et Desvougues (2006) indiquent qu'un support visuel aide à la compréhension. Hibbard *et al.* (2002) ont montré que dans 45% des cas, les individus choisissaient une assurance en santé de qualité moindre et que l'ajout de supports visuels réduisait le taux d'erreur à 16% seulement. L'utilisation de supports visuels peut donc aider à réduire les anomalies lors d'évaluations contingentes. Bateman (2009) réalise une étude empirique et se base sur la littérature scientifique dans le but de montrer que le traitement d'informations numériques par les individus est plus difficile ; les estimations pouvant ainsi être biaisées.

Afin d'éviter l'effet de structure ou le biais hypothétique (différence entre les préférences révélées et celles obtenues de façon hypothétique), il est nécessaire d'élaborer un scénario compréhensible par tous, précis, complet et crédible.

4.4.3 Choix de la méthode de révélation des préférences

Il existe différentes méthodes pour poser les questions relatives à la valorisation du bien. Le choix de la méthode n'est pas neutre et peut induire un biais instrumental.

A. Le système d'enchères

On propose un montant à l'individu. S'il accepte de le payer (refuse), on propose un second montant supérieur (ou inférieur), puis un troisième, et ainsi de suite. Cette méthode fut la plus utilisée lors des premières applications de l'évaluation contingente (Bonnieux, 1998). La principale critique faite à cette méthode est que les réponses dépendent principalement de la première enchère (Mitchell et Carson, 1989). En effet, si l'enchère de départ est élevée, le résultat obtenu sera d'autant plus élevé (biais instrumental lié à l'enchère de départ). Cette méthode engendre également un processus d'apprentissage qui peut s'avérer fastidieux pour l'individu mais qui lui permet également de se familiariser avec le bien à valoriser⁹¹.

B. La question ouverte

On demande aux individus le montant maximal qu'ils sont prêts à payer. Cette méthode est la plus simple à mettre en œuvre mais reste peu utilisée car elle pose des difficultés, notamment lorsque l'individu n'est pas familier avec le bien à évaluer. Il peut alors rencontrer des difficultés

⁹⁰ Bonnieux, F. (1998). "Principes, mises en œuvre et limites de l'évaluation contingente". Économie publique, 1: 47-90.

⁹¹ Bonnieux, F. (1998). "Principes, mises en œuvre et limites de l'évaluation contingente". Économie publique, 1: 47-90.

lorsqu'il tente de déterminer une valeur. Pour cette raison, on observe un fort taux de non réponses⁹². En outre, lorsque le même bien est évalué au moyen de la méthode des enchères et d'une question ouverte, le résultat obtenu est plus faible avec la question ouverte (Cumings et al., 1986). Il est conseillé de n'utiliser cette méthode que pour des biens familiers (Bonnieux, 1989 ; Mitchell et Carson, 1989).

C. La carte de paiement

On propose ex ante des montants ou des intervalles parmi lesquels l'individu doit choisir celui correspondant à son consentement à payer (ou l'incluant, dans le cas des intervalles). Cette méthode a été développée par Mitchell et Carson (1989). L'objectif est de réduire le biais instrumental des enchères et de réduire le taux de non réponse obtenu avec la méthode des questions ouvertes. Le choix des valeurs est subtil puisqu'il ne doit pas induire de biais instrumental. En effet, comme pour le système des enchères, les réponses peuvent être influencées par les montants proposés. Cette technique est surtout utilisée en marketing (Gouriéroux et al, 1998).

D. La question fermée ou technique du référendum

Un montant est proposé à l'individu qui doit indiquer s'il accepte de le payer. Cette méthode facilite grandement la tâche de l'individu qui n'a qu'à répondre oui ou non. Selon Arrow et al (1993), cette méthode est préférable car elle se rapproche d'une situation de marché où le consommateur accepte ou refuse l'échange selon le prix proposé.

Le choix de la valeur proposée reste un point primordial de l'enquête. Ces offres sont aléatoires et tirées selon une loi de probabilité de façon à prendre en compte toutes les valeurs possibles du consentement à payer. L'offre la plus élevée doit être refusée par 100% de la population et la plus faible acceptée par ces mêmes 100%. Ici encore on observe un biais instrumental : la distribution des valeurs pouvant influencer les réponses⁹³. Un autre défaut de la méthode est la perte d'information. Il n'existe pas d'indications quant au niveau exact du consentement à payer. Mitchell et Carson (1989) suggèrent de proposer un autre montant. Cette méthode se rapproche alors de la méthode des enchères et les réponses peuvent, ici encore, être influencées par les montants proposés. Le NOAA Panel recommande d'utiliser cette méthode.

E. Le classement contingent

L'individu doit classer des paniers de biens. Les hypothèses de cette méthode sont donc limitées : existence d'un ordre de préférence et cohérence des choix. Chaque panier doit également comprendre des biens privés. L'évaluation est ici implicite et correspond à un arbitrage entre

⁹² Luchini, S. (2002). "De la singularité de la méthode d'évaluation contingente". *Economie et Statistique*. **357**: 141-152.

⁹³ Bonnieux, F. (1998). "Principes, mises en œuvre et limites de l'évaluation contingente". *Économie publique*. **1**: 47-90.

quantités de biens publics et de biens privés. La principale limite de la méthode réside dans le fait qu'un individu n'est capable de classer qu'un nombre limité de paniers. Cette méthode semble peu utilisée. Elle n'est pas toujours présentée lorsque l'on parle des méthodes de révélation des préférences (Luchini, 2002).

A l'exception du NOAA Panel qui préconise l'utilisation du référendum, aucune méthode ne semble faire l'objet d'un consensus. Le choix de la méthode reste une étape délicate car il peut être à l'origine de biais divers, remettant en cause la validité des résultats.

Lorsque le questionnaire est prêt, il faut s'interroger sur son mode d'administration. La représentativité de l'échantillon et divers biais dépendent de cette étape.

4.4.4 Réalisation de l'enquête, traitement des données et validité du questionnaire

A. Questions socio-économiques

Lors d'une évaluation contingente, le questionnaire permet également de recueillir les caractéristiques socio-économiques des individus (âge, sexe, CSP, revenu moyen...) et des informations concernant leur pratique en lien avec la politique évaluée (la fréquentation d'un parc national dans le cas d'une politique ayant pour but la protection du parc par exemple). Ces informations peuvent être utilisées comme variables explicatives des différences de CAP. En effet, selon le NOAA Panel, le rapport final devrait comprendre une partie expliquant les CAP à l'aide de ces questions complémentaires (Arrow et al., 1993).

B. Mode d'interrogation

La méthode choisie pour réaliser l'enquête est importante car elle a un impact sur les résultats obtenus (du fait de l'existence de biais) et sur la validité de l'échantillon. Nous notons que l'impact n'est pas nécessairement le même que nous étudions la décision de financer le projet (CAP positif) ou le montant moyen du CAP (Maguire, 2009).

B.1. Enquête réalisée par un interviewer

Elle permet de bien définir le marché hypothétique et le bien étudié : des documents peuvent être présentés et l'enquête peut-être suffisamment longue pour que l'individu soit capable d'évaluer correctement le bien. Cette méthode est fréquemment recommandée (Arrow et al, 1993; Bonnieux, 1998). Les enquêtes réalisées par les interviewers ont comme principal inconvénient leur coût. Des solutions moins coûteuses sont alors recherchées.

B.2. Enquête téléphonique

Lorsque les questionnaires sont simplifiés, l'enquête par téléphone est une solution économique ; cependant, le matériel pouvant être utilisé lors de la présentation du scénario est limité.

B.3. Enquête postale

L'enquête postale, tout en restant économiques, permet de présenter un matériel assez complexe. Elle est un compromis acceptable même si elle ne permet pas de vérifier l'existence d'un effet d'ordre (les individus répondent aux questions dans l'ordre qu'ils désirent et cela peut influencer leurs réponses).

B.4. Enquête réalisée par mail

Depuis quelques années, la réalisation de questionnaires par mail ou sur internet se développe⁹⁴ (voir Iraguen et Ortuzar, 2004 ; Li et al., 2004 ; Schwappach et Strassmann, 2006 et Tsuge et Washida, 2003). Ce type de questionnaire a un avantage économique par rapport aux questionnaires réalisés par voie postale : il permet d'interroger des échantillons plus importants et de réaliser des analyses sur des sous-groupes. Les questionnaires par voie électronique peuvent également être réalisés dans un laps de temps relativement court (quelques jours) et les réponses obtenues en ligne peuvent immédiatement être traitées dans des feuilles de calcul ou intégrées à des bases de données. En plus des gains de temps et d'argent, les erreurs humaines sont alors réduites. Enfin, des questionnaires innovants peuvent être développés (supports visuels et audio, alertes lorsqu'on ne répond pas à des questions...)⁹⁵.

L'aspect économique et le matériel présenté ne sont pas les seuls éléments devant être pris en compte afin de faire son choix. L'existence de biais et la représentativité de l'échantillon doivent également être étudiées.

C. Représentativité de l'échantillon

L'échantillon utilisé doit être représentatif de la population étudiée. Avec les enquêtes postales et électroniques, une partie de la population (ne figurant pas dans les annuaires ou n'ayant pas accès à internet) ne peut pas être interrogée, ce qui peut remettre en cause la représentativité de l'échantillon. En outre, avec ce type d'enquêtes, il n'est pas possible de s'assurer qu'une seule personne a répondu au questionnaire.

Pour ces raisons, Arrow et al. (1993) suggèrent que ce type d'enquête devrait être utilisé s'il est associé à une méthode supplémentaire utilisée afin de croiser et de valider les résultats.

⁹⁴ Fleming, C. M. and Bowden, M. (2009). "Web-based surveys as an alternative to traditional mail methods." *Journal of Environmental Management*, **90**(1): 284-292.

⁹⁵ Fleming, C. M. and Bowden, M. (2009). "Web-based surveys as an alternative to traditional mail methods." *Journal of Environmental Management*, **90**(1): 284-292.

Dans le cas des enquêtes téléphonique, le choix de numéros au hasard rend l'échantillonnage approximatif⁹⁶. Le choix de l'échantillon et de sa taille est une étape difficile et technique. Arrow et al. conseillent de faire appel à un statisticien pour cette étape.

D. Le problème de l'incertitude

Le concept de l'incertitude des préférences est de plus en plus présent dans les évaluations contingentes (Akter, Bennet et Akhter, 2008). Li et Mattson (1995) indiquent qu'ignorer l'incertitude des préférences peut entraîner un biais. Des méthodes ont été développées afin d'intégrer l'incertitude dans les estimations. D'après des études empiriques, les estimations qu'elles génèrent sont souvent différentes (Shaikh et al., 2007). Akter, Bennet et Akhter (2008) tentent de savoir si l'utilisation de deux méthodes différentes (largement répandues dans le domaine de l'évaluation contingente) permet d'obtenir des résultats cohérents (« consistant »).

Les deux méthodes étudiées sont l'échelle numérique de la certitude (« numerical certainty scale » ou NCS) et la méthode des choix polychotomiques (PC).

- NCS : les questions fermées sont suivies d'une question numérique demandant aux individus d'indiquer leur niveau de certitude sur une échelle de 1 à 10. Les partisans de cette méthode indiquent que la mesure de l'incertitude est plus précise, car indiquée à partir d'une échelle plus large. Toutefois, les hypothèses de la méthode peuvent être discutées. En effet, on fait l'hypothèse que l'individu est parfaitement capable d'établir son propre niveau d'incertitude. On mesure l'incertitude des préférences car les individus ne sont pas certains de leur estimation du bien à évaluer. Penser qu'ils sont capables d'établir leur niveau de certitude concernant leur estimation peut sembler contradictoire. La seconde hypothèse implique que tous les individus interrogés interprètent l'échelle de la même façon. Il apparaît que certains individus répondent toujours en utilisant les valeurs les plus élevées tandis que d'autres utilisent les valeurs les plus basses (Roe et al., 1996).
- PC : les individus doivent indiquer leur niveau de certitude en choisissant parmi « Oui, certainement », « Probablement oui », « Peut-être que oui », « Peut-être que non », « Probablement non » et « Non, certainement ». Cette méthode fut utilisée pour la première fois par Ready et al. (1995) qui montrent que des individus répondent « oui » car cela ne leur demande pas de s'impliquer ni de réellement réfléchir à la question posée. En outre, cette méthode engendre une interprétation des mots qui peut varier d'un individu à l'autre, les niveaux de certitude exprimés étant alors différents.

Le débat portant sur l'utilisation de ces méthodes est toujours d'actualité.

⁹⁶ Arrow, K. et al. (1993). "Report of the NOAA Panel on Contingent Valuation".

Une fois les informations recueillies, il est nécessaire de traiter les données. Il existe deux techniques largement répandues dans le cas d'une échelle numérique de la certitude :

- le modèle d'incertitude asymétrique (ASUM) : on multiplie les OUI (=1) et les NON (=0) obtenus à la question fermée par les valeurs obtenues au moyen de l'échelle de certitude.
- le modèle d'incertitude symétrique (SUM) : on recode les OUI et les NON obtenus par leur niveau de certitude. Les OUI obtenus avec certitude sont recodés « 1 » et les NON obtenus avec certitude sont recodés « 0 ». Un OUI obtenu avec une certitude de 60% est codé 0,6 alors qu'un NON obtenu avec une certitude de 60% est codé $1 - 0,6 = 0,4$.

Dans le cas de choix polychotomiques, les réponses obtenues sont subjectives ainsi que leur interprétation. Il est possible de calibrer toutes les réponses positives (« Oui, certainement », « Probablement oui », « Peut-être que oui ») par 1 et les réponses négatives par 0.

On a été tenté de définir une relation causale entre le niveau d'incertitude des préférences et des variables indépendantes. Suite au modèle économétrique de Loomis et Ekstrand (1998), il existe un accord stipulant que les variables sont intuitives plutôt que théoriques : manque de connaissances, intérêt insuffisant, incapacité à faire un choix rapidement, la présence de biens substituts et complémentaires et la non compréhension de la contingence (Shaikh et al., 2007).

Étant donné l'absence de consensus concernant la méthode à utiliser pour mesurer et traiter l'incertitude, davantage de recherches s'interrogeant sur la capacité à mesurer adéquatement le niveau d'incertitude des répondants doivent encore être menées⁹⁷.

E. Analyse économétrique des consentements à payer

Après avoir obtenu les CAP des individus interrogés, il est nécessaire d'estimer un modèle des CAP tenant compte de facteurs explicatifs issus des questionnaires. Lucchini (2002) indique que le modèle utilisé est spécifique à la question de valorisation adoptée. Par exemple, dans le cas d'une question ouverte, la spécification économétrique est un modèle usuel dans lequel le CAP (variable à expliquer) est continue alors que lorsque l'on utilise la méthode du référendum, le modèle utilisé est un modèle à choix binaire.

Le traitement des valeurs nulles ou extrêmes peut poser des problèmes lors de l'analyse. En effet, comme l'indique Jorgensen et al. (1999), les répondants aux évaluations contingentes peuvent indiquer un CAP mais cette réponse peut être une réponse protestataire. Ces réponses protestataires sont en général supprimées de l'analyse car on fait l'hypothèse qu'elles ne donnent pas d'indication sur la réelle valorisation du bien par les individus. Ce traitement des valeurs nulles est à l'origine d'une controverse. On ne prend pas en compte le fait que les réponses protestataires et leur

⁹⁷ Akter, S., J. Bennett, Akhter, S. (2008). "Preference uncertainty in contingent valuation." *Ecological Economics*, 67(3): 345-351.

signification peuvent varier en fonction du bien évalué, de la méthode d'interrogation...⁹⁸ La question du traitement économétrique de ces réponses n'est toujours pas résolue.

F. Validité du questionnaire

D'après Arrow et al. (1993) il incombe à la personne réalisant le questionnaire de prouver sa validité. Il doit prouver que son étude ne comprend pas les différents problèmes cités ci-dessus. En effet, si une étude contient les éléments suivants, les résultats seront jugés peu fiables :

- Un important taux de non réponse (le questionnaire dans son ensemble ou le consentement à payer) ;
- Manque de compréhension de la tâche à effectuer par le répondant ;
- Les répondants ne jugent pas le scénario crédible ;
- Pas de justification des réponses OUI/NON.

Le meilleur moyen de vérifier la validité de l'évaluation est de comparer les résultats obtenus avec les montants effectivement payés sur un marché⁹⁹. Etre en mesure d'effectuer un tel test est rare. Toutefois, lorsqu'il a pu être réalisé, il a été observé que les valeurs hypothétiques étaient supérieures à celles observées sur le marché^{100,101,102}.

4.4.5 Existence de biais

A. Critiques communes aux méthodes d'enquêtes

A.1. Le biais de désirabilité sociale

Certains individus ont tendance à modifier leurs réponses car ils recherchent l'approbation sociale. Dans le cas de l'environnement, on observe une surévaluation intentionnelle des CAP (Fisher, 1993 ; Krosnick, 1999 ; Bonnieux, 1998 ; Legget et al., 2003). Il est possible d'identifier ce biais en comparant les résultats de deux questionnaires identiques, l'un réalisé avec un interviewer et l'autre sans. Il est alors nécessaire de garder les autres éléments qui pourraient affecter les réponses inchangés. Legget et al. (2003) étudient l'existence du biais de désirabilité sociale. D'après leur revue de la littérature, les évidences de la présence du biais de désirabilité sociale sont limitées (sept études faites de 1989 à 2002 comparent différents modes d'administration du questionnaire deux à deux). Toutefois, leur étude montre l'existence d'un biais lié au mode d'administration du questionnaire (le CAP moyen est de 7,51\$ lorsque le questionnaire est réalisé face à face et est de

⁹⁸ Jorgensen, B. S. *et al.* (1999). « Protest Responses in Contingent Valuation » *Environmental and Resource Economics* **14**(1) :131-150.

⁹⁹ Bayoumi, A. M. (2004). "The measurement of contingent valuation for health economics". *Pharmacoeconomics*. **22**(11): 691-700.

¹⁰⁰ Johannesson, M. *et al.* (1999). "Calibrating hypothetical willingness to pay responses". *Journal of Risk and Uncertainty*. **18**: 21-32.

¹⁰¹ Lundberg, L. *et al.* (1999). "Quality of life, health-state utilities and willingness to pay in patients with psoriasis and atopic eczema". *British Journal of Dermatology*. **141**: 1067-1075.

¹⁰² Onwujekwe, O. (2001). "Searching for a better willingness to pay elicitation method in rural Nigeria: the binary question with follow-up method versus the bidding game technique". *Health Economics*. **10**: 147-158.

5,71\$ pour un questionnaire auto-administré et le coefficient pour la variable en personne est positif et significatif).

A.2. Le biais lié à l'interviewer

Dans le même ordre d'idée, les individus peuvent modifier leurs réponses pour faire plaisir à l'interviewer¹⁰³. Legget *et al.* utilisent des variables dichotomiques pour évaluer si les interviewers ont un impact sur les résultats. Il semblerait, d'après leurs résultats, que l'interviewer peut influencer les résultats. Arrow *et al.* (1993) recommandent de tester les résultats afin de détecter ce type de biais.

Dans le cas d'enquêtes réalisées en personne, il est possible de limiter le biais de désirabilité sociale et le biais lié à l'interviewer en permettant à l'individu de répondre sur un bulletin de vote qu'il glissera dans une urne scellée, s'assurant ainsi que l'interviewer ne connaîtra pas sa réponse (Arrow *et al.*, 1993 ; Legget *et al.*, 2003).

A.3. Le biais d'autosélection

Lorsque l'on réalise une enquête au moyen d'un interviewer, il est recommandé de la réaliser sur le lieu de résidence et non pas sur site¹⁰⁴ (à l'entrée du parc naturel étudié dans l'étude par exemple). En effet, les individus interrogés sur le site sont susceptibles d'avoir un intérêt particulier pour l'objet à évaluer et ne sont plus représentatifs de la population concernée dans son ensemble. Le biais d'autosélection peut limiter la portée d'une enquête sur site.

Le biais d'autosélection concerne également les enquêtes postales et électroniques. Les taux de retour varient grandement mais peuvent dépasser 50% dans de bonnes conditions. On observe toutefois que le taux de réponse diminue avec l'âge et est plus faible chez les individus ayant un revenu inférieur ou un niveau d'instruction moins important.

B. Critiques communes aux méthodes d'élicitation des préférences déclarées

B.1. Biais hypothétique

Murphy *et al.* (2005) ont réalisé une revue de la littérature montrant que les valeurs hypothétiques sont supérieures aux valeurs réelles. D'autres articles ne montrent pas de différences significatives (Özdemir, Johnson et Hauber, 2009).

Le biais hypothétique existe car les personnes interrogées ne font pas face à un marché réel mais à un marché contingent. En général, les individus n'ont jamais eu à traduire leurs préférences en termes monétaires. Attribuer une valeur monétaire pour la première fois représente un handicap car aucun processus d'apprentissage n'a encore eu lieu (même si le répondant désire répondre le plus sérieusement possible). Les individus doivent assimiler en peu de temps toutes les informations

¹⁰³ Arrow, K. *et al.* (1993). "Report of the NOAA Panel on Contingent Valuation".

¹⁰⁴ Bonnieux, F. (1998). "Principes, mises en œuvre et limites de l'évaluation contingente". *Économie publique*. 1: 47-90.

(parfois nouvelles), et leurs réponses dépendent de leurs capacités cognitives. En outre, il faut que les personnes interrogées acceptent de payer directement mais aussi de se conformer à la méthode de paiement. Cet élément peut-être rejeté, notamment dans le cas de politiques de santé (le financement des soins de santé par l'Assurance Maladie accentue ce problème)¹⁰⁵.

Si par construction, un marché contingent ne peut fournir qu'une demande hypothétique, le signe du biais associé est indéterminé (Bonnieux, 1998). Encore une fois, la description d'un scénario aussi crédible et réaliste que possible semble être la seule solution pour limiter le biais hypothétique.

Schläpfer (2008) tente d'expliquer l'existence du biais hypothétique. Lorsqu'il s'agit d'un bien non familier, les répondants ne sont pas capables d'exprimer leurs préférences à partir des informations concernant le bien et la politique envisagée, à moins que le scénario ne leur fournisse des indices contextuels pouvant être exploités lors d'un processus de réflexion simplifié et heuristique. Cette hypothèse suggère que les préférences des individus sont « développées » lors de l'évaluation contingente (Payne *et al.*, 1999 ; Swait *et al.*, 2002). Une étude de Kahneman et Tversky (1984) présente une nouvelle maladie qui devrait tuer 600 personnes. Les individus doivent choisir entre les propositions « A et B » dans un questionnaire ou « C et D » selon les questionnaires. Les alternatives A et C sont identiques du point de vue des résultats, ainsi que les alternatives B et D. Il n'y a que le vocabulaire utilisé dans le scénario qui change (on parle de « vies sauvées » dans les cas A et B et de « vies perdues » dans les cas C et D), entraînant une différence certaine des résultats. Dans cet exemple, l'auteur montre que les individus n'ont pas la même perceptions des gains et des pertes et qu'il est possible d'altérer les choix des individus en fonction des facteurs contextuels (voir aussi Kahneman *et al.*, 1999 ; McFadden, 1999). Schläpfer (2008) recommande de fournir aux individus des informations contextuelles fiables (basée sur des avis d'experts par exemple) dans le but d'obtenir des estimations les plus fiables possibles (proches de la réalité).

Afin de réduire le biais hypothétique, il est également possible de communiquer à ce sujet auprès des répondants (*cheap talk*). Le scénario les informe de l'importance de l'étude et de l'existence du biais hypothétique (*cheap talk*)¹⁰⁶. Les premiers à utiliser cette méthode dans le domaine de la révélation des préférences furent Cummings et Taylor (1999). Leur scénario comprenait un texte expliquant le problème du biais hypothétique et demandant aux répondants de faire attention aux alternatives existantes et aux coûts d'opportunité. Le chercheur tente d'impliquer les individus et les encourage à fournir davantage d'efforts et d'attention. L'étude d'Özdemir *et al.* (2009) montre que

¹⁰⁵ Özdemir, S ; Johnson, F. R. ; Hauber, A. B. (2009). "Hypothetical bias, cheap talk, and stated willingness to pay for health care". *Journal of Health Economics*. **28**: 894-901.

¹⁰⁶ *Ibid.*

l'utilisation du cheap-talk permet de réduire le biais hypothétique mais permet également d'accroître la qualité des données lors d'études ayant pour but la révélation des préférences.

Un autre moyen préconisé pour éviter le biais hypothétique est le traitement de l'incertitude des préférences. Selon Champ et Bishop (2001), calibrer l'incertitude des préférences peut permettre d'éliminer le biais hypothétique. La littérature scientifique montre que l'utilisation de la méthode NCS (*cf.* Le problème de l'incertitude) représente une garantie contre le biais hypothétique (Champ *et al.*, 1997 ; Loomis et Ekstrand, 1998 ; Samneliev *et al.*, 2005 ; Shaikh *et al.*, 2007 et Chang *et al.*, 2007). Par exemple, un score de certitude inférieur ou égal à 5 sur une échelle allant de 1 à 10 indique que dans la vie réelle, les répondants ne paieraient pas pour la politique envisagée. Dans le cas de la méthode PC (*cf.* Le problème de l'incertitude), il n'existe pas d'évidences permettant d'affirmer que le biais hypothétique est supprimé. La convergence des réponses hypothétiques obtenues avec la méthode PC et les réponses observées dans la réalité dépend du choix des mots et du traitement des réponses.

B.2. Biais d'inclusion

Il a parfois été constaté que le CAP ne variait pas en fonction de l'ampleur de l'objet étudié (Kahneman et Knetsch, 1992 ; Diamond *et al.*, 1992). Protéger un millier d'oiseaux ou dix mille oiseaux n'aurait qu'un effet négligeable sur la valeur donnée¹⁰⁷. Les choix de l'individu ne semblent plus rationnels. En effet, il est raisonnable de penser que tant que le niveau de satiété n'a pas été atteint, une plus grande quantité d'un bien accroît l'utilité d'un individu. Cet élément devrait se traduire par un CAP supérieur (hormis lorsqu'on observe un biais d'inclusion). Le biais d'inclusion apparaît lorsque la définition du bien manque de précision ou de réalisme¹⁰⁸. Les résultats de l'évaluation contingente sont alors remis en cause. Lorsqu'on parle du biais d'inclusion, trois notions sont abordées :

- La notion d'effet d'envergure

Cet effet intervient lorsque les individus ne font pas de distinction entre les quantités d'un bien. Les individus attribuent une somme forfaitaire au bien étudié qui est l'expression de la satisfaction morale d'adhérer à une cause jugée juste. Diamond et Hausman (1993) se réfèrent à un don charitable pour expliquer cet effet (« *Warm glow* » : les montants mentionnés témoignent plus d'une approbation du projet que d'une réelle volonté de payer). La solution réside dans une spécificité adéquate du bien et du marché contingent.

- La notion d'effet d'ordre

¹⁰⁷ Desvousges, W.S. ; Smith, V. K. ; McGivney, M. P. (1992). "Comparison of alternative approaches for estimating recreation and related benefits for water quality improvements". U. S. Environmental Protection Agency.

¹⁰⁸ Arrow, K. *et al.* (1993). "Report of the NOAA Panel on Contingent Valuation".

Il y a un effet d'ordre lorsque la valeur attribuée à plusieurs biens dépend de l'ordre dans lequel ils sont présentés. Le CAP pour protéger une espèce menacée dépendra de sa place lors des évaluations.

- La notion de sous-additivité

On rencontre cette notion lorsque la valeur attribuée à un ensemble de biens est inférieure à la somme des valeurs attribuées à chacune de ses parties. La sous-additivité peut également être observée lorsque l'on agrège les CAP d'un même individu pour des biens différents, les évaluations ayant été obtenues indépendamment les unes des autres.

Le problème est présent dans le domaine de la santé au même titre que dans celui de l'environnement. Lorsqu'on interroge les individus sur plusieurs politiques séparément, il est probable que le CAP soit positif pour toutes les politiques. Cependant, lorsqu'on agrège les CAP obtenus, la somme obtenue peut représenter une importante fraction du revenu ou être supérieure à celui-ci. L'agrégation donne des valeurs anormalement élevées et la contrainte budgétaire n'est plus respectée. Même si les individus répondent sérieusement à un questionnaire portant sur une politique, les CAP sont souvent surestimés lorsqu'il existe un nombre important de politiques en jeu.

Lorsqu'il s'agit de biens publics¹⁰⁹ substituables, le CAP diminue à mesure que l'offre globale augmente en raison des rendements décroissants. Si l'individu ne tient pas compte des substituts, son CAP va être surestimé. La précision du scénario joue encore un rôle important.

Les effets d'ordre et de sous-additivité sont des effets généraux et ne s'appliquent pas uniquement à l'évaluation contingente. Ils sont liés aux relations de substitutions et à la décroissance de l'utilité marginale.

La réalisation du scénario est donc une étape primordiale. Il est nécessaire de s'assurer de sa bonne compréhension et ainsi, d'éviter les possibles biais. L'auteur de l'évaluation contingente doit ensuite décider de la méthode qu'il va employer pour la révélation des préférences.

B.3. Biais positif

Lorsqu'ils découvrent le scénario hypothétique, les individus ont tendance à porter leur attention sur les éléments concernant l'objet à évaluer et à exprimer une « certaine bienveillance » à son égard. En effet, à la lecture des informations données, trois faits peuvent être observés :

- Les individus ont tendance à favoriser la recherche d'informations concernant l'objet qu'ils ont à évaluer. ;

¹⁰⁹ Pour Samuelson, un bien public répond aux deux critères suivants :

- un critère de non-rivalité : cela signifie que la consommation de ce bien par un usager n'entraîne aucune réduction de la consommation des autres usagers ;
- un critère de non-exclusion : il est impossible d'exclure quiconque de la consommation de ce bien ; il est, par conséquent impossible de faire payer l'usage de ce bien.

Sénat, <http://www.senat.fr/rap/r03-233/r03-23319.html>, 15/01/21010.

- A la lecture du scénario, les lecteurs négligent les données concernant les alternatives présentées.
- Enfin, lorsque les informations sont ambiguës, elles sont interprétées comme des arguments en faveur de l'objet de l'étude.

La perception des informations et leur traitement sont déformés, ainsi que le jugement des individus. Ce phénomène peut être à l'origine d'un biais positif (Posavac, 2006).

Posavac *et al.* (2009) cherchent à savoir s'il est possible d'atténuer le biais positif au moyen d'un processus comparatif. Afin que les individus puissent prendre des décisions correspondant à leurs habitudes et à leurs préférences, il faut qu'ils aient connaissance des alternatives existantes. Une expérience à l'aide de cinq questionnaires est menée. Quatre d'entre eux portent sur l'évaluation d'un parc naturel américain en particulier. Les questionnaires sont attribués au hasard et le répondant en est informé. Le cinquième questionnaire porte sur l'évaluation des quatre parcs simultanément. Le scénario contient, dans tous les cas, les informations sur les quatre parcs et les politiques alternatives pouvant être menées dans les différents parcs.

L'analyse des résultats met en évidence l'existence d'un biais positif. Présenter les alternatives dans le scénario ne suffit donc pas à le supprimer. Une solution permettant de le limiter semble être la rédaction d'un questionnaire portant sur les différentes alternatives. Ceci pose toutefois un problème dans le cas de l'évaluation contingente : en effet, on ne se limite plus à une simple évaluation contingente portant sur un seul objet.

B.4. Les biais volontaires

On note l'existence du biais volontaire. Il résulte d'un comportement intentionnel des personnes interrogées. Schläpfer (2008) fait l'hypothèse que la probabilité et l'impact de réponses stratégiques dépendent de la différence entre le coût hypothétique et du coût effectif que l'individu interrogé aura à payer si la politique est implantée. Lorsque les coûts hypothétiques sont attribués au hasard, alors, les disparités entre coûts hypothétiques et coûts réels sont inévitables. L'ampleur des disparités dépend de la distribution des coûts si la politique est implantée et des valeurs proposées dans le questionnaire.

Le cas le plus étudié de biais volontaire est celui du passager clandestin. Le passager clandestin ne désire pas payer pour les autres. S'il pense que l'accès est gratuit, il peut tendre à surestimer son consentement à payer. Au contraire, s'il anticipe un accès payant, il sous-évalue son CAP.

Des techniques permettant de mettre en évidence de tels comportements peuvent être utilisées. Ainsi, une stratification de l'échantillon avec des modes de paiement différents est une solution simple à ce problème (Bohm, 1972). Dès lors que les individus sont convaincus que leur réponse ne

peut pas influencer le montant réel qu'ils auront à payer, il leur est difficile d'avoir un comportement stratégique (Bonnieux, 1998). Pour obtenir ce résultat, il est possible d'indiquer aux répondants que le montant à payer sera égal à la moyenne des réponses. D'après les travaux empiriques, d'autres biais volontaires, parfois difficile à contrôler, existent.

C. Critiques propres à l'évaluation contingente

C.1. Biais d'ancrage

Dans le cas de l'évaluation contingente, le biais d'ancrage existe lorsque les questions portent sur un montant explicite. La valeur proposée influence les réponses des individus. Modifier cette valeur entraîne généralement une variation des estimations obtenues. Du fait du biais d'ancrage, les évaluations contingentes réalisées à partir de questions fermées sont régulièrement remises en question (Flachaire et Hollard, 2005).

Pour observer l'existence du biais d'ancrage, il est possible de poser deux questions successives : une question fermée demandant aux individus de comparer leur estimation à une valeur fournie par l'expérimentateur (l'ancre) suivie d'une question ouverte interrogeant l'individu sur sa propre estimation. Il a été observé que les estimations des individus dépendent fortement de l'ancre fournie au préalable. Ce phénomène peut-être lié au fait que les individus doivent évaluer un bien auquel ils n'ont encore jamais attribué de valeur. L'ancre leur apparaît alors comme une information sur la valeur cible. Des expériences montrent que cette explication semble erronée. Il a été demandé à des individus de comparer leur estimation à des valeurs qu'ils fournissaient eux-mêmes (les deux derniers chiffres de leur numéro de téléphone par exemple)¹¹⁰. Le sujet est alors conscient du fait que l'ancre ne contient aucun élément d'information. Toutefois, l'effet d'ancrage persiste^{111,112}. En outre, il semblerait qu'un avertissement préalable ne semble pas suffire pour limiter le biais d'ancrage (Wilson *et al.* 1996).

C.2. Centrage et valeurs extrêmes

Un autre type de biais observé est lié au fait que certains individus ont tendance à systématiquement choisir les valeurs centrales, ou au contraire, les valeurs extrêmes. En effet, les ajustements des individus ne sont pas complets du fait de la contrainte de temps ou du caractère hypothétique (le répondant n'est pas encouragé à chercher son réel CAP car il n'y a aucune conséquence s'il est différent de ses préférences réelles). Roach, Boyle et Welsh (2002) tentent d'évaluer l'impact du phénomène de centrage et des valeurs extrêmes. Tout d'abord ils se positionnent en faveur de questions fermées à plusieurs valeurs (multiple bounded-question). Cette

¹¹⁰ Flachaire, E. and Hollard, G. (2005). "Une approche comportementale de l'évaluation contingente".

¹¹¹ Mussweiler, T. and Strack, F. (2001). "Considering the impossible: Explaining the effects of implausible anchors". *Social Cognition*. 19: 145-160.

¹¹² Flachaire, E. and Hollard, G. (2005). "Une approche comportementale de l'évaluation contingente".

méthode se rapproche de la carte de paiement, mais ici, les individus doivent donner leur avis pour chaque valeur proposée (dans la méthode de la carte de paiement, l'individu indique quel montant, parmi les valeurs proposées, correspond à son CAP). L'avantage de cette méthode est qu'elle évite de se concentrer sur une seule valeur (contrairement à une question fermée simple).

On note que la manipulation des valeurs proposées peut influencer les estimations. Dans le cas d'un effet de « centrage », les individus ont tendance à répondre négativement pour les premières valeurs et à changer leur réponse en réponse positive à partir des valeurs centrales. Si la plage de valeurs est plus importante, l'estimation devrait être modifiée (même lorsqu'il n'y a pas de « centrage »). En outre, ajouter des valeurs extrêmes ou en tronquer devrait également impacter les estimations. Dans leur étude empirique, Roach, Boyle et Welsh (2002) ne trouvent pas d'évidences concernant le phénomène de centrage. En revanche, ils montrent que tronquer des valeurs extrêmes influence les résultats.

Encore une fois, le choix des valeurs est important. Il faut minimiser les effets qu'il peut avoir sur les estimations.

4.5 Méthode des choix discrets

4.5.1 Origines et bases théoriques

Comme l'évaluation contingente, la méthode des choix discrets est une méthode d'élicitation des préférences. Elle est utilisée afin de permettre aux individus d'exprimer leurs préférences relatives à un bien complexe, une politique des soins de santé... Elle a d'abord été appliquée dans le domaine des transports puis de l'environnement. Lors de l'utilisation de la MCD, les individus sont interrogés sur leurs préférences au moyen de scénarios comprenant des options hypothétiques. Chaque option est décrite au moyen de plusieurs attributs et chaque attribut se voit octroyer différents niveaux. Lors de sa prise de décision, le répondant doit réaliser un arbitrage. Les choix sont ensuite analysés afin d'estimer dans quelles proportions les attributs contribuent à l'utilité totale¹¹³.

A. La MCD dans le secteur de la santé

Les premières utilisations dans le domaine de la santé ont eu lieu au début des années 90. A plusieurs reprises, la MCD fut critiquée (Bryan et Dolan, 2004 et Wainwright, 2003). Même si ces critiques ont été perçues comme justifiées, elles visaient davantage les applications de la MCD dans le secteur de la santé que la MCD elle-même¹¹⁴.

L'emploi de la MCD offre différents avantages, le plus important étant qu'elle permet d'obtenir de nouvelles sources de données disponibles pour la réalisation d'évaluations médico-économiques et la prise de décision (certaines de ces données ne pouvant pas être obtenues autrement)¹¹⁵. Elle permet d'estimer les préférences des individus. Les premières applications dans le secteur de la santé se sont concentrées sur cet aspect (Johnson *et al.*, 2000 ; Ryan et Hughes, 1997 ; Scott, 2001...). On cherchait avant tout à évaluer les résultats autres que les résultats de santé. Plus récemment, il a été observé que la MCD pouvait être utilisée afin d'étudier les intérêts d'une nouvelle politique et pour estimer les résultats de santé. La MCD permet d'obtenir des données pouvant servir pour les études coûts-bénéfices comme pour les analyses coûts-utilité. En effet, elle est de plus en plus utilisée pour estimer le CAP et la valeur monétaire des bénéfices (possible utilisation pour les études coûts-bénéfices). La possible utilisation de la MCD dans le calcul des QALYs est en cours d'exploration^{116,117,118}.

¹¹³ Lancsar, E. et J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* **26**(8): 661-677.

¹¹⁴ Lancsar, E. et C. Donaldson (2005). "Discrete choice experiments in health economics: distinguishing between the method and its application." *Eur J Health Econ* **6**(4): 314-6.

¹¹⁵ Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* **26**(8): 661-677.

¹¹⁶ Hakim, Z. *et al.* (1999). "Modelling the EuroQol data : a comparison of discrete conjoint and conditional preference modelling." *Health Economics*. **8**(2):103-16.

¹¹⁷ Ryan, M. *et al.* (2006). "Using discrete choice experiments to estimate a preference-based measure of outcome: an application to social care for older people." *Journal of Health Economics*. **25**(5): 927-44.

B. Les bases théoriques

Les bases théoriques de la méthode des choix discrets sont complexes car elles combinent différentes théories économiques¹¹⁹ : la théorie des utilités aléatoires (*random utility theory*), la théorie des choix probabilistes (*probabilistic choice theory*) et la théorie économique de Lancaster et le point de vue néoclassique (Lancaster, 1966 ; Manski, 1977).

4.5.2 Mise en place de la méthode des choix discrets

Le développement du questionnaire et du scénario est, comme dans le cas de l'évaluation contingente, une étape importante car on va déterminer la quantité d'information à incorporer, celle-ci pouvant influencer les résultats de l'étude¹²⁰. À ce stade, le chercheur doit identifier les questions de méthodologie pouvant nécessiter une réflexion approfondie.

Plusieurs auteurs ont étudié la mise en place d'une étude faisant appel à la MCD (Green et Srinivasan 1978 ; Gustafsson *et al.*, 2000 ; Louvière *et al.*, 2000 ; Ryan, 1996). Les différences observées concernent principalement le nombre d'étapes que comprend le processus. Nous allons nous baser sur le modèle de Ryan (1996), la littérature médico-économique se basant principalement dessus¹²¹ (Carlsson et Martinsson, 2003 ; Ratcliffe et Buxton, 1999 ; Ryan, 1996 ; Ryan, 1999). Cinq étapes sont donc étudiées:

- A. Étape 1 : Identification des attributs
- B. Étape 2 : Identification des niveaux
- C. Étape 3 : Design expérimental
- D. Étape 4 : Recueil des données
- E. Étape 5 : Analyse des données

Cependant, il est important que le chercheur ait, au préalable, identifié les points clés du problème de décision et les objectifs de l'étude. Le contexte de l'étude et la valeur économique à évaluer sont alors définis. Il est important que le contexte et les alternatives présentées dans l'étude soient compréhensibles. En outre, le répondant doit-être en capacité de croire que les alternatives présentées puissent exister en situation réelle (Bennett et Blamey, 2001).

A. Identification des attributs

¹¹⁸ Viney, R., E. Savage, et al. (2005). "Empirical investigation of experimental design properties of discrete choice experiments in health care." *Health Economics* **14**(4): 349-362.

¹¹⁹ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." *Health Economics Papers* University of Southern Denmark. Odense Denmark, 139p.

¹²⁰ Louvière, J.J., Hensher, D. A. et Swait, J. (2000). "Stated Choice Methods, analysis and application" Cambridge University Press, U.K.

¹²¹ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." *Health Economics Papers* University of Southern Denmark. Odense Denmark, 139p.

Il n'existe aucun consensus quant à la méthode à utiliser pour identifier les différents attributs (Louvière, 2000). Toutefois, il est essentiel que les attributs soient pertinents du point de vue de la problématique de décision. Par ailleurs, ils doivent être compréhensibles et avoir un intérêt pour les répondants (dans le cas contraire, leur variation n'influencera pas les réponses)¹²². Les attributs peuvent être identifiés à partir d'une revue de la littérature ou en faisant appel à des experts ou à des consommateurs¹²³. Keeney et Raiffa (1976) identifient cinq propriétés que doivent avoir les attributs :

- Exhaustivité (*completeness*) : les attributs couvrent l'ensemble des aspects importants de la problématique de décision ;
- Opérationnels : ils doivent signifier quelque chose pour les répondants ;
- Décomposables : l'aspect étudié doit pouvoir être divisé en deux éléments de dimension inférieure ;
- Non redondants : les attributs doivent être définis de façon telle qu'une double prise en compte d'un élément ne soit pas possible ;
- Taille minimum : il est important de maintenir la quantité d'attributs aussi petite que possible. Il n'existe pas de règle générale quant au nombre d'attributs devant être définis, toutefois, il semble qu'il y ait un consensus voulant que le nombre maximum d'attributs soit de huit¹²⁴ même si dans le domaine de la santé, des études présentent douze attributs (Hall, 2004) ont été observées. Plus l'étude comprend d'attributs, plus la tâche cognitive est complexe¹²⁵. En effet, si l'étude comprend un nombre d'attributs trop important, il est possible que le répondant n'en utilise qu'un sous ensemble afin de répondre aux questions¹²⁶. Les résultats risquent alors d'être biaisés.

Lancsar et Louvière (2008) indiquent qu'il n'est pas nécessaire de prendre en compte tous les attributs jugés important par chaque répondant. Toutefois, il faut distinguer les attributs dont l'omission aura un impact sur les conclusions : l'exclusion d'attributs importants peut biaiser l'étude (Kjaer, 2005). En l'absence de certains attributs, il est possible que les répondants fassent des hypothèses ou modifient leurs réponses¹²⁷.

¹²² Blamey *et al.* (2002). "Attribute causality in environmental choice modelling." *Environmental and Resource Economics*, 32(2): 269-286.

¹²³ Lancsar, E. and J. Louvière (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* 26(8): 661-677.

¹²⁴ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." *Health Economics Papers* University of Southern Denmark. Odense Denmark, 139p.

¹²⁵ Mangham, L. J., K. Hanson, et al. (2009). "How to do (or not to do) Designing a discrete choice experiment for application in a low-income country." *Health Policy and Planning* 24(2): 151-158.

¹²⁶ Mangham, L. J., K. Hanson, et al. (2009). "How to do (or not to do) Designing a discrete choice experiment for application in a low-income country." *Health Policy and Planning* 24(2): 151-158.

¹²⁷ Lancsar, E. and J. Louvière (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* 26(8): 661-677.

Il est également recommandé de porter attention aux attributs mutuellement dépendants et à ceux liés par une relation de causalité. Dans le cas des attributs mutuellement dépendants, ceux-ci peuvent être combinés en un unique attribut. Une partie de l'information sera alors perdue. Il est également possible de présenter un des attributs dans le texte introductif. Toutefois, supprimer un des attributs pourrait être à l'origine d'un biais lié à l'omission d'une variable.

Dans le cas d'une relation de causalité, Blamey *et al.* (2002) indiquent, suite à leur étude, que l'utilité est « transférée » dans le cas où l'attribut « causal » n'est pas présenté. Les résultats sont inchangés.

B. Identification des niveaux des attributs

La seconde étape consiste à attribuer des niveaux aux différents attributs. Il faut tout d'abord identifier si les attributs vont être présentés de façon qualitative ou quantitative (valeur absolue ou relative)¹²⁸. Il faut également s'interroger sur la nature des attributs :

- génériques : le même niveau est utilisé pour toutes les alternatives ;
- spécifique : certains attributs ou leurs niveaux sont modifiés en fonction des alternatives¹²⁹.

Par ailleurs, Ryan (1999) indique que les niveaux des attributs doivent être plausibles, recevables et impliquer un arbitrage de la part du répondant. Une attention particulière doit-être apportée aux attributs de risque (risque de morbidité, mortalité...) associés à différents états de santé car les individus ont des difficultés à interpréter les probabilités¹³⁰.

Dans le cas où les différents niveaux sont trop proches ou trop éloignés, le répondant peut percevoir les écarts comme étant insignifiants (respectivement trop importants). Il peut alors ne pas réaliser l'arbitrage nécessaire à l'obtention de résultats valides. En outre, il est important de tenir compte des niveaux proposés lors de l'interprétation des résultats. En effet, un coefficient non significatif ne signifie pas que l'attribut n'est pas important mais que celui-ci n'a pas influencé les choix du répondant¹³¹. Il est également conseillé d'éviter des niveaux qui seraient préférés dans toutes les situations (Green et Srinivasan, 1990).

Le nombre de niveaux peut influencer l'importance des attributs. Ce phénomène est connu sous le nom d'effet-attribut (*attribute-effect*). Même si l'intervalle étudié reste identique, si un nombre plus important de niveau est présenté, l'impact de l'attribut s'accroît (Ratcliffe et Longworth, 2002).

¹²⁸ Bennett, J. et R. K. Blamey (2001). "The choice Modelling approach to Environmental Valuation" Edward Elgar Publishing Limited, UK.

¹²⁹ Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* 26(8): 661-677.

¹³⁰ Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* 26(8): 661-677.

¹³¹ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." *Health Economics Papers* University of Southern Denmark. Odense Denmark, 139p.

Le phénomène peut-être minimisé en attribuant le même nombre de niveaux à chaque attribut (Curry, 1997).

Un type d'attribut particulier est l'« attribut-coût ». De nombreuses études ne le prennent pas en compte. Bryan *et al.* (1998) ont montré qu'une étude incluant un tel attribut engendré une plus grande quantité de données manquantes. Une certaine résistance de la part des répondants est plausible. En effet, des chercheurs ont indiqué que l'avantage de la MCD relativement à l'évaluation contingente était qu'on observait une moins grande probabilité de résistance étant donné le fait qu'on ne demande pas aux individus d'indiquer directement leur CAP. En outre, exprimer le coût renvoie aux problèmes rencontrés lors de l'étude de l'évaluation contingente lorsqu'il faut identifier le véhicule de paiement (il faut pouvoir présenter un véhicule de paiement crédible, réaliste, cohérent avec le bien étudié¹³² et permettant de minimiser les comportements stratégiques). Bateman *et al.* (2002) indiquent qu'il est important de formuler les questions de façon telle que les répondants soient incités à révéler leurs réelles préférences. Ainsi, le coût et le mode de paiement peuvent influencer les réponses (choisir une taxe ou une donation n'aura pas le même effet). Dans le cas d'un système de santé financé collectivement, inclure un coût peut-être perçu comme non réaliste. Enfin, il est recommandé de définir un niveau minimum suffisamment faible afin de ne pas surestimer le CAP des individus (de même, définir un niveau maximum suffisamment élevé pour ne pas sous-estimer le CAP). Malgré ces critiques, la prise en compte d'un tel attribut peut permettre d'estimer l'utilité marginale du revenu¹³³.

C. Design expérimental

La méthode recommandée pour définir le design expérimental ne cesse d'évoluer (DeShazo et Fermo, 2002 ; Street *et al.*, 2005 ; Adamowicz, 2006 ; Louvière, 2006). Cette étape consiste à définir les différents choix qui vont être proposés aux répondants. L'objectif est d'en minimiser le nombre tout en étant capable d'identifier les utilités. On manipule les différents attributs et leurs niveaux afin de tester les hypothèses (Louvière *et al.*, 2000).

Une méthode factorielle est utilisée (*factorial design*). Le nombre d'alternatives s'accroît de façon exponentielle lorsque le nombre d'attributs et leurs niveaux augmentent ($n^{\circ} \text{ d'alternatives} = n^{\circ} \text{ de niveaux} \times n^{\circ} \text{ d'attributs}$). Cette méthode soulève une question : est-il possible de présenter aux répondants toutes les alternatives ou faut-il en réduire le nombre ? Il est alors nécessaire de tenir compte de la quantité d'information et de propositions que le répondant peut gérer.

La méthode factorielle peut-être divisée en méthode factorielle totale ou partielle.

¹³² Smith, R. D. (2003). "Construction of the contingent valuation market in health care: a critical assessment." *Health Economics* 12(8): 609-628.

¹³³ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." *Health Economics Papers* University of Southern Denmark. Odense Denmark, 139p.

- Méthode factorielle totale :

L'ensemble des alternatives possibles sont présentées aux répondants ou à des sous-groupes (chaque sous-groupe devant se prononcer sur une fraction des alternatives). L'utilisation de cette méthode est plus aisée que ce qu'il n'y paraît en raison du recours possible à des sous-groupes. Cette méthode représente un intérêt au niveau de ses propriétés statistiques car elle garantit une indépendance totale des effets attribués à chacun des attributs.

- Méthode factorielle partielle :

Comme son nom l'indique, cette méthode implique de ne présenter qu'une fraction des propositions possibles tout en maintenant les propriétés de la méthode factorielle totale afin que les effets que l'on cherche à évaluer puissent être estimés de la façon la plus efficace possible. Il faut tout de même avoir conscience de la perte d'informations statistiques que cette méthode implique.

Le design expérimental influence le type de fonction d'utilités indirectes que l'on peut obtenir (Lancsar et Louvière, 2008). Il est important de s'assurer de l'efficacité de la méthode utilisée. Lors des premières utilisations de la MCD, la sélection des alternatives et la construction des paires étaient faites de façon non systématique (de façon randomisée par exemple). L'efficacité statistique n'était pas très importante¹³⁴. De nouveaux outils ont été développés. Les principes développés sont regroupés sous le nom de *D-efficiency (design efficiency)*. Quatre principes ont été identifiés par Huber et Zverina (1996) :

- Équilibre des niveaux : les différents niveaux d'un attribut apparaissent à fréquence égale.
- Orthogonalité : les niveaux doivent varier indépendamment les uns des autres.
- « Croisement minimum » : un niveau n'est pas présenté plusieurs fois dans une même proposition de choix.
- Équilibre des utilités : la prise en compte de ce principe peut poser problème car il implique une certaine connaissance des préférences des répondants. Il faut que les utilités des alternatives présentées soit approximativement de même niveau. Cela dans le but de s'assurer du fait que les répondants doivent effectuer un arbitrage pour répondre.

Lorsque l'on définit les différents choix et que l'on développe le questionnaire, il est important de tenir compte de la complexité cognitive de la tâche demandée aux répondants ainsi que des objectifs de la méthode (Louvière, 2000) :

- déterminer les fonctions d'utilité avec un maximum de précision

¹³⁴ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." Health Economics Papers University of Southern Denmark. Odense Denmark, 139p.

- tenter de reproduire une situation proche d'une situation de marché réaliste.

Enfin, on observe que demander aux individus de choisir entre plusieurs options peut poser problème car cela implique qu'ils vont nécessairement consommer le bien évalué¹³⁵. Il peut-être nécessaire d'inclure une alternative additionnelle : le choix de « ne choisir aucune des alternatives proposées » (alternative « statu quo »). Présenter une option statu quo est important lorsque l'étude porte sur une nouvelle technologie, chère, risquée...¹³⁶ Cela permet de prendre en compte le fait que les personnes puissent ne pas vouloir consommer le bien. Bateman *et al.* (2002) indiquent que l'exclusion du statu quo implique une violation de la mesure de bien-être pouvant entraîner un biais¹³⁷. Par ailleurs, inclure cette option permet d'estimer la valeur de « choisir une alternative » par rapport à « ne rien faire » (et non plus uniquement la variation de bien-être impliquée par le choix entre deux options).

Cependant, des problèmes potentiels existent : si un répondant estime que la tâche implique un important effort cognitif, il peut-être tenté de choisir cette option car cela lui évite d'avoir à faire un effort et non pas parce c'est le choix qui lui procure la plus grande utilité. On observe également un phénomène de biais de statu quo : les répondants ont tendance à choisir ce qu'ils connaissent même si cette option leur procure une utilité moindre (Salkeld, 2000). Enfin, cette option peut-être source de difficultés du point de vue économétrique et au niveau de l'interprétation^{138,139}. Dans le domaine environnemental, il y a un consensus qui veut que l'option « statu quo » soit comprise dans l'étude. Dans le secteur de la santé, les tentatives ont été sporadiques (Ryan et Skatun, 2004) toutefois, Lancsar et Louvière (2008) indiquent que du point de vue de l'évaluation, le comparateur d'intérêt est fréquemment un traitement faisant référence à une situation de statu quo. Dans le cas où le chercheur incorpore le choix statu quo, il faut s'interroger sur la signification de cette option pour les répondants¹⁴⁰. Une autre alternative consiste à inclure l'option « indifférent », toutefois elle implique des problèmes similaires¹⁴¹.

¹³⁵ Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* **26**(8): 661-677.

¹³⁶ Adamowicz, W. *et al.* (1998). "Introduction to Attribute-Based Stated Choice Models." US Department of Commerce, NOAA.

¹³⁷ Boyle, K. J. et S. Ozdemir (2009). "Convergent Validity of Attribute-Based, Choice Questions in Stated-Preference Studies." *Environmental & Resource Economics* **42**(2): 247-264.

¹³⁸ Bennett, J. et R. K. Blamey (2001). "The choice Modelling approach to Environmental Valuation" Edward Elgar Publishing Limited, UK.

¹³⁹ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." *Health Economics Papers* University of Southern Denmark. Odense Denmark, 139p.

¹⁴⁰ Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* **26**(8): 661-677.

¹⁴¹ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." *Health Economics Papers* University of Southern Denmark. Odense Denmark, 139p.

D. Recueil des données

La réalisation et l'administration du questionnaire sont des étapes clés. Il peut-être nécessaire de pré-tester son questionnaire auprès de groupes témoins. Cela permet notamment d'affiner son design expérimental, de voir s'il existe des problèmes de compréhension...¹⁴²

Pour le recueil des données, il est possible d'utiliser les mêmes méthodes que celles présentées dans le cas de l'évaluation contingente¹⁴³ (interviewer, téléphone, courrier, internet)¹⁴⁴. On observe, que les questionnaires administrés par voie postale sont fréquemment utilisés même s'ils engendrent un taux de réponse faible¹⁴⁵. Les interviews en personne sont également utilisées mais ne semble pas être les plus coût-efficaces (contrairement aux questionnaires en ligne)¹⁴⁶.

D.1. Le scénario

Lors de l'utilisation de la MCD, choisir une option peut-être une tâche complexe ou inconnue pour le répondant. Il est donc important de prendre en compte l'expérience et les connaissances que le répondant a du bien en question. A partir de ces éléments, il faut estimer la quantité d'information qu'il faut lui présenter. Cela permet d'éviter des situations dans lesquelles le répondant émet des hypothèses ou fait appel à des sources d'information externes et inconnues du chercheur¹⁴⁷. Lorsque l'on présente la situation aux répondants, il est suggéré de :

- spécifier les objectifs de l'étude ;
- introduire les choix qui vont être présentés dans le questionnaire ;
- indiquer le temps que cela devrait nécessiter ;
- assurer de la confidentialité des réponses,
- et présenter la MCD et les différents attributs (Bennett et Blamey, 2001).

Par ailleurs, certaines études ont émis l'hypothèse qu'il puisse y avoir un problème de validité concernant le premier choix réalisé (les répondants étant en train de se familiariser avec la tâche). Inclure un exemple dans l'introduction peut-être une solution.

D.2. Présentation des différents choix

On explique le contexte dans lequel les répondants doivent prendre leur décision. Comme dans le cas de l'évaluation contingente, il est important de rappeler aux répondants l'existence de leur

¹⁴² Mangham, L. J., K. Hanson, et al. (2009). "How to do (or not to do) Designing a discrete choice experiment for application in a low-income country." *Health Policy and Planning* **24**(2): 151-158.

¹⁴³ Bennett, J. et R. K. Blamey (2001). "The choice Modelling approach to Environmental Valuation" Edward Elgar Publishing Limited, UK.

¹⁴⁴ Dillman, D. et D.K. Bowker (2001). "Mail and internet surveys : the tailored design method." New York: Wiley.

¹⁴⁵ Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* **26**(8): 661-677.

¹⁴⁶ Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* **26**(8): 661-677.

¹⁴⁷ Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* **26**(8): 661-677.

contrainte budgétaire, de substituts et de biens complémentaires¹⁴⁸. Dans le cas où l'étude implique des dépenses publiques, il est recommandé de rappeler les coûts d'opportunité que cela peut engendrer. Enfin, le questionnaire doit-être construit de façon telle que le répondant ait le sentiment que les informations qu'il fournit vont être utilisées lors de la prise de décision afin de supprimer le sentiment que l'ensemble de l'étude est purement hypothétique (Bennett et Blamey, 2001).

D.3. L'échantillon

La population dont va être tiré l'échantillon dépend de l'objet de l'étude (des patients souffrant d'un cancer par exemple)¹⁴⁹. A partir de cette population, la taille de l'échantillon (définie à partir de méthodes statistiques) et la stratégie utilisée pour définir l'échantillon (randomisation, sous-groupes en fonction de caractéristiques prédéfinies...) vont pouvoir être déterminées. Bennett et Blamey (2001) indiquent que la taille minimum d'un sous-groupe doit-être de cinquante répondants. Toutefois, la taille optimale d'un échantillon est fortement dépendante du taux de réponse attendu.

D.4. Langage et émotions

Il est possible que les réponses dépendent des émotions impliquées dans le processus d'évaluation. Arana et Leon (2009) ont tenté d'évaluer le rôle des émotions lors de la valorisation de biens non marchands afin de comprendre comment les individus prennent leurs décisions économiques. Ils montrent que les émotions jouent un rôle hétérogène et influencent le comportement des individus. Ainsi, l'évaluation de biens non marchands ou de politiques publiques peut générer des estimations de bien-être différentes en fonction des réactions émotionnelles de la population. Cette conclusion peut-être utilisée afin d'expliquer les comportements des individus qui peuvent être dépendants du contexte (Swait *et al.*, 2002). De nouvelles recherches concernant les déterminants de l'effet de contexte et ces conséquences pourraient aider les chercheurs à mettre en place des questionnaires permettant l'obtention des préférences plus précises¹⁵⁰.

E. Analyse des données

L'analyse des données implique deux étapes : la définition du modèle probabiliste et l'interprétation des résultats. Une approche critique est nécessaire lors de l'interprétation des résultats car si cette dernière est erronée des mauvaises recommandations seront faites.

L'un des objectifs de la MCD est de comparer l'impact relatif des attributs (le temps d'attente des résultats d'un test est-il plus important que la précision du test par exemple). La plupart des études comparent l'impact relatif des attributs en étudiant la taille des paramètres estimés et s'ils sont significatifs. Lancsar *et al.* (2007) indiquent que ces paramètres ne sont pas directement comparables

¹⁴⁸ Bennett, J. et R. K. Blamey (2001). "The choice Modelling approach to Environmental Valuation" Edward Elgar Publishing Limited, UK.

¹⁴⁹ Louviere, J.J., Hensher, D. A. et Swait, J. (2000). "Stated Choice Methods, analysis and application" Cambridge University Press, U.K.

¹⁵⁰ Araña, J. E. et C. J. León (2009). "Understanding the use of non-compensatory decision rules in discrete choice experiments: The role of emotions." *Ecological Economics* 68(8-9): 2316-2326.

car les estimés des paramètres dépendent de « l'échelle d'utilité » sous-jacente : les estimés combinent l'impact relatif d'un attribut et « l'échelle d'utilité » associées à ces différents niveaux. Lancsar *et al.* (2007) ont donc présenté cinq méthodes utilisées afin de comparer l'impact relatif des attributs :

- analyse partielle de log-likelihood (*partial log-likelihood analysis*) ;
- taux marginal de substitution ;
- mesure Hicksienne du bien-être ;
- analyse probabiliste ;
- échelle des attributs.

Les quatre premières méthodes permettent de traiter l'impact relatif des attributs dans le cas d'une MCD traditionnelle. Les effets des attributs sont alors placés sur des échelles communes et comparables. La cinquième méthode utilise une méthode modifiée afin de décomposer les impacts d'un attribut. Chaque méthode présentant des avantages différents, le choix dépendra des autres objectifs de l'étude. Il peut-être avantageux d'utiliser une combinaison des méthodes¹⁵¹.

4.5.3 Biais

A. Présentation de différents biais

L'un des avantages des méthodes de révélation des préférences est qu'il est possible de « contrôler » l'étude. Toutefois, il se peut que les résultats ne correspondent pas à ce qui est observé dans la réalité. La nature hypothétique de l'étude peut alors être perçue comme un inconvénient. Dans le secteur de la santé, les critiques à l'encontre de la MCD indiquent que les préférences révélées sont sujettes à d'importants biais^{152,153}. Étant donné les similitudes entre la méthode d'évaluation contingente et la MCD, les deux méthodes ont en commun une importante partie de leurs biais (cf. Evaluation contingente, partie V.A et V.B).

La validité d'une MCD dépend de l'ampleur des biais qui lui sont associés ; c'est pourquoi il est important de les minimiser. Kjaer (2005) a tenté d'identifier les différents biais pouvant poser problème lors de l'utilisation de la MCD :

¹⁵¹ Lancsar, E., J. Louviere, et al. (2007). "Several methods to investigate relative attribute impact in stated preference experiments." *Social Science & Medicine* **64**(8): 1738-1753.

¹⁵² Cookson, R. (2003). "Willingness to pay methods in health care: a sceptical view." *Health Economics*, **12**(11): 891-894.

¹⁵³ Wainwright, D. M. (2003). "More 'con' than 'joint': problems with the application of conjoint analysis to participatory healthcare decision making" *Critical Public Health* **13**(4): 373-380.

1. Incitations à modifier les réponses	
Biais stratégique	Sous ou surestimation du CAP en raison du comportement des individus (free-rider par exemple)
Biais de justification	Désir de rationaliser ses réponses de la part du répondant. Possible lorsque le répondant pense que le chercheur possède des informations concernant son comportement en situation réelle
2. Erreur de spécification du scénario	
Effet d'envergure	Le répondant ne prend pas en compte les quantités.
Effet de sous-additivité	La valeur attribuée à différentes politiques évaluées simultanément est inférieure à la somme des valeurs attribuées séparément
Biais de statu quo	Préférence pour la situation de statu quo
Effet de contexte	Les préférences sont influencées par la communication (publicité par exemple) réalisée autour du projet
Effet d'ordre	Les répondants sont particulièrement sensibles au contexte
Biais d'information	Suffisamment d'informations doivent être fournies aux répondant afin qu'ils puissent réaliser leurs choix
Biais d'étendue	Le niveau des attributs influence les préférences
Biais lié au véhicule de paiement	Le véhicule de paiement influence les réponses
Biais d'erreur de spécification	La communication a été réalisée de telle sorte que les répondants ne comprennent pas les informations comme le chercheur l'aurait souhaité
Biais théorique	Violation de la théorie économique
3. Échantillon	
Biais de non réponse	Si le taux de non réponse est trop important, cela peut-être lié à la représentativité de l'échantillon et les résultats ne pourront pas être généralisables
Biais d'échantillonnage	La probabilité qu'un individu accepte d'être questionné est liée à la construction de l'étude : des individus ayant des aptitudes cognitives supérieures seront plus à-même de participer. La généralisabilité des résultats est alors sujette à caution

Tableau 2 : biais rencontrés

B. Réponses négatives

Certains répondants indiquent des valeurs nulles aux questions ouvertes et refusent la valeur qui leur est proposée lors de questions fermées. Ces réponses négatives reflètent-elles réellement des estimations égales à zéro ou sont-elles le reflet d'une objection émanant du répondant¹⁵⁴? Il est possible de demander aux individus répondants « non » d'indiquer pourquoi ils ont répondu ainsi.

¹⁵⁴ Dziegielewska, D. A. and R. Mendelsohn (2007). "Does "No" mean "No"? A protest methodology." *Environmental & Resource Economics* 38(1): 71-87.

L'étude des réponses permet d'identifier les raisons « légitimes » d'avoir un CAP faible (revenu trop faible, aucun intérêt...). D'autres raisons indiqueront un rejet de la méthode d'évaluation (information insuffisante, quelqu'un d'autre devrait financer le projet...). Certains auteurs indiquent que tronquer les réponses des individus témoignant d'un rejet de la méthode d'évaluation biaiserait l'échantillon (Jorgensen *et al.*, 1999 ; Jorgensen et Syme, 2000). Toutefois, ils ne présentent aucune alternative permettant de prendre en compte ces réponses sans biaiser les résultats. Dziegielewska et Mendelsohn présentent un nouveau modèle : lorsque les individus rejetant la méthode ont été identifiés, une fonction d'utilité indirecte peut-être spécifiée. Les paramètres du modèle sont estimés à partir de l'échantillon tronqué. Afin d'éviter un biais lié à un échantillon non représentatif, ils estiment le CAP de tous les répondants à partir des coefficients estimés pour l'ensemble de l'échantillon. En appliquant leur méthode à la pollution de l'air en Pologne, ils montrent que le traitement adéquat des réponses provenant d'individus rejetant la méthode influence les résultats. Une attention particulière doit donc être apportée à ce type de réponses.

4.5.4 Validité

Il est recommandé de voir si l'étude prend en compte l'ensemble des questions jugées importantes afin que les estimations obtenues soient valides. Cela renvoie aux choix des attributs, à leurs niveaux, à l'ordre dans lequel ils sont présentés et au contexte. Le recours à des groupes témoins et la réalisation d'études pilotes sont des éléments clés permettant d'améliorer les études. Par ailleurs, il est également conseillé d'inclure des tests permettant de vérifier la cohérence des réponses obtenues.

A. Test de validité

Afin d'identifier la présence de biais et de tester la validité de l'étude, il peut-être utile d'inclure d'autres choix. Par exemple, on peut inclure un choix pour lequel une alternative domine nécessairement. Des réponses « incorrectes » peuvent signifier que le répondant a un comportement irrationnel, qu'il n'a pas compris ce qui lui était demandé, ou bien qu'il a simplement réalisé une erreur. On peut également inclure le même choix à deux reprises afin de vérifier que le répondant apporte la même réponse dans les deux cas. Enfin, on peut tester la rationalité en testant l'axiome de transitivité. D'autres possibilités existent :

- Cadre : faire varier le cadre de l'étude selon les sous-groupes afin d'identifier un possible effet lié au contexte.
- Ordre : varier l'ordre des attributs, des choix ... pour tester un possible effet d'ordre.
- Niveaux : varier les niveaux afin de tester un effet niveau.

Comme pour l'évaluation contingente, on peut utiliser la méthode du « *cheap-talk* » et des questions permettant de prendre en compte l'incertitude^{155,156,157,158} (cf. Évaluation contingente). Toutefois, davantage d'études sont nécessaires afin de s'assurer de l'effet que cela peut avoir sur le biais hypothétique dans le cas de la MCD. Des questions portant sur le répondant et sur sa prise de décision peuvent également être posées. L'objectif étant de tester la validité, d'identifier des incohérences ou une possible incompréhension de la part du répondant (Bennett et Blamey, 2001). Enfin, il peut être utile d'inclure des questions socio-économiques (cf. Évaluation contingente). L'échantillon pourra alors être divisé en sous-groupes et on pourra tester l'hétérogénéité des préférences (en examinant les possibles interactions entre les caractéristiques personnelles et les attributs préférés).

B. Validité externe et validité interne

Permettant l'obtention de résultats hypothétiques, la MCD a souvent été critiquée. Sa capacité à fournir des résultats utiles lors de la prise de décision est remise en question. Lorsque d'autres mesures existent (en particulier un marché réel), il est important de voir si les estimés correspondent aux autres mesures. Cela n'est en général possible pour des biens privés uniquement (les prix de marché n'étant pas fixés pour les biens publics)¹⁵⁹. Dans le domaine des transports, Louvière et Woodsworth (1983) ont montré que la MCD pouvait être utilisée afin d'expliquer les choix dans le domaine des transports. Dans le cas de la santé, la MCD a difficilement été reliée aux choix réels observés (Ryan, 2004 ; Ryan et Gérard, 2003). Telser et Zweifel (2007) ont, quant à eux, montré que la validité externe de la MCD semblait être importante. On peut également observer la validité liée à la méthode utilisée (validité interne et cohérence).

¹⁵⁵ Champ, P. et R. Bishop (2001). "Donation payment mechanisms and contingent valuation: An empirical study of hypothetical bias." *Environmental and Resource Economics* **19**(4): 383-402.

¹⁵⁶ Cummings, R. et L. Taylor (1999). "Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method." *American Economic Review* **89**(3): 649-665.

¹⁵⁷ Lusk, J. (2003) "Effects of cheap talk on consumer willingness-to-pay for golden rice." *American Journal of Agricultural Economics*. **85**(4): 840-856.

¹⁵⁸ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." *Health Economics Papers* University of Southern Denmark. Odense Denmark, 139p.

¹⁵⁹ Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." *Health Economics Papers* University of Southern Denmark. Odense Denmark, 139p.

4.6 Transférabilité

Lors de prises de décision concernant des nouvelles technologies de santé, la connaissance des coûts d'opportunité et de l'efficacité sont des éléments jouant un rôle croissant. En outre, les enjeux liés au développement d'un nouveau produit sont les contraintes de ressources et budgétaires. Pour ces raisons, un nombre croissant de juridictions¹⁶⁰ requiert des données économiques lors de leur processus décisionnel (fixation des prix, des seuils de remboursement...) portant sur les technologies de santé¹⁶¹. Les évaluations médico-économiques ont pour but d'aider les décideurs publics dans leur tâche. La réalisation de telles études est souvent longue et coûteuse (Goeree, 2007). Il est donc nécessaire de s'interroger sur la possibilité d'utiliser les résultats d'évaluations économiques réalisées dans des contextes différents. Patients, coûts, facteurs épidémiologiques, etc. peuvent varier d'une juridiction à l'autre. La question est de savoir si les variations observées et leurs répercussions sur l'efficacité pratique et les coûts entraînent des différences significatives dans le rapport coût-efficacité (Wilke *et al.*, 1998). Jusqu'à aujourd'hui, il a été observé que les décideurs restaient peu enclins à utiliser des études médico-économiques provenant d'autres pays¹⁶².

4.6.1 Éléments empiriques, de définition et guidelines

A. Variabilité des données, des méthodes et des résultats

Bernie O'Brien¹⁶³ fut l'un des premiers à se préoccuper de la généralisabilité des données relatives aux coûts et à l'efficacité. Il souligne le fait que des différences observées au niveau démographique, épidémiologique, des pratiques cliniques, des incitations dont les professionnels de santé font l'objet, des prix relatifs et des coûts d'opportunité sont des éléments diminuant la possible transférabilité des études¹⁶⁴.

Barbieri *et al.* (2005) ont réalisé un travail de recherche empirique ayant pour but de mettre en évidence l'existence de différences au niveau des ratios coût-efficacité en Europe de l'Ouest. En observant des études réalisées dans cinq pays, ils concluent que la variabilité existe. En effet, dans dix-sept cas sur vingt-sept on observe des différences substantielles au niveau des ratios coûts-

¹⁶⁰ Nous parlerons indifféremment de juridictions ou de pays au cours de la synthèse.

¹⁶¹ Drummond, M., *et al.* (2009). "Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force Report". *Value in Health*. **12**(4): 409-418.

¹⁶² Li, S. C. (2007). "How to improve pharmacoeconomic data generalisability between different countries". *Expert Opinion on Pharmacotherapy*. **8**(10): 1409-1413.

¹⁶³ O'Brien, B. J. (1997). "A tale of two (or more) cities: Geographic transferability of pharmacoeconomic data". *American Journal of Managed Care*. **3**: S33-S39.

¹⁶⁴ O'Brien, B. J. (1997). "A tale of two (or more) cities: Geographic transferability of pharmacoeconomic data". *American Journal of Managed Care*. **3**: S33-S39.

efficacité incrémentaux (au moins le double) pouvant entraîner des prises de décisions différentes. D'après les auteurs, la source majeure de ces différences réside dans le type d'étude réalisé et l'ampleur de la variabilité dépend des décisions analytiques prises par le chercheur (grouper ou non les données lors de l'estimation des résultats).

Finalement, ils indiquent que la généralisabilité n'est pas « directe » : estimer si oui ou non une étude est généralisable à un autre contexte va dépendre du jugement du chercheur utilisant l'étude en question. En outre, l'étude montre que les différences coûts-efficacité d'un pays à l'autre ne sont pas systématiques (pas toujours dans le même sens et/ou dans la même proportion) même s'il semble que les études réalisées en France soient toujours plus coût-efficaces que celles réalisées au Royaume-Uni. Il est donc difficile de tirer des conclusions d'un pays à l'autre. Cet article, fréquemment cité dans la littérature portant sur la transférabilité témoigne de l'importance d'estimer la généralisabilité d'une étude avant de l'utiliser (dans son intégralité ou en partie).

Urdhal, Manca et Sculpher (2006) ont cherché à savoir comment la variabilité était gérée dans les évaluations publiées en lien avec l'ostéoporose. Quatre éléments ont été observés :

- Savons-nous si les décideurs qui exploiteront les résultats de l'étude sont clairement définis ? Connaître à qui est destinée l'étude est essentiel lorsque l'on souhaite savoir si le modèle et les données utilisées sont appropriés dans un autre contexte.
- Le modèle est-il présenté de façon transparente (question étudiée, structure et données prises en compte) ? La transparence du modèle est un pré-requis dans le cas où l'on estime la pertinence du modèle pour le décideur et si l'on cherche à évaluer la transférabilité du modèle à d'autres juridictions.
- Pouvons-nous estimer la pertinence des données utilisées dans le modèle pour le décideur ?
- Le modèle est-il robuste ? L'utilisation d'analyse de sensibilité permet de faire varier les différentes données et d'étudier les différentes valeurs que l'on peut observer au sein d'une même juridiction ou dans un autre.

Seulement 16% des études donnaient suffisamment d'informations permettant d'identifier la juridiction et la personne à qui l'étude était destinée. Toutefois, la majorité des études indiquaient la population de patients, la structure du modèle et les hypothèses clés. Enfin, peu d'analyses de sensibilité ont été réalisées. En conclusion, peu d'analyses font l'effort de traiter de la généralisabilité de leurs résultats. L'utilité que les études peuvent avoir pour d'autres décideurs semble donc limitée. Selon les auteurs, davantage de transparence devrait accroître la transférabilité.

B. Transférabilité et généralisabilité

Il n'existe pas de consensus quant à la définition de la généralisabilité et de la transférabilité. Les glossaires en économie de la santé ne font pas de distinction entre les deux termes¹⁶⁵ qui font références à la validité externe. Il est possible d'avoir plus de précisions si l'on considère le fait qu'il existe au moins deux façons pour un décideur d'utiliser une évaluation médico-économique provenant d'une autre juridiction :

- en appliquant les résultats à sa propre juridiction (résultats perçus comme étant pertinents par estimation ou par hypothèse) ;
- en utilisant la méthode présentée dans l'étude et les données jugées pertinentes tout en remplaçant les éléments méthodologiques et les données qui ne s'applique pas à la juridiction cible.

La généralisabilité ou la transférabilité s'applique aux coûts ainsi qu'aux effets cliniques¹⁶⁶. Lorsque des auteurs utilisent des définitions différentes pour les deux termes ceci sera précisé.

C. Les données

Lorsque l'on souhaite utiliser une étude provenant d'une autre juridiction (comparaison ou transférabilité), il est reconnu que les données économiques peuvent varier. En revanche, on accepte, en général, le fait que les données relatives aux effets de la technologie de santé sont les mêmes d'une juridiction à l'autre^{167,168}.

C.1. Données économiques

Un aspect de la généralisabilité a trait au traitement des données économiques recueillies dans le cadre d'un essai réalisé ailleurs qu'en France. En effet, il est généralement reconnu que les coûts et leur prise en compte (Mullins et Ogilvie, 1998) varient d'un pays à l'autre (témoignant des différences dans les modèles d'utilisation des ressources et du coût unitaire relatif¹⁶⁹). Par exemple, on observe des durées d'hospitalisation différentes pour une même intervention chirurgicale en fonction du pays et du système de santé. Une recommandation que l'on retrouve dans les guidelines (cf. Transférabilité - Les guidelines) comme dans la littérature scientifique est de présenter les données relatives à l'utilisation des ressources en quantité et en coûts. La plupart des guidelines

¹⁶⁵ Boulenger, S., *et al.* (2005). "Can economic evaluations be made more transferable?". *The European Journal Of Health Economics: HEPAC: Health Economics In Prevention And Care*. 6(4): 334-346.

¹⁶⁶ Goeree, R., *et al.* (2005). "Generalizability of economic evaluation: using results from other geographic areas or from multinational trials to help inform health care decision making in Canada". CCOHTA HTA Capacity Building Grants Program. Toronto (ON): Canadian Coordination Office for Health Technology Assessment. Grant N. 67

¹⁶⁷ Drummond, M. F., Manca, A., Sculpher, M. J. (2005). "Increasing the generalisability of economic evaluations: recommendations for the design, analysis and reporting of studies". *International Journal of Technology Assessment in Health Care*. 21(2): 165-171.

¹⁶⁸ Goeree, R., *et al.* (2007). "Transferability of economic evaluations: approaches and factors to consider when using results from one geographic area for another". *Current Medical Research And Opinion*. 23(4): 671-682.

¹⁶⁹ Ottawa: Canadian Agency for Drugs and Technologies in Health (2006). "Guidelines for the economic evaluation of health technologies" Canada [3rd Edition].

s'accordent également pour dire qu'il faut prendre en compte les coûts unitaires et marginaux. Ces recommandations permettent aux décideurs de voir si les pratiques observées s'appliquent à leur juridiction et il est également plus facile de remplacer les coûts de l'étude par ceux en pratique dans la juridiction cible. Étudier les coûts dans une étude est donc primordial lorsque l'on souhaite comparer des études ou transférer des résultats.

C.2. Données relatives aux effets de la technologie de santé

Avant d'exploiter une étude réalisée dans un autre contexte, il faut s'interroger sur la pertinence des résultats dans le pays cible. Afin de faciliter la comparaison et la transférabilité des études, il est conseillé de quantifier les gains de santé au moyen de mesures simples¹⁷⁰. En outre, le développement d'une nouvelle technologie de santé se traduit par des coûts d'opportunité en termes de bénéfices de santé pour une autre pathologie. A moins que les gains de santé et les coûts d'opportunités ne soient exprimés dans la même mesure, il n'est pas possible de quantifier les gains de santé nets. L'estimation des résultats ainsi que leur prise en compte varie d'une étude à l'autre (plusieurs guidelines – NICE, Canada... - recommandent d'utiliser les QALY). Pour estimer les résultats sur la santé, il est possible de se baser sur des instruments génériques (EQ-5D, *Health Utilities Index*...) ou de réaliser les estimations à partir des patients lors d'essais cliniques. Les QALY, bien que controversées, restent la mesure de santé la plus largement répandue¹⁷¹.

C.3. Données individuelles des patients

Lorsque des essais cliniques internationaux sont réalisés, il est recommandé d'analyser les données individuelles des patients (cela aide à évaluer la transférabilité de l'étude)¹⁷². Au moyen d'approches analytiques, on cherche à évaluer l'hétérogénéité au niveau des coûts, de l'utilisation des ressources, des taux de survie... Un deuxième élément est l'estimation de la différence observée au niveau de l'utilisation des ressources, des coûts... qui représente un intérêt pour le décideur mais qui n'ont pas été pris en compte dans l'essai clinique. Les méthodes analytiques utilisées vont varier en fonction des données disponibles¹⁷³. Un algorithme développé par Manca et Willan (2006) différencie les méthodes selon que la population de patients de l'essai clinique comprend ou non des patients de la juridiction d'intérêt. En outre, lorsque les données individuelles des patients sont disponibles, trois méthodes statistiques générales sont fréquemment utilisées (détection de

¹⁷⁰ Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontieres - Can we ever make economic evaluations generalisable across jurisdictions?". *Pharmacoeconomics*. **24**(11): 1087-1099.

¹⁷¹ Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontieres - Can we ever make economic evaluations generalisable across jurisdictions?". *Pharmacoeconomics*. **24**(11): 1087-1099.

¹⁷² Drummond, M., et al. (2009). "Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force Report". *Value in Health*. **12**(4): 409-418.

¹⁷³ Drummond, M., et al. (2009). "Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force Report". *Value in Health*. **12**(4): 409-418.

l'hétérogénéité, *fixed effects model* et *multilevel model*)¹⁷⁴. Ces approches font l'hypothèse que les données sont mesurées en quantités. Toutefois, dans la majorité des analyses coût-efficacité, ces quantités sont calculées à partir de modélisations et les approches ne sont pas toujours directement applicables. L'utilisation du *multilevel model* a comme avantage de fournir des estimés plus précis lorsque les données des patients sont regroupées par localisation. Le groupe de travail de l'ISPOR s'attend à des avancées méthodologiques dans le futur.

D. Les guidelines

Les questions de transférabilité et de comparabilité sont très peu abordées dans les guidelines. Seuls les guides canadien et belge abordent réellement le sujet de la généralisabilité. Une étude des guidelines a permis au groupe de travail de l'ISPOR¹⁷⁵ d'observer un consensus (16 guidelines sur 21 étudiés) sur le fait qu'il existe des problèmes relatifs au transfert des données économiques. En effet, les coûts pris en compte ne sont pas toujours les mêmes (cf. synthèse portant sur la prise en compte des coûts directs et indirects) et l'on observe une variabilité des données et des méthodes utilisées. Des preuves empiriques existent quant à la variabilité des résultats coûts-efficacité d'une juridiction à l'autre (cf. IA - Variabilité des données, des méthodes et des résultats). Malgré ces évidences, le problème de la généralisabilité n'est pas traité en profondeur.

Comme nous l'avons vu, il est fréquemment recommandé de présenter les quantités séparément afin d'accroître la transparence des études. En outre, certains guidelines demandent explicitement à ce qu'on utilise uniquement les coûts unitaires de la juridiction cible (en raison des différences de coûts observées d'une juridiction à l'autre). Le groupe de travail de l'ISPOR souhaiterait que les auteurs des guidelines justifient cette recommandation étant donné les difficultés additionnelles que cela engendre pour les auteurs des études.

D.1. Le guide canadien

Il recommande de justifier l'emploi de données ne provenant pas du Canada et de vérifier leur applicabilité au contexte canadien. En cas d'ajustement des données, il faut décrire et justifier les méthodes utilisées.

Il identifie deux situations :

- Les résultats peuvent ne pas être transférables car le point de départ de l'étude ne s'applique pas à la problématique de décision du Canada : dans le cas où la technologie étudiée et/ou le comparateur ne s'appliquent pas à la juridiction cible ou si la qualité

¹⁷⁴ Drummond, M., et al. (2009). "Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force Report". *Value in Health*. 12(4): 409-418.

¹⁷⁵ La question de la transférabilité étant un sujet lié à des enjeux importants (gains de temps et économies), l'ISPOR a mis en place un groupe de travail (task force) en 2004 devant s'interroger sur la transférabilité des évaluations économiques et des données utilisées. Ce groupe de travail, était composé d'universitaires, d'industriels et de personnes travaillant dans les agences de remboursement originaires d'Italie, des Pays-Bas, de Grande-Bretagne, de Pologne, de Suisse et des États-Unis.

méthodologique de l'étude ne répond pas aux critères de la juridiction cible (alors, le transfert des données de coûts n'est plus valide). Enfin, la population de patient peut différer d'une juridiction à l'autre, rendant impossible l'utilisation des informations coût-efficacité. Ces aspects sont appelés critères de knock-out général (*general knock-out criteria*).

- Les résultats de coût-efficacité sont transférables après ajustement uniquement. Ils peuvent porter sur des différences portant sur le mode d'administration du traitement (un traitement administré à l'hôpital dans une juridiction et chez le médecin de ville dans une autre), les coûts unitaires, etc. Il en résulte que des différences de pratiques médicales peuvent empêcher la transférabilité des résultats en l'absence d'ajustements préalables. En outre, les écarts rencontrés au niveau des prix unitaires font que de nouvelles estimations sont nécessaires. Il est possible de remplacer les coûts unitaires de l'étude d'origine par ceux du pays cible ou l'ajustement peut-être réalisé au moyen des parités de pouvoir d'achat. L'horizon temporel est également un élément à prendre en compte car il peut changer d'une juridiction à l'autre. Enfin, l'auteur cherchant à transférer une étude doit porter son attention sur les coûts pris en compte, la perspective adoptée... Ces critères sont des critères de knock-out spécifiques. Ils font référence à des points qui doivent être analysés avant le transfert des résultats. Dans la plupart des cas, des ajustements sont nécessaires pour rendre la transférabilité possible.

D.2. Le guide belge

Le guideline belge¹⁷⁶ autorise la transférabilité à plusieurs conditions :

- elle doit-être discutée par l'auteur de l'étude réalisée en Belgique ;
- si le recueil de données propres à la Belgique n'est pas possible ;
- si la population de patients étudiée dans l'étude originale est semblable à la population de patients belges ;
- et si les sources et les méthodes d'estimation sont présentées de façon transparente.

Le point de vue sociétal devra alors être abordé. En outre, le guide belge reprend le modèle de Welte (cf. Modèle développé par Welte et al. (2004) en annexe).

Étant donné les différences observées concernant le contexte de prise de décision, il n'est pas surprenant d'observer des variations dans les guidelines. Cependant, la question de savoir si toutes ces variations sont justifiées se pose¹⁷⁷. Certaines de ces variations sont attendues et légitimes alors

¹⁷⁶ Cleemput, I., et al. (2006). "Recommandations provisoires pour les évaluation pharmacoéconomiques en Belgique". KCE reports. **28B**.

¹⁷⁷ Cleemput, I., et al. (2006). "Recommandations provisoires pour les évaluation pharmacoéconomiques en Belgique". KCE reports. **28B**.

que d'autres ne sont pas toujours justifiables (*cf.* annexe 2)¹⁷⁸. La première catégorie fait références à des variations au niveau des recommandations justifiées du fait des différences observées d'un système de santé à l'autre (la taille de la population traitée ou les comparateurs par exemple). D'autres variations peuvent être justifiées en raison de différences d'opinion (méthodes utilisées pour estimer les coûts...). Enfin, on observe des variations ne pouvant pas être justifiées car elles sont jugées incohérentes avec les objectifs du système et ses contraintes. Lorsqu'un auteur souhaite apporter des changements à une étude, il faudrait qu'il justifie ses choix. Même s'ils reconnaissent l'existence de variations légitimes en fonction des différents systèmes de santé, Sculpher et Drummond (2006) recommandent une harmonisation des guidelines. L'objectif étant d'éviter de gaspiller des ressources rares en réalisant des études spécifiques pour des situations non justifiées.

4.6.2 Facteurs affectant la transférabilité et éléments à prendre en compte lors de l'étude de la transférabilité

A. Les facteurs

Une étude a été menée par Goeree *et al.* (2007) dans le but d'aider le Canada à mieux appréhender la généralisabilité¹⁷⁹ des études. Les auteurs cherchent à identifier, au moyen d'une revue de la littérature, les facteurs affectant la transférabilité d'études et l'utilisation de résultats provenant d'essais cliniques internationaux :

- Caractéristiques des patients (démographie – âge, sexe –, niveau d'éducation, statut socio-économique, facteurs de risques, historique médical, facteurs génériques, facteurs environnementaux, espérance de vie, revenu, type d'assurance...). Elles prennent également en compte leur attitude à l'encontre du traitement, l'observance, les incitations auxquelles ils peuvent faire face (copaiement, montants déductibles...). Les résultats d'une étude vont dépendre de l'attitude des patients¹⁸⁰.
- Caractéristiques de la pathologie (épidémiologie, *case-mix*, gravité de la maladie, comorbidité). Les pays varient en termes de démographie et de prévalence de la pathologie. De telles variations peuvent influencer les résultats coûts-efficacité d'une intervention de différentes façons. Par exemple, si l'on évalue un programme de santé ayant comme objectif le dépistage d'une maladie et que le dénominateur du ratio coût-efficacité comprend le

¹⁷⁸ Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontieres - Can we ever make economic evaluations generalisable across jurisdictions?". *Pharmacoeconomics*, **24**(11): 1087-1099.

¹⁷⁹ Dans l'étude, la généralisabilité nous renvoie à la possibilité d'utiliser des données récoltées auprès d'une population de patient particulière, dans des circonstances uniques, à un moment précis dans le temps et traitées d'une façon singulière et de les appliquer à une autre juridiction.

¹⁸⁰ Goeree, R., *et al.* (2007). "Transferability of economic evaluations: approaches and factors to consider when using results from one geographic area for another". *Current Medical Research And Opinion*, **23**(4): 671-682.

nombre de cas détectés, ce nombre sera bien plus grand dans un pays pour lequel le risque de base (la probabilité d'être atteint de la pathologie) est plus important¹⁸¹. Avant d'envisager la transférabilité d'une étude, il est nécessaire d'étudier ces paramètres afin de mettre en évidence toutes différences notables.

- Caractéristiques des professionnels de santé (pratiques cliniques, type de rémunération, etc.). Les résultats d'une analyse économique réalisée dans un pays peuvent ne pas être transférables car les pratiques cliniques observées dans le pays de départ diffèrent du pays cible. En outre, différents systèmes de santé ont des régulations et donc des incitations différentes (rémunération à l'acte vs capitation par exemple)¹⁸². Ces facteurs peuvent expliquer certaines variations au niveau des pratiques cliniques.
- Caractéristiques du système de santé (ressources disponibles, niveau de technologie et innovation, autres traitements disponibles, listes d'attente, disponibilité de génériques, de substituts, etc.). Une des données économiques menaçant la transférabilité est le niveau de prix relatif. Même si le niveau des prix varie d'un pays à l'autre, c'est la variation du prix d'une ressource par rapport à une autre qui pose problème. Par exemple, si tous les prix en vigueur aux États-Unis sont le double de ceux pratiqués au Canada, alors la comparaison coût-efficacité d'un traitement par rapport à l'autre reste la même dans les deux pays. Toutefois, il existe d'importantes différences de prix relatifs à travers le monde¹⁸³.
- Critères méthodologiques (prise en compte des coûts, perspective de l'étude, horizon temporel, taux d'actualisation, coûts d'opportunité).

Plus de soixante-dix facteurs peuvent avoir un impact sur la transférabilité des études ou l'analyse de données provenant d'essais cliniques internationaux¹⁸⁴. Les facteurs les plus fréquemment cités sont les prix absolus ou relatifs, les pratiques cliniques, l'efficacité du traitement et la démographie. Manca et Willan (2006) identifient les mêmes facteurs, à l'exception des facteurs méthodologiques. En revanche, ils prennent en compte d'autres facteurs socio-économiques (la volonté et la capacité d'un pays à utiliser ces ressources pour des soins de santé ou les préférences de la population envers la santé). Les préférences du consommateur sont également citées par O'Brien (1997). Il pense qu'il n'y a pas de raison pour que les préférences soit constantes d'une

¹⁸¹ O'Brien, B. J. (1997). "A tale of two (or more) cities: Geographic transferability of pharmacoeconomic data". *American Journal of Managed Care*, 3: S33-S39.

¹⁸² O'Brien, B. J. (1997). "A tale of two (or more) cities: Geographic transferability of pharmacoeconomic data". *American Journal of Managed Care*, 3: S33-S39.

¹⁸³ O'Brien, B. J. (1997). "A tale of two (or more) cities: Geographic transferability of pharmacoeconomic data". *American Journal of Managed Care*, 3: S33-S39.

¹⁸⁴ Goeree, R., et al. (2007). "Transferability of economic evaluations: approaches and factors to consider when using results from one geographic area for another". *Current Medical Research And Opinion*, 23(4): 671-682.

juridiction à l'autre. Il n'est donc pas certains que des QALY mesurées dans un pays soient directement applicables à un autre.

D'autres études ont cherché à mettre en évidence les facteurs pouvant avoir un impact sur la transférabilité (Welte *et al.* 2004 par exemple). Ces études vont également plus loin ; elles cherchent à estimer la transférabilité des études. Lorsqu'elle n'est pas possible, elles tentent de mettre en place une méthode permettant d'estimer quelles informations (données, paramètres) ou éléments de méthodologie peuvent être appliqué dans le pays cible.

B. Le décideur et les paramètres disponibles

L'approche utilisée pour gérer la transférabilité va dépendre de deux facteurs principaux : les exigences du décideur et les données disponibles¹⁸⁵. La perspective adoptée, les paramètres pris en compte, les comparateurs étudiés... dépendent fréquemment du décideur et du guideline du pays (s'il existe). Par exemple, le guide hollandais recommande de prendre le traitement standard comme unique comparateur. Quant au NICE, il recommande de ne prendre en compte que les coûts supportés par le NHS/PSS.

C. La perspective

Le choix de la perspective est un élément important lorsque l'on traite de la transférabilité. Les recommandations des guidelines relatives à la perspective de l'étude varient (cf. 3.2 - Les coûts). Alors que certains guidelines recommandent d'adopter le point de vue du payeur, d'autres préfèrent adopter une perspective sociétale¹⁸⁶. La plupart des gouvernements qui doivent financer les soins de santé préfèrent des études adoptant le point de vue du payeur afin d'estimer les impacts économiques sur leur budget. En revanche, lorsque l'on tente d'accroître la généralisabilité, c'est l'utilisation de la perspective sociétale qui est recommandée. Il est possible de satisfaire les deux perspectives en séparant les informations propres au payeur.

D. La problématique sur laquelle porte la prise de décision

Pour toute évaluation il est important de connaître les éléments sur lesquels porte la prise de décision^{187,188}. Cela va permettre aux auteurs de définir la population cible, les sous-groupes de patients, les comparateurs devant être pris en compte...¹⁸⁹ Il est possible que la problématique varie d'une juridiction à l'autre. En effet, dans le cas de médicaments, les licences diffèrent d'un marché à

¹⁸⁵ Drummond, M., *et al.* (2009). "Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force Report". *Value in Health*. **12**(4): 409-418.

¹⁸⁶ Drummond, M. F., Manca, A., Sculpher, M. J. (2005). "Increasing the generalisability of economic evaluations: recommendations for the design, analysis and reporting of studies". *International Journal of Technology Assessment in Health Care*. **21**(2): 165-171.

¹⁸⁷ Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontieres - Can we ever make economic evaluations generalisable across jurisdictions?". *Pharmacoeconomics*. **24**(11): 1087-1099.

¹⁸⁸ Drummond, M. F., Manca, A., Sculpher, M. J. (2005). "Increasing the generalisability of economic evaluations: recommendations for the design, analysis and reporting of studies". *International Journal of Technology Assessment in Health Care*. **21**(2): 165-171.

¹⁸⁹ Drummond, M., *et al.* (2009). "Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force Report". *Value in Health*. **12**(4): 409-418.

l'autre, affectant les utilisations cliniques du nouveau médicament et/ou des comparateurs. En outre, les mêmes médicaments ne sont pas utilisés dans toutes les juridictions, affectant le choix du comparateur. Il est donc important qu'un auteur redéfinisse la problématique propre à la juridiction cible sans se baser sur les études réalisées dans d'autres juridictions.

E. L'horizon temporel

Les textes « standards » portant sur la méthodologie recommandent d'utiliser un horizon temporel égal à la période pendant laquelle on observe des différences entre les coûts et les bénéfices des différentes options étudiées¹⁹⁰. Lorsque les études sont utilisées pour la prise de décision, il semblerait qu'une adaptation de l'horizon temporel puisse être nécessaire car ce dernier dépend des objectifs et des contraintes du système de santé et il est probable que la durée des budgets administratifs suggère une période plus courte (du moins pour la partie portant sur les coûts)¹⁹¹.

F. Les comparateurs

Lors d'une évaluation médico-économique, une étape clé est l'identification des comparateurs. L'un des éléments qui semble faire l'unanimité dans les guidelines (selon Mullins et Ogilvie, 1998) est le fait que le comparateur doit être la thérapie utilisée lors des pratiques standard ou la thérapie susceptible d'être remplacée par la nouvelle technologie. Dans la pratique, on ne sait pas toujours identifier cette dernière et les données permettant de comparer deux traitements ne sont pas forcément disponibles car un nombre conséquent d'essais cliniques comparent la technologie de santé à un placebo. Les comparateurs utilisés ne sont pas nécessairement disponibles dans tous les pays et s'ils le sont, ils ne sont pas toujours utilisés pour les mêmes types de patients. Toutefois, Sculpher et Drummond (2006) rappellent que l'objectif étant une utilisation des ressources efficace, toutes les options existantes doivent être prises en compte, même si elles ne sont pas utilisées dans la juridiction cible. Ils pensent qu'il est important de ne pas se limiter au comparateur « le plus commun » ou « le plus efficace » car laisser de côté des options pertinentes pourrait mener à une mauvaise prise de décisions (non adaptée à la situation réelle dans son ensemble). Lors du choix du comparateur, si une option n'est pas prise en compte, il est conseillé de justifier pourquoi elle a été exclue de l'analyse¹⁹².

¹⁹⁰ Drummond, M., Sculpher, M. J. et Torrance, G. W. (2005). "Methods for the economic evaluation of health care programmes." 3rd édition. Oxford : Oxford University Press.

¹⁹¹ Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontieres - Can we ever make economic evaluations generalisable across jurisdictions?". *Pharmacoeconomics*, **24**(11): 1087-1099.

¹⁹² Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontieres - Can we ever make economic evaluations generalisable across jurisdictions?". *Pharmacoeconomics*, **24**(11): 1087-1099.

Afin que le choix du comparateur soit pertinent pour toutes les juridictions, Drummond, Manca et Sculpher (2005) proposent de s'accorder sur l'utilisation de comparateur(s) reflétant les pratiques « normales » dans un ensemble de juridictions. Une approche alternative est de laisser les praticiens (ou les centres dans lesquels les essais cliniques sont réalisés) choisir leur propre comparateur¹⁹³.

G. Recommandations pour la présentation de la méthode et des résultats

Afin d'informer le décideur, les évaluations doivent comprendre toutes les données pertinentes. Ce qui est perçu comme « pertinent » peut varier d'une juridiction à l'autre en fonction des besoins du décideur¹⁹⁴. En effet, il est fréquent que des essais cliniques ne prennent pas en compte tous les paramètres et les comparateurs qui intéressent les décideurs. En outre, les décideurs ne sont pas toujours intéressés par les mêmes éléments. Ce phénomène est accentué par l'hétérogénéité qui caractérise les différentes juridictions. Le challenge méthodologique est alors d'utiliser toutes les preuves existantes tout en reflétant la variabilité géographique et en tentant de satisfaire les besoins du décideur¹⁹⁵.

Mullins et Ogilvie (1998) soulignent l'importance de standardiser les études médico-économiques. Le facteur primordial permettant d'accroître la comparabilité et la transférabilité des études est la transparence. En effet, la façon de présenter la méthode et les résultats peut aider à accroître la transférabilité des études, l'objectif général étant d'aider les utilisateurs des études à savoir ce qui est pertinent dans le cas de leur juridiction¹⁹⁶. Le niveau de détail et de complexité de certains modèles rend la communication de tous les éléments relatifs au modèle (structure, hypothèses, paramètres...) difficile. Un des éléments clés lorsque l'on présente le modèle est de montrer que chaque paramètre/donnée utilisé est pertinent dans le cadre de la juridiction étudiée. En outre, dans le cas où les paramètres ont été obtenus au moyen d'une « pré-analyse », il est conseillé de fournir tous les renseignements s'y rapportant¹⁹⁷. Lorsqu'il s'agit d'un essai clinique international, il est possible de présenter un tableau avec les données de chaque pays. Toutefois, la disponibilité de ces informations additionnelles est limitée par la taille des articles (spécialement lorsqu'ils sont publiés dans des revues).

¹⁹³ Cela peut-être fait à la condition que l'échantillon de praticiens soit représentatif.

¹⁹⁴ Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontieres - Can we ever make economic evaluations generalisable across jurisdictions?". *Pharmacoeconomics*. **24**(11): 1087-1099.

¹⁹⁵ Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontieres - Can we ever make economic evaluations generalisable across jurisdictions?". *Pharmacoeconomics*. **24**(11): 1087-1099.

¹⁹⁶ Drummond, M. F., Manca, A., Sculpher, M. J. (2005). "Increasing the generalisability of economic evaluations: recommendations for the design, analysis and reporting of studies". *International Journal of Technology Assessment in Health Care*. **21**(2): 165-171.

¹⁹⁷ Drummond, M. F., Manca, A., Sculpher, M. J. (2005). "Increasing the generalisability of economic evaluations: recommendations for the design, analysis and reporting of studies". *International Journal of Technology Assessment in Health Care*. **21**(2): 165-171.

H. Les analyses de sensibilité

L'utilisation d'analyses de sensibilité permet également d'accroître la comparabilité et la transférabilité. Elle permet d'estimer la variabilité des résultats. Une limite réside dans le fait que seul l'auteur de l'étude décide des variables devant être soumises à ce type d'analyses ainsi que des valeurs utilisées. Les variables et les valeurs d'intérêt pour la juridiction cible ne sont pas nécessairement abordées.

I. Comment accroître la transférabilité d'une étude

Il ressort des éléments présentés que la transférabilité d'une étude peut-être accrue. Une étude de qualité et favorisant la transparence sera en effet plus facilement exploitable par d'autres personnes. Différentes informations peuvent apparaître dans une étude pour faciliter sa transférabilité :

- Donner un maximum de renseignements concernant les données : les présenter de façon précise ainsi que leurs sources, l'origine des prix et des quantités utilisées, la devise, le modèle utilisé...
- Présenter le(s) comparateur(s) : pourquoi ont-ils été choisis, la posologie recommandée, les indications thérapeutiques...
- Une présentation des coûts et des quantités est recommandée.
- Différents éléments, tels que les coûts notamment dépendent de la perspective de l'étude. Il est donc conseillé de l'indiquer clairement.
- Enfin, il est conseillé de présenter des analyses de sensibilité lorsque cela est jugé nécessaire.
- Lorsqu'il n'est pas possible de faire figurer toutes ces informations dans une étude, indiquer au lecteur comment se les procurer peut-être conseillé (annexes, renseignements auprès de l'auteur...°).

4.6.3 Aperçu des méthodes existantes ayant trait à la généralisabilité

Lorsque l'on souhaite réaliser une étude médico-économique, le groupe de travail de l'ISPOR recommande de voir si une étude traitant du même sujet existe déjà dans une autre juridiction. En général, les études ne seront pas directement applicables à la juridiction cible. Il faut alors estimer si les résultats peuvent être adaptés afin de les rendre transférables¹⁹⁸. Différents modèles ont été développés dans ce but. Ces modèles de décision sont utilisés car ils permettent de mieux gérer la

¹⁹⁸ Drummond, M., et al. (2009). "Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force Report". *Value in Health*. **12**(4): 409-418.

variabilité géographique et sont un bon outil lorsque l'on souhaite rendre les résultats d'une étude réalisée dans une juridiction applicables à une autre¹⁹⁹.

A. Modèle développé par Welte et al. (2004)

Welte *et al.* reconnaissent l'existence de variabilité et développent un modèle permettant d'estimer la transférabilité d'une étude. Tout d'abord il est spécifié que les études doivent être transparentes dans le but d'accroître leur transférabilité.

Les auteurs ont ensuite tenté de définir les facteurs pouvant affecter la transférabilité au moyen d'une analyse de la littérature. Pour être pris en compte, les facteurs doivent posséder quatre caractéristiques :

- influencer les résultats de l'évaluation médico-économique ;
- subir des variations internationales (varier en fonction de la juridiction, de la localisation) ;
- être mesurable ;
- pouvoir être distingué des autres facteurs pris en compte.

Les auteurs ont également cherché à voir par quels moyens ces facteurs influençaient la transférabilité.

A.1. Les critères de knock-out

Lorsque l'on rencontre des critères de *knock-out* dans une étude, elle n'est plus transférable. Toutefois, dans certains cas, il est possible d'utiliser la méthodologie et/ou des paramètres et des données. La transférabilité de l'étude va être évaluée au moyen d'une *checklist* élaborée par les auteurs. Les critères affectant la transférabilité sont regroupés en deux catégories.

- Les critères généraux de knock-out

Ces critères s'appliquent lorsque la technologie évaluée n'est pas comparable à celle concernée par le problème de décision dans la juridiction cible, lorsque le comparateur utilisé n'est pas pertinent dans la juridiction cible et lorsque la qualité de l'étude n'est pas suffisante (elle ne répond pas aux standards de la juridiction cible). Si l'un de ces éléments est observé, la transférabilité est jugée impossible et la meilleure option reste la réalisation d'une nouvelle étude.

- Les critères de knock-out spécifiques

Chaque facteur de transférabilité identifié par les auteurs peut devenir un critère de knock-out lorsqu'il ne peut pas être évalué en raison d'un manque de données ou si la différence observée entre les deux juridictions est jugée trop importante.

¹⁹⁹ Urdahl, H., Manca, A. and Sculpher, M. J. (2006). "Assessing generalisability in model-based economic evaluation studies - A structured review in osteoporosis". *Pharmacoeconomics*, **24**(12): 1181-1197.

Après avoir identifié les critères de *knock-out* et les avoir hiérarchisé, Welte *et al.* (2005) ont voulu identifier le moyen par lequel ces critères influençaient les résultats (directement ou au travers de déterminants clés).

A.2. Facteurs de transférabilité et déterminants clés

Quatorze facteurs de transférabilité ont été identifiés (*cf.* annexe 4). Une variation des valeurs d'un paramètre peut affecter le ratio coût-efficacité directement (le choix du taux d'actualisation affecte la valeur des coûts futurs et des effets observés) ou par le biais d'un ou de plusieurs autres éléments clés (une différence observée au niveau des pratiques cliniques a un impact sur les économies d'échelle, l'efficacité, les pertes de productivité).

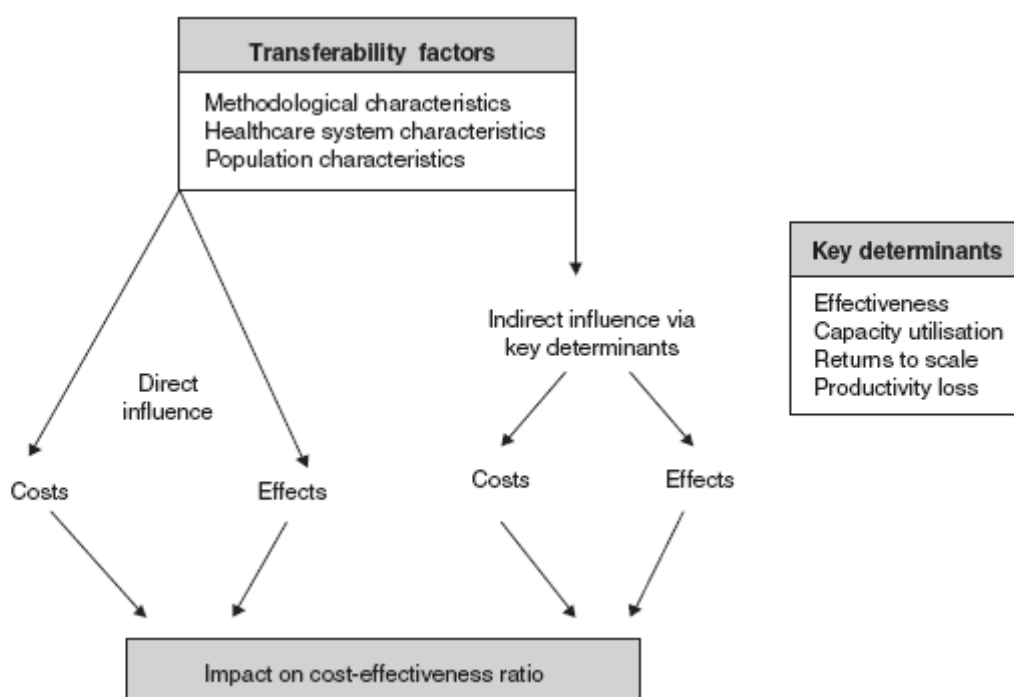


Figure 4 : L'influence des facteurs de transférabilité sur les résultats coût-efficacité
Welte, R., *et al.* (2004). "A decision chart for assessing and improving the transferability of economic evaluation results between countries". *Pharmacoeconomics*. 22(13): 857-876.

Les déterminants clés définis par les auteurs (efficacité, économies d'échelles, perte de productivité, capacité d'utilisation) ont une influence sur les résultats de l'étude et doivent être analysés lorsque l'on cherche à estimer l'influence d'un facteur de transférabilité sur les résultats d'une étude.

A.3. Arbre de décision

Après avoir identifié les critères et les moyens par lesquels ils affectent les résultats de l'étude (directement ou par le biais de déterminants clés), les auteurs ont développé un arbre de décision permettant de mettre en évidence la possible transférabilité d'une étude (*cf.* Annexe 5)

L'utilisation de l'arbre de décision permet de décider si oui ou non une étude est exploitable dans une autre juridiction. Dans le cas où l'on a passé les critères de *knock-out* généraux et spécifiques, il faut identifier les ajustements qui peuvent être nécessaires.

A.4. Ajustements possibles

Selon les auteurs, les ajustements sont nécessaires dès lors que les différences observées (pratiques médicales, prix relatifs, incidence/prévalence de la pathologie) entre le pays d'origine et le pays cible sont importantes²⁰⁰. Toutefois, tous les autres facteurs de transférabilité observés peuvent nécessiter des ajustements au niveau des données et/ou du modèle (utilisation des données locales, modification du modèle de départ). Ces ajustements sont possibles si et seulement si tous les paramètres pertinents sont donnés et que la méthodologie est présentée de façon transparente (permettant de pouvoir l'appliquer à un autre contexte ou d'y apporter des modifications).

Remplacer les données spécifiques à la juridiction d'origine par celles propres à la juridiction cible permet d'accroître la transférabilité car les différences que l'on peut observer lors de l'étude des facteurs de transférabilité seront alors bien moins importantes (exemple, lors de la modification des pratiques cliniques).

A.5. Estimation de nouvelles données

Il est possible de réaliser une nouvelle estimation monétaire des ressources utilisées au moyen des coûts unitaires du pays cible dans le but de supprimer les différences existantes au niveau de certains éléments clés (économies d'échelle, ressources disponibles et d'autres facteurs liés à ces éléments) et ici encore, d'accroître la transférabilité. S'il n'est pas possible de remplacer tous les paramètres pertinents de l'étude par leur valeur respective dans le pays cible, il est tout de même conseillé de modifier le paramètre le plus sensible *a minima*. Pang (2002) conseille d'identifier les paramètres les plus critiques par le biais d'experts ou de revue de la littérature. Il est également suggéré d'utiliser des analyses de sensibilité (réaliser une analyse de sensibilité univariée pour tous les paramètres) : tous les paramètres sensibles dans l'analyse d'origine doivent alors être remplacés par des paramètres du pays cible.

A.6. Inflation, devises différentes et/ou pouvoir d'achat

Afin que les résultats soient exploitables par les décisionnaires de la juridiction cible, il peut-être nécessaire d'apporter des modifications du fait de l'utilisation de devises différentes, de l'inflation et/ou de différences de pouvoir d'achat.

Les études provenant d'autres pays sont en général présentées dans une devise autre que celle du pays cible ou dans la même devise mais avec un pouvoir d'achat pouvant varier. Lorsque l'on

²⁰⁰ Aucune indication n'est donnée quant aux différences jugées « importantes ».

désire convertir une devise, la parité de pouvoir d'achat est recommandée par Welte *et al.* (2004) (les taux de change étant soumis à la spéculation). Il est conseillé d'utiliser la parité de pouvoir d'achat basée sur le PIB fourni par l'OCDE (plutôt que la parité de pouvoir d'achat développée à partir des produits pharmaceutique).

Dans le cas de l'inflation (liée au fait que les études sont utilisées dans des pays différents et/ou à des dates différentes) on peut utiliser un indice des prix basé sur l'économie en général ou sur le système de santé en particulier. Les indices des prix basés sur le système de santé ont comme avantage théorique le fait qu'ils peuvent mieux représenter le développement des prix des biens médicaux. Toutefois ils ne sont pas nécessairement représentatifs de leur secteur et/ou disponibles. Pour cette raison, les auteurs recommandent l'utilisation de l'indice des prix basé sur le PIB.

Si des ajustements doivent être apportés du fait de l'inflation et de l'utilisation de différentes devises (ou différents pouvoir d'achat), il est recommandé de tout d'abord déflater les données présentées dans l'étude d'origine puis de modifier la devise en vue d'une meilleure transférabilité (il est préférable d'apporter le maximum de modifications à l'étude d'origine puis de transférer les résultats).

A.7. Évaluer la transférabilité des résultats suite aux ajustements

Si le modèle utilisé (modifié ou non) s'applique au pays cible et que tous les paramètres ont été remplacés par des valeurs correspondant au pays, alors les résultats seront parfaitement applicables.

La transférabilité des résultats ajustés peut être étudiée en réappliquant la *checklist* ou au moyen d'une analyse de sensibilité multivariée probabiliste. Il est recommandé de faire évaluer la variabilité des paramètres non-substituables par un panel d'expert (ou au moyen d'une revue de la littérature).

Le modèle développé par Welte *et al.* (2004) doit non seulement permettre d'évaluer la transférabilité mais doit également permettre à l'auteur d'identifier les ajustements nécessaires pour permettre son amélioration. Il permet de prioriser les ajustements en fonction de leur importance. En appliquant leur modèle à différentes études, les auteurs ont montré qu'il était efficace et facile d'utilisation. Toutefois, ils reconnaissent que leur modèle n'est encore qu'un premier pas vers une transférabilité plus rapide et transparente étant donné ses limites :

- Le modèle se limite aux facteurs de transférabilité les plus importants. Il est possible que des facteurs jugés moins importants le deviennent avec le temps.
- Le modèle n'a été appliqué qu'à peu de situations (trois), il est donc possible que des problèmes propres à d'autres cas n'aient pas été identifiés.

- La *checklist* développée relève d'une démarche pragmatique et non pas purement scientifique. Dans le cas où l'on dispose de suffisamment de temps, les auteurs conseillent l'utilisation d'un modèle afin d'évaluer l'incertitude associée au transfert.

Le groupe de travail de l'ISPOR utilise le modèle de Welte afin d'évaluer la transférabilité des études.

B. Evaluation du modèle de Welte par Knies et al. (2009)

Knies *et al.* (2009) ont testé le modèle de Welte afin d'évaluer la transférabilité d'une sélection de quatorze études provenant de divers pays aux Pays-Bas.

Sept études ne remplissaient pas les critères généraux de knock-out (score $\leq 10/19$) et ont été jugées comme étant de qualité insuffisante pour envisager toute transférabilité. Au regard des critères spécifiques, six études sur sept ont été jugées comme étant partiellement transférables.

A partir de ces résultats, une comparaison a été réalisée entre deux groupes d'études (A : toutes les études et B : études jugées transférables à partir du modèle de Welte) et une étude réalisée aux Pays-Bas. Les résultats semblent montrer que le modèle de Welte permet d'obtenir une meilleure estimation des coûts alors que les prédictions sont moins bonnes au niveau des effets de la technologie de santé. Les différences observées entre les deux groupes (causées par les critères de knock-out généraux et spécifiques) permettent aux auteurs d'affirmer que le modèle de Welte influence les résultats finaux. On note que le modèle met l'accent sur les coûts (il est en général reconnu que les coûts varient plus d'une juridiction à l'autre, ce qui n'est pas toujours le cas pour les données cliniques). Enfin, il semblerait que plus une étude est de qualité méthodologique satisfaisante, plus la probabilité qu'elle soit jugée transférable augmente (même conclusion dans l'article de Boulenger *et al.* (2005)). En utilisant le modèle afin d'évaluer les études, Knies *et al.* (2009) ont remarqué que certains facteurs semblent se recouper rendant l'utilisation du modèle plus compliquée.

C. Modèle développé par Mason et Mason (2006)

Mason et Mason travaillent sur le thème de la généralisabilité afin de résumer les avancées méthodologiques récentes et de donner des critères d'évaluation de la généralisabilité.

Les auteurs soulignent le fait que les termes généralisabilité et transférabilité sont souvent utilisés comme synonymes, toutefois, ils les distinguent :

- La généralisabilité comprend les trois dimensions de la qualité technique, l'applicabilité au contexte et la transférabilité.
- La transférabilité renvoie à la possibilité d'utiliser les résultats provenant d'une autre juridiction.

En se basant sur les trois dimensions de la généralisabilité telle qu'il la définit, les auteurs indiquent comment l'estimer.

C.1. Qualité technique

Afin d'envisager la généralisabilité d'une étude, les auteurs recommandent d'estimer sa qualité technique : lorsque l'on évalue un nouveau traitement, il faudrait analyser l'acquisition de connaissances à son sujet pour voir si les données ont été recueillies de façon méthodique. Trois niveaux de validité sont alors analysés :

- Validité : comprendre le mode d'action de la nouvelle technologie, ses effets potentiels, son fonctionnement.
- Validité interne : on évalue les performances de la nouvelle technologie dans des conditions de recherche idéalisées (essais cliniques). Un essai clinique réalisé comme il se doit permet d'obtenir des preuves au niveau de la validité interne (les résultats sont une observation non biaisée des effets du nouveau traitement sur les participants à l'essai).
- Validité externe : on cherche à connaître les performances en situation de vie réelle.

La qualité technique fait donc référence à la qualité des différentes étapes de recherches et aux méthodes utilisées pour réaliser l'analyse économique (l'objectif étant de minimiser les biais systématiques et les erreurs randomisées). Notons que l'utilisation de modèles est nécessaire pour extrapoler les résultats à partir des données observées. La qualité technique est donc en partie basée sur des principes largement acceptés dans le domaine de l'évaluation médico-économique. Elle joue un rôle important car les résultats étudiés (ratio coût-efficacité incrémental) sont basés sur les variations de coûts et de bénéfices. Si l'observation de ces variations résulte d'une analyse biaisée, alors la prise de décision risque de ne pas être adéquate. Dans le cas où des recherches rigoureuses ont été réalisées, une analyse répondra au critère de validité interne mais sera-t-elle applicable dans un autre contexte ?

C.2. Applicabilité

Afin que les résultats obtenus à partir de l'essai clinique soient exploitables dans un contexte de prise de décision, des extrapolations doivent être réalisées. Par exemple, afin d'aider lors de la prise de décision concernant le remboursement, il peut-être nécessaire d'extrapoler les résultats à un groupe de patients plus large comprenant des patients plus gravement malades, moins malades, plus jeunes, plus vieux... Les analyses économiques, en appliquant les « ratios risque » à des patients ayant un « risque initial » (*baseline risk*) différent s'adaptent au *case-mix* et à la gravité de la pathologie d'une façon appropriée au point de vue épidémiologique.

Des extrapolations dans le temps sont également nécessaires. En effet, les résultats ne sont observés que durant la période de l'essai clinique et la prise de décision affecte une période plus importante. Ces extrapolations font l'objet de modélisation. Il existe des recommandations de bonne pratique qui sont communes à différentes juridictions quant à la modélisation : estimer les paramètres à partir des données observées («*best evidence*»), choisir un comparateur approprié, inclure les coûts et bénéfices, gérer l'incertitude de façon appropriée... Toutefois, même si les recommandations sont parfois semblables, les résultats ne sont pas nécessairement transférables.

C.3. Transférabilité

Comme nous l'avons déjà observé, il est tout à fait possible (même fort probable) que les résultats coûts-efficacité obtenus lors d'une étude varient en fonction du contexte (du fait d'éléments tels que la gravité de la pathologie, la variation des prix, la comorbidité...). Il est alors nécessaire d'estimer la transférabilité.

Dans la partie incertitude, nous avons vu que l'utilisation du ratio coût-efficacité incrémental peut ne pas être adaptée (limites statistiques) lorsque des valeurs négatives sont observées. Un moyen de pallier ces limites et l'utilisation du bénéfice monétaire net (MNB) qui permet de traiter les coûts et les bénéfices de façon linéaire. Utiliser une régression permet alors d'estimer la transférabilité des résultats grâce à l'analyse de sous-groupes et l'identification de différences systématiques parmi les patients et en fonction de leur localisation. Par exemple, Manca *et al.* (2005) utilisent un modèle multi-niveaux afin de calculer des estimations spécifiques des ratios coûts-efficacité en fonction de la localisation. Leur analyse tient compte du regroupement de patients dans des centres d'essais cliniques ou dans des centres localisés dans le même pays. Un modèle hiérarchique (*multilevel model*) permet ainsi d'accroître la précision des estimations de coûts-efficacité.

Enfin, certaines juridictions sont plus averses au risque que d'autres ou ont différentes contraintes budgétaires. Il est possible que deux juridictions prennent des décisions différentes tout en se basant sur les mêmes résultats. Les pressions budgétaires vont renforcer l'aversion au risque et une nouvelle technologie peut-être rejetée au nom de l'incertitude alors que la véritable raison poussant au rejet est la pression sur les coûts. Pour accroître la transférabilité et pour éviter la prise de décision biaisée, il est nécessaire d'indiquer de façon explicite les coûts d'opportunité ainsi que les standards utilisés pour estimer les résultats.

Selon Mason et Mason (2006), ce n'est qu'après avoir observé ces dimensions (qualité technique, applicabilité et transférabilité) que l'on peut estimer si une étude est généralisable à d'autres juridictions.

D. Modèle développé par Boulenger et al. (2005)

Les auteurs font une distinction entre généralisabilité et transférabilité. La généralisabilité indique dans quelle mesure les résultats s'appliquent à un autre contexte (est-ce que le même traitement aura les mêmes effets dans un autre contexte). Les études sont jugées généralisables si aucun ajustement de l'analyse n'est requis et transférables si les données locales (coûts unitaires...) peuvent être substitués et l'analyse retravaillée dans le contexte du pays cible²⁰¹.

L'objectif des auteurs est de développer une checklist permettant :

- d'estimer la quantité d'informations aidant à la transférabilité. Ils estiment cette étape comme étant primordiale si l'on souhaite estimer la transférabilité ;
- d'estimer le niveau de transférabilité et de généralisabilité de l'Angleterre à la France d'une étude médico-économique ;
- d'identifier les raisons majeures empêchant la transférabilité et la généralisabilité ;
- d'identifier ce qui pourrait être réalisé dans les études futures pour accroître la transférabilité et la généralisabilité des résultats ;
- et de souligner comment les bases de données pourraient assister les utilisateurs à estimer la généralisabilité et la transférabilité.

D.1. Méthodologie

Une *checklist* a été élaborée à partir des modèles du NHS EED et de la base CODECS. Elle a été appliquée à des sous-groupes puis modifiée afin d'être le plus claire, compréhensible et efficace possible. Le changement majeur fut l'ajout de la réponse « pas d'information » dans les choix de réponses. La version finale comprend six catégories :

- sujet et éléments clés de l'étude ;
- caractéristiques de la méthode utilisée afin de mesurer les résultats cliniques ;
- la mesure des bénéfices (du point de vue de la santé) ;
- les coûts ;
- actualisation ;
- et commentaires de l'auteur.

Pour chaque question, les réponses proposées sont « Oui » ; « Partiellement » ; « Non » ; « Pas d'information » et « Ne s'applique pas ».

Étant donné la quantité de questions, il a été décidé de développer une sous-checklist ne prenant en compte que les items les plus importants au regard de la transférabilité. En effet, une étude peut

²⁰¹ Nixon, J., et al. (2009). "Guidelines for completing the EURONHEED transferability information checklists". The European Journal Of Health Economics. HEPAC: Health Economics In Prevention And Care. **10**(2): 157-165.

avoir un bon score dans son ensemble mais ne pas remplir certaines conditions essentielles à la transférabilité.

L'étude menée par les auteurs est basée sur des articles publiés et dont le contexte est la France et/ou l'Angleterre (*a minima*). Elle tente d'identifier les raisons principales entraînant un manque de transférabilité ou de généralisabilité.

D.2. Observations

Les auteurs observent que les études ayant un score inférieur à 50% ont perdu des points dans les parties « Population étudiée », « Efficacité » et « Coûts » en raison du manque d'informations concernant l'échantillon de l'étude et le calcul des résultats. Quant aux coûts, ils sont rarement présentés séparément et les prix unitaires ainsi que les sources des données relatives à la consommation des ressources ne sont pas donnés.

L'étude a donc mis en évidence le fait que le manque d'information et de transparence nuit à la comparabilité et à la transférabilité des études. Toutefois, des limites existent. En effet, après avoir appliqué la *checklist* et obtenu un score, il est nécessaire de relier ce score au niveau « réel » de transférabilité (qu'est-ce qu'un score de 68% implique réellement). En outre, des études ont un score global « satisfaisant » mais ne répondent pas aux critères de certaines sections en particulier.

On note tout de même une corrélation entre le score obtenu à la sous-checklist (transférabilité uniquement) et celui obtenu en répondant à l'ensemble des questions. Ainsi, les études jugées transférables semblent également répondre aux critères méthodologiques. Boulenger *et al.* (2005) recommandent de valider la *checklist* en l'appliquant à un échantillon d'étude plus important et en réalisant de nouvelles recherches.

D.3. Pour aller plus loin : Nixon et al. (2009)

En 2009, les auteurs ont publié un article²⁰² complétant leur recherche. L'article met l'accent sur l'importance de la qualité d'une étude. L'idée est qu'une étude ne peut pas être transférable si elle est de qualité insuffisante et si les données ne sont pas complètes. En effet, les auteurs ont montré dans leur étude précédente²⁰³ que le manque de précisions concernant les données de l'étude était un obstacle à la transférabilité.

Nixon *et al.* (2009) s'interrogent sur la possibilité de pondérer les questions. Ils indiquent que leurs *checklists* incorporent implicitement des éléments de pondération :

- Le nombre de questions portant sur un item reflète son importance lors de l'évaluation d'une étude (il y a 11 questions sur 42 portant sur la méthode de prise en compte des coûts

²⁰² Nixon, J., *et al.* (2009). "Guidelines for completing the EURONHEED transferability information checklists". *The European Journal Of Health Economics*. HEPAC: Health Economics In Prevention And Care. **10**(2): 157-165.

²⁰³ Boulenger, S., *et al.* (2004). "Guidelines for completing the EURONHEED transferability checklist". EURONHEED Working Paper.

et la présentation des coûts) et la sous-checklist témoigne de l'importance de certaines questions.

- En outre, même si toutes les questions de la checklist sont traitées de façon identique, les auteurs attribuent des scores de 0, 0,5 et 1 afin de mettre en évidence le fait que l'auteur de l'étude a partiellement traité une question (score de 0,5).
- Enfin, les items qui ne s'appliquent pas à l'étude analysée sont exclus du calcul du score, évitant ainsi de le faire baisser lorsqu'une question ne s'applique pas. Le score reflète donc de façon plus appropriée les informations fournies dans l'étude ainsi que la généralisabilité et/ou la transférabilité.

La nouvelle étude renforce les premières conclusions : les auteurs ont montré l'importance de la qualité méthodologique et des données lorsque l'on souhaite exploiter une étude. Il sera possible d'envisager la généralisabilité et/ou transférabilité d'une étude si et seulement si les critères de qualité sont satisfaits.

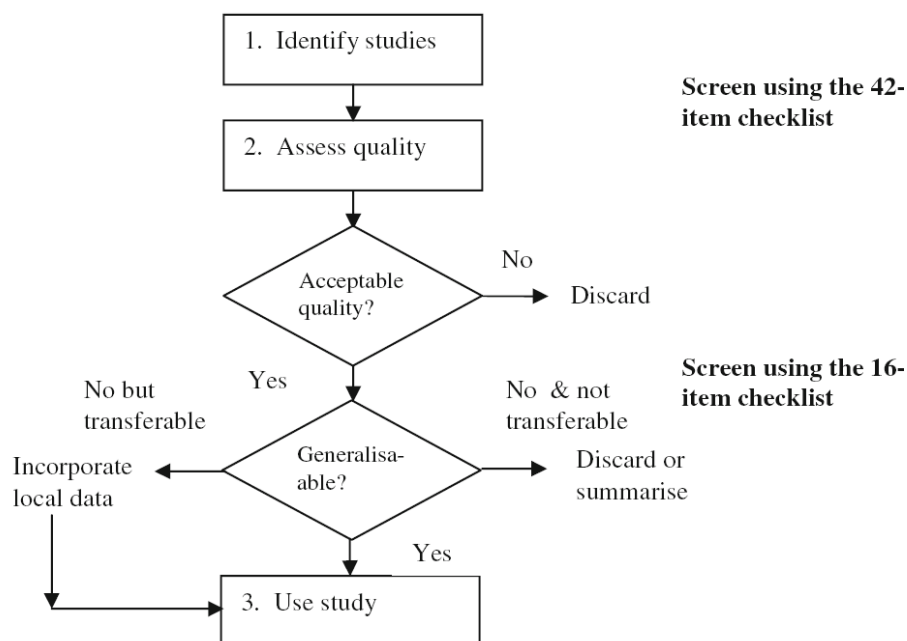


Figure 5 : Modèle de décision suggéré - Nixon, J., *et al.* (2009). "Guidelines for completing the EURONHEED transferability information checklists". *The European Journal Of Health Economics*. *HEPAC: Health Economics In Prevention And Care*. 10(2): 157-165.

E. Modèle développé par Späth et al. (1999)

Les auteurs cherchent à évaluer la transférabilité d'une étude au sein du même système de santé (hôpital, médecine de ville,...) en étudiant le cas de la France. Leur étude est réalisée en trois étapes : une recherche bibliographique des études médico-économiques (234 articles trouvés sur la période

1982-1996 mais seulement 30 répondants aux critères de sélection), une évaluation critique des études basée sur quatre critères et une analyse de la transférabilité des études.

E.1. Critères d'évaluation des études

Quatre critères ont été sélectionnés par les auteurs. Ils sont considérés comme étant essentiels lorsque l'on souhaite estimer la validité interne et externe d'une étude. Ils ont également l'avantage de pouvoir être étudiés rapidement.

Tout d'abord la perspective de l'analyse doit pouvoir être identifiée car elle influence les coûts et les résultats pris en compte. Dans un second temps, au moins deux options doivent être comparées car les décideurs qui souhaitent une allocation optimale des ressources doivent choisir parmi différents traitements. Les thérapies évaluées doivent être décrites de façon compréhensible afin que les utilisateurs des études puissent savoir si les technologies étudiées sont présentes dans l'environnement qui les intéresse. Enfin, les thérapies doivent exister dans le système de santé français.

Suite à la recherche bibliographique, trente et un articles correspondant à vingt-six études différentes ont été sélectionnés. Seuls six de ces articles remplissent les quatre critères d'évaluation.

E.2. Evaluation de la transférabilité

Späth *et al.* (1999) ont sélectionné 5 indicateurs prenant en compte trois dimensions (le cadre dans lequel l'étude va être utilisée, la transférabilité des données relatives aux effets sur la santé et la transférabilité des données portant sur l'utilisation des ressources).

- Utilisateurs potentiels : afin d'identifier la perspective, les auteurs étudient la perspective mentionnée par l'auteur, les coûts incluent dans l'étude et les sources des données de coûts.
- Caractéristiques de la population traitée : les résultats d'une étude sont influencés par la population traitée. Il est donc conseillé de voir si ce sont les mêmes catégories de patients (âge et historique médical) qui vont recevoir le traitement dans l'ensemble du système de santé.
- Résultats de santé : ils comprennent l'efficacité du traitement, l'efficience du traitement, et les préférences.
- Utilisation des ressources : comme les pratiques varient d'un pays à l'autre et au sein du même système de santé, les ressources utilisées sont susceptibles de varier. L'objectif de cette étape est de voir si l'ensemble des ressources utilisées est listé de façon exhaustive et quantifié.
- Prix unitaire et taux d'actualisation : ici, les auteurs regardent si l'ensemble des coûts, leur source et les taux d'actualisation utilisés sont indiqués dans l'étude.

Après avoir appliqué ces critères d'éligibilité, il s'est avéré qu'aucune des études n'était transférable.

E.3. Commentaires et limites

L'une des limites de l'analyse de Späth *et al.* (1999) et qu'ils ne prennent en compte que les articles publiés. Il semble possible qu'un biais de publication existe sans toutefois pouvoir évaluer son importance. La seconde limite renvoie au fait que les auteurs étudient uniquement les données publiées. Or, en raison d'un manque de place, toutes les données pertinentes ne figurent pas dans les articles. Des informations additionnelles sont souvent disponibles dans des rapports complémentaires. Avoir accès à ces rapports pourrait accroître la transférabilité des résultats.

Les auteurs recommandent aux utilisateurs potentiels des études de vérifier si les données s'appliquent effectivement au cas qu'ils souhaitent étudier. Si ce n'est pas le cas, Späth *et al.* conseillent de les adapter. Aucune information n'est donnée quant à la méthode à appliquer. Enfin, le groupe de travail de l'ISPOR rappelle que certains paramètres ont besoin d'être spécifique à la juridiction (prix, et *baseline risk* dans certains cas). Au contraire, les effets du traitement et les risques relatifs sont probablement plus facilement transférables. Il est tout de même recommandé de justifier le fait qu'on utilise des données d'autres juridictions car les preuves de généralisabilité des pratiques cliniques et de l'état de santé sont mixtes. Selon Drummond *et al.* (2009) davantage de recherches sont encore nécessaires.

F. Les études sont-elles transférables ?

Au regard des informations données, des modèles développés pour évaluer la transférabilité et des études empiriques, il semble que les évaluations médico-économiques ne soient pas encore transférables. Il est reconnu que les données médicales sont plus aisément transférables que les données économiques. Les études peuvent être en partie transférables, mais les résultats obtenus ne sont pas applicables tels quels à une autre juridiction.

4.6.4 Essais cliniques internationaux

Les essais cliniques sont un bon moyen d'obtenir des estimations non biaisées de l'efficacité clinique et de l'impact de la technologie de santé sur l'utilisation des ressources, cela malgré la nature artificielle du protocole²⁰⁴.

Un nombre croissant d'essais cliniques sont maintenant internationaux. Ils ont l'avantage de permettre de travailler sur des échantillons plus importants, sur une période plus courte et

²⁰⁴ Cook, J. R., *et al.* (2003). "Assessing the appropriateness of combining economic data from multinational clinical trials". *Statistics In Medicine*. **22**(12): 1955-1976.

permettent d'augmenter la représentativité. Néanmoins, les résultats peuvent être plus difficiles à interpréter.

Même si les pratiques cliniques (les praticiens peuvent influencer l'efficacité, notamment lorsqu'il ne s'agit pas d'un médicament²⁰⁵) et les patients diffèrent d'un centre à l'autre (les caractéristiques des patients ne sont pas toujours distribuées de façon homogène au sein des centres²⁰⁶), il est généralement reconnu que les effets biologiques d'un médicament ne doivent pas varier. On devrait donc être capables de grouper les données cliniques et d'estimer les effets du traitement. Dans ce cas, si une interaction avec un pays existe, elle peut rester non détectée étant donné la petite taille d'échantillons dans certains centres.

Dans le cas des données économiques, on fait l'hypothèse que l'existence de différences systématiques d'un pays à l'autre empêche les chercheurs de grouper les données²⁰⁷. En effet, les variations observées au niveau des pratiques cliniques et du mode de paiement peuvent affecter l'utilisation des ressources. En outre, même si les taux d'utilisation sont similaires, les prix relatifs des ressources diffèrent d'un pays à l'autre. Ces différences peuvent affecter les résultats coût-efficacité des traitements.

A. Existence d'interaction pays-traitement

A partir de ces éléments, Cook *et al.* (2003) cherchent à savoir si les éléments économiques diffèrent d'un pays à l'autre, si les interactions avec les pays diffèrent en magnitude uniquement ou si elles peuvent également être de sens opposé. Enfin, ils tentent de savoir quand il est approprié de regrouper les données économiques des différents pays. Leur étude est réalisée à partir d'un essai clinique (*the Scandinavian Simvastatin Survival Study* ou 4S).

La présence d'hétérogénéité des effets du traitement parmi différents sous-groupes de patients est une interaction entre le traitement et la variable utilisée pour créer les sous-groupes de patients. Ici les auteurs souhaitent voir si une interaction pays-traitement existe. La mise en évidence d'une telle interaction va être réalisée dans le cas de la mortalité, l'hospitalisation et le rapport coût-efficacité. Dans le cas de la mortalité, l'impact pour chaque pays est proche de celui observé pour l'ensemble de l'étude. Il n'y a donc pas d'interaction. Il en est de même pour l'hospitalisation. Le taux relatif d'hospitalisation des patients recevant le traitement par rapport aux patients recevant le placebo est semblable pour chaque pays et pour l'étude dans son intégralité. Ainsi, il est possible d'utiliser l'ensemble des données pour calculer l'estimé et cela devrait permettre aux auteurs

²⁰⁵ Manca, A. and Willan, A. R. (2006). "'Lost in translation' - Accounting for between-country differences in the analysis of multinational cost-effectiveness data". *Pharmacoeconomics*, **24**(11): 1101-1119.

²⁰⁶ Manca, A. and Willan, A. R. (2006). "'Lost in translation' - Accounting for between-country differences in the analysis of multinational cost-effectiveness data". *Pharmacoeconomics*, **24**(11): 1101-1119.

²⁰⁷ O'Brien, B. J. (1997). "A tale of two (or more) cities: Geographic transferability of pharmacoeconomic data". *American Journal of Managed Care*, **3**: S33-S39.

d'obtenir un résultat proche de la réalité dans chaque pays. Dans le cas du ratio coût-efficacité, on note que les ratios sont différents en fonction des pays²⁰⁸. Les variables économiques telles que l'utilisation des ressources et les coûts devraient donc être analysées en gardant à l'esprit la forte probabilité d'interaction pays-traitement.

Ces résultats soutiennent l'idée de Drummond, Manca et Sculpher (2005) qui pensent que les résultats des analyses coût-efficacité réalisées à partir d'essais cliniques internationaux ne sont pas directement applicables aux pays en question. Des modélisations additionnelles sont nécessaires afin d'adapter les résultats à un pays en particulier²⁰⁹. Plusieurs raisons justifient cela²¹⁰ :

- Les décisionnaires sont spécifiques à un pays et ne sont intéressés que par les résultats et les données pertinents pour leur juridiction.
- Le pays qui nous intéresse n'a pas forcément participé à l'essai clinique.
- Même si le pays a participé à l'essai clinique, la présence de facteurs spécifiques aux pays peut affecter la variabilité géographique des résultats.

Des méthodes peuvent être utilisées afin que les résultats puissent être appliqués à un pays.

B. Méthodes pour rendre les estimés des ratios coût-efficacité spécifiques à un pays

Etant donné les éléments cités précédemment et le temps et les efforts nécessaires pour réaliser un essai clinique international, il semble judicieux (du point de vue des industriels et des gouvernements) de supporter l'utilisation de méthodes facilitant la transférabilité des résultats. En permettant aux décideurs d'estimer si les résultats des essais cliniques sont pertinents dans la situation qui les intéresse, elles jouent un rôle clé. En effet, les décisions de remboursement étant prises localement, des informations globales provenant des essais cliniques internationaux ne sont pas pertinentes. En outre, les systèmes de santé n'ont pas toujours la capacité financière de financer des essais cliniques. Enfin, ces méthodes évitent de dépenser des ressources dans des essais cliniques similaires (réalisé dans un seul pays).

Caro *et al.* (2000) reconnaissent qu'il est possible d'observer une variabilité d'un pays à l'autre des facteurs de risques pouvant entraîner des résultats coût-efficacité différents en fonction des localisations. En outre, lorsque l'on cumule les différences pouvant être observées concernant la distribution des facteurs de risques des patients (pris individuellement : *individual-patient level risk*

²⁰⁸ Cook, J. R., *et al.* (2003). "Assessing the appropriateness of combining economic data from multinational clinical trials". *Statistics In Medicine*, **22**(12): 1955-1976.

²⁰⁹ Buxton, M. J., Drummond, M. *et al.* (1997). "Modelling in economic evaluation : an unavoidable fact of life." *Health Economics*, **6**(3) : 217-227.

²¹⁰ Manca, A. and Willan, A. R. (2006). "'Lost in translation' - Accounting for between-country differences in the analysis of multinational cost-effectiveness data". *Pharmacoeconomics*, **24**(11): 1101-1119.

factors) et les différences propres aux facteurs des pays (système de santé, part du PIB dépensée pour des soins de santé...) cela limite la généralisabilité des effets relatifs du traitement observés au cours de l'essai clinique. Malgré ces observations, l'hypothèse concernant la généralisabilité des effets relatifs d'un traitement n'est que très rarement étudiée en profondeur²¹¹.

Wilke *et al.* (1998) furent parmi les premiers à étudier les analyses statistiques des essais cliniques internationaux. Ils réalisent une étude basée sur différentes hypothèses :

- Une analyse regroupant les données de tous les pays.
- Une analyse regroupant les données de tous les pays réalisée individuellement pour chaque pays en prenant en compte les prix de chaque pays. Une faible variabilité des résultats a été observée.
- Une analyse reliant les effets des effets observés pour l'ensemble des pays aux coûts de chaque pays pris individuellement. La différence de résultats est plus notable que dans l'exemple précédent.
- Une analyse prenant en compte les données individuelles de chaque pays.

Le problème de cette approche est que l'auteur de l'étude doit réaliser un choix entre regrouper les données ou utiliser les données individuelles des pays. Les limites d'une analyse regroupant les données sont liées à la variabilité des données. Cependant, séparer toutes les données peut-être problématique dans le cas où le nombre de patients originaires du pays cible était faible relativement aux autres pays. En outre, il peut-être suggéré que les données provenant de pays différents peuvent tout de même partager des similitudes et qu'il peut y avoir des avantages à tenter de les prendre en compte.

En outre, lorsque l'on étudie les essais cliniques internationaux auxquels le pays cible a participé, on note que les chercheurs font l'hypothèse que le « risque de base » (*baseline risk*) peut varier d'un pays à l'autre alors que les effets relatifs du traitement sont généralisables²¹². Dans ce cas, il est conseillé d'utiliser un modèle en particulier (*events-based model*) développé à partir des éléments généralisables de la pathologie et des pronostics et d'utiliser les données individuelles des patients obtenues grâce à l'essai clinique afin d'estimer la probabilité que les « évènements cliniques » ayant un impact sur l'utilisation des ressources et/ou sur la santé (QOL) surviennent. Alors, les effets relatifs du traitement sont appliqués au pays d'intérêt (en fonction de son risque de base). Notons

²¹¹ Manca, A. and Willan, A. R. (2006). "'Lost in translation' - Accounting for between-country differences in the analysis of multinational cost-effectiveness data". *Pharmacoeconomics*, **24**(11): 1101-1119.

²¹² O'Connell, D., et al. (2001). "Results of clinical trials and systematic trials: to whom do they apply?" *The advanced handbook of methods in evidence based healthcare*. 57-72

que cette méthode peut également être utilisée dans le cas où le pays cible a participé à l'essai clinique et que les données individuelles des patients sont disponibles pour l'ensemble des pays uniquement.

D'autres auteurs (Manca et Willan, 2006 par exemple) ont étudié l'utilisation de modèles hiérarchiques (*multilevel models*). Il offre la possibilité d'analyser toutes les données tout en permettant d'avoir des résultats spécifiques à chaque pays. Il offre ainsi plus de flexibilité que la méthode étudiée par Wilke *et al.* (1998). Le modèle se base sur l'hypothèse que les groupes (centres/pays) sont sélectionnés de façon randomisée à partir de la population. Il est particulièrement utilisé dans le cas où l'on souhaite expliquer la variabilité observée d'un pays à l'autre lors de l'étude des résultats de coût-efficacité.

Il peut être utilisé de deux façons²¹³ :

- Il permet d'estimer les résultats coût-efficacité en se basant sur les résultats de tous les pays. Dans ce cas, l'incertitude est accrue.
- Il est utilisé afin d'estimer des résultats spécifiques à chaque pays. L'auteur de l'étude établira par la suite dans quelles mesures les résultats sont cohérents.

Calculer des estimés spécifiques pour chaque localisation est facilité grâce au modèle hiérarchique. Une critique fréquemment faite à l'encontre du modèle hiérarchique est qu'il peut-être inapproprié lorsque l'on s'attend à observer des différences systématiques d'un pays à l'autre.

Etant donné le besoin d'obtenir des résultats coût-efficacité robustes lors d'une prise de décision spécifique à une juridiction, ce n'est qu'en prenant en compte les spécificités des pays et des patients que l'hétérogénéité peut-être entièrement étudiée. Manca *et al.* (pas encore publié) ont utilisé le modèle hiérarchique afin de ré-analyser les résultats de l'étude ATLAS (*Assessment of Treatment with Lisinopril and Survival*). En comparant les résultats à ceux obtenus avec la méthode de Wilke *et al.* il semble que les nouveaux estimés soient plus proches de la moyenne de l'échantillon total.

²¹³ Manca, A., *et al.* (2007). "Cost-effectiveness analysis using data from multinational trials: The use of bivariate hierarchical modeling". *Medical Decision Making*. **27**(4): 471-490.

5. Annexes

Annexe 1 – Tableau récapitulatif des recommandations des comparateurs

	Tech. la plus susceptible d'être remplacée	Tech. la plus utilisée	Guidelines de pratiques cliniques	Tech. la plus efficace, la moins coûteuse, la plus coût-efficace	Prise en compte du futur
France					
NICE		OUI	OUI		
Allemagne					
Pays-Bas		OUI	OUI	OUI	OUI
Autriche	OUI (2 nd choix)		OUI	OUI (2 nd choix)	
Suède		OUI			
Nouvelle-Zélande	OUI	OUI			OUI
Canada	OUI				OUI
Pologne	OUI	OUI		OUI (2 nd choix)	
Belgique	OUI		OUI (2 nd choix)		

Tableau 3 : Recommandations concernant le choix du comparateur

Annexe 2 – Niveau de variation attendu dans les études médico-économiques

Table I. Levels of justification for differences in international guidelines of system-level decision makers on methods for healthcare economic evaluation. The table assumes that the healthcare system is using economic evaluation to identify interventions and programmes for which use is consistent with maximising some health-focused objective function, subject to a budget constraint

Requirements for economic evaluation	Variation expected and justified	Variation may be justified by differences in opinion about appropriateness	Variation difficult to justify
Objectives and constraints	Dimensions and weights in the objective function (hence, maximand in evaluations) Budgetary arrangements (hence, cost perspective) Existence of other constraints	Monetary vs non-monetary valuation of health benefits Methods of valuation of costs and health effects Sources of preference data	Inconsistent perspective on costs Failure to require a generic measure of health
Specifying the decision problem	Specific details of population, subpopulations and options to compare		Partial set of comparators based on, for example, the 'most clinically effective' options Unclear specification of the (sub)populations of interest
Use of evidence	In some parameter values to reflect geographical heterogeneity	Comprehensiveness of evidence search methods	Failure to require inclusion of all relevant evidence Failure to require systematic identification of evidence
Decision making under uncertainty	Measure of opportunity cost (e.g. threshold cost effectiveness) Size of the effective population for value of information methods Decision maker's ability to mandate or commission future research		Need for full specification of parameter and structural uncertainty Presentation of decision uncertainty

Tableau 4 : Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontières - Can we ever make economic evaluations generalisable across jurisdictions?". *Pharmacoeconomics*. **24**(11): 1087-1099.

Annexe 3 – Facteurs à prendre en compte lorsque l'on souhaite évaluer la qualité d'une méthode des choix discrets

Table I. Checklist of factors to consider in undertaking and assessing the quality of a discrete choice experiment

1. Conceptualizing the choice process	Was a choice rather than ranking, rating task used? What type of choice was used: binary response, pairs, multiple options? Was a generic or labelled choice used? Was an opt-out, neither or status quo option included? If a forced choice was used, was a justification provided? Was the task incentive compatible?
2. Attribute selection	How were they derived and validated? Was the number of attributes appropriate? Was the coverage appropriate? What form was used: generic or alternative specific? Was price included? If so, was an appropriate payment vehicle used? Was risk included? If so, was it appropriately communicated?
3. Level selection	How were they derived and validated? Was the number of levels per attribute appropriate? Was an appropriate range used? Were the levels evenly spaced?
4. Experimental design	What type of design was used? Full factorial? Fractional factorial? If fractional, which effects are identified: main effects; main effects + higher order interactions? How were the profiles generated and allocated to choice sets? What are the properties of the design? What is the efficiency of the design? Was identification checked (e.g. is the variance-co-variance matrix block diagonal)? Was the design blocked into versions? If so, how were choice sets allocated to versions? Were the resulting properties of the versions checked? Were respondents randomly allocated to versions? How many choice sets were considered per respondent? If some profiles were implausible – how was implausibility defined and how was it addressed?
5. Questionnaire design	Was an appropriate level of background and contextual information provided? Were the task instructions appropriate? Was the medium used to communicate attribute/level information (e.g. words, pictures, multi-media) appropriate?
6. Piloting	Was coverage of attributes and levels checked? Was understanding and complexity checked? Was the length and timing checked?
7. Population/study perspective	Appropriate for research question?
8. Sample and sample size	Were inclusion/exclusion criteria explicit? Was sample size appropriate for model estimation?
9. Data collection	What recruitment method was used? How were data collected (e.g. mail, personal interview, web survey)? What was the response rate? Were incentives used to enhance response rates?
10. Coding of data	Was coding explicitly discussed? Was the coding appropriate for effects to be estimated?
11. Econometric analysis	Were the estimation methods appropriate given experimental design and type of choice response? Was the functional form of the indirect utility functions appropriate given the experimental design? Were alternative specific constants included? Were sociodemographics and other co-variables included? Was goodness of fit considered?
12. Validity	Was internal or external validity investigated? Were answers for any respondents deleted and if so on what basis?
13. Interpretation	Was the interpretation appropriate given coding of data? Were results in line with <i>a priori</i> expectations? Were relative attribute effects compared using a common and comparable metric?
14. Welfare and policy analysis	Was willingness to pay estimated using welfare theoretic compensating variation? Was probability analysis undertaken? Were marginal rates of substitution calculated?

Tableau 5 : Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." *Pharmacoeconomics* **26**(8): 661-677.

Annexe 4 – Facteurs de transférabilité

Transferability factor	Key determinants that might be influenced	Direct influence on:	Effort to check correspondence between study and decision country ^a	Transferability might be improved by adjustment of:	
				resource utilisation	resource valuation
Methodological characteristics					
Perspective		Costs and effects	Very low	Yes	Yes
Discount rate		Costs and effects	Very low	Yes	Yes
Medical cost approach		Direct medical costs	Very low	No	Yes
Productivity cost approach		Productivity costs	Very low	No	Yes
Healthcare system characteristics (supply of technology)					
Absolute and relative prices in healthcare		Direct medical costs	Medium–very high ^b	No	Yes
Practice variation	Effectiveness, capacity utilisation, returns to scale, productivity loss	Costs and effects	Medium–very high ^b	Yes	Yes
Technology availability	Capacity utilisation, returns to scale, productivity loss	Costs	Medium–very high ^b	Yes	Yes
Population characteristics (demand for technology)					
Disease incidence/prevalence	(i) Non-screening programmes: Capacity utilisation, returns to scale	(i) Costs	Medium–very high ^b	(i) No	Yes
	(ii) Screening programmes: Effectiveness, ^c capacity utilisation, returns to scale, productivity loss ^d	(ii) Costs and effects ^c		(ii) Yes ^c	
Case-mix	Effectiveness, capacity utilisation, returns to scale, productivity loss	Costs and effects	Medium–very high ^b	Yes	Yes
Life expectancy	Effectiveness	Costs and effects	Low	No	No
Health-status preferences	Effectiveness	Effects	Medium–very high ^b	No	No
Acceptance, compliance, incentives to patients	Effectiveness, capacity utilisation, returns to scale, productivity loss	Costs and effects	Medium–very high ^b	Yes	Yes
Productivity and work-loss time	Productivity loss	Productivity costs	Low–very high ^b	Yes	Yes
Disease spread ^d	Effectiveness, capacity utilisation, returns to scale, productivity loss	Costs and effects	Medium–very high ^b	Yes	Yes

a Indication of average effort to check correspondence between study country and decision country estimated from the experience with the case studies: **very low** = no further data are necessary in addition to the original study; **low** = additional data necessary which can be very easily accessed via database; **medium** = additional data necessary which can be accessed with some effort via literature or databases; **high** = additional data and analysis of the data necessary; **very high** = new study required to obtain the necessary data.

b Depending whether studies exist that address the respective transferability factor.

c For screening tests with a specificity <1 and subsequent treatment with negative adverse effects.

d Infectious diseases with population-dependent characteristics.

Tableau 6 : Facteurs de transférabilité, comment ils influencent la transférabilité et ajustements possibles.
 Welte, R., *et al.* (2004). "A decision chart for assessing and improving the transferability of economic evaluation results between countries". *Pharmacoeconomics*. 22(13): 857-876.

Annexe 5 – L'arbre de décision dans le modèle de Welte

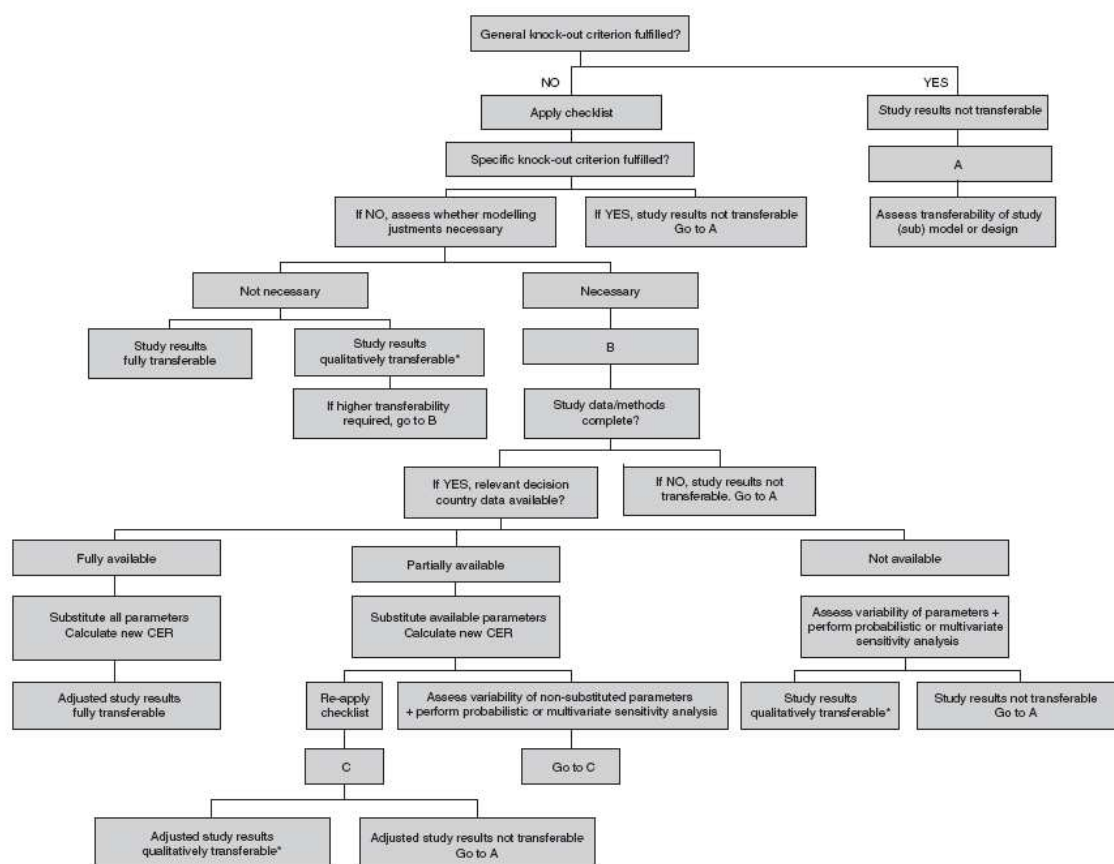


Figure 6 : L'arbre de décision selon Welte *et al.* (2004)

Welte, R., *et al.* (2004). "A decision chart for assessing and improving the transferability of economic evaluation results between countries". *Pharmacoeconomics*. **22**(13): 857-876.

Annexe 6 – Tableau comparatif : Nice, IQWiG et France

	France	NICE	IQWiG
Comparateur	Faire l'inventaire des différentes stratégies. Choisir (en argumentant) celle que l'on va prendre en compte.	Choisir les comparateurs pertinents : étudier en priorité les technologies utilisées le plus fréquemment et les technologies recommandées par les guides de pratiques cliniques (y compris les guides développés par le NICE). Il y a souvent plusieurs comparateurs à prendre en compte car les pratiques cliniques peuvent varier au sein du NHS et parce que les technologies les plus utilisées ne correspondent pas toujours aux meilleures pratiques cliniques.	Toutes les alternatives cliniques pertinentes.
Coûts	Distinction entre coûts directs médicaux et coûts directs non médicaux. Prendre en compte les coûts directs médicaux. Il est recommandé d'utiliser les tarifs conventionnels et les dernières données publiées par les caisses d'Assurance Maladie comme base de valorisation des coûts en médecine de ville. Dans le cas d'une perspective sociétale on y ajoute les parts non remboursées et les dépassements d'honoraires. Les coûts directs non médicaux : coûts supportés par le patient et sa famille (dépense ou temps perdu). Le guideline ne recommande pas une méthode en particulier pour les valoriser (propension à payer des patients, valorisation du nombre d'heures de travail auxquelles l'aidant doit renoncer, valorisation au coût horaire correspondant au coût d'une ressource équivalente sur le marché. Les coûts indirects doivent être pris en compte si l'on pense que leur impact est important (nombre de personnes concernées ou durée des conséquences de la maladie dans la vie des patients). Ils peuvent être valorisés au moyen de la méthode du capital humain, des frictions costs ou en utilisant la théorie du bien-être. La méthode des coûts de frictions est plus réaliste mais requiert un important travail empirique. La théorie du bien-être recommande de prendre en compte la valeur du temps perdu dans l'évaluation des résultats d'une action de soins, et pas dans	Prendre en compte les coûts à la charge du NHS et du PSS uniquement. Valoriser les coûts en prenant en compte les prix pertinents pour le NHS/PSS. Utiliser des analyses de sensibilité pour tenir compte de la variation des prix. Prendre en compte les coûts unitaires. Prendre en compte les coûts liés à la pathologie étudiée au cours des années de vie gagnées grâce à la technologie. Les coûts sans rapport avec la technologie ou la pathologie étudiée doivent être exclus de l'étude. Si l'implantation de la nouvelle technologie nécessite le développement de nouvelles infrastructures, il faut tenir compte de ces coûts. Présenter séparément les coûts supportés par d'autres instances du gouvernement. Prendre en compte les coûts à la charge des patients s'ils sont remboursés par le NHS/PSS. Ne pas en tenir compte autrement.	Distinction entre coûts directs médicaux et coûts directs non médicaux. Prendre en compte les coûts directs médicaux. Les coûts directs non médicaux doivent être pris en compte dans les cas pertinents. Les coûts indirects ne sont pas pris en compte. Ils doivent être considérés séparément.

	les coûts.		
Incertitude	<p>Dans le cadre d'analyses utilisant des données issues de différentes sources, il convient d'effectuer une analyse de sensibilité sur les variables et les hypothèses retenues. Celles-ci doivent être clairement identifiées ainsi que leur intervalle de variation. En cas de modèle complexe, une analyse en termes de scénarios correspondant aux situations la plus et la moins favorable peut être effectuée.</p> <p>Dans le cadre d'analyses stochastiques où les données de coût et d'effet des traitements de chaque patient sont disponibles, il convient de recourir à la méthode paramétrique fondée sur le théorème de Fieller et l'approche non-paramétrique par le bootstrap. On étudiera préalablement la représentativité des données utilisées. Une approche alternative et plus informative pour prendre en compte l'incertitude consiste à représenter la courbe d'acceptabilité des stratégies. Celle-ci représente la probabilité que la stratégie étudiée soit coût-efficace en fonction de différentes valeurs-plafond du ratio coût-résultat, ces dernières variant entre 0 et l'infini.</p>	<p>Incertitude liée aux choix méthodologiques (hypothèses du modèle) peut-être explorée en utilisant différents scénarios. Incertitude liée aux sources des paramètres peut être explorée au moyen d'analyses de sensibilité. Il faut justifier le choix des valeurs utilisées lors de l'analyse.</p> <p>Lorsque la source la plus appropriée a été identifiée, de l'incertitude liée à la précision des valeurs des paramètres existe. On recommande l'utilisation d'analyses probabilistes dans ce cas.</p>	<p>Incertitude liée aux paramètres et incertitude liée aux choix méthodologiques (hypothèses du modèle).</p> <p>Utiliser des analyses de sensibilité pour gérer l'incertitude : analyses de sensibilité déterministe univariée et multivariée et analyses probabiliste (Monte-Carlo).</p> <p>On recommande de ne pas remplacer les analyses de sensibilité univariées par des analyses de sensibilité probabilistes. Elles doivent être menées en complément.</p> <p>Lorsque l'on analyse les résultats, il faut porter attention à l'impact qu'ils peuvent avoir sur la frontière d'efficience.</p>
Transférabilité	<p>La transférabilité est mentionnée lorsque l'on traite du taux d'actualisation.</p> <p>Le postulat de transférabilité peut être justifié. Le décideur doit être en mesure d'évaluer l'état de santé d'une population à partir des états de santé des individus qui la composent.</p> <p>Ne renvoie pas à la transférabilité d'études réalisées dans une autre juridiction.</p>	Non mentionnée.	<p>Il faut prendre en compte les conditions locales liées à l'épidémiologie, aux ressources disponibles, l'accès aux soins de santé, pratiques cliniques, les méthodes de remboursements...</p> <p>Les procédures et les résultats dépendent de la problématique, des données disponibles.</p>

Tableau 7 : Tableau comparatif – NICE, IQWiG et France

Bibliographie

1. Guidelines internationaux

- Agence canadienne des médicaments et des technologies de la santé. (2006). "Lignes directrices de l'évaluation économique des technologies de la santé au Canada." **3** : 34.
- Cleemput, I.; Crott, R. ; Vrijens, F. ; Huybrechts, M. ; Van Wilder, P. ; Ramaekers, D. (2006). "Pharmacoeconomic guidelines." KCE reports. **28** : 20
- Collège des Economistes de la Santé. (2003). "Guide méthodologique pour l'évaluation économique des stratégies de santé."
- College voor zorgverzekeringen. (2006). "Guidelines for pharmacoeconomic research, updated version." Page 7.
- IQWiG (2009). General Methods for the Assessment of the Relation of Benefits to Costs.
- NICE. (2008). "Guide to the methods of technology appraisal."
- Orlewska, E. et P. Mierzejewski. (2003). "Polish guidelines for conducting pharmacoeconomic evaluations".
- Pharmaceutical Benefits Board. (2003). "General guidelines for economic evaluations from the Pharmaceutical Benefits Board." **3** : 2.
- Pharmaceutical Management Agency. (2007). "Prescription for Pharmacoeconomic Analysis." Page 45.
- Walter, E. ; Zehetmayr, S. (2006). "Guidelines on Health Economic Evaluation." Institut für Pharmaökonomische Forschung. Page 5.

2. Compareur

- Cox, H. L., K. B. Laupland, et al. (2006). "Economic evaluation in critical care medicine." Journal Of Critical Care **21**(2): 117-124.
- Draborg, E. et al. (2005). "International comparison of the definition and practical application of health technology assessment." International Journal of Health Technology Assessment in Health Care **21**(1) : 89-95.
- Drummond, M., Sculpher, M. J. et Torrance, G. W. (2005). "Methods for the economic evaluation of health care programmes." 3rd edition. Oxford : Oxford University Press.
- Edwards, S. J. et al. (2009). "Indirect comparisons of treatments based on systematic reviews of randomised controlled trials." The International Journal of Clinical Practice **63**(6): 841-54.
- EUCOMED, (2001). "Health Technology Assessment for Medical Devices in Europe – What has to be considered?"
- Song et al. (2003). "Validity of indirect comparisons for estimating efficacy of competing interventions: empirical evidence from published meta-analyses." BMJ **326**: 472-6.
- Sullivan et al. (2009) "Health Technology Assessment in Health-Care Decisions in the United States." Value in Health **12**(S2) : S39-S44.

3. Prise en compte des coûts

- Brouwer, W.B.F., Rutten, F.F.H., Koopmanschap, M.A.,(2001). "Costing in economic evaluations." dans : Drummond, M.F., McGuire, A. (Eds.), *Economic Evaluation in Health Care: Merging Theory with Practice*. Oxford University Press, New York.
- Claxton, K. ; Sculpher, M. ; Drummond, M. (2002). "A rational framework for decision making by the National Institute for Clinical Excellence." The Lancet. **360**: 711-15.

Feenstra, T. L. ; van Baal, P. H. M. ; Gandjour, A. ; Brouwer, W. B. F. (2008). "Future costs in economic evaluation. A comment on Lee." Journal of Health Economics. **27**: 1645-1649.

Gandjour, A. (2006). "Consumption costs and earnings during added years of life - a reply to Nyman." Health Economics. **15**: 315-317.

Garattini, L. ; Scopelliti, D. ; Mantovani, L. (1995). "A proposal for Italian guidelines in Pharmacoeconomics." PharmacoEconomics. **7**(1) : 1-6.

Garber, A.M., Phelps, C.E. (1997). "Economic foundations of cost-effectiveness analysis. » Journal of Health Economics. **16**, 1–31.

Graf von der Schulenburg, J. M. ; Greiner W. et al. (2008). "German recommendations on health economic evaluation: third and updated version of the Hanover consensus." Value in Health. **11**(4): 542.

Launois, R. ; Vergnenègre, A. ; Garrigues, B. (2003). "Notions et mesure des coûts en fonction de la perspective choisie- Costs, costs and more costs: which one should we use?" Bulletin du Cancer. **90**(11) : 946-954.

Lee, R. H. (2008). "Future costs in cost effectiveness analysis." Journal of Health Economics. **27**: 809-818.

Liljas, B. ; Karlsson, G. S. and StÅlhammar, N-O. (2008). "On future non-medical costs in economic evaluations." Health Economics. **17**: 579-591.

Lundin, D. and Ramsberg, J. (2008). "Comment on survival consumption costs - A reply to Nyman." Health Economics. **17**: 293-297.

Meltzer, D. (2008). "Response to "Future costs and the future of cost-effectiveness analysis."" Journal of Health Economics. **27**: 822-825.

Meltzer D. (1997) "Accounting for future costs in medical cost effectiveness analysis." J Health Econ. **16**(1): 33–64.

Menon, D. ; Schubert, F. ; Torrance, G. W. (1996). "Canada's New Guidelines for the Economic Evaluation of Pharmaceuticals." Medical Care. **34**(12) : DS77-DS86.

Nyman, J. A. (2006). "More on survival consumption costs in cost-utility analysis." Health Economics. **15** : 319-322.

Nyman, J. A. (2004). "Should the consumption of survivors be included as a costs in cost-utility analysis?" Health Economics. **13** : 417-427.

O'Brien, J. A. ; Jacobs, L. M. ; Pierce, D. (2000). "Clinical practice guidelines and the cost of care." International Journal of Technology Assessment Health Care. **16** : 1077-1091.

O'Brien, B. ; Briggs, A. H. (2002). "Analyses of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods." Statistical Methods and Medical Research. **11**: 455-68.

Oliva, J. et al. (2005). "Indirect costs of cervical and breast cancer in Spain." European Journal of Health Economics. **6**: 390-313.

Oostenbrink, J. B. ; Koopmanschap, M. A. ; Rutten, F. F. H. (2002). "Standardisation of costs - The dutch manual for costing in economic evaluations." PharmacoEconomics. **20**(7): 443-454.

Richardson, J. R. J. and Abel Olsen, J. (2006). "In defence of societal sovereignty: a comment on Nyman 'the inclusion of survivor consumption in CUA'." Health Economics. **15**: 311-313.

Russell, L.B. (1986). "Is Prevention Better Than Cure?" The Brookings Institution, Washington, DC.

Smith, M. W. and Barnett, P. G. (2003). "Direct measurement of health care costs." Medical Care Search and Reviews. **60**(3): 74S-91S.

Sultan Taïeb, H., Tessier, P. et Béjean, S. (2009). "Capital humain et coûts de friction : quels critères de choix pour l'évaluation des pertes de production ?", Revue Economique. **60**(2) : 297-306.

Thorsten, T. ; Makela, M. (1999). "Chancing Professional Practice: Theory and Practice of Guidelines Implementation." Danish Institute for Health Services Research.

Van Baal, P. H. M. ; Feenstra, T. L. ; Hoogenveen, R. T. ; Ardine de Witt, G. ; Brouwer, W. B. F. (2007). "Unrelated medical care in life years gained and the cost utility of primary prevention: in search of a 'perfect' cost-utility ratio." Health Economics. **16** : 421-433.

Weinstein, M.C. (1980). "Clinical Decision Analysis". Saunders, Philadelphia.

4. Incertitude

Ades, A. E. et Cliffe, S. (2002). "Markov Chain Monte Carlo estimation of a multiparameter decision model: Consistency of evidence and the accurate assessment of uncertainty." Med Decision Making. **22**: 359-371.

Brigs, A. H. (2000). "Handling uncertainty in cost-effectiveness models". Pharmacoeconomics. **17**(5): 479-500.

Briggs, A. H. (2004). "Statistical approaches to handling uncertainty in health economic evaluation." European Journal of Gastroenterology and Hepatology. **16**: 551-561.

Buxton, M. J. (2001). "Implications of the appraisal function of the National Institute for Clinical Excellence." Value in Health. **4**(3) : 212-216.

Carayanni, V. (2009). "Handling uncertainty in the cost effectiveness healthcare evaluations. A review of statistical approaches." Communications in Statistics- Theory and Method. **38**(8) : 1224-140.

Claxton, K. et Posnett, J. (1996). "An economic approach to clinical trial design and research priority-setting." Health Economics. **5**:513–524.

Drummond, M. et al. (1997). "Methods for the economic evaluation of health care programmes." New-York : Oxford University Press.

Fenwick, E.; Claxton, K. et Sculpher, M. (2001). "Representing uncertainty: the role of cost-effectiveness acceptability curves". Health Economics Letters. **10**: 779-787.

Ishak J. ; Gestisios D. et Caro J. (2009). "When does value of information analysis add value?" ISPOR Europe.

O'Brien, B. J. et Briggs, A. H. (2002). "Analysis of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods". Statistical Methods in Medical Research. **11**: 455-468.

Riou Frana L. et Launois R. (2005). "Prise en compte de l'incertitude dans l'valuation des technologies de sant. Simulation de Monte Carlo, Familles de distribution de probabilit, calibrage des distributions." REES France.

Sculpher, M. et Claxton, K. (2005). "Establishing the cost effectiveness of new pharmaceuticals under conditions of uncertainty. When is there sufficient evidence?" Value in Health. **8**(4): 433-446.

Siani, C. et de Peretti, C. (2006). "Decision-making with the incremental cost-effectiveness ratio under uncertainty." SAS Information dcision patient. **9** (111-145).

Siani, C. et de Peretti, C. (2006). "Fieller's method performance in problematic cases for decision-making." Health and System Science. 1-24.

Siani, C. et Moatti, JP. (2003). "Le traitement de l'incertitude dans les valuations conomiques de stratgies sanitaires." Revue d'Epidmiologie et de Sant Publique. **51**(2) : 255-276.

Stinnett, A. et Mullahy, J. (1998) "Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis." Med Decis Making. **18** (2 suppl):S68–S80.

Tambour, M. ; Zethraeus, N. et Johannesson, M. (1998). "A note on confidence intervals in cost-effectiveness analysis." Int J Technol Assess Health Care. **14**:467–471.

Zethraeus, N. ; Johannesson, M. ; Jönsson, B. ; Löthgren, M. et Tambour, M. (2003). "Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies." Pharmacoeconomics. **21**(1): 39-48.

5. Méthode des Choix Discrets

- Adamowicz, W. *et al.* (1998). "Introduction to Attribute-Based Stated Choice Models." US Department of Commerce, NOAA.
- Adamowicz, W. (2006). "Frontiers in stated preference methods: an introduction." Environmental and Resource Economics **34**: 1–6.
- Araña, J. E. *et al.* (2009). "Understanding the use of non-compensatory decision rules in discrete choice experiments: The role of emotions." Ecological Economics **68**(8-9): 2316-2326.
- Bateman, I. J. *et al.* (2002). "Economic Evaluation with Stated Preference Techniques, A Manual" 1 edn, Edward Elgar Publishing Limited, Cheltenham.
- Bennett, J. *et al.* (2001). "The choice Modelling approach to Environmental Valuation" Edward Elgar Publishing Limited, UK.
- Blamey *et al.* (2002). "Attribute causality in environmental choice modelling." Environmental and Resource Economics **32**(2): 269-286.
- Boyle, K. J. *et al.* (2009). "Convergent Validity of Attribute-Based, Choice Questions in Stated-Preference Studies." Environmental & Resource Economics **42**(2): 247-264.
- Bryan, S. *et al.* (1998). "Magnetic resonance imaging for thinvestigation of knee injuries: An investigation of preferences." Health Economics **7**(7): 595-603.
- Bryan, S. *et al.* (2004). "Discrete Choice Experiments in Health Economics: For Better or for Worse?" The European Journal of Health Economics **5**(3): 199-202.
- Carlsson, F. *et al.* (2003). "Design techniques for stated preference methods in health economics." Health Economics **12**(4): 281-294.
- Champ, P. *et al.* (2001). "Donation payment mechanisms and contingent valuation: An empirical study of hypothetical bias." Environmental and Resource Economics **19**(4): 383-402.
- Cookson, R. (2003). "Willingness to pay methods in health care: a sceptical view." Health Economics **12**(11): 891-894.
- Cummings, R. *et al.* (1999). "Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method." American Economic Review **89**(3): 649-665.
- Curry, J. (1997). "After the Basic: Keeping Key issues in mind makes conjoint analysis easier to apply" Marketing Research **9**: 6-11.
- DeShazo, J.R. *et al.* (2002). "Designing choice sets for stated preference methods: the effects of complexity on choice consistency." J Environ Econ Manage **44**: 123-43.
- Dillman, D. *et al.* (2001). "Mail and internet surveys : the tailored design method." New York: Wiley.
- Dziegielewska, D. A. and R. Mendelsohn (2007). "Does "No" mean "No"? A protest methodology." Environmental & Resource Economics **38**(1): 71-87.
- Green, P. E. *et al.* (1990). "Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice." Journal of Marketing **54**(4): 3-19.
- Green, P. E. *et al.* (1978). "Conjoint Analysis in Consumer Research - Issues and Outlook2" Journal of consumer research **5**(2): 103-123.
- Gustafsson, A., Herrmann, A., *et al.* (2000). "Conjoint Measurement - Methods and Application" Springer-Verlag, Berlin, Germany.
- Hakim, Z. *et al.* (1999). "Modelling the EuroQol data : a comparison of discrete conjoint and conditional preference modelling." Health Economics **8**(2):103-16.
- Hall, J., R. Viney, *et al.* (2004). "Using stated preference discrete choice modeling to evaluate health care programs." Journal of Business Research **57**(9): 1026-1032.
- Huber, J. *et al.* (1996). "The importance of utility balance in efficient choice designs" Journal of Marketing Research **33**(3): 307-317.
- Johnson, F. *et al.* (2000). "Willingness to pay for improved respiratory and cardiovascular health: A multiple-format, stated-preference approach" HealthEconomics **9**(4): 295-317.

Jorgensen, B.S. *et al.* (1999). "Protest responses in contingent valuation." Environ Resour Econ **14**(1):131–50.

Jorgensen, B.S. et G.J. Syme (2000). "Protest responses and willingness to pay: attitude toward paying for storm water pollution abatement." Ecol Econ **33**(2):251–265.

Keeney, R. L. & H. Raiffa (1976). "Decisions with Multiple Objectives - Preferences and Value Tradeoffs" John Wiley & Sons, New York.

Kjaer, T. (2005). "A review of the discrete choice experiment – with emphasis on its application in health care." Health Economics Papers University of Southern Denmark. Odense Denmark, 139p.

Lancaster, K. J. (1966). "A new approach to consumer theory" Journal of Political Economy **74**(2): 132-157.

Lancsar, E. et C. Donaldson (2005). "Discrete choice experiments in health economics: distinguishing between the method and its application." Eur J Health Econ **6**(4): 314-6.

Lancsar, E. and J. Louviere (2006). "Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences?" Health Economics **15**(8): 797-811.

Lancsar, E. and J. Louviere (2008). "Conducting discrete choice experiments to inform Healthcare decision making." Pharmacoeconomics **26**(8): 661-677.

Lancsar, E., J. Louviere, et al. (2007). "Several methods to investigate relative attribute impact in stated preference experiments." Social Science & Medicine **64**(8): 1738-1753.

Louviere J.J. (2006). "What you don't know might hurt you: some unresolved issues in the design and analysis of discrete choice experiments." Environmental and Resource Economics **34**: 173–88.

Louviere, J.J. (2000). "Why stated preference discrete choice modelling is NOT conjoint analysis" Memetrics White Paper.

Louviere, J.J., Hensher, D. A. et Swait, J. (2000). "Stated Choice Methods, analysis and application" Cambridge University Press, U.K.

Louviere, J.J. et G. Woodsworth (1983). "Design and analysis of simulated choice or allocation experiments: an approach based on aggregate data" Journal of Marketing Research **20**: 350–67.

Lusk, J. (2003) "Effects of cheap talk on consumer willingness-to-pay for golden rice." American Journal of Agricultural Economics. **85**(4): 840-856.

Mangham, L. J., K. Hanson, et al. (2009). "How to do (or not to do) Designing a discrete choice experiment for application in a low-income country." Health Policy and Planning **24**(2): 151-158.

Manski, C. H. (1977). "The Structure of Random Utility Models" Theory and Decision **8**: 229-254.

Ratcliffe, J. et L. Longworth (2002). "Investigating the structural reliability of a discrete choice experiment within health technology assessment." International Journal of Technology Assessment in Health Care **8**(1): 139-144.

Ratcliffe, J. et M. Buxton (1999). "Patients' preferences regarding the process and outcomes of lifesaving technology - An application of conjoint analysis to liver transplantation" International Journal of Technology Assessment in Health Care **15**(2): 340-351.

Ryan, M. (2004). "Deriving welfare measures in discrete choice experiments: a comment to Lancsar and Savage (1)." Health Economics **13**(9): 909-912.

Ryan, M. (2004). "Discrete choice experiments in health care - NICE should consider using them for patient centred evaluations of technologies." British Medical Journal **328**(7436): 360-361.

Ryan, M. (1999). "A role for conjoint analysis in technology assesment in health care?" International Journal of Technology Assessment in Health Care **15**(3): 443-457.

Ryan, M. et K. Gerard (2003). "Using discrete choice experiments to value health care programmes: current practice and future research reflections." Applied Health Economics and Health Policy **2**(1): 55-64.

Ryan, M. et J. Hughes (1997). "Using conjoint analysis to assess women's preferences for miscarriage management" Health Economics **6**(3): 261-273.

- Ryan, M. et D. Skatun (2004). "Modelling non-demanders in choice experiments" Health Economics **13**(4): 397-402.
- Ryan, M. *et al.* (2006). "Using discrete choice experiments to estimate a preference-based measure of outcome: an application to social care for older people." Journal of Health Economics. **25**(5): 927-44.
- Salkeld, G. *et al.* (2000). "The veil of experience. Do consumers prefer what they know best?" Health Economics **9**(3): 267-270.
- Scott, A. (2001). "Eliciting GPs' preferences for pecuniary and non-pecuniary job characteristics." Journal of Health Economics **20**: 329-347.
- Smith, R. D. (2003). "Construction of the contingent valuation market in health care: a critical assessment." Health Economics **12**(8): 609-628.
- Street, D. J. *et al.* (2005). "Quick and easy choice sets: Constructing optimal and nearly optimal states choice experiments." International Journal of Research in Marketing **22**: 459-470.
- Swait, J. *et al.* (2002). "Context dependence and aggregation in disaggregate choice analysis" Marketing Letters **13**(3): 195-205.
- Telser, H. et P. Zweifel (2007). "Validity of discrete-choice experiments evidence for health risk reduction." Applied Economics **39**(1): 69-78.
- Viney, R., E. Savage, et al. (2005). "Empirical investigation of experimental design properties of discrete choice experiments in health care." Health Economics **14**(4): 349-362.
- Wainwright, D. M. (2003). "More 'con' than 'joint': problems with the application of conjoint analysis to participatory healthcare decision making" Critical Public Health **13**(4): 373-380.

6. Evaluation contingente

- Ahlheim, M. *et al.* (2007). "Using citizen expert groups in environmental valuation — lessons from a CVM study in Northern Thailand." Hohenheimer Diskussionsbeiträge. 283.
- Akter, S. ; Bennett, J. ; Akhter, S. (2008). "Preference uncertainty in contingent valuation." Ecological Economics. **67**(3): 345-351.
- Araña, J. and León, C. (2002). "Willingness to pay for health risk reduction in the context of altruism". Health Economics. **11**: 623-635.
- Arrow, K. *et al.* (1993). "Report of the NOAA Panel on Contingent Valuation".
- Arrow, K. (1963). "Social Choice and Individual Values". Yale University Press, New Haven.
- Bateman, I. J. *et al.* (2009). "Reducing gain-loss asymmetry: A virtual reality choice experiment valuing land use change." Journal of Environmental Economics and Management. **58**(1): 106-118.
- Bayoumi, A. M. (2004). "The measurement of contingent valuation for health economics". Pharmacoeconomics. **22**(11): 691-700.
- Becker, G. (1993). "Nobel lecture: the economic way of looking at behavior". Journal of Political Economy. **101**(3): 385-409.
- Bergstrom, J. ; Dillman, B. ; Stoll, J. (1985). "Public environment amenity benefits of private land: the case of prime agricultural land". Southern Journal of Agricultural Economics. **17**: 139-49.
- Bohm, P. (1972). "Estimating demand for public goods an experiment". European Economic Review. **3**: 111-130.
- Bonnieux, F. (1998). "Principes, mises en œuvre et limites de l'évaluation contingente". Economie publique. **1**: 47 - 90.
- Brookshire, D. and Coursey, D. (1987). "Measuring the value of a public good: an empirical comparison of elicitation procedures". American Economic Review. **77**: 554-556.
- Carson, R. T. and Groves, T. (2007). "Incentive and informational properties of preference questions". Environmental and Resource Economics. **37**: 181-210.

Champ, P. A. and Bishop, R. C. (2001). "Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias". Journal of Environmental Economics and Management. **19**: 383–402.

Champ, P. A. *et al.* (1997). "Using donation mechanisms to value non-use benefits from public goods". Journal of Environmental Economics and Management. **33**: 151–162.

Champ, P. A. *et al.* (2002). "Contingent valuation and incentives". Land Economics. **78**(4): 591–604.

Chanel, O. *et al.* (2004). "Les consentements à payer pour des programmes de prévention sanitaire incluent-ils de l'altruisme?" Revue Économique. **55**(5): 923–946.

Chang, J. I. ; Yoo, S. H. ; Kwak, S. J. (2007). "An investigation of preference uncertainty in the contingent valuation study". Applied Economics Letters. **14**: 691–695.

Cummings, R. G. ; Brookshire, D. S. ; Schulze, W. D. (1986). "Valuing environmental goods: a state of the arts assessment of the contingent method". Totowa NJ: Rowman and Allanheld editors.

Cummings, R. G. and Taylor, L. O. (1999). "Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method". American Economic Review. **89**: 649–665.

Desvousges, W.S. ; Smith, V. K. ; McGivney, M. P. (1992). "Comparison of alternative approaches for estimating recreation and related benefits for water quality improvements". U. S. Environmental Protection Agency.

Diamond, P. A. and Hausman, J. A. (1992). "On contingent valuation measurement of nonuse values". Paper presented at the Cambridge Economics Inc. Symposium, Contingent Valuation : A Critical Assessment. Washington, D. C.

Diamond, P. A. *et al.* (1992). "Does contingent valuation measure preferences? Experimental evidence." Paper presented at the Cambridge Economics Inc. Symposium, Contingent Valuation: A Critical Assessment. Washington, D. C.

Drummond, M. F. *et al.* (1997). "Méthode d'évaluation économique des programmes de santé". Edition Economica pour la traduction française. 2^{ème} édition.

Etzioni, A. (1988). "The moral dimension: toward a new economics". The Free Press, New York.

Fisher, R. J. (1993). "Social desirability bias and the validity of indirect questioning". Journal of Consumer Research. **20**: 303–315.

Flachaire, E. and Hollard, G. (2005). "Une approche comportementale de l'évaluation contingente".

Fleming, C. M. and Bowden, M. (2009). "Web-based surveys as an alternative to traditional mail methods." Journal of Environmental Management. **90**(1): 284–292.

Frör, O. (2008). "Bounded rationality in contingent valuation: Empirical evidence using cognitive psychology." Ecological Economics. **68**(1-2): 570–581.

Gouriéroux, C. (1998). "Aspects statistiques de la méthode d'évaluation contingente". Etudes et Recherches en Economie Publique. **1**: 125–159.

Hibbard, J. H. *et al.* (2002). "Strategies for reporting health plan performance information to consumers: evidence from controlled studies". Health Services Research. **37**(2): 291–313.

Iraguen, P. and Ortuzar, J. (2004). "Willingness-to-pay for reducing fatal accident risk in urban areas: an Internet-based web page stated preference survey." Accident Analysis and Prevention. **36**: 513–524.

Johanesson, M. *et al.* (1999). "Calibrating hypothetical willingness to pay responses". Journal of Risk and Uncertainty. **18**: 21–32.

Jorgensen, B. S. *et al.* (1999). "Protest Responses in Contingent Valuation." Environmental and Resource Economics **14**(1):131–150.

Kahneman, D. and Knetsch, J. (1992). "Valuing public goods: the purchase of moral satisfaction". Journal of Environmental Economics and Management. **22**: 57–70.

Kahneman, D. and Tversky, A. (1984). "Choices, values, and frames". American Psychologist. **39**: 341–350.

- Krosnick, J. A. (1999). "Maximising measurement quality: principles of good questionnaire design". In Measures of Political Attitudes, ed. J. P. Robinson, P. R. Shaver and L. S. Wrightsman, New-York : Academic Press.
- Lamiraud, K. *et al.* (2009). "The impact of information on patient preferences in different delivery patterns: A contingent valuation study of prescription versus OTC drugs." Health Policy. **93**(2-3): 102-110.
- Lee, S. J. *et al.* (1998). "The impact of risk information on patients' willingness to pay for autologous blood donation". Medical Care. **36**: 1162-73.
- Leggett, C. G. *et al.* (2003). "Social desirability bias in contingent valuation surveys administered through in-person interviews." Land Economics. **79**(4): 561-575.
- Li, H. *et al.* (2004). "Would developing country commitments affect US households' support for a modified Kyoto Protocol?" Ecological Economics. **48**: 329-343.
- Li, C. and Mattson, L. (1995). "Discrete choice under preference uncertainty: an improved structural model for contingent valuation". Journal of Environmental Economics and Management. **28**: 256-269.
- Loomis, J. and Ekstrand, E. (1998). "Alternative approaches for incorporating respondent uncertainty when estimating willingness to pay: the case of the Mexican Spotted Owl". Ecological Economics. **27**: 29-41.
- Luchini, S. (2002). "De la singularité de la méthode d'évaluation contingente". Economie et Statistique. 357-358.
- Lundberg, L. *et al.* (1999). "Quality of life, health-state utilities and willingness to pay in patients with psoriasis and atopic eczema". British Journal of Dermatology. **141**: 1067-1075.
- Macmillan, D. ; Hanley, N. ; Lienhoop, N. (2006). "Contingent valuation: environmental polling or preference engine?" Ecological Economics. **60**: 299-307.
- Maguire, K. B. (2009). "Does mode matter? A comparison of telephone, mail, and in-person treatments in contingent valuation surveys." Journal of Environmental Management. **90**(11): 3528-3533.
- Mathews, K. E. ; Freeman, M . L. ; Desvousges, W. H. (2006). "How and how much? The role of information in stated choice questionnaires" in B. Kanninen (Ed.), Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice, The Economics of Non-Market Goods and Services, Springer, Dordrecht. **8**: 111-133.
- McFadden, D. (2001). "Economic choices". American Economic Review. **91**(3): 351-378.
- Mitchell, R. C. and Carson, R. T. (1989). "Using surveys to value public goods: the contingent valuation method". Washington, DC: Resources for the Future.
- Munro, A. and Hanley, N. (2001). "Information, uncertainty and contingent valuation." In: Bateman, I. ; Willis, K. editors. Valuing environmental preferences: theory and practice of the contingent valuation method in the US, EU, and developing countries. Oxford: Oxford University Press. 258-79.
- Murphy, J. J. *et al.* (2005). "Is cheap talk effective at eliminating hypothetical bias in a provision point mechanism?" Environmental and Resource Economics. **30**: 327-343.
- Mussweiler, T. and Strack, F. (2001). "Considering the impossible: Explaining the effects of implausible anchors". Social Cognition. **19**: 145-160.
- Olsen, J.A. et Smith, R. (2001). "Theory versus practice: A review of willingness to pay in health and health care." Health Economics. **10**(1):39-52.
- Onwujekwe, O. (2001). "Searching for a better willingness to pay elicitation method in rural Nigeria: the binary question with follow-up method versus the bidding game technique". Health Economics. **10**: 147-158.
- Özdemir, S ; Johnson, F. R. ; Hauber, A. B. (2009). "Hypothetical bias, cheap talk, and stated willingness to pay for health care". Journal of Health Economics. **28**: 894-90.

Payne, J. W. ; Bettman, J. R. ; Schkade, D. A. (1999). "Measuring constructed preferences: towards a building code". Journal of Risk and Uncertainty. **19** (1–3): 243–270.

Plott, C. R. and Zeiler, K. (2005). "The willingness to pay / willingness to accept gap, the endowment effect, subject misconceptions and experimental procedures for eliciting valuations". American Economic Review. **95**.

Posavac, S. S. *et al.* (2009). "On assuaging positive bias in environmental value elicitation." Journal of Economic Psychology. **30**(3): 482-489.

Ready, R. ; Whitehead, J. ; Blomquist, G. (1995). "Contingent valuation when respondents are ambivalent". Journal of Environmental Economics and Management. **29**: 181–197.

Roach, B. ; Boyle, K. J. ; Welsh, M. (2002). "Testing bid design effects in multiple-bounded, contingent valuation questions". Land Economics. **78**(1): 121-131.

Roe, B. ; Boyle, K. ; Teisl, M. (1996). "Using conjoint analysis to derive estimates of compensating variation". Journal of Environmental Economics and Management. **31**: 145–159.

Rose, J. M. et Masiero, L. (2010). "A comparison of prospect theory in WTP preference space." Quaderni della facoltà di Scienze economiche dell'Università di Lugano.

Samnaliev, M. ; Stevens, T. H. ; More, T. (2005). "A comparison of alternative certainty calibration techniques in contingent valuation". Ecological Economics. **57**: 507–519.

Sayman, S. and Öncüler, A. (2005). "Effects of study design characteristics on the wta-wtp disparity: a meta analytical framework". Journal of Economic Psychology. **26**: 289-312.

Schläpfer, F. (2008). "Contingent valuation: A new perspective." Ecological Economics. **64**(4): 729-740.

Schogren, J. *et al.* (1994). "Resolving differences in willingness to pay and willingness to accept." American Economic Review. **84**: 255-270.

Schwappach, D. and Strassmann, T. (2006). "Quick and dirty numbers'? The reliability of a stated-preference technique for the measurement of preferences for resource allocation". Journal of Health Economics. **25**: 432-448.

Shaikh, S. L. ; Sun, L. ; van Kooten, G. C. (2007). "Treating respondent uncertainty in contingent valuation: a comparison of empirical treatments". Ecological Economics. **62**: 115–125.

Shiell, A. and Rush, B. (2003). "Can willingness to pay capture the value of altruism? An exploration of Sen's notion of commitment". Journal of Socio-Economics. **32**: 647-660.

Swait, J. *et al.* (2002). "Context dependence and aggregation in disaggregate choice analysis". Marketing Letters. **13**(3): 195–205.

Tsuge, T. and Washida, T. (2003). "Economic valuation of the Seto Inland Sea by using an Internet CV survey". Marine Pollution Bulletin. **47**: 230-236.

Vatn, A. (2009). "An institutional analysis of methods for environmental appraisal." Ecological Economics. **68**(8-9): 2207-2215.

Wilson, T. D. *et al.* (1996). "A new look at anchoring effects: basic anchoring and its antecedents". Journal of Experimental Psychology. **4**: 387-402.

7. Transférabilité

Barbieri, M., *et al.* (2005). "Variability of cost-effectiveness estimates for pharmaceuticals in Western Europe: Lessons for inferring generalizability". Value in Health. **8**(1): 10-23.

Boulenger, S., *et al.* (2005). "Can economic evaluations be made more transferable?". The European Journal Of Health Economics: HEPAC: Health Economics In Prevention And Care. **6**(4): 334-346.

Caro, J. J., *et al.* (2000). "Are the WOSCOPS clinical and economic findings generalizable to other populations? A case study for Belgium". Acta Cardiologica. **55**(4): 239-246.

Cook, J. R., *et al.* (2003). "Assessing the appropriateness of combining economic data from multinational clinical trials". Statistics In Medicine. **22**(12): 1955-1976.

Drummond, M., *et al.* (2009). "Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force Report". Value in Health. **12**(4): 409-418.

Drummond, M., Manca, A., Sculpher, M. J. (2005). "Increasing the generalisability of economic evaluations: recommendations for the design, analysis and reporting of studies". International Journal of Technology Assessment in Health Care. **21**(2): 165-171.

Drummond, M., Sculpher, M. J. et Torrance, G. W. (2005). "Methods for the economic evaluation of health care programmes." 3rd edition. Oxford : Oxford University Press.

Goeree, R., *et al.* (2005). "Generalizability of economic evaluation: using results from other geographic areas or from multinational trials to help inform health care decision making in Canada". CCOHTA HTA Capacity Building Grants Program. Toronto (ON): Canadian Coordination Office for Health Technology Assessment. Grant N. 67

Goeree, R., *et al.* (2007). "Transferability of economic evaluations: approaches and factors to consider when using results from one geographic area for another". Current Medical Research And Opinion. **23**(4): 671-682.

Knies, S., *et al.* (2009). "The transferability of economic evaluations: Testing the model of Welte". Value in Health. **12**(5): 730-738.

Li, S. C. (2007). "How to improve pharmacoeconomic data generalisability between different countries". Expert Opinion on Pharmacotherapy. **8**(10): 1409-1413.

Manca, A., *et al.* (2007). "Cost-effectiveness analysis using data from multinational trials: The use of bivariate hierarchical modeling". Medical Decision Making. **27**(4): 471-490.

Manca, A. and Willan, A. R. (2006). "Lost in translation' - Accounting for between-country differences in the analysis of multinational cost-effectiveness data". Pharmacoeconomics. **24**(11): 1101-1119.

Mason, J. M. and Mason, A. R. (2006). "The generalisability of pharmacoeconomic studies. Issues and challenges ahead". Pharmacoeconomics. **24**(10): 937-945.

Mullins, C. D. and Ogilvie, S. (1998). "Emerging standardization in pharmacoeconomics". Clinical Therapeutics. **20**(6): 1194-1202.

Nixon, J., *et al.* (2009). "Guidelines for completing the EURONHEED transferability information checklists". The European Journal Of Health Economics. HEPAC: Health Economics In Prevention And Care. **10**(2): 157-165.

O'Brien, B. J. (1997). "A tale of two (or more) cities: Geographic transferability of pharmacoeconomic data". American Journal of Managed Care. **3**: S33-S39.

O'Connell, D., *et al.* (2001). "Results of clinical trials and systematic trials: to whom do they apply?" The advanced handbook of methods in evidence based healthcare. 57-72

Pang, F. (2002). "Design, analysis and presentation of multinational economic studies - The need for guidance". Pharmacoeconomics. **20**(2): 75-90.

Sculpher, M. J. and Drummond, M. F. (2006). "Analysis sans frontieres - Can we ever make economic evaluations generalisable across jurisdictions?". Pharmacoeconomics. **24**(11): 1087-1099.

Späth, H. M., *et al.* (1999). "Analysis of the eligibility of published economic evaluations for transfer to a given health care system. Methodological approach and application to the French health care system". Health Policy (Amsterdam, Netherlands). **49**(3): 161-177.

Urdahl, H., Manca, A. et Sculpher, M. J. (2006). "Assessing generalisability in model-based economic evaluation studies - A structured review in osteoporosis". Pharmacoeconomics. **24**(12): 1181-1197.

Welte, R., *et al.* (2004). "A decision chart for assessing and improving the transferability of economic evaluation results between countries". Pharmacoeconomics. **22**(13): 857-876.

Welte, R., *et al.* (2003). "A user-friendly tool for evaluating and improving the transferability of economic evaluation results between countries". Value in Health. **6**(6): 616-617.

Willke, R. J., *et al.* (1998). "Estimating country-specific cost-effectiveness from multinational clinical trials". Health Economics. **7**(6): 481-493.