

QUELLES MÉTHODES POUR ÉLICITER LES UTILITÉS EN SANTÉ ET REFLÉTER AU MIEUX LES PRÉFÉRENCES DES QUÉBÉCOIS?

THOMAS PODER, PH. D.
HOSEIN AMERI, PH. D.



Les rapports de projet sont destinés plus spécifiquement aux partenaires et à un public informé. Ils ne sont ni écrits à des fins de publication dans des revues scientifiques ni destinés à un public spécialisé, mais constituent un médium d'échange entre le monde de la recherche et le monde de la pratique.

Project Reports are specifically targeted to our partners and an informed readership. They are not destined for publication in academic journals nor aimed at a specialized readership, but are rather conceived as a medium of exchange between the research and practice worlds.

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du gouvernement du Québec, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Quebec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the government of Quebec, and grants and research mandates obtained by its research teams.

Les partenaires du CIRANO – CIRANO Partners

Partenaires Corporatifs - Corporate Partners

Autorité des marchés financiers
Banque de développement du Canada
Banque du Canada
Banque Nationale du Canada
Bell Canada
BMO Groupe financier
Caisse de dépôt et placement du Québec
Énergir
Hydro-Québec
Intact Corporation Financière
Investissements PSP
Manuvie
Mouvement Desjardins
Power Corporation du Canada
Pratt & Whitney Canada
VIA Rail Canada

Partenaires gouvernementaux - Governmental partners

Ministère des Finances du Québec
Ministère de l'Économie, de l'Innovation et de l'Énergie
Innovation, Sciences et Développement Économique Canada
Ville de Montréal

Partenaires universitaires - University Partners

École de technologie supérieure
École nationale d'administration publique
de Montréal
HEC Montreal
Institut national de la recherche scientifique
Polytechnique Montréal
Université Concordia
Université de Montréal
Université de Sherbrooke
Université du Québec
Université du Québec à Montréal
Université Laval
Université McGill

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web. CIRANO collaborates with many centers and university research chairs; list available on its website.

© Novembre 2025. Thomas Poder et Hosein Ameri. Tous droits réservés. *All rights reserved.* Reproduction partielle permise avec citation du document source, incluant la notice ©. *Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.*

Les idées et les opinions émises dans cette publication sont sous l'unique responsabilité des auteurs et ne représentent pas les positions du CIRANO ou de ses partenaires. *The observations and viewpoints expressed in this publication are the sole responsibility of the authors; they do not represent the positions of CIRANO or its partners.*

ISSN 1499-8629 (version en ligne)

Quelles méthodes pour éliciter les utilités en santé et refléter au mieux les préférences des Québécois ?

Thomas Poder^{}, Hosein Ameri[†]*

Résumé/Abstract

Dans ce rapport, nous comparons empiriquement sept techniques d'élicitation pour évaluer les utilités associées aux états de santé décrits par le Short-Form 6-Dimension version 2 (SF-6Dv2) et où le pari ordinaire (standard gamble – SG) est considérée comme l'approche de référence. Ce faisant, cela nous permet d'estimer un ensemble de valeurs d'utilité pour le SF-6Dv2 à utiliser au Québec et d'introduire une nouvelle approche basée sur le choix dichotomique multiple borné (multiple bounded discrete choice, MBDC). Diverses techniques d'estimation économétrique ont été utilisées et toutes les analyses ont été réalisées à partir de données recueillies auprès de personnes francophones âgées de 18 ans et plus résidant au Québec, Canada, en 2016 et 2018.

L'approche du choix expérimental discret (discrete choice experiment – DCE) combinée avec un best-worst scaling (DCEBWS) apparaît comme étant une alternative cohérente et comparable au SG. Il ressort également de nos analyses que l'approche MBDC est réalisable et permet de générer des données d'utilité représentatives des préférences des répondants. L'ensemble de référence de valeurs d'utilité du SF-6Dv2 retenu pour le Québec est issu du DCEBWS et varie de -0,683 pour le pire état de santé (555655) à 1 pour une santé parfaite (111111). La dimension de la douleur présente ici les plus fortes diminutions de coefficients en termes d'ampleur, ce qui indique sa grande influence dans l'établissement des valeurs d'utilité des états de santé.

In this report, we empirically compare seven elicitation techniques for evaluating the utilities associated with health states described by the Short-Form 6-Dimension version 2 (SF-6Dv2), where the standard gamble (SG) is considered the reference approach. This allows us to estimate a SF-6Dv2 value set of health utilities for use in Quebec and to introduce a new approach based on multiple bounded discrete choice (MBDC). Various econometric estimation techniques were used, and all analyses were performed on data collected from French-speaking individuals aged 18 and over residing in Quebec, Canada, in 2016 and 2018.

The discrete choice experiment (DCE) approach combined with best-worst scaling (DCEBWS) appears to be a coherent and comparable alternative to the SG. Our analyses also show that the MBDC approach is feasible and allows for the generation of utility data representative of respondents' preferences. The reference SF-6Dv2 value set selected for Quebec is derived from DCEBWS and ranges from -0.683 for the worst health condition (555655) to 1 for perfect health

^{*} Département de gestion, évaluation et politique de santé à l'École de santé publique de l'Université de Montréal, CIRANO

[†] Département de gestion, évaluation et politique de santé à l'École de santé publique de l'Université de Montréal

(111111). The pain dimension shows the largest decreases in coefficients in terms of magnitude, indicating its significant influence in establishing the utility values for health conditions in Quebec.

Mots-clés/Keywords : Année de vie ajustée par la qualité; pari ordinaire; choix expérimental discret; choix dichotomique multiple borné; utilité en santé / Quality-adjusted life year; standard gamble; discrete choice experiment; multiple bounded discrete choice; health utility

Codes JEL/JEL Codes : I10, I31.

Pour citer ce document / To quote this document

Poder, T. G., & Ameri, H. (2025). Quelles méthodes pour éliciter les utilités en santé et refléter au mieux les préférences des Québécois ? (2025RP-26, Rapports de projets, CIRANO.)

<https://doi.org/10.54932/UNEK2485>

Résumé

Contexte et objectifs : Une version récente du système de classification abrégé à six dimensions (short-form 6-dimension, SF-6D) a été élaborée (SF-6Dv2). Les objectifs de ce rapport sont : 1) de comparer empiriquement sept techniques d'élicitation pour évaluer les utilités associées aux états de santé décrits par le SF-6Dv2; 2) d'introduire l'approche du choix dichotomique multiple borné (multiple bounded discrete choice, MBDC); 3) d'estimer un ensemble de valeurs d'utilité pour le SF-6Dv2 à utiliser au Québec; et 4) d'évaluer l'effet de l'importance explicite attribuée par les participants aux dimensions du SF-6Dv2 sur l'ordre de grandeur de ces dimensions dans l'ensemble de valeurs.

Méthodes : Divers modèles d'estimation économétrique ont été utilisés pour estimer les données du pari ordinaire (standard gamble, SG). Différents types de choix expérimentaux discrets (discrete choice experiment, DCE), c'est-à-dire le DCE avec durée (DCETTO), l'échelle de choix meilleur-pire avec durée (best-worst-scaling, BWSTTO), le DCE combiné avec le BWS (DCEBWS), le DCE avec mort immédiate comme choix (DCEDOD) et le DCE avec mort immédiate comme attribut (DCEPOD), ont été évalués à l'aide de modèles logit conditionnels. Des modèles semblables à ceux du SG ont été utilisés pour l'approche MBDC. L'ensemble de référence de valeurs d'utilité pour le SF-6Dv2 a été estimé à l'aide du modèle logit conditionnel. Enfin, l'étude a examiné si la contribution relative de chaque dimension de la santé aux scores d'utilité de santé demeure stable lorsque l'importance explicite est intégrée aux données collectées dans un DCETTO. Toutes les analyses ont été réalisées à partir de données recueillies auprès de personnes francophones âgées de 18 ans et plus résidant au Québec, Canada, en 2016 et 2018.

Résultats : L'approche DCEBWS apparaît comme une alternative cohérente et comparable à celle du SG. L'approche DCEDOD augmente les valeurs d'utilité et produit un intervalle de valeurs plus restreint que l'approche DCEBWS, tandis que l'approche DCEPOD diminue les valeurs d'utilité et génère un ensemble de valeurs plus large que l'approche DCEBWS. L'approche MBDC est réalisable et permet de générer des données d'utilité représentatives des préférences des répondants. Elle offre ainsi le potentiel de développer des ensembles de valeurs d'utilité complets pour les instruments basés sur les préférences. L'ensemble de référence de valeurs d'utilité du SF-6Dv2 retenu pour le Québec est issu du DCEBWS et varie de -0,683 pour le pire état de santé (555655) à 1 pour une santé parfaite (111111). La dimension de la douleur présente ici les plus fortes diminutions de coefficients en termes d'ampleur, ce qui indique sa grande influence dans l'établissement des valeurs d'utilité des états de santé. Par ailleurs, les résultats montrent que l'importance explicite attribuée aux six dimensions du SF-6Dv2 par les participants ne modifie pas significativement l'importance relative des dimensions dans les estimations.

Recommandations : Les recherches futures devraient se concentrer sur plusieurs domaines clés afin d'améliorer l'élicitation des utilités en santé. Premièrement, l'approche DCEBWS mérite d'être explorée plus avant en tant qu'approche de rechange cohérente et comparable à celle du SG. Deuxièmement, l'effet de l'intégration de la mort immédiate comme attribut dans les DCE, notamment en comparant des méthodes comme DCEDOD et DCEPOD, nécessite une étude plus approfondie afin de comprendre comment les différentes présentations de la mort immédiate

modifient les valeurs d'utilité et les perceptions des participants. Troisièmement, l'approche MBDC est prometteuse pour générer des données fiables en matière de préférences et devrait être affinée et testée pour diverses pathologies, en la comparant à des approches traditionnelles comme le SG et le DCE. Quatrièmement, les ensembles de valeurs d'utilité du SF-6Dv2 générées pour le Québec permettront de mieux évaluer les interventions de santé dans un cadre d'évaluation économique. Enfin, l'importance explicite attribuée aux dimensions du SF-6Dv2 par les participants n'est pas suffisamment significative en termes statistiques et d'amplitude pour être incluse de façon automatique dans l'estimation des ensembles de valeurs d'utilité en santé.

Liste des acronymes

ACU : analyse coût-utilité
AHP : analytic hierarchy process (processus de hiérarchisation analytique)
AVAQ : année de vie ajustée par la qualité
BWS_{TTO} : best-worst scaling with TTO (échelle de choix meilleur-pire avec durée)
DCE : discrete choice experiment (choix expérimentaux discrets)
DCE_{BWS} : DCE followed by a best-worst scaling (DCE avec échelle de choix meilleur-pire)
DCE_{DOD} : DCE_{TTO} followed by the dominated option and immediate dead (DCE avec mort immédiate comme choix)
DCE_{POD} : DCE with probability of dying (DCE avec mort immédiate comme attribut)
DCE_{TTO} : DCE with time trade-off (DCE avec durée)
EQ-5D : EuroQol 5-dimension
EUT : expected utility theory (théorie de l'utilité espérée)
GLS : generalized least squares (moindres carrés généralisés)
HUI : health utilities index
MBDC : multiple bounded discrete choice (choix dichotomique multiple borné)
MCO : moindres carrés ordinaires
QALY : quality-adjusted life-year
QVRS : qualité de vie reliée à la santé
RUT : random utility theory (théorie de l'utilité aléatoire)
SF-6D : short-form 6-dimension
SF-6Dv2 : short-form 6-dimension version 2
SG : standard gamble (pari ordinaire)
TTO : time trade-off (arbitrage temporel)
VAS : visual analogue scale (échelle visuelle analogue)
WTD : worse than dead (pire que la mort)

Liste des tableaux

Tableau 1. Structure du SF-6Dv2 (dimensions et niveaux)
Tableau 2. Comparaison des modèles de SG, DCE_{TTO}, BWS_{TTO} et DCE_{BWS}
Tableau 3. Comparaison des modèles de DCE_{TTO}, DCE_{BWS} et DCE_{POD}
Tableau 4. Comparaison des modèles de DCE_{TTO}, DCE_{BWS} et DCE_{DOD}
Tableau 5. Comparaison des modèles de SG et MBDC
Tableau 6. Estimation du modèle de référence pour les utilités issues du SF-6Dv2
Tableau 7. Comparaison indirecte des sept méthodes d'élicitation

Liste des annexes

Annexe 1. Exemple de carte d'un choix expérimental discret avec durée (DCE_{TTO})
Annexe 2. Exemple de carte de choix pour l'échelle de choix meilleur-pire avec durée (BWS_{TTO})
Annexe 3. Exemple de carte d'un choix expérimental discret avec mort immédiate comme attribut (DCE_{POD})
Annexe 4. Exemple de carte d'un choix expérimental discret avec la mort immédiate comme choix (DCE_{DOD})
Annexe 5. Exemple de carte de choix avec le pari ordinaire (SG)
Annexe 6. Exemple de carte de choix dichotomique multiple borné (MBDC)

1. Introduction

Les années de vie ajustées par la qualité (AVAQ ou QALY pour *quality-adjusted life-year*) sont une mesure essentielle utilisée dans l'analyse coût-utilité (ACU). Basées sur les préférences des individus, les AVAQ permettent de comparer et d'évaluer efficacement diverses interventions de santé en fonction non seulement de l'espérance de vie, mais aussi de la qualité de vie reliée à la santé (QVRS) de ces années. L'utilisation des AVAQ est fréquemment recommandée comme unité de mesure des résultats en matière de santé dans les lignes directrices d'évaluation économique de pays comme le Canada (1, 2). Une AVAQ est calculée en multipliant la durée passée dans un état de santé particulier par la QVRS de cet état, à laquelle est attribuée une valeur comprise entre 0 (représentant la mort) et 1 (représentant une santé parfaite). Le calcul des AVAQ permet de déterminer les préférences pour différents états de santé, qui peuvent être mesurées par des méthodes directes ou indirectes (1). La méthode directe comprend des approches telles que le pari ordinaire (*standard gamble*, SG), l'arbitrage temporel (*time trade-off*, TTO), l'échelle visuelle analogue (*visual analogue scale*, VAS) et le choix expérimental discret (*discrete choice experiment*, DCE). Les méthodes indirectes correspondent à l'utilisation d'instruments multidimensionnels basés sur les préférences tels que l'*EuroQol 5-dimension* (EQ-5D), le *short-form 6-dimension* (SF-6D) et le *health utilities index* (HUI) (3).

Les méthodes d'élicitation directe des préférences, bien qu'elles fournissent des données détaillées et individualisées, peuvent être très gourmandes en ressources et en temps. Pour relever ces défis, des instruments basés sur les préférences indirectes ont été développés. Au lieu de mener des exercices détaillés d'élicitation des préférences pour chaque étude, ces instruments utilisent des questionnaires demandant aux répondants de décrire leur QVRS selon plusieurs dimensions générales (par exemple, mobilité, douleur, anxiété, fonctionnement physique, vitalité, santé mentale) avec différents niveaux de sévérité. Les réponses à ces questionnaires sont ensuite converties en valeurs d'utilité à l'aide d'un « ensemble de valeurs » dérivé de données de population à grande échelle et utilisant des méthodes courantes d'élicitation directe des préférences, notamment les SG, le TTO et le DCE (4-6).

L'approche SG s'appuie sur la théorie de l'utilité espérée (*expected utility theory*, EUT), qui suppose que les individus évaluent les bénéfices et les risques associés à chaque option et choisissent celle qui, selon eux, leur apportera le plus grand avantage (7). Compte tenu de l'incertitude inhérente associée à la prise de décision dans le domaine clinique, l'approche SG est largement considérée comme la méthode de référence pour la mesure des utilités en santé. Cependant, elle est amplement affectée par des contraintes cognitives compte tenu de sa complexité. De plus, des facteurs comme l'aversion au risque peuvent influencer les résultats du SG, contribuant à des phénomènes comme l'effet plafond (4, 8). Cet effet fait référence à la tendance des répondants à attribuer un score d'utilité plus élevé avec le SG qu'avec d'autres approches (9). Pour ces raisons, l'adoption d'approches ordinales alternatives, en particulier le DCE, suscite un intérêt croissant (10). Le DCE offre un moyen plus accessible de saisir les préférences en santé, ce qui rend cette approche particulièrement intéressante pour des populations et des contextes décisionnels divers, notamment les enquêtes en ligne. Cette approche est ancrée dans la théorie de l'utilité aléatoire (*random utility theory*, RUT), qui suppose que les individus choisissent l'option offrant l'utilité globale la plus élevée en tenant compte à la fois des composantes systématiques et aléatoires de l'utilité (11).

Au cours de la dernière décennie, la plupart des études ont utilisé le DCE standard pour générer des ensembles de valeurs d'utilité pour les instruments de mesure de l'AVAQ, avec ou sans considération de durée (DCE_{TTO}) (12). Dans cette approche, les participants sélectionnent un état de santé ayant leur préférence parmi un ensemble d'au moins deux états de santé. L'intégration de la durée comme une caractéristique associée à l'état de santé permet aux chercheurs d'explorer la façon dont les individus valorisent la survie supplémentaire dans leur processus décisionnel. L'inclusion de la durée fournit également des informations sur la préférence relative pour un état de santé par rapport à un état de mort, permettant un ancrage dans une échelle de 0 à 1 pour l'analyse (13). La tâche de comparaison par paires, bien que moins exigeante sur le plan cognitif que le SG, peut toutefois amener les participants à accorder moins d'attention aux options offertes (14). Par ailleurs, le concept d'état de mort dans une enquête n'est pas explicitement présenté aux participants, mais est déduit comme représentant un état de santé de durée nulle.

L'échelle de choix meilleur-pire avec durée (*best-worst-scaling*, BWS_{TTO}) est une approche relativement récente qui aborde plus directement la question du décès et fournit davantage d'informations, allant au-delà de ce que propose le DCE standard. Elle présente aux répondants plusieurs options pour indiquer la meilleure et la pire option, offrant ainsi une compréhension plus complète des préférences et des processus décisionnels (15). Dans cette approche, les répondants se voient généralement présenter deux états de santé ainsi qu'un troisième scénario décrivant la « mort immédiate », puis sont invités à indiquer la meilleure et la pire option parmi les états de santé envisagés (14, 16).

Cependant, cette approche pose un problème important : les participants expriment souvent une forte aversion au risque de décès, car ils comparent explicitement l'état de mort à d'autres états de santé, ce qui les conduit à choisir ce qu'ils auraient indirectement considéré comme des états de santé pires que la mort (*worse than dead*, WTD) dans d'autres modèles. Ce choix réduit ainsi l'intervalle des valeurs d'utilité, car moins d'états de santé sont considérés comme pires que la mort (17, 18). Cette réduction d'intervalle a des implications importantes pour la prise de décision en matière de santé et la valeur relative des interventions ayant une incidence sur la mortalité et la morbidité. Dans une telle situation, les interventions améliorant la QVRS peuvent être relativement moins valorisées qu'une approche produisant un éventail de valeurs plus large. Pour relever ce défi et mettre à profit les avantages potentiels du BWS_{TTO}, certains chercheurs ont utilisé la combinaison du DCE_{TTO} et du BWS_{TTO} (ci-après dénommée DCE_{BWS}) pour estimer des ensembles de valeurs pour des instruments tels que le SF-6Dv2 (19) et l'EQ-5D-5L (20). Leurs résultats ont montré que l'intervalle de valeurs obtenu était plus étroit que celui obtenu avec le DCE_{TTO} seul, mais plus large qu'avec BWS_{TTO} seul.

Compte tenu de ces différences méthodologiques, nous avons effectué une comparaison empirique de cinq approches utilisées pour le DCE (DCE_{TTO}, BWS_{TTO}, DCE_{BWS}, DCE_{DOD} et DCE_{POD}) avec le SG afin de déterminer l'approche la plus appropriée pour estimer un ensemble de valeurs d'utilité du SF-6Dv2 au Québec. Le DCE_{TTO} avec une probabilité de mourir immédiatement comme attribut (DCE_{POD}) et le DCE_{TTO} suivi de l'option dominée avec l'option de décès immédiat (DCE_{DOD}) sont deux approches utilisées dans cette étude pour intégrer l'état de mort dans les DCE. Nous avons également introduit et comparé une nouvelle technique d'éllicitation, soit le choix dichotomique multiple borné (MBDC), à l'approche SG. Un ensemble de valeurs d'utilité de référence a ensuite été estimé pour le SF-6Dv2 au Québec. Finalement, l'importance explicite attribuée par les participants à chaque dimension du SF-6Dv2 a été évaluée. Le SF-6Dv2 est la version la plus récente du SF-6D, dont le système descriptif définit la santé en six dimensions (tableau 1) : fonctionnement physique (*physical functioning*, PF), limitations des activités (*role limitations*, RL), fonctionnement social (*social functioning*, SF), douleur (*pain*, PA), santé mentale (*mental health*, MH) et vitalité (*vitality*, VT) (19, 21).

Tableau 1. Structure du SF-6Dv2 (dimensions et niveaux)

Dimensions	Niveaux
<i>Fonctionnement physique (PF)</i>	
PF1	Pas du tout limité dans les activités exigeant un effort physique important
PF2	Un peu limité dans les activités exigeant un effort physique important
PF3	Un peu limité dans les activités exigeant un effort physique modéré
PF4	Très limité dans les activités exigeant un effort physique modéré
PF5	Très limité pour prendre un bain ou m'habiller
<i>Limitations des activités (RL) : combinaison du rôle physique et moral</i>	
RL1	Jamais accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes
RL2	Rarement accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes
RL3	Parfois accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes
RL4	La plupart du temps accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes
RL5	Tout le temps accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes
<i>Fonctionnement social (SF)</i>	
SF1	Activités sociales jamais limitées
SF2	Activités sociales rarement limitées
SF3	Activités sociales parfois limitées
SF4	Activités sociales la plupart du temps limitées
SF5	Activités sociales tout le temps limitées
<i>Douleur (PAIN)</i>	
PAIN1	Aucune douleur
PAIN2	Douleurs très légères
PAIN3	Douleurs légères
PAIN4	Douleurs moyennes
PAIN5	Douleurs intenses
PAIN6	Douleurs très intenses
<i>Santé mentale (MH)</i>	
MH1	Jamais très nerveux ou triste et démoralisé
MH2	Rarement très nerveux ou triste et démoralisé
MH3	Parfois très nerveux ou triste et démoralisé
MH4	La plupart du temps très nerveux ou triste et démoralisé
MH5	Tout le temps très nerveux ou triste et démoralisé
<i>Vitalité (VIT)</i>	
VIT1	Jamais épuisé et vidé
VIT2	Rarement épuisé et vidé
VIT3	Parfois épuisé et vidé
VIT4	La plupart du temps épuisé et vidé
VIT5	Tout le temps épuisé et vidé

Dans le DCE_{POD}, l'attribut de probabilité de décès immédiat a été ajouté au DCE_{TTO}. Cette probabilité a été intégrée à chaque option d'état de santé du DCE_{TTO} en fonction du choix antérieur du participant. Lorsque le participant préfère un profil de santé à un autre, une probabilité de décès plus élevée est attribuée à l'option préférée dans le DCE_{POD} qui suit

le DCE_{TTO}. L'étude de DCE_{POD} inclut des durées de survie courantes de 1, 4, 7 et 10 ans, ainsi que des probabilités de décès immédiat fixées à 10, 20, 30, 40 et 50 % (12).

Dans le DCE_{DOD}, chaque participant se voit proposer des choix dichotomiques. Comme dans le DCE_{POD}, l'exercice se déroule en deux étapes (22). Il débute par un DCE_{TTO} où chaque état de santé est associé à une durée. Les participants doivent choisir entre les deux profils de santé hypothétiques. Après avoir fait leur choix, le DCE_{DOD} leur est présenté et on leur demande de choisir entre l'option dominée (c.-à-d. l'une des options du choix initial) et la mort immédiate. Par exemple, si un participant a initialement choisi l'option A plutôt que l'option B (c.-à-d. l'option B est dominée), il doit ensuite exprimer une préférence entre B et la mort immédiate.

Le SG a été mené selon l'approche recommandée par Brazier *et al.* (4, 23), également connue sous le nom de protocole de Sheffield. Par ailleurs, nous avons introduit une nouvelle approche pour éliciter les utilités en santé, soit le MBDC, avec l'hypothèse qu'elle atténue les difficultés cognitives associées au SG (24). Bien que couramment utilisée en économie de l'environnement et en recherche sur les politiques publiques, l'approche MBDC n'a pas encore été appliquée aux études de préférences en santé. L'approche MBDC améliore les formats de choix dichotomiques traditionnels en fournissant un format de question « à bornes multiples » avec des options de réponse incertaines. Elle propose aux répondants une fourchette de montants d'enchères (c.-à-d. probabilité de succès) et leur demande d'indiquer leur acceptation ou leur rejet pour chaque enchère, fournissant ainsi des informations plus riches sur leurs préférences. En présentant aux répondants des choix contraints dans des limites spécifiées, le MBDC atténue certains des défis cognitifs associés aux méthodes traditionnelles telles que le SG (4, 25, 26). Dans le cadre du MBDC, les participants ont deux choix : le « choix A », un état de santé intermédiaire décrit par le SF-6Dv2, et le « choix B », un résultat incertain avec deux états de santé possibles (24). Le « choix B » implique une santé parfaite avec une probabilité P et le pire état de santé avec une probabilité 1-P. Les probabilités varient de 0 à 100 %, et les participants doivent choisir entre : « absolument oui », « probablement oui », « pas sûr », « probablement non » et « absolument non ». Les probabilités sont ensuite systématiquement variées, incluant des combinaisons telles que 100/0 et 0/100. Par exemple, les participants comparent un état de santé certain, mais imparfait (choix A) à un traitement offrant une chance de santé parfaite ou une détérioration immédiate à l'état le plus défavorable (choix B). En fonction de la probabilité de succès du traitement, les participants expriment leur préférence. Les réponses sont ensuite utilisées pour « enchaîner » les valeurs d'états de santé, les plaçant ainsi sur une échelle allant de 0, représentant la mort, à 1, représentant la pleine santé. Les valeurs sont ajustées à l'aide de la formule : $MBDC_{adj} = MBDC + (1 - MBDC) * P$, où $MBDC_{adj}$ est la valeur ajustée, MBDC est la valeur de l'état de santé intermédiaire et P est la valeur fournie par les participants pour le pire état de santé. Le point médian entre « probablement oui » et « probablement non » a été utilisé pour estimer les valeurs lorsque les répondants n'indiquent pas de point d'indifférence clair (23).

Finalement, l'influence de l'importance explicite accordée par les participants aux six dimensions du SF-6Dv2 est évaluée en incluant leur pondération. La pondération de chaque dimension est ici attribuée à l'aide du processus de hiérarchisation analytique (*analytic hierarchy process*, AHP) (27).

2. Collecte et analyse des données

Les états de santé pour les différentes approches ont été sélectionnés à l'aide d'une procédure orthogonale dans le but de produire un plan avec une faible erreur D. Le plan a été construit sous forme de paires basées sur des matrices quasi orthogonales à l'aide du logiciel SAS. Lors des deux enquêtes, les participants ont été sélectionnés aléatoirement pour être affectés à l'une ou l'autre des approches utilisées, soient le SG, le MBDC ou divers types de DCE. Dans le DCE_{TTO}, dix paires de choix (c.-à-d. choix A ou B) séquentielles ont été présentées aux participants. Dans le BWS_{TTO}, ceux-ci se voyaient présenter trois triplets de choix (c.-à-d. choix A, B ou mort immédiate). Dans le DCE_{BWS}, 13 ensembles de choix ont été attribués aléatoirement aux participants (c.-à-d. dix premiers ensembles de choix sous forme de DCE_{TTO}, suivis de trois ensembles de choix sous forme de BWS_{TTO}). Dans le DCE_{POD}, les participants ont effectué dix exercices de DCE_{TTO}, chacun suivi d'un exercice de DCE_{POD}. Dans le DCE_{DOD}, les participants ont effectué dix exercices de DCE_{TTO}, chacun suivi d'un choix répété entre l'option dominée (c.-à-d. A ou B) et le décès immédiat. Dans le cadre du SG, les participants ont réalisé six exercices : cinq portant sur des états de santé intermédiaires et un portant sur le pire état de santé possible ou la mort immédiate (selon que leur préférence était pour le pire état de santé ou la mort immédiate dans une question précédente). Les participants au MBDC ont été invités à évaluer dix états de santé intermédiaires, puis à effectuer une onzième évaluation visant à comparer le pire état de santé (défini par 555655 dans le SF-6Dv2) à la mort immédiate. Si le pire état était préféré à la mort immédiate, il devenait le choix A; s'il était jugé pire que la mort immédiate,

le choix A était défini comme la mort immédiate. Les annexes 1 à 6 fournissent des exemples de cartes de choix utilisées pour chacune des méthodes présentées ici.

Les données ont été recueillies à l'aide d'une technique d'échantillonnage par quotas auprès de personnes francophones âgées de 18 ans et plus résidant au Québec, Canada, en 2016 et 2018. Sur les 5 028 personnes invitées, 1 979 ont été réparties aléatoirement pour effectuer des DCE, 1 468 pour le SG et 1 176 pour le MBDC. Pour modéliser les données collectées avec SG et MBDC, des modèles de moindres carrés ordinaires (MCO), de moindres carrés généralisés (*generalized least squares*, GLS), GLS Tobit et Tobit hétéroscédastique ont été utilisés. Pour les DCE, un modèle logit conditionnel développé par Bansback *et al.* a été utilisé (13). La comparaison des approches a été évaluée sur les plans de l'acceptabilité, des niveaux de significativité, de la magnitude des coefficients associés aux niveaux des dimensions, du pourcentage d'états de santé considérés comme pires que la mort et de l'intervalle de valeurs estimées. L'acceptabilité a été évaluée en comparant les taux de complétion des exercices, les temps de réalisation de ceux-ci, la difficulté autodéclarée à répondre aux exercices de choix et la qualité autodéclarée des réponses fournies. La magnitude des coefficients des niveaux des dimensions de chaque modèle a été évaluée en fonction de l'ampleur de leurs plus grands décrets (c.-à-d. les niveaux les plus sévères de chaque dimension). Cet indicateur a été utilisé pour classer les dimensions selon leur importance, particulièrement en comparant la taille du plus fort coefficient de chaque dimension. Concernant le pourcentage d'états de santé considérés comme pires que la mort, il s'agit de la proportion des 18 750 états de santé possibles décrits par les dimensions et niveaux du SF-6Dv2 (tableau 1) qui prennent une valeur négative à la suite des estimations économétriques effectuées. Finalement, l'intervalle de valeurs estimées correspond respectivement aux scores attribués à l'état de santé le plus grave et au meilleur état de santé. Pour l'outil utilisé, le SF-6Dv2, cela décrit, d'une part, le pire état de santé, soit celui décrit par les niveaux les plus sévères au tableau 1 (c.-à-d. 555655), et, d'autre part, une santé parfaite, correspondant aux meilleurs niveaux (c.-à-d. 111111). Considérant que nous avons utilisé des modèles ancrés sur l'échelle d'AVAQ où 0 correspond à la mort et 1 à une santé parfaite (1), la valeur d'intervalle supérieure sera ici toujours égale à 1.

3. Résultats

L'ensemble de données comprend 714 répondants pour les exercices de SG; 585 pour ceux du MBDC; 1 153 pour ceux du DCE_{TTO}, du BWS_{TTO} et du DCE_{BWS}; 348 pour ceux du DCE_{DOD}; et 369 pour ceux du DCE_{POD}. Le tableau 2 indique un intervalle de valeurs d'utilité plus étroit pour l'approche SG que pour le DCE_{TTO}, le BWS_{TTO} et le DCE_{BWS} (28). Le BWS_{TTO} produit également moins d'états de santé pires que la mort que le DCE_{TTO} et le DCE_{BWS}. À l'inverse, dans l'approche SG, seul 0,08 % des états de santé ont été estimés pires que la mort. Parmi ces quatre approches, la dimension de la douleur dans le SF-6Dv2 est apparue systématiquement comme ayant le plus fort effet sur le niveau d'utilité associé aux états de santé, tandis que la dimension de la vitalité a eu le plus faible effet dans trois approches (SG, BWS_{TTO} et DCE_{BWS}) (28). Les résultats montrent un classement similaire des valeurs pour les états de santé autorapportés les plus fréquents dans le DCE_{TTO}, le BWS_{TTO} et le DCE_{BWS}, ce qui est différent pour les valeurs rapportées avec l'approche SG (tableau 2).

Tableau 2. Comparaison des modèles de SG, DCE_{TTO}, BWS_{TTO} et DCE_{BWS}

	SG	DCE _{TTO}	BWS _{TTO}	DCE _{BWS}
Intervalle d'utilité	(-0,143 à 1)	(-1,063 à 1)	(-0,505 à 1)	(-0,637 à 1)
% d'états pires que la mort	0,08 %	36,11 %	5,63 %	11,61 %
Ordre des dimensions	PA-PF-SF-RL-MH -VT	PA-VT-RL-PF-SF-MH	PA-PF-MH-RL-SF-VT	PA-PF-MH-SF-RL-VT
Décrément le plus fort (coefficient)	Niveau 6 de PA (-0,221)	Niveau 6 de PA (-0,435)	Niveau 6 de PA (-0,356)	Niveau 6 de PA (-0,368)
Valeur d'utilité des cinq états de santé autorapportés les plus fréquents (%)				
111133	0,936 (0,45)	0,675 (0,92)	0,933 (0,46)	0,822 (0,70)
131334	0,832 (1,22)	0,338 (0,85)	0,637 (0,66)	0,527 (0,69)
213411	0,791 (0,54)	0,502 (0,84)	0,733 (0,61)	0,664 (0,74)
213232	0,761 (0,66)	0,421 (0,75)	0,678 (0,58)	0,588 (0,73)
222231	0,886 (1,01)	0,399 (0,90)	0,662 (0,53)	0,541 (0,72)
Acceptabilité des approches				
	SG	DCE _{TTO}	BWS _{TTO}	
Durée de complétion (en secondes) (DS)	109,49 (90,17)	42,90 (43,97)	33,73 (41,51)	-
Niveau de difficulté autorapporté à répondre (n et %)				
Très facile	48 (6,63)	72 (6,24)	-	Valeur de P 0,156
Facile	203 (28,04)	307 (26,63)	-	
Neutre	168 (23,20)	290 (25,15)	-	

Difficile	260 (35,91)	439 (38,07)	-	
Très difficile	45 (6,22)	45 (3,90)	-	
Niveau de qualité autorapporté des réponses (n et %)				
Très bon	161 (22,24)	239 (20,73)	-	Valeur de P
Bon	401 (55,39)	662 (57,42)	-	0,706
Moyen	155 (21,41)	242 (20,99)	-	
Mauvais	3 (0,41)	7 (0,61)	-	
Très mauvais	4 (0,55)	3 (0,26)	-	

DS, déviation standard.

Les critères d'acceptabilité des approches indiquent une durée moyenne passée à effectuer le SG de 109,49 secondes, ce qui est statistiquement significativement plus long que les durées observées dans les approches DCE_{TTO} (42,90 secondes) et BWS_{TTO} (33,73 secondes). Aucune différence significative n'a été observée dans la difficulté à répondre et la qualité des réponses autorapportées entre les approches SG et DCE. On note également que les répondants, s'ils trouvent les exercices de choix relativement difficiles à réaliser, ont cependant jugé que la qualité de leurs réponses était le plus souvent bonne.

L'approche DCE_{POD} a été comparée au DCE_{TTO} et au DCE_{BWS} pour les mêmes participants (tableau 3). L'intervalle de valeurs d'utilité est plus large (-2,063 à 1) avec DCE_{POD} par rapport aux approches DCE_{TTO} et DCE_{BWS}. De plus, les états de santé pires que la mort, soit ceux présentant des valeurs négatives, ont été observés plus fréquemment dans DCE_{POD} que dans DCE_{TTO} et DCE_{BWS}. Le niveau 5 de la dimension de santé mentale au SF-6Dv2 présente la réduction la plus significative sur la valeur d'utilité pour tous les modèles. Les quatre états de santé autorapportés par les participants présentant la somme minimale du code décrivant les états de santé (c.-à-d. les meilleurs états de santé) ont dans chaque approche des valeurs positives, tandis que les pires états autorapportés ont systématiquement des valeurs négatives.

Tableau 3. Comparaison des modèles de DCE_{TTO}, DCE_{BWS} et DCE_{POD}

	DCE _{TTO}	DCE _{BWS}	DCE _{POD}
Intervalle d'utilité	(-1,860 à 1)	(-1,473 à 1)	(-2,335 à 1)
% d'états pires que la mort	58,03	56,77	76,53
Ordre des dimensions	MH-PA-PF-VT-RL-SF	MH-PA-VT-PF-RL-SF	MH-RL-VT-PA-PF-SF
Décrément le plus fort (coefficient)	Niveau 5 de MH (-0,903)	Niveau 5 de MH (-0,806)	Niveau 5 de MH (-0,946)
Valeurs d'utilité des quatre meilleurs états de santé autorapportés			
111133	0,509	0,567	0,401
111121	0,933	0,852	0,680
112121	0,802	0,793	0,571
331131	0,475	0,543	0,117
Valeurs d'utilité des quatre pires états de santé autorapportés			
523654	-1,631	-1,211	-1,980
345654	-1,350	-0,931	-1,715
535645	-1,150	-0,975	-1,545
554542	-0,859	-0,568	-1,304

Concernant l'approche DCE_{DOD}, comparée au même ensemble de participants, son intervalle de valeurs d'utilité est plus étroit que pour le DCE_{TTO} et le DCE_{BWS} (tableau 4). DCE_{DOD} produit également un pourcentage plus faible d'états de santé pires que la mort par rapport au DCE_{TTO} et au DCE_{BWS}. La douleur apparaît systématiquement comme la dimension ayant le plus d'effet dans les trois approches, suivie par PF, MH, RL, VT et SF (22). Le niveau 6 de PA présente ainsi la réduction la plus significative de la valeur d'utilité dans tous les modèles. Les résultats montrent un schéma identique de classement des valeurs d'utilité pour les états de santé les plus fréquemment autorapportés dans les approches DCE_{TTO}, DCE_{BWS} et DCE_{DOD}. Ce schéma cohérent indique que les valeurs estimées pour les états 111133 et 222231 étaient respectivement les plus élevées et les plus basses parmi les cinq principaux états de santé autorapportés.

Tableau 4. Comparaison des modèles de DCE_{TTO}, DCE_{BWS} et DCE_{DOD}

	DCE _{TTO}	DCE _{BWS}	DCE _{DOD}
Intervalle d'utilité	(-1,578 à 1)	(-1,150 à 1)	(-0,921 à 1)
% d'états pires que la mort	47,19	33,73	20,01
Ordre des dimensions	PA-PF-MH-RL-VT-SF	PA-PF-MH-RL-VT-SF	PA-PF-MH-RL-VT-SF

Décrément le plus fort (coefficient)	Niveau 6 de PA (-0,587)	Niveau 6 de PA (-0,502)	Niveau 6 de PA (-0,540)
Valeur d'utilité des cinq états de santé autorapportés les plus fréquents (%)			
111133	0,791 (0,78)	0,796 (0,67)	0,910 (0,42)
131334	0,701 (0,80)	0,631 (0,65)	0,750 (0,49)
213411	0,738 (0,89)	0,713 (0,76)	0,793 (0,56)
213232	0,545 (1,01)	0,539 (0,73)	0,681 (0,57)
222231	0,285 (0,76)	0,401 (0,67)	0,504 (0,53)

Le tableau 5 présente les résultats des estimations pour le MBDC par rapport au SG. Aucune différence significative n'est observée dans les difficultés autorapportées et la qualité des réponses entre les approches, bien que le MBDC compte moins de participants exclus et que ceux-ci soient moins enclins à signaler des difficultés à répondre (24). De plus, le temps de réponse dans le groupe MBDC est significativement plus court (68,1 contre 99,8 secondes). L'intervalle de valeurs d'utilité généré par le MBDC est plus large que celui généré par le SG, et le nombre d'états de santé pires que la mort dans le MBDC (0,91 %) est plus élevé que pour le SG (0,08 %). Dans les deux cas, le nombre d'états de santé pires que la mort est bien moindre que dans les approches DCE. La dimension de la douleur au SF-6Dv2 est identifiée comme la plus importante, tandis que la dimension de la vitalité possède l'effet le plus faible. Les quatre états de santé décrits par le SF-6Dv2 étant les plus fréquemment autorapportés comme les meilleurs comprennent uniquement les niveaux 1, 2, 3 ou 4. Malgré des variations dans leurs valeurs estimées, tous ces états ont des valeurs de 0,75 ou plus. À l'inverse, certains des états de santé autorapportés comme étant les pires présentent des valeurs d'utilité négatives.

Tableau 5. Comparaison des modèles de SG et MBDC

		SG	MBDC
Intervalle d'utilité		(-0,137 à 1)	(-0,372 à 1)
% d'états pires que la mort		0,08 %	0,91 %
Ordre des dimensions		PA-PF-SF-RL-MH -VT	PA-SF-PF-RL-MH-VT
Décrément le plus fort (coefficient)		Niveau 6 de PA (-0,252)	Niveau 6 de PA (-0,266)
Valeurs d'utilité des quatre meilleurs états de santé autorapportés			
111133		0,926 (0,45)	0,948 (0,92)
131334		0,808 (1,22)	0,801 (0,85)
213232		0,754 (0,66)	0,827 (0,75)
222231		0,871 (1,01)	0,754 (0,90)
Valeurs d'utilité des quatre pires états de santé autorapportés			
535645		0,147 (0,61)	-0,049 (1,23)
535654		0,054 (0,32)	-0,070 (0,42)
554542		0,230 (0,45)	0,145 (1,40)
523654		0,193 (0,87)	0,068 (0,13)
Acceptabilité des approches			
	SG	MBDC	Valeur de P
Durée de complétion (en secondes) (DS)	99,80 (88,77)	68,12 (46,98)	<0,003
Niveau de difficulté autorapporté à répondre (n et %)			
Très facile	47 (6,58)	37 (6,32)	Valeur de P 0,255
Facile	200 (35,99)	191 (32,65)	
Neutre	166 (23,25)	127 (21,71)	
Difficile	257(28,01)	206 (35,21)	
Très difficile	44 (6,16)	24 (4,10)	
Niveau de qualité autorapporté des réponses (n et %)			
Très bon	161 (22,54)	115 (19,66)	Valeur de P 0,205
Bon	394 (55,18)	334 (57,09)	
Moyen	152 (21,29)	131 (22,39)	
Mauvais	3 (0,42)	5 (0,85)	
Très mauvais	4 (0,56)	-	

DS, déviation standard.

Le tableau 6 présente les résultats de la combinaison de dix DCE_{TTO} et de trois BWS_{TTO} pour 1 153 participants, correspondant à l'approche DCE_{BWS} (29). Il s'agit d'un modèle parcimonieux ancré utilisant un logit conditionnel qui permet de calculer les valeurs d'utilité du SF-6Dv2 au Québec. Ainsi, les coefficients générés par le modèle peuvent servir à dériver un ensemble de valeurs d'utilité pour le SF-6Dv2. Cet ensemble de valeurs permet aux chercheurs et aux décideurs politiques de calculer la valeur d'utilité de chaque état de santé (Hi) :

$$U(H_i) = 1 - (0,103PF_2 + 0,152PF_3 + 0,225PF_4 + 0,338PF_5 + 0,057RL_2 + 0,079RL_3 + 0,079RL_4 + 0,228RL_5 + 0,036SF_2 + 0,070SF_3 + 0,134SF_4 + 0,246SF_5 + 0,097PA_2 + 0,137PA_3 + 0,173PA_4 + 0,253PA_5 + 0,417PA_6 + 0,033MH_2 + 0,112MH_3 + 0,143MH_4 + 0,253MH_5 + 0,033VT_2 + 0,064VT_3 + 0,140VT_4 + 0,201VT_5)$$

Par exemple, la valeur d'utilité de l'état de santé 223221 (c.-à-d. je suis un peu limité lors d'activités exigeant un effort physique important, j'accomplis rarement moins de choses que je l'aurais voulu au travail ou dans mes autres activités quotidiennes, mes activités sociales sont parfois limitées, j'ai de très légères douleurs, je suis rarement triste, démoralisé ou très nerveux, je ne suis jamais épuisé et vidé) est calculée comme suit :

$$1 - (0,103 + 0,057 + 0,070 + 0,097 + 0,033 + 0) = 0,64$$

Tableau 6. Estimation du modèle de référence pour les utilités issues du SF-6Dv2

Niveaux	Modèle		
	β	ES	95 % IC
PF2	-0,103**	0,020	(-0,146 -0,065)
PF3	-0,152**	0,020	(-0,195 -0,114)
PF4	-0,225**	0,020	(-0,266 -0,185)
PF5	-0,338**	0,021	(-0,382 -0,298)
RL2	-0,057**	0,015	(-0,088 -0,028)
RL3	-0,079**	0,015	(-0,110 -0,050)
RL4	-0,079**	0,015	(-0,110 -0,050)
RL5	-0,228**	0,015	(-0,261 -0,201)
SF2	-0,036	0,019	(-0,077 -0,002)
SF3	-0,070**	0,019	(-0,109 -0,034)
SF4	-0,134**	0,018	(-0,171 -0,100)
SF5	-0,246**	0,020	(-0,287 -0,207)
PA2	-0,097**	0,018	(-0,136 -0,062)
PA3	-0,137**	0,020	(-0,179 -0,099)
PA4	-0,173**	0,019	(-0,213 -0,136)
PA5	-0,253**	0,024	(-0,302 -0,207)
PA6	-0,417**	0,023	(-0,465 -0,372)
MH2	-0,033*	0,016	(-0,070 -0,001)
MH3	-0,112**	0,021	(-0,156 -0,072)
MH4	-0,143**	0,018	(-0,179 -0,107)
MH5	-0,253**	0,017	(-0,289 -0,219)
VT2	-0,033*	0,016	(-0,068 -0,004)
VT3	-0,064**	0,017	(-0,100 -0,033)
VT4	-0,140**	0,017	(-0,177 -0,107)
VT5	-0,201**	0,017	(-0,236 -0,167)
U(555655)	-0,683		
WTD (%)	13,01		
#non significatif	1		

ES, erreur standard; IC, intervalle de confiance; PF, physical functioning; RL, role limitations; SF, social functioning; PA, pain; MH, mental health; VT, vitality; WTD, worse than dead.

L'ordre d'importance des dimensions du tableau 6 est : PA, VT, RL, PF, SF et MH. Parmi ces dimensions, la dimension de la douleur produit les plus fortes baisses dans le calcul des scores d'utilité. Cela suggère que la douleur produit l'effet négatif le plus significatif sur la QVRS par rapport aux autres dimensions. Cependant, la répartition des problèmes signalés par les participants au questionnaire SF-6Dv2 révèle que les problèmes de santé les plus souvent signalés réfèrent à la dimension de la vitalité (29). Ce résultat souligne l'importance de mettre en œuvre des stratégies ciblées visant les problèmes de vitalité et de fatigue dans les soins de santé, car la prise en compte des préoccupations liées à la vitalité pourrait entraîner des améliorations significatives sur la QVRS. De plus, la répartition des dimensions du questionnaire SF-6Dv2, effectuée par les participants dans un simple exercice de classement, montre que la vitalité occupe la première place. Cependant, la répartition explicite attribuée par les participants ne modifie pas de façon statistiquement significative l'ordre de grandeur des dimensions lorsqu'une pondération est incluse dans un processus de hiérarchisation analytique.

Le tableau 7 résume les principales différences entre les différentes approches testées. Si certaines comparaisons sont directes, comme entre le DCE_{TTO} et le DCE_{DOD} ou le DCE_{POD}, ou entre le DCE_{TTO} et le DCE_{BWS}, la plupart des comparaisons sont indirectes, car ce ne sont pas les mêmes répondants qui ont répondu aux différents exercices de choix. Toutefois, dans ce dernier cas, la comparaison reste valable dans la mesure où l'allocation du répondant dans l'une ou l'autre méthode a été faite au hasard (c.-à-d. allocation randomisée).

Tableau 7. Synthèse comparative des sept méthodes d'élicitation des états de santé décrits par le SF6Dv2

Méthode	Caractéristiques	Intervalle d'utilité	Proportion d'états pires que la mort
DCE _{TTO}	Un choix parmi deux états de santé décrits par SF-6Dv2 avec durée	(-1,860 à 1)	36,11 % - 58,03 %
BWS _{TTO}	Deux choix parmi trois états de santé décrits par SF-6Dv2 avec durée ou la mort	(-0,505 à 1)	5,63 %
DCE _{BWS}	DCE _{TTO} + BWS _{TTO}	(-1,473 à 1)	11,61 % - 56,77 %
DCE _{POD}	Un choix parmi deux états de santé décrits par SF-6Dv2 avec durée et probabilité de mourir	(-2,335 à 1)	76,53 %
DCE _{DOD}	Un choix parmi deux états de santé décrits par SF-6Dv2 avec durée ou la mort	(-0,921 à 1)	20,01 %
SG	Un choix entre un état de santé intermédiaire décrit par SF-6Dv2 et un traitement risqué	(-0,143 à 1)	0,08 %
MBDC	Indiquer le niveau de certitude dans son choix entre un état de santé intermédiaire et un traitement risqué pour différentes probabilités de succès ou d'échecs.	(-0,372 à 1)	0,91 %

Note : Pour DCE_{TTO}, DCE_{BWS} et SG, les intervalles d'utilité et de proportion d'états pires que la mort correspondent aux plus grands écarts parmi les résultats rapportés dans les tableaux précédents.

4. Recommandations

Cette étude avait deux principaux objectifs, le premier étant de comparer sept méthodes d'élicitation des préférences en santé pour calculer des AVAQ, le second étant d'évaluer la faisabilité du MBDC comme méthode de rechange au SG. Les principaux résultats indiquent que chacune des méthodes produit des résultats cohérents en termes de décrets (c.-à-d. les niveaux les plus sévères ont les effets négatifs les plus importants) et de significativité statistique. Toutefois, des différences notables sont observées dans les intervalles d'utilité, les proportions d'états de santé pires que la mort (c.-à-d. pourcentage d'états ayant une valeur négative parmi les 18 750 états possibles décrits par le SF-6Dv2), l'importance relative des dimensions et, dans certains cas, la facilité d'utilisation. Il en ressort que le DCE_{BWS} présente des valeurs d'utilité situées à mi-chemin des extrêmes trouvés avec le DCE_{TTO} et le BWS_{TTO} et surtout le SG. Le pari ordinaire (SG) étant la méthode de référence de par ses fondements théoriques, il apparaît qu'il serait plus approprié de trouver une méthode ne produisant pas des résultats trop différents en ce qui concerne la valeur d'utilité attribuée au pire état de santé et la proportion des états pires que la mort. Cependant, le SG étant critiqué pour produire des valeurs d'utilité trop élevées et une proportion d'états pires que la mort trop faible en raison de l'aversion au risque des répondants, il n'est pas non plus recommandé d'en être trop proche. La question de l'aversion au risque, en particulier au risque de mourir, apparaît ainsi centrale. C'est pourquoi nous avons testé deux autres façons de considérer la mort dans les méthodes d'élicitation : DCE_{POD} et DCE_{DOD}. L'ajout d'un attribut de probabilité de mourir dans la description des états de santé avec le DCE_{POD} réduit le problème d'aversion au risque, mais génère une augmentation manifeste des valeurs d'utilité négatives et accroît de manière irréaliste (76 %) la proportion d'états de santé pires que la mort. À l'inverse, le DCE_{DOD}, en suivant une procédure en deux étapes, où à la suite d'un choix en DCE_{TTO} il est demandé au participant de faire un choix entre la mort et l'option dominée précédemment, permet de générer des valeurs d'utilité plus élevées que le DCE_{POD} et une proportion plus faible d'états pires que la mort (20 %). Ne sachant pas quelle est la réelle proportion d'états de santé pires que la mort parmi les 18 750 décrits par le SF-6Dv2, il est difficile de déterminer par ce critère si le DCE_{DOD} est plus approprié que le DCE_{BWS} ou le BWS_{TTO}. D'autres critères, tels que le nombre de coefficients statistiquement significatifs, le nombre d'incohérences dans la progression de la sévérité des niveaux, ou encore la stabilité dans l'ordre d'importance des dimensions, sont alors à considérer. La méthode du DCE_{DOD} produisant davantage d'incohérences que le DCE_{BWS}, ce dernier est à privilégier, d'autant plus qu'il permet de combiner la simplicité du DCE_{TTO} avec l'ajout d'information que produit le BWS_{TTO}.

Mis à part l'utilisation d'une méthode ordinale (comme le DCE), une autre manière de pallier les difficultés rencontrées avec le SG (par exemple, forte aversion au risque et ses corollaires sur les valeurs élicitées, difficulté avec le processus itératif de la méthode, durée élevée de l'exercice de choix, taux d'abandon élevé) pourrait également être de simplifier l'utilisation du SG tout en conservant la puissance de ses assises théoriques. Pour cette raison, nous avons adapté le SG en utilisant le MBDC. Les résultats de notre étude indiquent que cette approche est faisable, plus rapide et d'un niveau de difficulté comparable au SG traditionnel. En plus de générer davantage de points d'observation pour un coût d'enquête similaire, le MBDC produit légèrement plus d'états de santé pires que la mort (0,91 % vs 0,08 %) et la valeur inférieure de son intervalle d'utilité est plus éloignée de zéro que celle obtenue par le SG traditionnel, tout en restant moins éloignée que celle du DCE_{BWS} ou du BWS_{TTO}.

Les recherches futures devraient se concentrer sur plusieurs domaines clés afin d'améliorer l'élicitation des utilités en santé. Premièrement, l'approche DCE_{BWS} mérite d'être explorée plus avant en tant que solution cohérente et comparable à l'approche du SG, afin de garantir sa robustesse et sa fiabilité. Notamment, bien qu'encore très utilisé dans le domaine, mais de plus en plus contesté, l'attribut de durée pourrait être retiré pour tester si les résultats se maintiennent, tout en utilisant différentes méthodes d'ancrage dans l'échelle de 0 à 1 utilisée pour les AVAQ. Notamment, l'attribut de durée est contesté pour entraîner la production de préférences non linéaires dans le temps, ce qui est contraire à la théorie utilisée. Cette modification permettrait notamment d'intégrer le DCE dans des approches hybrides d'estimations conjointes. Par ailleurs, les chercheurs devraient continuer à étudier l'incidence de l'intégration de la mort immédiate comme attribut ou choix dans les DCE, notamment en comparaison avec l'approche TTO, afin de mieux comprendre comment cette inclusion influence la prise de décision dans les études de préférences visant à évaluer des états de santé. Le potentiel de l'approche MBDC à générer des données fiables en matière de préférences justifie également un examen plus approfondi. Les études futures devraient viser non seulement à affiner cette approche, notamment par l'utilisation d'embranchements conditionnels pour les probabilités de succès du traitement, ce qui réduirait la durée de l'exercice, mais aussi à explorer son application à diverses pathologies et à comparer ses résultats à ceux issus d'approches plus traditionnelles comme le TTO et le DCE. L'ensemble de valeurs d'utilité pour le SF-6Dv2, développé à partir de la méthode DCE_{BWS}, est recommandé ici pour le calcul des AVAQ dans les analyses coût-utilité des interventions ou programmes de santé dans la mesure où il s'agit de l'approche apportant le meilleur équilibre dans la représentation des dimensions, la présence d'états de santé pires que la mort et la performance des estimations. Toutefois, le MBDC pourrait également être une bonne avenue. À l'inverse le DCE_{TTO} seul et le DCE_{POD} ne sont pas recommandés. Enfin, l'importance explicite attribuée aux dimensions du SF-6Dv2 par les participants n'est pas suffisamment significative d'un point de vue statistique pour être incluse dans l'estimation des ensembles de valeurs d'utilité en santé.

Finalement, les avancées méthodologiques rapportées dans ce document seront utiles pour aider à développer un indicateur servant à suivre les évolutions dans le temps (par exemple, à la suite d'une réforme) et l'espace (par exemple, inégalités géographiques) de l'état de santé de la population québécoise. À ce jour, la plupart des indicateurs de suivi sont unidimensionnels (par exemple, l'espérance de vie, le taux de mortalité infantile) et, parfois, ne mesurent qu'indirectement la santé de la population (par exemple, temps d'attente aux urgences, nombre de patients par médecin). La combinaison de la qualité de vie mesurée par les AVAQ, basée sur les préférences locales, et des tables de mortalité produites par les statistiques provinciales permettrait d'avoir une idée plus précise de l'état de santé réel de la population québécoise et de son évolution au gré des mutations vécues dans la société.

Références

1. Weinstein MC, Torrance G, McGuire A. QALYs: the basics. *Value in health*. 2009;12:S5-9.
2. CDA. Guidelines for the economic evaluation of health technologies: 4th edition. Canada's Drug Agency. 2017.
3. Green C, Brazier J, Deverill M. Valuing health-related quality of life: a review of health state valuation techniques. *Pharmacoeconomics*. 2000;17:151-65.
4. Brazier J, Ratcliffe J, Saloman J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation: Oxford university press; 2017.
5. Roy LM, Neill A, Swampy K, Auger J, Campbell SM, Chatwood S, Al Sayah F, Johnson JA. Preference-based measures of health-related quality of life in indigenous people: a systematic review. *Quality of life research*. 2024;33(2):317-33.
6. Weernink MG, Janus SI, van Til JA, Raisch DW, van Manen JG, IJzerman MJ. A systematic review to identify the use of preference elicitation methods in healthcare decision making. *Pharmaceutical medicine*. 2014;28:175-85.
7. Mongin P. Expected utility theory. In *Handbook of Economic Methodology*. Edited by J. Davis, W. Hands, and U. Maki. Edward Elgar: London. 1997. p. 342-350.
8. Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health economics*. 2002;11(5):447-56.
9. Abellán Perpiñán JM, Sánchez Martínez FI, Martínez Pérez JE, Méndez I. Lowering the 'floor' of the SF-6D scoring algorithm using a lottery equivalent method. *Health economics*. 2012;21(11):1271-85.
10. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. *Pharmacoeconomics*. 2008;26:661-77.
11. Lancaster KJ. A new approach to consumer theory. *Journal of political economy*. 1966;74(2):132-57.
12. Wang H, Rowen DL, Brazier JE, Jiang L. Discrete choice experiments in health state valuation: a systematic review of progress and new trends. *Applied health economics and health policy*. 2023;21(3):405-18.
13. Bansback N, Brazier J, Tsuchiya A, Anis A. Using a discrete choice experiment to estimate health state utility values. *Journal of health economics*. 2012;31(1):306-18.
14. Flynn TN, Louviere JJ, Peters TJ, Coast J. Best–worst scaling: what it can do for health care research and how to do it. *Journal of health economics*. 2007;26(1):171-89.
15. Clark MD, Determann D, Petrou S, Moro D, de Bekker-Grob EW. Discrete choice experiments in health economics: a review of the literature. *Pharmacoeconomics*. 2014;32:883-902.
16. McIntosh E, Louviere J, editors. Separating weight and scale value: an exploration of best-attribute scaling in health economics. Health Economists' Study Group meeting: Odense. 2002.
17. Mulhern B, Norman R, Brazier J. Valuing SF-6Dv2 in Australia using an international protocol. *Pharmacoeconomics*. 2021;39:1151-62.

18. Norman R, Mulhern B, Lancsar E, Lorgelly P, Ratcliffe J, Street D, Viney R. The use of a discrete choice experiment including both duration and dead for the development of an EQ-5D-5L value set for Australia. *Pharmacoeconomics*. 2023;41(4):427-38.
19. Brazier JE, Mulhern BJ, Bjorner JB, Gandek B, Rowen D, Alonso J, Vilagut G, Ware JE; SF-6Dv2 International Project Group. Developing a New Version of the SF-6D Health State Classification System From the SF-36v2: SF-6Dv2. *Medical care*. 2020;58(6):557-565.
20. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life research*. 2011;20:1727-36.
21. Mulhern BJ, Bansback N, Norman R, Brazier J; SF-6Dv2 International Project Group. Valuing the SF-6Dv2 Classification System in the United Kingdom Using a Discrete-choice Experiment With Duration. *Medical care*. 2020;58(6):566-573.
22. Ameri H, Pöder TG. Assessing the Direct Impact of Death on Discrete Choice Experiment Utilities. *Applied health economics and health policy*. 2025;23(2):319-27.
23. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *Journal of health economics*. 2002;21(2):271-92.
24. Pöder TG, Ameri H. A novel approach for health state valuation: Multiple bounded dichotomous choice compared to the traditional standard gamble. *Social science & medicine*. 2024;357:117173.
25. Alberini A, Boyle K, Welsh M. Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty. *Journal of environmental economics and management*. 2003;45(1):40-62.
26. Welsh MP, Bishop RC. Multiple bounded discrete choice models. Benefits & costs transfer in natural resource planning, Western Regional Research Publication, W-133, Sixth Interim Report, Department of Agricultural and Applied Economics, University of Georgia. 1993:331-52.
27. Saaty TL. Decision making with the analytic hierarchy process. *International journal of services sciences*. 2008;1(1):83-98.
28. Ameri H, Pöder TG. Comparison of four approaches in eliciting health state utilities with SF-6Dv2. *The European journal of health economics*. 2024:1-16.
29. Ameri H, Pöder TG. Valuing SF-6Dv2 using a discrete choice experiment in a general population in Quebec, Canada. *International journal of health policy and management*. 2024;13:8404.

Annexes

Annexe 1. Exemple de carte d'un choix expérimental discret avec durée (DCE_{TTO})

Description de santé A		Description de santé B	
Carte 453215		Carte 224612	
Vous vivez dans l'état de santé suivant puis vous mourrez :		Vous vivez dans l'état de santé suivant puis vous mourrez :	
4 années de vie		10 années de vie	
Très limité dans les activités exigeant un effort physique modéré		Un peu limité dans les activités exigeant un effort physique important	
Tout le temps accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes		Rarement accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes	
Activités sociales parfois limitées		Activités sociales la plupart du temps limitées	
Douleurs très légères		Douleurs très intenses	
Jamais très nerveux ou triste et démoralisé		Jamais très nerveux ou triste et démoralisé	
Tout le temps épuisé et vidé		Rarement épuisé et vidé	

Quelle situation préférez-vous?

Description de santé A ☐
Description de santé B ☐

Annexe 2. Exemple de carte de choix pour l'échelle de choix meilleur-pire avec durée (BWS_{TTO})

À noter que le DCE_{BWS} utilise une carte de DCE_{TTO} (voir annexe 1) suivie de la même carte avec l'ajout de l'option de la mort (comme ci-dessous).

Description de santé A		Description de santé B		Description de santé C
Carte 453215		Carte 224612		Mort immédiate et sans douleur
Vous vivez dans l'état de santé suivant puis vous mourrez :		Vous vivez dans l'état de santé suivant puis vous mourrez :		
4 années de vie		10 années de vie		
Très limité dans les activités exigeant un effort physique modéré		Un peu limité dans les activités exigeant un effort physique important		
Tout le temps accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes		Rarement accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes		
Activités sociales parfois limitées		Activités sociales la plupart du temps limitées		
Douleurs très légères		Douleurs très intenses		
Jamais très nerveux ou triste et démoralisé		Jamais très nerveux ou triste et démoralisé		
Tout le temps épuisé et vidé		Rarement épuisé et vidé		

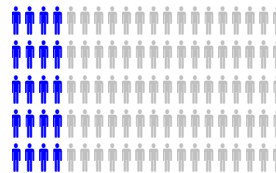
Quelle option est la meilleure?

État de santé A ☐ État de santé B ☐ Mort immédiate C ☐

Quelle option est la pire?

État de santé A ☐ État de santé B ☐ Mort immédiate C ☐

Annexe 3. Exemple de carte d'un choix expérimental discret avec mort immédiate comme attribut (DCE_{POD})

Description de santé A		Description de santé B	
Carte 453215		Carte 224612	
Vous vivez dans l'état de santé suivant puis vous mourrez :		Vous vivez dans l'état de santé suivant puis vous mourrez :	
4 années de vie		10 années de vie	
Très limité dans les activités exigeant un effort physique modéré		Un peu limité dans les activités exigeant un effort physique important	
Tout le temps accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes		Rarement accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes	
Activités sociales parfois limitées		Activités sociales la plupart du temps limitées	
Douleurs très légères		Douleurs très intenses	
Jamais très nerveux ou triste et démoralisé		Jamais très nerveux ou triste et démoralisé	
Tout le temps épuisé et vidé		Rarement épuisé et vidé	
20 % de probabilité de mourir immédiatement (20 chances sur 100)		10 % de probabilité de mourir immédiatement (10 chances sur 100)	

Quelle situation préférez-vous?

Description de santé A ☐

Description de santé B ☐

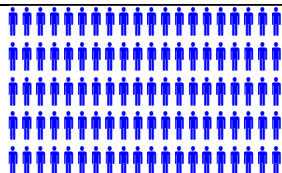
Annexe 4. Exemple de carte d'un choix expérimental discret avec la mort immédiate comme choix (DCE_{DOD})

Description de santé A		Description de santé B
<i>Carte 224612</i> Vous vivez dans l'état de santé suivant puis vous mourrez :		Mort immédiate et sans douleur
10 années de vie		
Un peu limité dans les activités exigeant un effort physique important		
Rarement accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes		
Activités sociales la plupart du temps limitées		
Douleurs très intenses		
Jamais très nerveux ou triste et démoralisé		
Rarement épuisé et vidé		

Quelle situation préférez-vous?

- Description de santé A ☐
- Description de santé B ☐

Annexe 5. Exemple de carte de choix avec le pari ordinaire (SG)

Choix A : État de santé certain à 100 % 100 chances sur 100 <i>Carte 113411</i>		
Vous vivez pendant 10 ans dans l'état de santé suivant puis vous mourrez :		
Pas du tout limité dans les activités exigeant un effort physique important		
Jamais accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes		
Activités sociales parfois limitées		
Douleurs moyennes		
Jamais très nerveux ou triste et démoralisé		
Jamais épuisé et vidé		

Afin d'améliorer votre état de santé tel que décrit ci-dessus, un traitement est disponible.

Choix B : 100 % de chance de retrouver un état de santé parfait et 0 % de chance de tomber dans le pire état de santé possible.			
100 % de chance de succès 100 chances sur 100 <i>Carte 111111</i>		0 % de chance d'échec 0 chance sur 100 <i>Carte 555655</i>	
Vous vivez pendant 10 ans dans l'état de santé suivant puis vous mourrez :		Vous vivez pendant 10 ans dans l'état de santé suivant puis vous mourrez :	
Pas du tout limité dans les activités exigeant un effort physique important		Très limité pour prendre un bain ou m'habiller	
Jamais accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes		Tout le temps accomplir moins de choses que vous l'auriez voulu dans vos activités quotidiennes	
Activités sociales jamais limitées		Activités sociales tout le temps limitées	
Aucune douleur		Douleurs très intenses	
Jamais très nerveux ou triste et démoralisé		Tout le temps très nerveux ou triste et démoralisé	
Jamais épuisé et vidé		Tout le temps épuisé et vidé	

Quelle situation préférez-vous?

Choix A

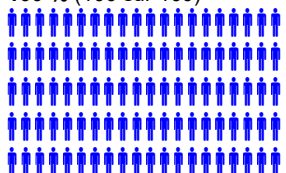
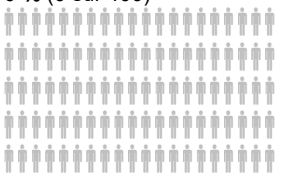
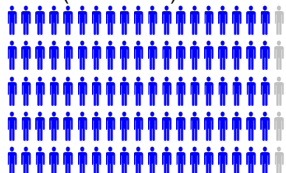
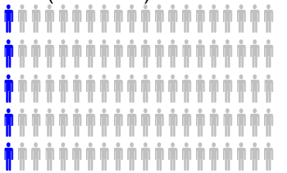
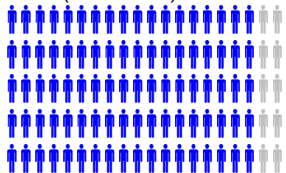
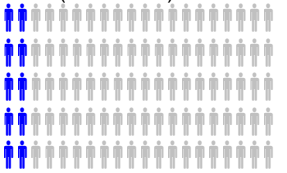
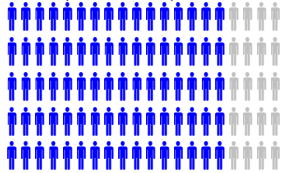
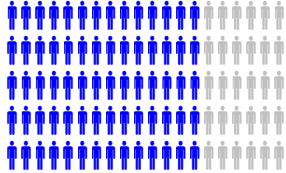
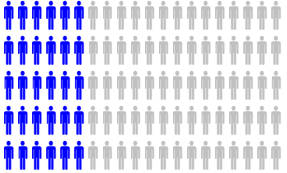
Choix B (avec traitement)

Je n'arrive pas à choisir.

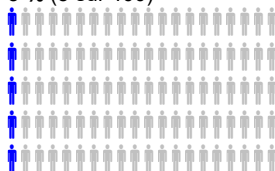
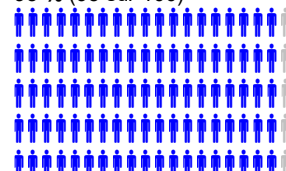
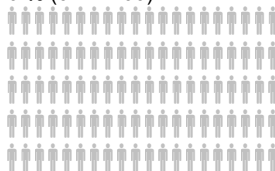
Annexe 6. Exemple de carte de choix dichotomique multiple borné (MBDC)

On prend une carte de choix du SG (voir annexe 5) et l'on fait varier les probabilités dans un tableau au lieu de la carte elle-même.

Chaque état de santé est évalué en présentant une carte de choix de type SG suivie du tableau ci-dessous :

Probabilité de <u>succès</u> du traitement	Probabilité d' <u>échec</u> du traitement	Je choisirais le traitement offert (choix B) :				
		Définitivement oui	Probablement oui	Indifférent	Probablement non	Définitivement non
100 % (100 sur 100) 	0 % (0 sur 100) 					
95 % (95 sur 100) 	5 % (5 sur 100) 					
90 % (90 sur 100) 	10 % (10 sur 100) 					
80 % (80 sur 100) 	20 % (20 sur 100) 					
70 % (70 sur 100) 	30 % (30 sur 100) 					

60 % (60 sur 100) 	40 % (40 sur 100) 					
50 % (50 sur 100) 	50 % (50 sur 100) 					
40 % (40 sur 100) 	60 % (60 sur 100) 					
30 % (30 sur 100) 	70 % (70 sur 100) 					
20 % (20 sur 100) 	80 % (80 sur 100) 					
10 % (10 sur 100) 	90 % (90 sur 100) 					

5 % (5 sur 100) 	95 % (95 sur 100) 					
0 % (0 sur 100) 	100 % (100 sur 100) 