

De rol van tijdsvoorkeur bij de waardering van gezondheid

Arthur Attema

Om zoveel mogelijk gezondheidswinst per euro te realiseren, is het noodzakelijk om gezondheidswinst te waarderen. Economen bestuderen meestal marktprijzen om de waardering van een goed te bepalen, maar de gezondheidszorgmarkt is een bijzondere markt, waarin alternatieve methoden nodig zijn om gezondheid te waarderen. Het is van belang dat deze methoden de voorkeuren van mensen zo nauwkeurig mogelijk weergeven. Alleen in dat geval kan men gegeven het budget zoveel mogelijk welvaart creëren. Tijdsvoorkeur speelt hierbij een grote rol. Dit artikel beschrijft de invloed van tijdsvoorkeur op het waarderen van gezondheidsuitkomsten en bespreekt een aantal ontwikkelingen in de waarderingsmethodologie.

1 Inleiding

Om met de beschikbare middelen zoveel mogelijk welvaart te realiseren, is het belangrijk deze middelen zo efficiënt mogelijk te gebruiken. Dit geldt ook voor de gezondheidszorg, waarin bijvoorbeeld de overheid met het zorgbudget zoveel mogelijk gezondheidswinst wil behalen. Hiervoor is er een representatieve maatstaf van deze gezondheidswinst nodig. Over het algemeen verkiezen economen de *revealed preference* methode om de waardering van een goed te bepalen. Dat wil zeggen, er wordt gekeken naar de keuzes die een consument maakt en aan de hand daarvan worden de voorkeuren van de consument en de daaruit voortvloeiende waardering van een goed achterhaald. Op het gebied van de gezondheidszorg is het echter niet eenvoudig om op een dergelijke manier een waardering op te stellen. Gezondheid, als uitkomst van het gebruik van gezondheidszorg, is namelijk niet meetbaar met behulp van *revealed preferences*. Het is bijvoorbeeld vaak lastig om de baten van verbeteringen in gezondheid ver-

oorzaakt door een operatie te bepalen. Er zijn daarom alternatieven nodig om een waardering te bepalen.

Bij gebrek aan marktwerking kunnen opgegeven voorkeuren (*stated preferences*) worden gebruikt. Hierbij gaat het vooral om het gebruik van directe onthullingmethoden zoals enquêtes en experimenten. Voor de waardering van gezondheid zijn op dit gebied diverse methoden ontwikkeld. De mate waarin deze methoden een correcte afspiegeling van de onderliggende voorkeuren van individuen geven, hangt af van de validiteit van enkele veronderstellingen over het gedrag van deze individuen. Dit betreft onder andere de manier van besluitvorming onder onzekerheid en over de tijd. Bij de waardering van gezondheid spelen twee aspecten een belangrijke rol, namelijk de waardering van een bepaalde gezondheidstoestand op zichzelf en het tijdstip waarop deze gezondheidstoestand optreedt.

Dit artikel zal ingaan op de ontwikkelingen op het gebied van de methodologie voor de waardering van gezondheid. In het bijzonder bespreek ik vernieuwende inzichten betreffende het meten van tijdsvoorkeur en het gebruik hiervan bij de waardering van gezondheidstoestanden. Daartoe beschrijf ik in sectie 2 eerst de meest gebruikte onthullingsmethoden voor de waardering van gezondheidstoestanden. Vervolgens zal ik nader ingaan op de *time tradeoff* (TTO) methode, een veelgebruikte *stated preference* methode voor het meten van gezondheidstoestanden. Aangezien tijdsvoorkeur hierbij een belangrijke rol speelt, zal ik in sectie 3 eerst de nodige aandacht aan dit onderwerp besteden, waarna ik een nieuwe methode om tijdsvoorkeur voor levensjaren te meten presenteer en met behulp van een voorbeeld aantoon hoe deze methode gebruikt kan worden om een betere schatting van het nut van gezondheidstoestanden te verkrijgen. In sectie 4 zullen enkele resultaten die met deze methoden zijn verkregen worden gerapporteerd. Tot slot volgt er een conclusie in sectie 5.

2 Voorkeuren voor gezondheidstoestanden

Alvorens een aantal *stated preferences* methoden voor het meten van voorkeuren voor gezondheid te introduceren, is het nuttig om eerst het meest gebruikte model voor het kwantificeren van effecten in de gezondheidseconomie te beschrijven. Dit is het *Quality Adjusted Life Year* (QALY)-model. Dit model vertegenwoordigt voorkeuren van mensen voor gezondheid als volgt:

$$V = \sum_{t=0}^T \delta(t)u(x_t) \quad (1)$$

waarbij $\delta(t)$ het gewicht is dat aan iedere periode wordt toegekend, $u(x_t)$ de voorkeur voor een bepaalde gezondheidstoestand x tijdens een bepaalde periode t weergeeft, T de laatste periode is en V het totale nut gedurende de T periodes. Dit model geldt in principe voor alle hieronder te bespreken meetmethoden.

Het is gebruikelijk om de *stated preferences* te meten door middel van een representatieve steekproef uit de bevolking. Er wordt echter ook wel eens aangeraden om een steekproef van patiënten te gebruiken, omdat deze ervaring met de betreffende gezondheidstoestand hebben en hier dus een beter oordeel over kunnen vellen. Er is een discussie gaande over welke steekproef het beste gebruikt kan worden om de waardering voor een gezondheidstoestand te bepalen (bijv. Dolan en Kahneman 2008), maar over het algemeen wordt een representatieve steekproef aanbevolen aangezien het, in verband met onder andere preventief gedrag, ook van belang is hoe de rest van de bevolking gezondheid waardeert. Bovendien zouden patiënten de ernst van hun toestand eventueel strategisch kunnen overdrijven, om bijvoorbeeld meer prioriteit aan hun ziekte toe te laten kennen.

Er zijn drie voorname *stated preferences* methoden om gezondheid te waarderen. Ten eerste is er de *Visual Analogue Scale*. Dit is een soort thermometer met aan het ene uiteinde de slechtst denkbare gezondheidstoestand en aan het andere uiteinde de best denkbare gezondheidstoestand, waarop een individu kan aangeven hoe hij een bepaalde gezondheidstoestand waardeert. De slechtst denkbare gezondheidstoestand wordt vaak gedefinieerd als dood en de best denkbare gezondheidstoestand als perfecte gezondheid, zodat de methode beter vergelijkbaar is met de hieronder beschreven *Standard Gamble* (SG) methode en de TTO-methode. In vergelijking (1) wordt de waarde van $u(x_t)$ voor de betreffende gezondheidstoestand hiermee direct gemeten. Over het algemeen speelt tijdsvoorkeur geen rol in deze methode, tenzij specifiek wordt gevraagd om te veronderstellen dat men zich gedurende enige tijd in deze toestand bevindt. Deze methode wordt echter weinig aanbevolen door economen doordat er bij deze methode geen keuzes gemaakt hoeven te worden, maar is wel populair onder niet-economen die zich bezighouden met zorgonderzoek.

Een tweede methode is de SG-methode, welke uitgaat van een risicovolle situatie. In deze methode wordt aan een individu een bepaalde risicovolle behandeling voorgelegd. Hierbij is er een kans p dat de operatie slaagt en het individu weer volledig gezond is en nog t jaar in perfecte gezondheid leeft. Er is ook een kans $1-p$ dat de operatie mislukt en het individu

overlijdt. Er wordt vervolgens aan het individu gevraagd om aan te geven hoe groot de kans p moet zijn om deze loterij equivalent te maken aan de status quo, dat wil zeggen niet opereren en nog t jaar in de huidige imperfecte gezondheidstoestand x blijven. Indien verwacht nut wordt verondersteld en het nut van perfecte gezondheid door normalisering op 1 ($U(PG)=1$) wordt gezet en het nut van dood op 0 ($U(D)=0$), dan volgt uitgaande van het QALY-model dat:

$$\sum_{t=0}^T \delta(t)u(x) = p \sum_{t=0}^T \delta(t)u(PG) + (1-p) \sum_{t=0}^T \delta(t)u(D) \Leftrightarrow u(x) = p \quad (2)$$

Doordat de duur van de gezondheidstoestanden hetzelfde is voor alle toestanden (T jaar) en doordat, indien het QALY-model (1) geldig is, het nut van een gezondheidstoestand onafhankelijk is van de duur ervan, speelt tijdsvoorkeur geen rol bij deze methode (zoals te zien aan het wegvallen van $\delta(t)$ uit (2)). Het aantal jaar dat nog te leven is in de te waarderen gezondheidstoestand zou bij deze methode dus niet uit moeten maken voor de waardering.

De verwachte nutstheorie is de normatieve theorie op het gebied van besluitvorming bij onzekerheid. In verband met deze theoretische basis was de SG-methode aanvankelijk erg populair onder economen. Er zijn echter twee problemen met de SG-methode. Het eerste probleem is de veronderstelling dat de SG-methode kan worden geëvalueerd met behulp van de verwachte nutstheorie. Er bestaat namelijk veel empirisch bewijs dat mensen zich niet volgens deze theorie gedragen (Starmer 2000). In de medische besluitvorming zijn falsificaties van de verwachte nutstheorie gevonden door onder anderen Bleichrodt (2001) en Oliver (2003). Indien de verwachte nutstheorie niet geldig is, zullen de resulterende utiliteiten niet overeenkomen met de werkelijke voorkeuren van mensen. In het bijzonder hebben mensen vaak moeite om kansen te evalueren. Objectieve kansen worden vaak subjectief gewogen. Kleine kansen krijgen doorgaans een te hoog gewicht, terwijl grotere kansen een te laag gewicht krijgen (Kahneman en Tversky 1979). Een concaaf verloop van de nutsfunctie en het onderwegen van kansen leiden beide tot risicomijdend gedrag. Als er in de meetprocedure geen rekening met kansweging wordt gehouden, zullen de resulterende nutswaarden in de gebruikelijke varianten van de SG-methode te hoog zijn (Wakker en Stiggelbout 1995). Er wordt dan ten onrechte een hoge utiliteitswaarde aan het leven in een bepaalde gezondheidstoestand toegekend, terwijl er in feite sprake is van het onderwegen van kansen en een lagere utiliteitswaarde van de gezondheidstoestand.

Een tweede zwak punt van deze methode is het gebruik van de uitkomst dood in de vraagstelling. De uitkomst dood wordt door respondenten vaak

als problematisch beschouwd, aangezien zij het moeilijk vinden om zich in te beelden dat zij binnen een week zullen sterven of aversie tegen dood hebben. Hierdoor ontstaat vaak extreem risicomijdend gedrag.

Een andere veelgebruikte methode is de TTO-methode, welke zich op het tijdsaspect in plaats van het risicoaspect richt. In een TTO dienen individuen een afweging te maken tussen kwaliteit van leven en levensduur. Een typische TTO-procedure omvat een afweging tussen het leven in een bepaalde imperfecte gezondheidstoestand gedurende tien jaar en het leven in perfecte gezondheid gedurende een periode van minder dan tien jaar, waarna in beide gevallen dood volgt. De hoeveelheid tijd die mensen bereid zijn op te offeren om in perfecte gezondheid te geraken geeft dan de waarde van de imperfecte gezondheidstoestand aan. De standaard TTO-methode is gestoeld op het lineaire QALY-model. Dit model vertegenwoordigt voorkeuren van mensen voor gezondheid als volgt:

$$V = \sum_{t=0}^T u(x_t) \quad (3)$$

Het lineaire QALY-model is dus een opsomming van het nut tijdens elke periode, waarbij elke periode evenveel gewicht krijgt. Dit model wordt voornamelijk gebruikt vanwege de eenvoud. Het is immers niet nodig een schatting van tijdsvoorkeuren te verkrijgen. Op basis van dit model levert de TTO-methode de volgende vergelijking op:

$$\sum_{t=0}^z u(x) + \sum_{t=z}^T u(D) = \sum_{t=0}^y u(PG) + \sum_{t=y}^T u(D) \quad (4)$$

met y het aantal jaar in perfecte gezondheid dat de respondent even hoog waardeert als z jaar in een bepaalde imperfecte gezondheidstoestand x , $u(D)$ het nut van dood en $u(PG)$ het nut van perfecte gezondheid. Vervolgens kan $u(D)$ door normalisering op 0 worden gezet en $u(PG)$ op 1, zodat de vergelijking aanzienlijk kan worden vereenvoudigd:

$$z * u(x) = y \leftrightarrow u(x) = \frac{y}{z} \quad (5)$$

De waarde voor $u(x)$ die resulteert, is de QALY-score van gezondheidstoestand x . De schattingen van QALY's in gezondheidseconomische evaluaties zijn vaak gebaseerd op dergelijke TTO-metingen.

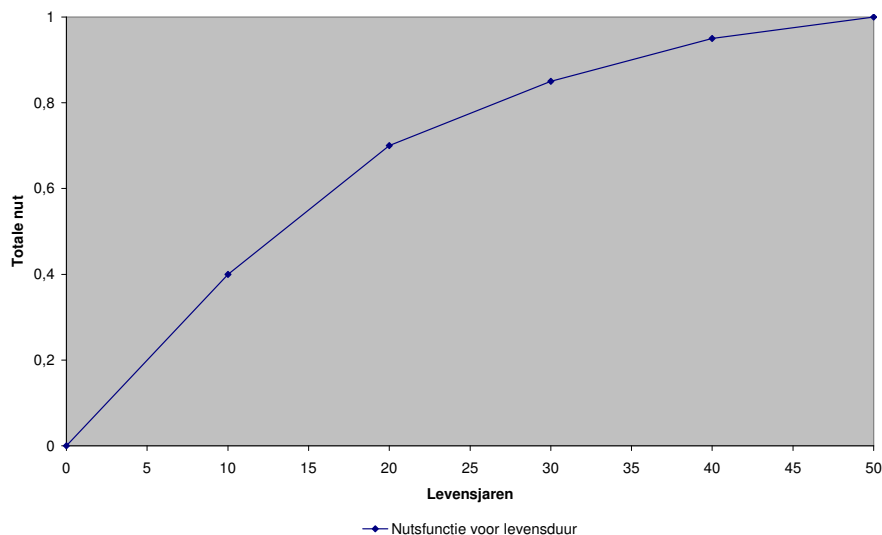
Ondanks haar populariteit is de TTO-methode echter ook niet gevrijwaard van problemen. In het bijzonder hangt de methode in sterke mate af van tijdsvoorkeur. In dit artikel richt ik mij op de afwijking die het veronderstellen van lineair nut van levensduur veroorzaakt. Indien de nutsfunc-

tie voor levensduur niet lineair is maar concaaf, worden toekomstige levensjaren verdisconteerd, waardoor de toekomstige jaren die in een TTO worden afgeruild minder waarde hebben. Indien men bij het berekenen van de TTO-scores eenvoudigweg het aantal toekomstige levensjaren dat individuen af willen ruilen gebruikt, zoals bij het gebruik van het lineaire QALY-model, zal het nut dat aan een imperfecte gezondheidstoestand wordt toegekend niet juist zijn. Om een betere schatting van de ware waardering van een gezondheidstoestand te verkrijgen is derhalve een correctie voor tijdsvoorkeur nodig.

3 Tijdsvoorkeur

Kosten en baten van een investering treden vaak op verschillende momenten in de tijd op. *Tijdsvoorkeur*, de mate waarin nut dat op verschillende tijdstippen wordt ervaren van elkaar verschilt, speelt daarom vaak een belangrijke rol. In het bijzonder geldt dat men baten meestal liever eerder dan later ontvangt, en dat men kosten liever later dan eerder betaalt. Dit geldt ook voor gezondheidsbaten en -kosten.

Mensen geven hun toekomstige levensjaren vaak niet dezelfde waarde. Jaren in de verre toekomst worden doorgaans verdisconteerd ten opzichte van jaren in de nabije toekomst. Er is dus sprake van tijdsvoorkeur voor levensjaren, hetgeen meestal wordt gemeten met behulp van een nutsfunctie voor levensduur. Deze functie geeft het totale nut dat men aan een leven ontleent op verschillende tijdstippen weer. Figuur 1 geeft een voorbeeld. De horizontale as beschrijft de totale levensduur en de verticale as beschrijft het cumulatieve nut dat hieruit voortvloeit. De grafiek in Figuur 1 heeft een concaaf verloop, wat overeenkomt met de voorkeuren van het merendeel van de bevolking. Een concave nutsfunctie voor levensduur is equivalent aan een positieve tijdsvoorkeur, ofwel het verdisconteren van toekomstige levensjaren. Het eerste jaar geeft immers meer nut dan het tweede jaar, het tweede jaar geeft meer nut dan het derde jaar, enzovoort. Meer specifiek betekent bijvoorbeeld een nutsfunctie van de machtsvorm ($v(t) = t^y + c$) dat toekomstige levensjaren (t) hyperbolisch verdisconteerd worden, terwijl een nutsfunctie van de exponentiële vorm ($v(t) = \exp(-rt) + c$) een constante verdiscontering impliceert.

Figuur 1 Nutsfunctie voor levensduur

Tijdsvoorkeur speelt een belangrijke rol in de TTO-methode. Individuen kunnen bij deze methode hun tijdsvoorkeur in het meetproces opnemen, zoals hieronder wordt uitgelegd. De gebruikers van deze methode veronderstellen veelal echter dat men niet verdisconteert. De verkregen QALY-waarden worden dus als niet-verdisconteerde waarden beschouwd en, indien ze in de toekomst optreden, regelmatig met behulp van een standaarddiscontovoet nogmaals verdisconteerd, om een contante waarde van QALY's te berekenen (Gravelle e.a. 2007). Er vindt met andere woorden een dubbele verdiscontering van QALY's plaats als men geen rekening houdt met de tijdsvoorkeuren van respondenten. Het gevolg is een onderschatting van het nut dat een bepaalde gezondheidstoestand oplevert. Men zou dit kunnen voorkomen door de verkregen QALY-waarden direct te gebruiken in economische evaluaties, zonder eerst de contante waarde te berekenen. Omdat individuen verschillende discontostructuren hebben, is het echter aan te bevelen om deze discontostructuren eerst te meten, en de TTO-waarden hiervoor te corrigeren, zodat een robuuste schatting van een gezondheidstoestand kan worden verkregen. Daarna kan men indien gewenst alsnog een contante waarde berekenen. Om gezondheid goed te kunnen waarderen, is het daarom essentieel om tijdsvoorkeur zo nauwkeurig mogelijk te schatten.

Het meten van tijdsvoorkeur voor levensjaren. Kennis van de nutsfunctie voor levensduur is onder andere belangrijk voor het corrigeren van TTO-waarden voor tijdsvoorkeur. De gebruikelijke manier om het nut van levensduur te meten is door middel van de *Certainty Equivalence* (zekerheidsequivalent) methode (CE-methode). Deze methode vertoont gelijkheid met de SG-methode. Bij de CE-methode wordt een besluitnemer weer geconfronteerd met een risicovolle behandeling en een zekere uitkomst. De operatie heeft een gegeven kans p om te slagen, waarna het individu nog T jaar in gezondheidstoestand x te leven heeft. Er is ook een kans $1-p$ dat de operatie mislukt en het individu direct sterft. Indien er geen operatie plaatsvindt, blijft het individu nog $\tau < T$ jaar leven in gezondheidstoestand x . Er wordt vervolgens aan het individu gevraagd om aan te geven hoeveel toegevoegde levensjaren τ met zekerheid hij of zij als equivalent aan de risicovolle behandeling beschouwt. Hierdoor ontstaat, als we zodanig normaliseren dat $\sum_{t=0}^T \delta(t) = 1$, de vergelijking

$$\sum_{t=0}^{\tau} \delta(t)u(x) = p \sum_{t=0}^T \delta(t)u(x) + (1-p) \sum_{t=0}^T \delta(t)u(D) \Leftrightarrow \sum_{t=0}^{\tau} \delta(t) = p. \quad (6)$$

De CE-methode kent logischerwijs soortgelijke problemen als de SG-methode. De extreme risicoaversie die door de uitkomst onmiddellijke sterfte en het onderwegen van kansen wordt veroorzaakt, betekent in de CE-methode dat er een hoge verdiscontering van toekomstige levensjaren plaatsvindt.

Om bovengenoemde problemen te omzeilen, hebben Attema e.a. (2007) een nieuwe methode voorgesteld om het nut van levensduur te meten. Deze methode vermijdt de uitkomst dood en is risicovrij, waardoor zij niet getroffen wordt door falsificaties van verwacht nut. Bovendien is de methode op keuzes gebaseerd.

In de risicovrije methode van Attema e.a. (2007) worden respondenten met twee gezondheidsscenario's geconfronteerd. In het eerste scenario begint de respondent in een goede gezondheidstoestand (γ). Na een bepaalde tijd ($t_{0,5}$) zal hij een slechtere gezondheid (β) hebben, die hij gedurende de rest van de totale periode (T) zal blijven houden. In het tweede scenario verkeert de respondent eerst in de slechtere gezondheid β en krijgt vanaf tijdstip $t_{0,5}$ een betere gezondheid (γ), waarin hij de resterende tijd blijft verkeren. Vervolgens wordt de waarde van de periode $t_{0,5}$ onthuld die de respondent indifferent maakt tussen deze scenario's. Deze waarde geeft het punt aan waarbij de periode tussen nu en $t_{0,5}$ evenveel nut geeft als de periode tussen $t_{0,5}$ en T . Als t kleiner dan de helft van T is, dan is er sprake van

concaaf nut (d.w.z. positieve discontering). Als deze procedure wordt herhaald, gebruikmakend van eerdere antwoorden, kan er meer informatie over de discontofunctie worden verkregen. De methode omvat geen risico, waardoor kansweging geen versturende factor is in de meetprocedure.

Attema e.a. (2007) hebben het nut van levensduur met de risicovrije methode experimenteel gemeten. Hieronder volgt een voorbeeld van de meetprocedure. In de eerste vraag in de risicovrije methode moest de respondent een keuze maken tussen de volgende scenario's. In scenario A verkeerde hij $t_{0,5}=25$ jaar in perfecte gezondheid en daarna 25 jaar in een gezondheidstoestand met regelmatige rugklachten. In scenario B verkeerde hij eerst 25 jaar in een gezondheidstoestand met regelmatige rugklachten en vervolgens 25 jaar in perfecte gezondheid. Als hij hier voor A koos, dan was de keuze in de volgende vraag als volgt. In scenario A verkeerde hij eerst $t_{0,5}=13$ jaar in perfecte gezondheid en daarna 37 jaar met regelmatige rugklachten. In scenario B verkeerde hij eerst 13 jaar met regelmatige rugklachten en vervolgens 37 jaar in perfecte gezondheid. Scenario A werd met andere woorden minder aantrekkelijk gemaakt en scenario B juist aantrekkelijker. Indien de respondent in de eerste vraag voor B had gekozen werd het omgekeerde gedaan. Deze procedure werd een aantal keer herhaald waarna een schatting van het punt $t_{0,5}$ werd gemaakt waarbij de respondent indifferent was tussen A en B.

Het kan worden aangetoond dat het niet nodig is om te weten wat het nut van het leven met rugklachten is om op deze manier het nut van levensduur te achterhalen (Attema e.a. 2007). Ook maakt deze methode het mogelijk om nut op een nonparametrische manier te achterhalen, dat wil zeggen, er is geen schatting van een bepaalde parametrische functie noodzakelijk.

Stel dat een respondent bij bovenstaand voorbeeld indifferent was tussen A en B als hij bij A de komende 20 jaar in perfecte gezondheid leefde en de daaropvolgende 30 jaar met rugklachten, terwijl hij bij B de komende 20 jaar rugklachten had en de 30 jaar daarna een perfecte gezondheid ($t_{0,5}=20$). Attema e.a. (2007) tonen aan dat het nut van de komende 20 jaar dan gelijk is aan 0,5, gegeven dat het nut van 0 jaar gelijk aan 0 is en het nut van 50 (=T) jaar gelijk aan 1. Vervolgens werd de respondent gevraagd te kiezen tussen:

- A: 10 jaar in perfecte gezondheid en daarna 10 jaar met rugklachten*
- B: 10 jaar met rugklachten en daarna 10 jaar in perfecte gezondheid*

Met behulp van dergelijke keuzes werd dan weer een nieuw indifferentiepunt, $t_{0,25}$, bepaald. Stel dat dit 8 jaar was. Het nut van 8 jaar was dan gelijk aan 0,25. Op deze manier werden 5 punten gemeten. Voor de betreffende

respondent waren dat: $v(3) = 0,125$; $v(8) = 0,25$; $v(20) = 0,5$; $v(34) = 0,75$; $v(40) = 0,875$. Hoe meer punten er onthuld werden, hoe meer informatie er over de nutsfunctie beschikbaar kwam. De volledige nutsfunctie kon benaderd worden door lineaire interpolatie tussen deze punten of het schatten van een bepaalde parametrische vorm die de beste fit gaf voor deze waarden.

Bij de CE-methode omvatte de eerste vraag een keuze tussen het met zekerheid leven van 25 jaar in perfecte gezondheid en een loterij met 50% kans op het leven van nog 50 jaar in perfecte gezondheid en 50% kans om direct te sterven. Als voor A werd gekozen, werd deze optie minder aantrekkelijk gemaakt. De respondent moest toen kiezen tussen 13 jaar in perfecte gezondheid met zekerheid en dezelfde loterij als in de eerste vraag. Het vervolg was te vergelijken met de procedure in de risicovrije methode. Attema e.a. (2007) tonen aan dat deze twee methoden dezelfde nutsfunctie zouden moeten opleveren als de respondenten zich volgens de verwachte nutstheorie gedragen.

Attema en Brouwer (2008a) laten zien hoe de risicovrije methode kan worden aangewend om TTO-scores te corrigeren voor de invloed van tijdsvoorkeur. De gangbare TTO-meting gebruikt het QALY-model zonder tijdsvoorkeur om tot een schatting van het nut van een gezondheidstoestand te komen. Als een respondent dan aangeeft dat hij onverschillig is tussen het leven van nog 5 jaar in perfecte gezondheid en het leven van nog 10 jaar met een bepaalde handicap, dan wordt het nut van de handicap vastgesteld op $u(x) = 5/10 = 0,5$. Deze procedure houdt echter geen rekening met het verdisconteren van toekomstige levensjaren. Als latere levensjaren minder gewicht krijgen dan eerdere levensjaren, dan zal de procedure een te lage schatting van het nut van een gezondheidstoestand opleveren.

Attema en Brouwer (2008a) veronderstellen het algemene QALY-model dat verdisconteren niet uitsluit (vergelijking 1). Als we in dat model $v(t)$ definiëren als

$$v(t) = \sum_{i=0}^t \delta_i,$$

dan verkrijgen we de volgende vergelijking voor de TTO-oefening:

$$v(10) * u(x) = v(5) * u(PG) + (v(10) - v(5)) * u(D) \quad (7)$$

Als we $u(D)$ weer normaliseren tot 0 en $u(PG)$ tot 1, dan kan deze vergelijking worden vereenvoudigd tot:

$$u(x) = \frac{v(5)}{v(10)} \quad (8)$$

Met behulp van de risicovrije methode kunnen zij vervolgens de functie $v(t)$ schatten en daarmee de tijdsvoorkeur voor levensjaren en het nut van de gezondheidstoestand van elkaar scheiden.

Aan de hand van een voorbeeld zullen we hier laten zien hoe men op deze manier een gecorrigeerde TTO-score kan verkrijgen. Stel dat de eerder beschreven voorbeeldresponent indifferent is tussen het leven van 5 jaar in perfecte gezondheid en 10 jaar met een handicap. De ongecorrigeerde TTO-score is dan gelijk aan $5/10 = 0,5$. Maken we echter gebruik van de schatting van zijn tijdsvoorkeur voor levensjaren, dan berekenen we eerst wat het nut van 5 jaar en het nut van 10 jaar is voor deze responent. We hebben geen meting van het nut van deze aantallen jaren, maar wel van het nut van 3, 8 en 20 jaar. Daarom benaderen we het nut van 5 en 10 jaar met behulp van lineaire interpolatie. Het nut van 5 jaar is dan (Attema en Brouwer 2008a):

$$v(5) = u(3) + \frac{5-3}{8-3} * (u(8) - u(3)) = 0,125 + \frac{2}{5} * 0,125 = 0,175 \quad (9)$$

Het nut van 10 jaar is:

$$v(10) = v(8) + \frac{10-8}{20-8} * (v(20) - v(8)) = 0,25 + \frac{2}{12} * 0,25 = 0,292 \quad (10)$$

De gecorrigeerde TTO-score is nu $\frac{v(5)}{v(10)} = \frac{0,175}{0,292} = 0,6$. Omdat deze responent toekomstige levensjaren verdisconteert, ontstaat daarom een te lage TTO-score (0,5 in plaats van 0,6) als men hier geen rekening mee houdt.

4 Resultaten

In een experiment hebben Attema e.a. (2007) de risicovrije methode en de CE-methode met elkaar vergeleken (*within subjects*). De opzet van het experiment maakte het mogelijk om de verkregen waarden direct met elkaar te vergelijken, zonder eerst bepaalde nutsfuncties te hoeven schatten. Dat wil zeggen, de functies konden niet-parametrisch geschat worden. De CE-

methode resulteerde in een meer concave nutsfunctie dan de risicovrije methode (*Friedman test*, $p < 0.05$).

Om een indicatie te geven hoe deze resultaten er in vergelijking met eerdere resultaten uitzagen, hebben Attema e.a. (2007) de analyse ook uitgevoerd met gebruik van enkele parametrische schattingen. Als beide nutsfuncties bijvoorbeeld met een machtsfunctie geschat werden, was de gemiddelde parameterschatting voor de CE-methode significant lager dan die van de risicovrije methode. In verband met uitschieters hebben zij de niet-parametrische *Wilcoxon signed ranks test* gebruikt voor een vergelijking van de schattingen en rapporteer ik hier de mediane machtschattingen: 0,44 voor de CE-methode en 0,62 voor de risicovrije methode ($p < 0.05$). Dit verschil kon verklaard worden door afwijkingen van verwacht nut. Als namelijk rekening gehouden werd met kansweging, gebruikmakend van eerder geschatte kanswegingsparameters (Tversky en Kahneman, 1992), dan was er geen significant verschil meer tussen de uitkomsten van deze twee methoden (de nieuwe mediaan van de geschatte macht voor de CE-methode was 0,57; $p = 0,80$).

Deze resultaten suggereren dat het mogelijk is het nut van levensduur te meten zonder de problematische uitkomst dood te hoeven gebruiken. Dit lijkt een belangrijk voordeel te zijn, aangezien de respondenten in een enquête aangaven dat ze de risicovrije methode als eenvoudiger te beantwoorden beschouwden dan de CE-methode ($p < 0.01$). Bovendien gaven zij aan dat zij zich de situatie in de risicovrije methode beter konden voorstellen dan de situatie in de CE-methode (Attema e.a. 2007).

Eerdere studies hebben pogingen ondernomen om waarderingen gemeten met de TTO-methode te corrigeren voor de nutsfunctie voor levensduur (o.a. Pliskin e.a. 1980, Van Osch e.a. 2004). Deze studies gebruikten hier echter allen de CE-methode voor, met de hierboven genoemde problemen. Attema en Brouwer (2008a) hebben derhalve de risicovrije methode aangewend om TTO-scores voor het nut van levensduur te corrigeren. Zij vonden gecorrigeerde TTO-scores die significant hoger waren dan de ongecorrigeerde waarden (gemiddelde resp. 0,75 en 0,80, $p < 0.01$). De correctie met behulp van de risicovrije methode was echter minder groot dan de correctie met behulp van de CE-methode, wat waarschijnlijk vooral veroorzaakt werd door kansweging in de CE-methode.

5 Conclusie

Om het beschikbare overheidsbudget zo efficiënt mogelijk te gebruiken, is het belangrijk om de kosten en baten van de alternatieve toepassingen zo

nauwkeurig mogelijk te schatten. Bij gebrek aan marktprijzen voor gezondheid zijn alternatieve meetmethoden benodigd om de baten van gezondheidsinterventies te schatten. Hoewel er veel vooruitgang is geboekt bij de ontwikkeling van deze methoden, kennen zij nog altijd de nodige problemen. In dit artikel zijn de voornaamste problemen besproken en zijn suggesties voor verbeteringen aangedragen.

De resultaten van Attema e.a. (2007) bevestigen de resultaten van eerdere studies in andere domeinen. Stalmeier en Bezembinder (1999) vonden dat het nut van diverse gezondheidstoestanden gemeten in een onzekere en een zekere situatie verschillend was, maar dat dit verschil verdween wanneer gecorrigeerd werd voor kansweging. Hetzelfde gold voor de studie van Abdellaoui e.a. (2007) in het geval van nut voor geld.

Het is dus zaak om te achterhalen wat de invloeden zijn die afwijkingen van rationaliteit veroorzaken en hoe groot die invloeden precies zijn. Een mogelijkheid om dit te bewerkstelligen is door een bepaalde vorm van correctie toe te passen, wat Attema e.a. (2007) bij het corrigeren van de CE-methode voor kansweging deden. Het meten van deze invloeden vergt doorgaans veel werk, zodat het de moeite waard is om naar alternatieven te zoeken die deze invloeden in het geheel vermijden. Een voorbeeld van een dergelijk alternatief is de hier beschreven risicovrije methode.

Er is echter nog veel onderzoek naar het verbeteren van meetmethoden nodig, aangezien er naast de problemen rond tijdsvoorkeur nog de nodige andere problemen bestaan. In het geval van de CE-methode en de SG-methode valt te denken aan de eerder genoemde weging van kansen, terwijl bij de TTO-methode eigenschappen zoals procedurele invariantie (Attema en Brouwer 2008b) en constante proportionele afwegingen (Attema en Brouwer 2008c) problematisch zijn. Bovendien speelt in al deze methoden verliesafkeer een grote rol (Bleichrodt e.a. 2003).

Auteur

De auteur is verbonden aan het instituut Beleid en Management van de Gezondheidszorg (iBMG) van het Erasmus MC. Correspondentieadres: Postbus 1738, 3000 DR Rotterdam. E-mail: a.attema@erasmusmc.nl. Internet: <http://www.bmg.eur.nl/personal/attema>

Literatuur

- Abdellaoui, M., C. Barrios en P.P. Wakker, 2007, Reconciling introspective utility with revealed preference: experimental arguments based on prospect theory, *Journal of Econometrics*, vol. 138(2): 356-78.
- Attema, A.E., H. Bleichrodt en P.P. Wakker, 2007, A new method for measuring the utility of life duration, Erasmus Universiteit Rotterdam.
- Attema, A.E. en W.B.F. Brouwer, 2008a, The correction of TTO-scores for utility curvature using a risk-free utility elicitation method, *Journal of Health Economics*, te verschijnen.
- Attema, A.E. en W.B.F. Brouwer, 2008b, Can we fix it? Yes we can! But what? A new test of procedural invariance in TTO-measurement, *Health Economics*, vol. 17(7): 877-85.
- Attema, A.E. en W.B.F. Brouwer, 2008c, On the (not so) constant proportional tradeoff in TTO-measurement, ingediend voor publicatie.
- Bleichrodt, H., 2001, Probability weighting in choice under risk: an empirical test, *Journal of Risk and Uncertainty*, vol. 23(2): 185-98.
- Bleichrodt, H., J.L. Pinto en J.M. Abellan-Perpinan, 2003, A consistency test of the time trade-off, *Journal of Health Economics*, vol. 22(6): 1037-52.
- Dolan, P. en D. Kahneman, 2008, Interpretations of utility and their implications for the valuation of health, *The Economic Journal*, vol. 118: 215-34.
- Gravelle, H., W.B.F. Brouwer, L.W. Niessen, M.J. Postma en F.F.H. Rutten, 2007, Discounting in economic evaluations: stepping forward towards optimal decision rules, *Health Economics*, vol. 16(3): 307-17.
- Kahneman, D. en A. Tversky, 1979, Prospect theory: an analysis of decision under risk, *Econometrica*, vol. 47(2): 263-91.
- Oliver, A.J., 2003, The internal consistency of the standard gamble: tests after adjusting for prospect theory, *Journal of Health Economics*, vol. 22(4): 659-74.
- Osch, S.M.C. van, P.P. Wakker, W.B. van den Hout en A.M. Stiggelbout, 2004, Correcting biases in standard gamble and time tradeoff utilities, *Medical Decision Making*, vol. 24(5): 511-17.
- Pliskin, J.S., D.S. Shepard en M.C. Weinstein, 1980, Utility functions for life years and health status, *Operations Research*, vol. 28(1): 206-24.
- Stalmeier, P.F.M. en T.G.G. Bezembinder, 1999, The discrepancy between risky and riskless utilities: a matter of framing?, *Medical Decision Making*, vol. 19(4): 435-47.
- Starmer, C., 2000, Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk, *Journal of Economic Literature*, vol. 38(2): 332-82.
- Tversky, A en D. Kahneman, 1992, Advances in prospect theory: cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, vol. 5(4): 297-323.
- Wakker, P.P. en A.M. Stiggelbout, 1995, Explaining distortions in utility elicitation through the rank-dependent model for risky choices, *Medical Decision Making*, vol. 15(2): 180-86.