2 Ethische aspecten van kostenutiliteitsanalyse[1]

De kostenutiliteitsanalyse (KUA) is ontwikkeld als instrument voor de prioritering van voorzieningen over de volle breedte van de gezondheidszorg. Voor het kwantificeren van de opbrengst van zeer diverse voorzieningen wordt daarbij als maatstaf meestal het Quality Adjusted Life Year *(QALY) gehanteerd. De discussie over mogelijk onrechtvaardige verdelingseffecten van dit instrumentarium is niet nieuw, maar wel actueel nu het in het Verenigd Koninkrijk daadwerkelijk wordt gebruikt bij de toelating van nieuwe gezondheidszorgvoorzieningen in de* National Health Service*. Op de vraag of en hoe KUA een moreel aanvaardbare rol kan spelen in besluitvorming over de verdeling van schaarse middelen in de gezondheidszorg, worden verschillende antwoorden gegeven. De bespreking daarvan in dit signalement is van belang met het oog op de voorgenomen opzet van een nationale beoordelingsstructuur voor beslissingen over het verzekeringspakket.*

2.1 Inleiding

Om bij een toenemende vraag de betaalbaarheid van de gezondheidszorg in stand te houden, is in alle westerse landen bewaking van de zorguitgaven een belangrijke doelstelling van het overheidsbeleid. Maatregelen voor gereguleerde toelating van voorzieningen tot het verstrekkingenpakket worden alom als noodzakelijk gezien. De discussie gaat dan ook allang niet meer over de vraag of aan die toelating beperkingen moeten gesteld ('rationing'), maar over hoe dat verantwoord kan gebeuren (Maynard 2001).

Met het oog op dergelijke beslissingen hebben gezondheidseconomen het instrument van de kostenutiliteitsanalyse (KUA) ontwikkeld. Doel daarvan is de zeer uiteenlopende effecten van alle voorzieningen binnen de gezondheidszorg op uniforme wijze af te kunnen zetten tegen de kosten. Het verschil met de kosteneffectiviteitsanalyse (KEA) – hoewel die laatste term ook wel in overkoepelende zin wordt gebruikt – is dat de KEA zich beperkt tot vergelijking van de effectiviteit van behandelvormen met een gelijksoortige uitkomst. Om, wat bij KUA de bedoeling is, beslissingen te kunnen baseren op een gezondheidszorgbrede vergelijking, is een gemeenschappelijke maatstaf nodig in termen waarvan uiteenlopende gezondheidstoestanden uniform kunnen worden gewaardeerd. Daarvoor zijn in de gezondheidseconomische literatuur verschillende voorstellen gedaan. Als theoretisch best onderbouwde en praktisch meest ontwikkelde uitkomst-

1 Opgesteld door dr WJ Dondorp, secretaris Beraadsgroep Gezondheidsethiek &- recht

maat geldt de QALY (*Quality Adjusted Life Year*). Varianten zijn de DALY (*Disability Adjusted Life Year*), de HYE (*Healthy Year Equivalent*) en de SAVE (*Saved Young Life Equivalent*).

De discussie over KUA en zijn mogelijke rol als prioriteringsinstrument heeft een belangrijke normatieve dimensie. Maximalisatie van in QALY's uitgedrukte kostenutiliteit leidt tot een verdeling van de beschikbare middelen die niet vanzelfsprekend ook de meest rechtvaardige is. Aan dat aspect is in de relevante gezondheidseconomische en ethische literatuur al sinds de introductie van het QALY-begrip in de jaren zeventig veel aandacht besteed. Maar wat dat moet betekenen voor de mogelijke rol van KUA bij de prioritering van gezondheidszorgvoorzieningen blijft een lastige vraag.

De discussie daarover is weer actueel nu in het Verenigd Koninkrijk (Engeland en Wales) een structuur is opgezet waarbinnen op basis van KUA-informatie daadwerkelijk over de introductie van nieuwe voorzieningen wordt besloten. Een centrale rol daarin heeft het in 1999 ingestelde 'National Institute for Clinical Excellence' (NICE[2]), dat vergaande bevoegdheden heeft gekregen om te bepalen welke voorzieningen binnen het budget van de *National Health Service* (NHS) beschikbaar moeten komen (Rawlins 2004). Omdat die Britse aanpak ook in ons land als model wordt gezien, is er alle reden om aandacht te vragen voor de stand van zaken in de discussie over de mogelijke rol van KUA en de ethische aspecten daarvan.

Dit signalement is als volgt opgebouwd. Na een korte schets van de achtergrond (2.2) en van de discussie over de verdelingseffecten van KUA (2.3) worden twee antwoorden besproken op de vraag hoe kan worden voorkomen dat toepassing van KUA tot onrechtvaardigheid leidt. Het eerste antwoord (in 2.4) komt neer op een voorstel tot verbreding van KUA, zodanig dat ook overwegingen van rechtvaardigheid daarin worden verdisconteerd. Het tweede antwoord (in 2.5) betekent een beperking van de rol van KUA. Het houdt in dat KUA-informatie de basis vormt voor prioriteringsbeslissingen, maar dat daarnaast ook andere overwegingen een rol kunnen spelen in de besluitvorming. Deze beide voorstellen worden vervolgens geconfronteerd met de recente analyse van de Gezondheidsraad over de mogelijke rol van KUA (2.6). Het signalement sluit af met agendapunten (2.7).

2.2 Kostenutiliteitsanalyse

De term 'kostenutiliteitsanalyse' (KUA) staat in het gangbare gezondheidseconomische spraakgebruik voor een doelmatigheidsanalyse die voor het kwantificeren van de opbrengst van een gezondheidszorgvoorziening een enkele maat hanteert die

1 toepasbaar is op de hele gezondheidszorg
2 uitdrukking geeft aan de waarde die aan het totaal van alle gezondheidseffecten van die voorziening wordt toegekend (De Neeling 2004).

2 Sinds 1 april 2005: National Institute for Health and Clinical Excellence (NICE): www.nice.org.uk

Methode

Meestal is die maat de QALY. De centrale notie achter de QALY is dat gezondheidszorgvoorzieningen effecten kunnen hebben op zowel de levensduur als de levenskwaliteit van mensen en dat die beide aspecten tegelijk in de evaluatie betrokken moeten worden. De QALY-maatstaf maakt dat mogelijk door bij de telling van het aantal gewonnen levensjaren een wegingsfactor te hanteren voor de kwaliteit daarvan. Bijvoorbeeld: 10 gewonnen levensjaren met een kwaliteitsscore van 0,9 leveren evenveel QALY's op als 9 jaar in volledige gezondheid.

Om de kwaliteit van gewonnen levensjaren te kunnen vangen in een cijfer, worden in de regel twee stappen gezet. Ten eerste worden op basis van vragenlijstonderzoek bij patiënten gezondheidstoestanden beschreven. Ten tweede kennen respondenten uit de algemene bevolking aan die toestanden een waardering toe. Voor die waarderingsmeting wordt gebruik gemaakt van psychometrische methoden of aan de respondenten voorgelegde hypothetische keuzesituaties. Voorbeelden van dat laatste zijn *standard gamble* (welk overlijdensrisico zou men nog willen aanvaarden in ruil voor de garantie bij overleving volledig genezen te zijn?) en *time trade-off* (een hoeveel korter leven in goede gezondheid zou men nog te verkiezen vinden boven een langer leven met de ziekte of aandoening in kwestie?).

Op basis van QALY-uitkomsten kan men vervolgens de (incrementele) kostenutiliteit van uiteenlopende voorzieningen vergelijken. Men kijkt dan naar de meerkosten gedeeld door de (in QALY's uitgedrukte) meeropbrengst, in vergelijking met een relevant alternatief. Meestal is dat de gebruikelijke zorg. Door op deze manier zichtbaar te maken hoe verschillende voorzieningen zich op een kostenutiliteitsschaal ('QALY league table') tot elkaar verhouden, wordt het mogelijk – dat is althans de claim – om prioriteringsbeslissingen van een objectieve onderbouwing te voorzien. Iedere uit de collectieve middelen voor de gezondheidszorg te besteden euro kan dan in principe zo worden ingezet dat die de meeste gezondheid (preciezer: gezondheidsgerelateerde levenskwaliteit) oplevert.

Toepassing

In 1997 heeft de Wetenschappelijke Raad voor het Regeringsbeleid er in zijn rapport 'Volksgezondheidszorg' voor gepleit deze aanpak te hanteren bij besluitvorming over het verstrekkingenpakket (WRR 1997). Waar het gaat om preventieve voorzieningen zou daarbij te streven zijn naar maximalisering van (in DALYs uitgedrukte) gezondheidswinst. Bepleit werd om verschillende preventieve programma's in volgorde van afnemende kostenutiliteit te financieren, tot het beschikbare budget is uitgeput. In de curatieve zorg zou gestreefd moeten worden naar een egalisering van kostenutiliteitsverhoudingen over alle categorieën voorzieningen heen. Daarbij zou voor het realiseren van één DALY, ongeacht waar in de gezondheidszorg die wordt verkregen, steeds eenzelfde maximumbedrag moeten gelden.

In het Verenigd Koninkrijk wordt een dergelijke aanpak van gezondheidszorgbrede prioritering inmiddels in de praktijk gebracht. Het eerder genoemde NICE baseert zijn 'appraisals' waar mogelijk op informatie over de incrementele kostenutiliteit (uitgedrukt in kosten per QALY) van een nieuwe voorziening. Daarbij ligt de aanvaardbaarheidsgrens voor opname in de NHS bij ratio's van £25.000 - £35.000/QALY[3]. Die grens wordt overigens niet gepresenteerd als een absolute drempel; wel geldt dat bijzondere redenen nodig zijn voor de aanvaarding van technologieën met een nog hoger (en dus ongunstiger) uitvallende kostenutiliteitsratio (Rawlins 2004).

Reden voor de instelling van NICE was de constatering dat de besluitvorming over de introductie van nieuwe voorzieningen in het Verenigd Koninkrijk weinig consistent was (wat leidde tot geografische toegangsverschillen, het zogenoemde 'postcode prescribing') en wetenschappelijk onvoldoende onderbouwd. Behalve kritiek heeft NICE inmiddels veel waardering gekregen, en het heeft ook internationaal een sterke wetenschappelijke reputatie opgebouwd. De Britse aanpak wordt elders als model gezien en 'may prove to be one of Britain's greatest cultural exports, along with Shakespeare, Newtonian physics, the Beatles, Harry Potter, and the Teletubbies' (Smith 2004).

Ook in Nederland is bepleit NICE als voorbeeld te nemen bij het streven naar een meer doelmatige inzet van nieuwe medische technologie (Rutten 2002). De Gezondheidsraad deed dat in zijn advies *Contouren van het basispakket* (Gezondheidsraad 2003). Volgens de raad kunnen de door NICE gehanteerde aanpak en de daarmee in het Verenigd Koninkrijk opgedane ervaringen bruikbare aanknopingspunten bieden bij de opzet van een nationale beoordelingsstructuur voor pakketbeslissingen. Zo'n beoordelingskader ontbreekt in ons land, al zijn elementen ervan bij verschillende organisaties terug te vinden. Hoewel de uiteindelijke beslissing over de toelating tot het pakket een zaak van de overheid blijft, zouden de adviezen dan wel zwaar moeten wegen.

Wat betreft de rol van KUA als instrument voor die beoordeling en advisering is de Gezondheidsraad echter terughoudend, in ieder geval waar het zou gaan om prioritering van voorzieningen over de volle breedte van de gezondheidszorg. Een dergelijke aanpak, zoals aanbevolen door de WRR en in de praktijk gebracht door NICE, noemt de commissie 'vooralsnog precair en wetenschappelijk betwistbaar' (Gezondheidsraad 2003). Voor deze stellingname baseert zij zich op een uitvoerige achtergrondstudie bij haar advies (De Neeling 2003). Het daarin gegeven overzicht van de actuele stand van de wetenschappelijke discussie maakt duidelijk dat de methodologie van de kostenutiliteitsanalyse nog veel haken en ogen kent, die nopen tot voorzichtigheid bij de toepassing.

In zijn advies *Grenzen van de Gezondheidszorg* uit 1986 had de Gezondheidsraad zich overigens al eerder in diezelfde zin uitgelaten: de resultaten van KUA zouden '*behulpzaam* kunnen zijn bij het verhogen van de kwaliteit van de besluitvorming, maar meer niet' (Gezondheidsraad 1986). Een rangordening naar mate van kostenutiliteit zou immers kunnen leiden tot willekeur. Daarbij

3 In een recent consultatiedocument worden lagere bedragen genoemd: £20.000 - £30.000 (NICE 2005).

doelde de Raad op onevenwichtige of inconsistente toepassing van dat instrument, maar benadrukte ook dat kostenutiliteit niet de enige maatstaf kan zijn voor de opname van voorzieningen in het pakket.

2.3 KUA en rechtvaardigheid

Prioritering van gezondheidszorgvoorzieningen veronderstelt een opvatting over verdelende rechtvaardigheid. Waar prioritering is gericht op maximalisatie van in QALY's uitgedrukte gezondheidswinst, is dat het utilisme (McKie 1998). Daarbij geldt als verdelingsprincipe de regel van Bentham: 'everybody to count for one, nobody for more than one'. Hier wil dit zeggen dat, in de calculatie van gezondheidswinst en -verlies, ieders QALY's even zwaar wegen. Als rechtvaardigheidstheorie is het utilisme echter omstreden. Overwegingen die samenhangen met andersoortige intuïties over wat een rechtvaardige verdeling zou zijn, worden – zo luidt de kritiek – bij voorbaat ondergeschikt gemaakt aan het streven naar maximalisatie van het nagestreefde goed, in deze context: gezondheid of: gezondheidgerelateerde levenskwaliteit.

Verdelingseffecten

In de gezondheidseconomische en ethische literatuur is veel aandacht besteed aan mogelijk onrechtvaardige verdelingseffecten van KUA. De belangrijkste thema's in die discussie kunnen als volgt worden samengevat (het onderstaande overzicht is grotendeels ontleend aan De Neeling 2003):

— Door hun kortere levensverwachting hebben ouderen (*ceteris paribus*) minder QALY's te winnen dan jongeren. Voorzieningen waarvan met name ouderen gebruik zullen maken komen daardoor al snel lager uit in een op kostenutiliteit gebaseerde rangorde. Verdelingsbeslissingen op basis van KUA bevoordelen dus jongeren boven ouderen. Over de vraag of dat gerechtvaardigd is wordt verschillend gedacht. Welke morele betekenis moet worden gehecht aan het feit dat iedereen mag verwachten de verschillende levensfasen van jeugd tot ouderdom te doorlopen? Volgens het zogenoemde 'fair innings' argument laat de kortere levensverwachting die mensen boven een bepaalde leeftijd nog hebben, alleen maar zien dat zij in beginsel hun aandeel in de kansen die het leven biedt hebben gehad.
— Een gelijksoortig effect doet zich voor op de kwaliteitsdimensie, bij mensen met chronische ziekten, handicaps en permanente functiebeperkingen: zij hebben bij een bepaald aantal jaren levensverlenging minder QALY's te winnen dan mensen zonder zulke ziekten en beperkingen. Dat komt doordat hun initiële gezondheidstoestand de marge bepaalt voor de voor hen te behalen gezondheidswinst. Herstel van volledige gezondheid zit er al bij voorbaat niet in. Prioritering op basis van KUA kan er dan toe leiden dat voorzieningen voor juist deze patiëntengroep minder hoog scoren in termen van kostenutiliteit. Mensen die op dergelijke voorzieningen zijn aangewezen, worden daardoor benadeeld.
— KUA let niet op de ernst van aandoeningen. Weliswaar speelt het begrip behoefte (*need*) een centrale rol, maar dat wordt in de gezondheidseconomische literatuur opgevat in termen van

iemands vermogen om van behandeling te profiteren ('capacity to benefit'). In dat opzicht ziet KUA geen verschil tussen een verandering in de gezondheidstoestand van een ernstig zieke en een in termen QALY's even grote verandering in de toestand van een vrijwel gezond persoon. De vraag valt te stellen of dat rechtvaardig is. De KUA-systematiek houdt ook geen rekening met wat wel de 'rule of rescue' is genoemd: de overtuiging dat kosten noch moeite gespaard mogen worden wanneer iemand hier en nu van een dreigende dood te redden is.

— Minder vaak bediscussieerd, maar niet minder juist, is de vaststelling dat, behalve leeftijd, ook andere demografische kenmerken, zoals geslacht, ras en sociaal- economische status, krachtige voorspellers zijn van het resultaat van een medische behandeling, zodat een verdeling van middelen op basis van KUA op gespannen voet staat met het streven naar een voor iedereen, ongeacht deze kenmerken, gelijke toegankelijkheid van gezondheidszorg.
— Verdelingseffecten vloeien ook voort uit de wijze waarop in KUA het aggregatieprobleem wordt opgelost. KUA belichamen, door de in QALY's gemeten gezondheidsbaten van verschillende personen bij elkaar op te tellen, een specifiek – en discutabel – antwoord op vragen als hoe de geredde levens van enkelingen afgewogen dienen te worden tegen de genezen keelontstekingen van velen.
— Door schaaleffecten bij de productie van medicijnen etc. wordt behandeling van een zeldzame aandoening duurder dan die van een veelvoorkomende. In de KUA-systematiek leidt dat (*ceteris paribus*) tot een lagere prioritering van voorzieningen voor mensen die lijden aan een zeldzame aandoening.
— De vraag of en hoe in de KUA-systematiek de gevolgen moeten worden betrokken van ziekte en behandeling voor iemands vermogen deel te nemen aan het arbeidsproces, is meer dan een louter technische kwestie. Meetellen van die productiviteitskosten of-opbrengsten kan immers leiden tot een relatief ongunstiger beoordeling van voorzieningen die met name van belang zijn voor mensen buiten het arbeidsproces, zoals ouderen, gehandicapten of kinderen.
— Het verdisconteren van de verloren vrije tijd van patiënten en informele zorgverleners, het waarderen van onbetaald werk, de wijze van omgaan met leeftijds- en geslachtsafhankelijke verschillen in beloning: het zijn onderdelen van het kostenonderzoek in KUA waar verschillend mee omgesprongen wordt, met onmiskenbare verdelingseffecten als gevolg.

Prioritering van voorzieningen op grond van kostenutiliteitsverhoudingen, zo luidt hier de conclusie, leidt tot een verdeling van gezondheidswinst die niet vanzelfsprekend ook de meest rechtvaardige is. In de meer recente gezondheidseconomische literatuur wordt dat ook onderkend. Men spreekt van de noodzaak van '*equity-efficiency trade-offs*', wat wil zeggen dat het nodig kan zijn om op grond van rechtvaardigheidsoverwegingen af te zien van maximalisatie van gezondheidswinst (Garber 1996). Van een strikt utilistische benadering is dan geen sprake meer.

Implicaties voor de rol van KUA

Wat betekent dat voor KUA? Er zijn hier in grote lijnen twee benaderingen (Powers 2000). Volgens de eerste moeten rechtvaardigheidsoverwegingen afzonderlijk in de besluitvorming aan de

orde worden gesteld. Zo is bepleit in economisch evaluatieonderzoek steeds expliciet aan te geven bij welke subpopulaties de gezondheidsbaten van een voorziening terecht komen, zodat besluitvormers inzicht houden in de verdelingseffecten van op kostenutiliteit gebaseerde keuzes (Sassi 2001). Op grond van door hen relevant geachte rechtvaardigheidsoverwegingen kunnen zij daar dan voor corrigeren. Welke overwegingen dat zijn, en welk gewicht eraan moet worden toegekend, is een kwestie van maatschappelijk en moreel beraad. Deze benadering impliceert een inperking van de mogelijke rol van KUA bij de prioritering van gezondheidszorgvoorzieningen. Daarover gaat het in paragraaf 1.5.

Veel gezondheidseconomen vinden dat onbevredigend (Williams 2001). Zij zien als bezwaar dat de besluitvorming daarmee aan precisie zou verliezen: 'the great deficiency of that approach is that it leaves considerations of equity floating without any conception of how important they would be' (Menzel 1999). Het door hen bepleite alternatief is verbreding van KUA op basis van gekwantificeerde informatie over de *equity trade-offs* die mensen bereid zijn te aanvaarden (Dolan 2002). Als eerste wordt in paragraaf 1.4 die benadering besproken.

2.4 Verbreding: KUA met ingebouwde 'equity weights'

Voor die verbreding van KUA zijn meetinstrumenten ontwikkeld van het type *person trade-off* (PTO*)*. Dat wil zeggen dat respondenten wordt gevraagd de in hypothetische keuzesituaties aan verschillende personen of groepen toevallende *benefits* tegen elkaar af te wegen. Bij sommige van de voor de QALY voorgestelde alternatieven maakt de PTO deel uit van de constructie van de maatstaf (DALY, SAVE) waarmee de opbrengst van voorzieningen wordt vergeleken (Murray 1997, Nord 1992). Dat samenlopen, in een enkele maat, van informatie over de waardering van gezondheidswinst en preferenties ten aanzien van verdelingsprincipes is echter meteen ook de reden waarom de onderbouwing van die alternatieven, in vergelijking met de QALY, als minder stevig geldt (De Neeling 2003).

'Societal value measurement'

De meest verdedigde poging tot verbreding van KUA verandert niets aan de wijze waarop QALY's worden berekend, maar houdt in dat op basis van PTO-informatie wegingsfactoren ('equity weights') worden opgesteld voor QALY's die terechtkomen bij verschillende categorieën personen: jongeren of ouderen, ernstig zieke of vrijwel gezonde personen, etc (Dolan 2005, Garber 1996, Menzel 1999). Bij een door de Gezondheidsraad over de ethische aspecten van KUA georganiseerde bijeenkomst van Nationale Ethiek Commissies in landen van de EU (december 2004) vertelde de Britse gezondheidseconoom Dolan over zijn (door NICE gefinancierde) onderzoek op dit gebied:

> Citizens are capable of making judgements about whether one QALY to a twenty year old is worth five QALY's to a fifty year old. They are the kind of questions that are funded by NICE to now take forward: think of a treatment that would improve the health of a young person by a small amount or

> an old person by a large amount, and think about trading-off the health gains between those two, think about the point at which the small benefit equals the large benefit. I think people are capable of expressing views over that. And that kind of preference information (...) is the sort of information that should (I make a normative statement here) be used to inform public policy (Gezondheidsraad 2005).

Er zijn met die verbreding van KUA twee soorten problemen: is het mogelijk via dergelijke 'societal value measurement' betrouwbare informatie te krijgen over wat mensen als rechtvaardig beschouwen? En als dat kan, wat volgt daar dan eigenlijk uit voor de vraag wat wel of niet een rechtvaardige prioritering van gezondheidszorgvoorzieningen zou zijn?

Dat eerste probleem wordt door auteurs als Dolan nadrukkelijk onderkend. Mede gelet op de complexiteit van de keuzesituaties die aan respondenten worden voorgelegd, waarbij niet gemakkelijk in één beoordelingsperspectief te vangen morele overwegingen in het geding zijn, is het nog maar de vraag in hoeverre via PTO betrouwbare preferentie-informatie te verkrijgen is. Er moet terdege rekening gehouden worden met beïnvloeding van de meting (de gevonden preferenties) door de formulering van de vraag of door andere aspecten van de meetprocedure.

Ubel wijst in dat verband op de uitkomsten van onderzoek naar dergelijke effecten en waarschuwt dat 'we should be cautious about interpreting preferences for fairness versus efficiency until such framing effects are more clearly understood' (Ubel 2000). Het zou overigens een misverstand zijn te denken dat het hier uitsluitend om meettechnieken zou gaan. Het probleem is ook van conceptuele aard: onder welke condities geven de antwoorden van respondenten inzicht in hun werkelijke preferenties? (Powers 2000).

Morele verantwoording

Maar het belangrijkste probleem is de verantwoording van de aan die antwoorden te verbinden normatieve conclusies. Met de suggestie dat de gezochte wegingsfactoren kunnen worden afgeleid van wat de meeste respondenten blijken te vinden (de keuze van de 'median voter'; Williams 2001), blijven we in utilistisch vaarwater. De in moreel opzicht beste uitkomst is dan die waarin zoveel mogelijk mensen hun voorkeuren inzake verdelingsbeslissingen vervuld zien.

Tegen deze 'aggregatieve' benadering is (onder meer door Daniels & Sabin) ingebracht dat ze geen recht doet aan de fundamentele verschillen tussen morele overtuigingen en preferenties (Daniels 1997, Daniels 2002). Morele overtuigingen ontlenen hun gewicht aan de redenen waarop ze zijn gebaseerd, preferenties (zoals gehanteerd in de utilistische calculus) aan het loutere feit dat iemand een voorkeur heeft voor A boven B. Over preferenties valt niet van mening te verschillen, over de redenen achter morele overtuigingen wel.

Het probleem van de aggregatieve benadering is dat die redenen er niet meer toe doen. Dat de meerderheid het in moreel opzicht bij het verkeerde eind zou kunnen hebben, kan dan niet eens

zinvol meer aan de orde worden gesteld. Om dezelfde reden schiet die benadering tekort als democratische procedure voor het beslechten van morele conflicten: 'it ignores the way in which the search for reasons we can agree on plays a central role in our deliberations about what is right' (Daniels 1997).

Dat ook volgens Dolan en zijn collega's een morele analyse niet uitsluitend op het meten van preferenties kan zijn gebaseerd, blijkt uit het antwoord op de vraag of bij hun aanpak niet het gevaar bestaat dat in de samenleving bestaande vooroordelen jegens sociale groepen of negatieve beeldvorming rond specifieke aandoeningen de in KUA op te nemen wegingsfactoren gaan bepalen:

> (..) of course those social preferences are not elicited in a policy or moral vacuum. We first need – and this is a normative question preceding any survey – to establish which attributes are policy and morally relevant. So for example we would not ask members of the UK general population to choose between gay and straight, black and white. In that sense there has already some laundering been done. We are not just going to the public and ask them: tell us what you think about anything. We ask them to think about what we consider to be the normative framework of attributes that we should focus upon (Dolan in Gezondheidsraad 2005).

Daarmee wordt echter niet meer gezegd dan dat de aanvaardbaarheid van de aan respondenten te ontlokken preferenties begrensd wordt door in de samenleving geldende morele regels: 'public preferences can be unjust, and therefore the way we incorporate them into health care rationing should be constrained by principles of justice' (Ubel 2000). Binnen die begrenzing is iedere uitkomst even goed als elke andere.

In een poging tot nadere morele onderbouwing van de in KUA op te nemen wegingsfactoren voor QALY's hebben Edgar *et al* bepleit niet te volstaan met het voorleggen van PTO-vragen aan individuele respondenten, maar hen te betrekken in een normatief discours, als 'subjects who can debate, defend and revise their preferences' (Edgar 1998). Een dergelijke benadering wordt ook toegepast bij de deels op uitkomsten van PTO-vragen gebaseerde DALY (Murray 1997). Het idee is om via een vorm van groepsdeliberatie antwoorden op dergelijke vragen te verkrijgen die in de samenleving breed worden gedeeld. Omdat echter niet alle betrokkenen (bij verdelingsbeslissingen in de gezondheidszorg zijn dat alle burgers) zich adequaat in dergelijke procedures vertegenwoordigd kunnen weten, blijft een op de uitkomsten daarvan gebaseerde kwantificering van rechtvaardigheidsoverwegingen moreel aanvechtbaar (Powers 2000, Daniels 2002).

De Amerikaanse ethicus Brock onderstreept dat de voor KUA relevante verdelingsvragen 'represent deep and long-standing divisions in moral and political philosophy about which there is not now, and may never be, anything approaching consensus' (Brock 2004). Met de eerder aangehaalde Daniels & Sabin (Daniels 2002), en net als Van der Wilt (Gezondheidsraad 2005), ziet Brock als enige begaanbare route dergelijke waardeconflicten tot voorwerp te maken van transparante en (afhankelijk van context en schaal) democratische of participatoire deliberatieve

besluitvormingsprocedures. De PTO-methode kan daarbij een nuttig hulpmiddel zijn, namelijk door 'structuur en precisie' te geven aan de in het geding zijnde opvattingen en waarden. Maar dit instrument kan niet in de plaats komen van morele deliberatie over het aan elk daarvan te geven gewicht (Brock 2004).

2.5 Inperking: KUA als basis voor besluitvorming

Bij zijn besluitvorming over prioritering is het Britse NICE tot nu toe uitgegaan van de klassieke KUA-regel dat aan iedere QALY hetzelfde gewicht toekomt. Rawlins en Culyer, voorzitter en voormalig vice-voorzitter van het instituut, verantwoorden dat beleid als volgt:

> The decision to give no differential weight is the result of a social value judgement that an additional adjusted life year is of equal importance to each person' (Rawlins 2004).

Dat wil niet zeggen dat de beoordelingen van NICE zijn gebaseerd op een strikt mechanische toepassing van het criterium van kostenutiliteit. Bijzondere redenen, waaronder contextbepaalde overwegingen van rechtvaardigheid, kunnen een grond zijn om voorzieningen met een op zichzelf onaanvaardbare (incrementele) kostenutiliteitsratio toch te vergoeden. Rawlins en Culyer noemen als voorbeeld een situatie waarin door eerder beleid bij patiënten een gerechtvaardigde verwachting was gewekt een bepaald geneesmiddel (vergoed) te kunnen krijgen (Rawlins 2004).

Aanvullende overwegingen

Recent is echter door NICE een discussie gestart over de vraag of aanpassing van de klassieke KUA misschien toch wenselijk is. In een dit voorjaar (April 2005) gepubliceerd consultatiedocument valt op dat de zojuist geciteerde (en op zijn minst discutabele) rechtvaardiging voor het afwijzen van gewichtsdifferentiatie ontbreekt. Weliswaar wordt het inbouwen van 'equity weights' in KUA, zoals voorgesteld door hierboven genoemde auteurs, nog steeds van de hand gewezen, maar nu omdat dit 'at best premature' zou zijn en wellicht zelfs onmogelijk gezien 'real difficulties in obtaining and legitimising public preferences' (NICE 2005). Uit deze afwijzing van een in KUA ingebouwde kwantitatieve differentiatie volgt echter niet dat maatschappelijke waardeoordelen geen rol kunnen spelen bij de besluitvorming. Of, wanneer en waarom dat wenselijk zou zijn, is de centrale vraag in het hier aangehaalde consultatiedocument.

De met de publicatie van dit document gestarte discussie moet uiteindelijk leiden tot richtlijnen voor het zo consistent mogelijk betrekken (of niet betrekken) van dergelijke waardeoordelen bij door NICE op te stellen aanbevelingen. Dat past ook in het streven naar grotere publieke verantwoording:

> While the members of NICE's advisory bodies are appointed for their competence in making scientific value judgements neither they, nor the board, have any legitimacy to impose their own social value judgements on the NHS and the patients that the service seeks to serve (NICE 2005).

Het consultatiedocument noemt in dat verband als centrale voorwaarde voor legitieme besluitvorming dat die is gebaseerd op redenen waarvan de relevantie door 'fair-minded people' kan worden onderschreven. Andere voorwaarden zijn: transparantie en de aanwezigheid van een mechanisme voor revisie en het maken van bezwaar (Daniels 2002). Met deze expliciete verwijzing naar het ideaal van 'deliberatieve democratie', wordt meteen ook afstand genomen van de suggestie dat het voor legitieme besluitvorming voldoende zou zijn om af te gaan op de uitkomsten van 'societal value measurement'.

Om invulling te geven aan dit streven is enkele jaren geleden een 'Citizens Council' ingesteld, bestaande uit een zo breed mogelijk samengestelde groep van dertig personen. De *Citizens Council* heeft tot taak NICE desgevraagd van advies te dienen in normatieve kwesties. Hij heeft zich inmiddels gebogen over de vraag welke mogelijke attributen van patiënten (zoals leeftijd) of condities (zoals ernst) wel of niet een reden kunnen zijn om voorzieningen anders te beoordelen dan uitsluitend op hun kostenutiliteit. Zoals in het consultatiedocument wordt aangegeven, was het oordeel van de *Citizens Council* terughoudend. Na extensief beraad zag de groep alleen reden om ruimte te maken voor een afwijkende beoordeling bij levensreddende ingrepen (*rule of rescue*), comorbiditeit en 'conditions associated with stigma' (NICE 2005).

De (voorlopige) aanbevelingen van NICE zelf zijn nog aanzienlijk terughoudender. Zo luidt de aanbeveling voor de *rule of rescue* dat bij eventuele toepassing 'grote voorzichtigheid' moet worden betracht. En in geval van condities met een maatschappelijk stigma zou afwijking van een op KUA gebaseerde beoordeling alleen te overwegen zijn als sprake is van een psychosociale belasting die niet ook al in die analyse is verdisconteerd. Ruimte voor 'special consideration' zou er juist wèl moeten zijn waar het gaat om innovatieve interventies die leiden tot belangrijke gezondheidswinst bij tot dan toe onbehandelbare aandoeningen (NICE 2005).

Wetenschappelijke en maatschappelijke waarden

Bij de instelling van de *Citizens Council* kwam de Britse patiëntenvereniging met scherpe kritiek op de beperkte rol van die groep. Een woordvoerder sprak van een 'token gesture', zolang het publiek niet daadwerkelijk betrokken wordt in het hele proces van de besluitvorming (Gulland 2002). Die kritiek raakt aan het door NICE gehanteerde onderscheid tussen wetenschappelijke en maatschappelijke waarden (Rawlins 2004, NICE 2005):

> Scientific value judgements are concerned with interpreting the significance of the available scientific, technical and clinical data. Social value judgements relate to society rather than basic or clinical science: they take account of the ethical principles, preferences, culture and aspirations that should underpin the nature and the extent of the care provided by the NHS. Nevertheless, the distinction is not absolute: measuring and understanding value judgements have a scientific dimension but do not form part of this document (NICE 2005)

Het is op zijn minst opmerkelijk dat hier niet op volgt dat het onderscheid ook in de tegenovergestelde richting moet worden gerelativeerd: die wetenschappelijke interpretatie van de feiten is vaak gebaseerd op normatieve keuzes. Weliswaar wordt elders in het consultatiedocument opgemerkt dat (maatschappelijke) waardeoordelen een rol spelen in de wetenschappelijke constructie van QALY's en kostenutiliteitsverhoudingen, en ook dat NICE zich bewust is dat daarbij aan relevante aspecten kan worden voorbijgezien, maar meer dan een zijdelingse constatering is dat niet.

Die constatering blijft zonder gevolgen voor hoe de rol van KUA wordt gezien, namelijk als 'a necessary, but not sufficient, basis for decision-making'. De in het consultatiedocument aan de orde gestelde discussie betreft de invulling van dat 'not sufficient': het gaat daarin om de vraag of er wel of niet reden kan zijn in sommige gevallen naast KUA-informatie nog andere overwegingen in de besluitvorming te betrekken. De rol van KUA als basis voor de besluitvorming staat echter niet ter discussie.

2.6 Verborgen verdelingseffecten

De opvatting dat KUA als basis voor de besluitvorming kan gelden, lijkt in moreel opzicht minder problematisch dan het voorstel tot verbreding via 'equity weights'. De vraag hoe de bij een bepaald budget te behalen gezondheidswinst moet worden verdeeld, kan immers expliciet tot voorwerp worden gemaakt van maatschappelijke deliberatie en besluitvorming. Beide benaderingen gaan echter voorbij aan de verborgen verdelingseffecten die onvermijdelijk samenhangen met de theoretische en methodologische onvolkomenheden en onzekerheden van KUA.

Onvolkomenheden en onzekerheden van KUA

Daarop is gewezen in de twee jaar geleden door de Gezondheidsraad gepubliceerde achtergrondstudie en de mede daarop gebaseerde conclusies in het advies van de Commissie 'Contouren van het basispakket' (Gezondheidsraad 2003). De achtergrondstudie bestaat uit een uitvoerige bespreking van de relevante wetenschappelijke literatuur (De Neeling 2003). Die laat zien dat sprake is van een groot aantal nog onopgeloste problemen, zowel wat betreft de theoretische onderbouwing van de QALY als meest courante maatstaf voor de opbrengst van gezondheidszorgvoorzieningen, als wat betreft de verschillende elementen in de constructie van de daarop gebaseerde analyse.

Het gaat hier onder meer om:
- problemen bij het integreren van 'kwaliteit' en 'kwantiteit' van leven en bij het aggregeren van QALY's van verschillende personen,
- de betrouwbaarheid van de meetinstrumenten en de vergelijkbaarheid van de daarmee verkregen uitkomsten, zowel bij de beschrijving van gezondheidstoestanden als bij de waardering daarvan,

— de vraag aan wie de waarderingsvragen het beste kunnen worden gesteld (patiënten of steekproef uit de algemene bevolking) en welke theoretische en praktische betekenis moet worden gehecht aan de verschillen tussen de antwoorden die dat oplevert,
— de interculturele (on)vergelijkbaarheid van meetresultaten en de constatering dat de voorkeuren van respondenten deels in interactie met de vraagstelling hun uiteindelijke vorm krijgen ('framing effects'),
— de vraag welke kosten wel en niet in KUA moeten worden betrokken, hoe ze moeten worden geschat en van welk kostenbegrip (gemiddeld of marginaal) daarbij moet worden uitgegaan,
— de vraag of voor toekomstige QALY's eenzelfde discontopercentage moet worden gehanteerd als voor toekomstige kosten.

De analyse in de achtergrondstudie laat zien waarom het in de Britse discussie gehanteerde onderscheid tussen wetenschappelijke en maatschappelijke (morele) waarden problematisch is. Achter tal van nog onbeantwoorde theoretische en methodologische kwesties gaan bij nadere beschouwing normatief geladen keuzes schuil, zoals ook wordt benadrukt door de al eerder aangehaalde Brock:

> There are important ethical and value choices to be made in constructing and using the measures; the choices are not merely technical, empirical, or economic, but moral and value choices as well. Each requires explicit attention by health policy-makers (Brock 2004).

Het valt daarom niet vol te houden dat de vraag naar de ethische aspecten van KUA pas begint bij de op zichzelf overzichtelijke kwestie van de verdeling van in QALY's uitgedrukte gezondheidswinst. Dan blijft namelijk buiten beschouwing dat de onzekerheden over de juiste methode voor het kwantificeren van de kostenutiliteitsverhouding zelf al allerlei verborgen verdelingseffecten kunnen hebben:

> Een gezondheidstoestandmeting die ongevoelig is voor de ziektespecifieke effecten van een bepaalde handeling, een steekproef die het subtiele verschil tussen de gezondheidstoestanden voor en na de behandeling onvoldoende op waarde weet te schatten, een waarderingsmethode die door de *reluctance to give up life* van de respondenten het verschil in utiliteit tussen twee toestanden onvoldoende tot uitdrukking brengt: al die en vele andere factoren kunnen ertoe leiden dat een KUA de gezondheidswinst die een bepaalde gezondheidszorgvoorziening teweeg brengt, onder- of juist overschat (De Neeling 2003).

Als op grond van die informatie wordt besloten een voorziening niet te financieren, wordt de patiëntengroep die er anders baat bij zou hebben gehad, benadeeld. Maar omdat van de meeste onzeker- en onvolkomenheden van KUA niet te zeggen is hoe ze precies zullen uitpakken, zijn dergelijke verdelingseffecten niet goed in kaart te brengen.

Conclusies

De achtergrondstudie loopt uit op de volgende conclusies:

— Aanvaard moet worden dat voor de vergelijking van de doelmatigheid van gezondheidszorgvoorzieningen geen algemeen toepasbare technische middelen voorhanden zijn. Het is daarom noodzakelijk dat besluitvormers zelf een afweging maken waarin alle relevante gevolgen en overwegingen betrokken zijn.

— Afbakening van deelgebieden in de gezondheidszorg lijkt noodzakelijk om de complexiteit van de besluitvorming hanteerbaar te maken, al betekent dat onvermijdelijk dat de verdeling over het geheel van de gezondheidszorg ongeregeld blijft.

— Bij prioriteitsstelling binnen afzonderlijke deelgebieden zullen de onvolkomenheden en onzekerheden van KUA minder zwaar wegen naarmate de diversiteit van de vergeleken voorzieningen, gezondheidsproblemen, patiëntenpopulaties en gezondheidsuitkomsten geringer is. Overigens zal in meer homogene domeinen vaak ook kunnen worden volstaan met op meer specifieke uitkomstmaten gebaseerde kosteneffectiviteitanalyses. Vergelijking in termen van QALY's is dan vaak niet nodig.

— Gelet op de legitimiteit van de besluitvorming zal het echter vooral zaak zijn om een transparante besluitvormingsprocedure te ontwerpen, waarin de partijen die de gezondheidszorg van binnenuit kennen, zorgverleners en patiënten voorop, actief participeren. KUA kan in dat besluitvormingsproces een hulpmiddel zijn om overzicht en overeenstemming te bereiken.

2.7 Agendapunten

— Met het oog op de, mede naar het voorbeeld van het Britse NICE, in ons land op te zetten 'beoordelingsstructuur voor pakketbeslissingen', zijn deze conclusies onverminderd relevant. Onderkend moet worden dat KUA niet in de plaats kan komen van een transparant, deliberatief en democratisch besluitvormingsproces.

— Daarmee is overigens het belang van KUA en de verdere ontwikkeling daarvan niet afgedaan. Voor zover KUA het mogelijk maakt in de vergelijking van de opbrengsten van voorzieningen ook kwaliteit-van-leven overwegingen een plaats te geven, kan zij bijdragen aan een verbreding en verscherping van het overzicht van wat er in de besluitvorming over prioritering op het spel staat.

— Evaluatie van de KUA-praktijk is van belang om meer inzicht te verkrijgen in de daadwerkelijke rol die KUA, als hulpmiddel bij de besluitvorming, kan spelen bij beslissingen over prioritering in de gezondheidszorg. Bij het (multidisciplinaire) onderzoek daarnaar dient ook de ethiek nadrukkelijk te worden betrokken.

Literatuur

Brock DW. Ethical issues in the use of cost effectiveness analysis for the prioritisation of health care resources. In: Anand S, e.a., editors. Public Health, ethics, and equity. OUP; 2004: 201-223.

Daniels N, Sabin J. Limits to health care: fair procedures, democratic deliberation, and the legitimacy problem for insurers. Philos Public Aff 1997; 4: 303-50.

Daniels N, Sabin J. Setting limits fairly. Can we learn to share medical resources? Oxford: OUP, 2002.

De Neeling JN. Kostenutiliteitsanalyse. Den Haag: Gezondheidsraad; 2003; publicatie nr A03/01.

De Neeling JN. Kostenutiliteitsanalyse; onzekerheden beperken toepasbaarheid. Ned Tijdschr Geneeskd 2004; 148(33): 1655-1656.

Dolan P, Olsen JA. Distributing health care; economics and ethical issues. Oxford: OUP; 2002.

Dolan P, Shaw R, Tsuvchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. Health Econ 2005; 14: 197-208.

Edgar A, Salek S, Shickle D, e.a. The Ethical QALY: Ethical issues in healthcare resource allocations. Haslemere: Euromed Communications, 1998.

Garber AM, Weinstein MC, Torrance GW, Kamlet MS. Theoretical foundations of cost-effectiveness analysis. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, red. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996: 25-53.

Gezondheidsraad. Grenzen van de gezondheidszorg. Den Haag: Gezondheidsraad, 1986; publicatie nr 1986/ 29.

Gezondheidsraad. Contouren van het basispakket. Den Haag: Gezondheidsraad; 2003; publicatie nr 2003/02.

Gezondheidsraad. Ethical Aspects of Cost-utility Analysis. 4th Forum of National Ethics Councils in EU Countries, Amsterdam, 22 December 2004. Presentations and discussion. The Hague: Health Council of the Netherlands, 2005; publicatie nr A05/ 01.

Gulland A. NICE proposals for citizens council condemned by patients. National Institute for Clinical Excellence. BMJ 2002; 325(7361): 406.

Maynard A. Ethics and health care 'underfunding'. J Med Ethics 2001; 27(4): 223-227.

McKie J, Richardson J, Singer P, e.a. The allocation of health care resources. An ethical evaluation of the 'Qaly' approach. Aldershot: Ashgate/Dartmouth, 1998.

Menzel P, Gold MR, Nord E, Pinto-Prades JL, Richardson J, Ubel P. Toward a broader view of values in cost-effectiveness analysis of health. Hastings Cent Rep 1999; 29(3): 7-15.

Murray CJ, Acharya AK. Understanding DALYs (disability-adjusted life years). J Health Econ 1997; 16: 703-730.

National institute for Health and Clinical Excellence. Social value judgements. Guidelines for the Institute and its advisory bodies. Draft for consultation. April 2005. http://www.nice.org.uk/pdf/CC_SVJ_0405.pdf

Nord E. An alternative to QALYs: the saved young life equivalent (SAVE). BMJ 1992; 305(6858): 875-877.

Powers M, Faden R. Inequalities in health, inequalities in health care: four generations of discussion about justice and cost-effectiveness analysis. Kennedy Inst Ethics J 2000; 10(2): 109-127.

Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. BMJ 2004; 329(7459): 224-227.

Rutten FFH, Brouwer WBF. Meer zorg bij beperkt budget; een pleidooi voor een betere inzet van het doelmatigheidscriterium. Ned Tijdschr Geneeskd 2003; 146(47): 2254-2258.

Sassi F, Archard L, Le-Grand J. Equity and the economic evaluation of healthcare. Health Technol Assessment 2001; 5(3): 1-138.

Smith R. The triumph of NICE. BMJ 2004; 329.

Ubel PA. Pricing life. Why it's time for health care rationing. Cambridge, MA: MIT Press, 2000.

Wetenschappelijke Raad voor het Regeringsbeleid. Volksgezondheidszorg. Den Haag: Sdu, 1997.

Williams A. How economics could extend the scope of ethical discourse. J Med Ethics 2001; 27(4): 251-255.