



UvA-DARE (Digital Academic Repository)

QALY's: een gevoeligheidsanalyse: brief uit Patagonië

den Hartogh, G.

Publication date

2007

Document Version

Final published version

Published in

NVBe nieuwsbrief

[Link to publication](#)

Citation for published version (APA):

den Hartogh, G. (2007). QALY's: een gevoeligheidsanalyse: brief uit Patagonië. *NVBe nieuwsbrief*, 14(1), 16-18. <http://www.nvbe.nl/nieuws/NVBe%2014-1-2007.pdf>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Thema De prijs van een levensjaar

QALY's: een gevoeligheidsanalyse

Brief uit Patagonië

Govert den Hartogh

Beste leden van de NVBE-Nieuwsbrief redactie,

Bij ons in Patagonië, waar de gezondheidszorg dankzij het gure klimaat en het wegvallen van de ozonlaag 30% van het BNP opslokt, en de gezondheidseconomie dan ook verder is ontwikkeld dan waar ook ter wereld, is onlangs een groot onderzoek gedaan naar de bruikbaarheid van de QALY voor het prioriteren van zorgvoorzieningen. Als uitgangspunt hebben wij daarbij de vertaling in het Patagonisch genomen van de achtergrondstudie 'Kostenutiliteitsanalyse' die J.N.D. de Neeling in 2003 schreef bij het rapport van jullie Gezondheidsraad *Contouren van het basispakket*. In die studie somt de Neeling overzichtelijk de verschillende beslissingen op die moeten worden genomen om de QALY als maat voor gezondheidswinst te kunnen gebruiken. In het verleden is voor sommige van die beslissingen wel eens nagegaan welke invloed zij in feite op de uitkomsten van een berekening in QALY's hebben. Ons onderzoek beoogde een soortgelijke gevoeligheidsanalyse te bieden voor alle beslissingen samen.

Wat hebben wij precies gedaan? Wij hebben voor elk van de beslissingen in kwestie een aantal mogelijke keuzen onderscheiden en zijn nagegaan wat het effect van die keuzen op de uitkomst was. Daarbij hebben we steeds bij alle keuzen voor de overige beslissingen uit het lijstje van De Neeling dezelfde keuze gemaakt, en wel de in de literatuur meest gangbare. Voor elke beslissing hebben we de beide meest extreme uitkomsten genomen en daartussen de ratio vastgesteld.

Natuurlijk, bij de meeste beslissingen kunnen meer keuzemogelijkheden worden onderscheiden dan wij hebben gedaan, en de invloed van bepaalde parameters op het eindresultaat kan verschillen als men andere parameters anders kiest. Bovendien, eerlijk gezegd noemt de Neeling meer controversiële aannames van het rekenen met QALY's dan wij zelfs in Patagonië, waar we de overige 70% van het BNP aan wetenschappelijk onderzoek besteden, konden onderzoeken. Maar modellering betekent nu eenmaal vereenvoudiging en in de wetenschap is modellering onvermijdelijk. Wij zijn daarom gewoon uitgegaan van de veronderstelling dat al deze mogelijk vertroebelende factoren elkaar wel zo'n beetje zouden opheffen. In Patagonië staat dat bekend als het 'ripples-in-the-pond-postulate'.

Gewoonlijk wordt aangenomen dat de waardering van een gezondheidstoestand in een

bepaalde periode onafhankelijk is van de duur van die periode en van de gezondheidstoestand daarvoor of daarna. Iedereen begrijpt natuurlijk dat dit in feite niet zo is. Wij hebben daarom twee groepen vrouwen met mammacarcinoom na hun borstamputatie gevolgd, waarbij we tegen de ene groep vertelden dat ze 99% kans hadden op volledig herstel, tegen de andere dat ze 70% kans hadden op recidief. Wat bleek? De eerste groep waardeerde het eerste volledige jaar na de ingreep gemiddeld op 0.8, de tweede op 0.4. Een verhouding dus van 1:2.

Ja, ik begrijp wat jullie willen vragen: bij ons bestaan ook ethische commissies, en die hebben soms de neiging om in zo'n geval moeilijk te doen. Hoe wij daarmee omgaan zal straks nog blijken.

Om de waarde van bepaalde gezondheidstoestanden te kunnen meten, moeten we ze eerst beschrijven op een manier die onderlinge vergelijking mogelijk maakt. Daartoe moeten we verschillende dimensies van 'kwaliteit van leven' onderscheiden. Daarbij treedt een dilemma op. Als je weinig dimensies onderscheidt, en die alleen in heel algemene termen beschrijft, loop je het risico van een bepaalde gezondheidstoestand een heleboel relevante aspecten te missen, met name aspecten die specifiek zijn voor een bepaalde kwaal. Maar als je veel dimensies onderscheidt, wordt het steeds moeilijker om daaraan in de onderlinge vergelijking een bepaald gewicht toe te kennen. Er zat in onze onderzoeksgroep een ethicus - later meer over hem - en die beweerde zelfs dat die dimensies 'incommensurabel' zouden zijn. Terwijl natuurlijk echte wetenschap begint met het inzicht dat alle dingen op aarde en in de hemel gemeten en dus vergeleken kunnen worden. Meten is weten! Aangemoedigd door het recente rapport van jullie Raad voor de Zorg hebben wij honderd zorgvoorzieningen geselecteerd, van longtransplantaties tot de inrichting van een snoezelkamer. Voor het eerst in de wereldgeschiedenis hebben we dus ook voorzieningen uit de care-sector meegenomen. We hebben de effecten van al die voorzieningen beschreven met behulp van een set van drie dimensies (hedonische kwaliteit, mobiliteit en zelfredzaamheid) en één van duizend dimensies. De effecten op de einduitkomst lieten een ratio zien van 1:3.

Het is overigens om te beginnen al de vraag welke gevolgen van een bepaalde voorziening we als aspecten van de gezondheidstoestand

beschouwen en welke niet. In een volgend onderzoek hebben we daarom een vliegveld gesloten dat voor een miljoen Patagoniërs geluidsoverlast veroorzaakte. De daarmee weggenomen ziektelast die verkeersgeluid door vervroegd overlijden veroorzaakt had (via de bekende causale keten stress, hoge bloeddruk en hart- en vaatziekte) bleek neer te komen op 420 DALY's, maar als we daar ernstige hinder en slaapverstoring bij optelden, kwamen we uit op 2300 DALY's, dus vijfmaal zoveel. Omdat DALY's QALY's in spiegelbeeld zijn, komt dat neer op een verhouding van 1:5.

Een bekend discussiepunt is of we de waardering van de beschreven effecten over moeten laten aan een representatieve steekproef uit de bevolking in het algemeen, of die moeten toevertrouwen aan patiënten die de gezondheidstoestand in kwestie hebben meegemaakt. Patiënten lijden in het algemeen aan het zure-druiven-syndroom, en, bovendien, de bevolking moet voor de voorzieningen betalen, dus die mag ook bepalen wat belangrijk is, ik las dat laatst nog in jullie onvolprezen Nieuwsbrief. Maar goed, dat is mijn standpunt, de kwestie blijft controversieel. Wij bedachten daarom een preventief programma voor osteoporose bij vrouwen in de overgang. De geschatte kostenutiliteit per gewonnen QALY bleek uiteen te lopen van 25000 harde Patagonische guldens voor vrouwen die geen heupfractuur hadden meegemaakt tot 100.000 voor ervaringsdeskundigen. Een ratio dus van 1:4.

Een volgend onderzoek betrof de discontovoet. Onze ethicus beweerde weliswaar dat de enige mogelijke reden voor discontering gelegen is in de kans dat de wereld vergaat, en dat discontering voor de rest neerkomt op gebrek aan prudentie of discriminatie van toekomstige generaties. Maar wij zijn zoals gezegd democraten en bovendien harde wetenschappers en wij kijken dus gewoon welke preferenties mensen in feite hebben. We vonden een spectrum van discontovoeten van onder de nul tot meer dan honderd procent per jaar. Dat levert natuurlijk een onaanvaardbare ratio op en daarom hebben we de bevolking ingedeeld in decielen. Aangezien de gemiddelde overlevingsduur van mensen die met de gezondheidszorg in aanraking komen 23,4 jaar is, zijn we nagegaan wat de effecten van de discontovoet per deciel op deze termijn waren. Bij de laagste discontovoet was een perfecte gezondheidstoestand over 23,4 jaar nog 0,8 QALY waard, bij de hoogste 0,2. Opnieuw een ratio dus van 1:4.

Voor ons laatste onderzoek moesten we verder kijken dan de QALY lang is. Stel: medische behandeling kan de ene patiënt nog drie volledig gezonde levensjaren erbij geven, en de andere patiënt nog 20 jaren kwakkelen op

kwaliteit 0.5, allebei tegen de prijs van driehonderdduizend euro. Volgens het voorstel van de RvZ heeft de tweede patiënt recht op die behandeling, de eerste niet. Daaruit blijkt dat de RvZ de interpretatie van 'everyone counts for one and nothing for more than one' voor ogen heeft die Bentham met deze woorden bedoelde. Een alternatieve interpretatie zou echter zijn: voor de eerste patiënt zijn die drie extra jaren zijn topprioriteit, voor de tweede die twintig jaren, en iedereen heeft evenveel recht om zijn topprioriteit gerealiseerd te krijgen. Dus als we per persoon maar 240.000 euro kunnen uitgeven, moeten ze allebei loten, waarbij ze allebei een kans van 8 op 10 krijgen. Om het effect van beide manieren van aggregeren over individuen te meten hebben we de QALY-methode vergeleken met een andere manier van meten van gezondheidswinst: Willingness to Pay. Terwille van de vergelijkbaarheid moesten we natuurlijk bepaalde dingen aannemen: hoeveel Patagonische guldens een QALY mag kosten, resp. welk budget de mensen krijgen van wie de bereidheid om voor een voorziening te betalen wordt gemeten. En we moesten de bevolking weer in decielen verdelen. Het is allemaal wat te ingewikkeld om hier uit te leggen, maar het komt aan op de gevonden ratio, en die was 1:10.

Om op basis van deze deelonderzoeken tot een conclusie te komen hebben we de validiteit van de QALY gedefinieerd als de totale gevoeligheid van de meetuitkomsten voor de onderzochte beslissingen. Die is gelijk aan het product van de gevonden ratio's. Dus: $1/2 \times 1/3 \times 1/5 \times 1/4 \times 1/4 \times 1/10 = 1/4800$. Kijk, dat is nu wetenschap!

Wat is nu het belang van dit resultaat? Er is wel bewerd dat de resultaten van het gebruik van de QALY, gezien de vele aannames waarop dat gebruik berust, volstrekt willekeurig is. Een soort jij-bak natuurlijk, want als er *iets* willekeurig is, dan is het zoals wij allen weten om het stellen van prioriteiten over te laten aan de kwalitatieve oordelen van een stelletje experts. Maar omdat deze opvatting zo hardnekkig is, hebben wij het waarheidsgehalte daarvan nu voor het eerst wetenschappelijk onderzocht. En wat blijkt: de claim is onjuist. De gevonden marge is weliswaar aanzienlijk, dat geef ik toe, maar toch: rekenen met QALY's is *niet* arbitrair. Als dat wel zo was, zouden we immers een ratio van 1: oneindig gevonden moeten hebben. Natuurlijk is het nu een zaak voor de politiek om te bepalen hoe men met een bandbreedte van 1: 4800 wenst om te gaan. Ons advies: bepaal voor elke controversiële aanname de keuze via de kruis-of-munt methode en laat de rest over aan het ripples-to-the-pond-postulate.

Onze ethicus was niet overtuigd. In zijn commentaar op het eindrapport schreef hij dat

elk van onze deelonderzoeken zelf weer berustte op aannames en vereenvoudigingen, waarvan we de effecten eerst maar eens in een meta-gevoeligheidsanalyse zouden moeten onderzoeken. Hij voorzag een regressus die uiteindelijk wel degelijk tot een ratio van 1:oneindig zou leiden. Hij zei bovendien niet te snappen waarom men zo'n moeite zou doen om kwalitatieve oordelen over prioriteiten te vermijden, als alle metingen uiteindelijk toch onvermijdelijk berustten op niet minder kwalitatieve waarderingen van gezondheidstoestan- den of van aspecten daarvan.

Deze meneer had ongetwijfeld op de mid- delbare school geen economie in zijn pakket gehad. Laat staan wiskunde. Een typische alfa

zonder enig benul van wetenschappelijke me- thoden. Wij hebben daarom vastgesteld dat hij leed aan het Syndroom van Snow. Precies: de C.P. Snow van *The Two Cultures*, in Patagonië verplichte literatuur voor alle meerderjarige burgers. Onze ethicus wordt daarvoor nu ge- dwongen behandeld in de Plasterk-kliniek. Zie over onze behandelmethode: <http://www.niemandhoeftbangtezijnvoorcijfers.p at>.

Prof.dr. Govert den Hartogh is als hoogleraar ethiek verbonden aan de Faculteit Geestesweten- schappen, Leerstoelgroep Ethiek van de Universi- teit van Amsterdam