

Evaluatie zonder grenzen

Complexiteit en 'transferabiliteit'
van economische evaluatiemodellen

instituut Beleid & Management Gezondheidszorg



Oratie 11 november 2011

Prof.dr. Hans Severens



Erasmus

ERASMUS UNIVERSITEIT ROTTERDAM
INSTITUUT BELEID & MANAGEMENT
GEZONDHEIDSZORG

Evaluatie zonder grenzen

Complexiteit en ‘transferabiliteit’ van economische evaluatiemodellen

Prof.dr. J.L. (Hans) Severens

Rede

in verkorte vorm uitgesproken bij de aanvaarding van het ambt
van hoogleraar Evaluatieonderzoek in de gezondheidszorg
aan het instituut Beleid & Management Gezondheidszorg
van de Faculteit der Geneeskunde en Gezondheidswetenschappen,
Erasmus Universiteit Rotterdam, op 11 november 2011



*'Seek the wisdom of the ages, but look
at the world through the eyes of a child'*

RON WILD

Colofon

Evaluatie zonder grenzen

Complexiteit en 'transferabiliteit' van economische evaluatiemodellen

Prof.dr. J.L. (Hans) Severens, 11 november 2011

978-94-90420-17-8

Oplage:

600

Productiebegeleiding:

Marketing & Communicatie iBMG

Vormgeving en druk:

B&T Ontwerp en advies (www.b-en-t.nl)

Evaluatie zonder grenzen

Complexiteit en 'transferabiliteit' van economische evaluatiemodellen

Mijnheer de rector magnificus, waarde collega's, familie en vrienden,

Vandaag aanvaard ik met het uitspreken van deze rede het ambt van hoogleraar Evaluatieonderzoek in de gezondheidszorg aan het instituut Beleid & Management Gezondheidszorg van de Erasmus Universiteit van deze stad aan de Maas. Zoals u weet hield ik, nu zo'n 10 jaar geleden, een stuk stroomopwaarts aan dezelfde Maas in een ander spelerstenuel al eens een oratie. De academische mores leren dat een hoogleraar een inaugurale rede alleen bij een eerste benoeming houdt, maar er is de afgelopen 10 jaar in en rond mijn vakgebied zoveel gebeurd dat ik deze gelegenheid graag aangrijp. Het verheugt me dan ook dat u hier allen vanmiddag bent, ongeacht of dit uw eerste of tweede keer is, en ongeacht of u de 11^e van de 11^e viert.

Casus ATAC-trial

Dames en heren, ik wil het onderwerp van vanmiddag graag bij u introduceren met behulp van een casus. In 2002 verscheen er in het toonaangevende medisch wetenschappelijke tijdschrift The Lancet een publicatie waarin de eerste resultaten van de ATAC-trial werden geopenbaard (Baum et al., 2002). Aan de ATAC-studie namen postmenopauzale vrouwen deel met hormoongevoelige borstkanker waarbij de tumor operatief verwijderd kon worden. Patiënten werden gerandomiseerd, ofwel willekeurig ingedeeld, naar één van drie ondersteunende behandelopties: behandeling met het tot dan toe gangbare middel tamoxifen, de experimentele behandeling met het product

anastrozole of een behandeling met beide geneesmiddelen tegelijk. Dergelijke ondersteunende of adjuvante behandeling wordt gegeven na primaire behandeling, in dit geval chirurgie, om kans op terugkeer van ziekte te verkleinen. In totaal namen meer dan 9000 patiënten aan de studie deel en zij werden gedurende een periode van 3 jaar behandeld en gevolgd. De resultaten waren positief, waarmee ik bedoel dat er een statistisch significant verschil werd gevonden tussen de behandelopties. Ten opzichte van de groep vrouwen behandeld met het conventionele middel tamoxifen was er in de anastrozole behandelde groep sprake van een hogere driejaars ziektevrije overleving, respectievelijk 87,4 en 89,4%. Voor de combinatiebehandeling kon geen verbetering in ziektevrije overleving ten opzichte van alleen tamoxifen worden vastgesteld. Daarnaast was er in de anastrozole groep sprake van minder ongewenste bijwerkingen.

Op basis van deze studie concludeerden de onderzoekers dat het nieuwe geneesmiddel effectief en veilig was en dus een goede optie voor de behandeling van deze patiënten. De Amerikaanse vereniging van klinische oncologie beschouwde dit bewijs echter niet afdoende om de standaard tamoxifen therapie te vervangen. De Federal Drug Administration (FDA) nam gezien dit laatste een positief besluit om anastrozole als additionele optie tot de gezondheidszorgmarkt toe te laten (Hillner 2004).

De Nederlandse tegenhanger van de FDA, het College ter Beoordeling van Geneesmiddelen (CBG) besloot vlak daarna anastrozole voor deze indicatie te registreren voor de Nederlandse gezondheidszorgmarkt, waardoor patiënten met vroege borstkanker toegang kregen tot dit geneesmiddel. Financiële toegankelijkheid werd eenvoudig gerealiseerd door, op basis van het argument van therapeutische gelijkwaardigheid, het product op te nemen in een vergoedingscluster van soortgelijke middelen. Verzekeraars betalen dan een maximumprijs voor de vergelijkbare producten. In Nederland kan er alleen een unieke, meestal hogere prijs worden gevraagd indien een therapeutische meerwaarde ten opzichte van bestaande middelen is aangetoond.

In tegenstelling tot deze casus van therapeutische gelijkwaardigheid verschijnen er innovatieve geneesmiddelen op de markt waarvoor voor een specifieke behandelingsindicatie geen goed alternatief bestaat. In dat geval zal er, om een product in Nederland in aanmerking te laten komen voor vergoeding, meer dan alleen veiligheid en werkzaamheid aangetoond moeten zijn. In dergelijk gevallen verlangt onze minister bewijsvoering

betreffende de doelmatigheid of wel kosten-effectiviteit van het nieuwe geneesmiddel. In dit kader wordt wel gesproken over de 4^e horde in de besluitvorming.

Resultaten uit de ATAC-trial zijn in verschillende landen gebruikt in doelmatigheidsonderzoek ten behoeve van beslissingen over vergoeding of richtlijnen klinisch handelen. Ik presenteer u deze casus omdat het elementen uit mijn vakgebied laat zien waar ik het over wil hebben, namelijk: complexiteit en 'transferabiliteit', ofwel overdraagbaarheid van economische modellen. Ik zal dan ook regelmatig op deze casus terugrijpen.

Economische evaluatie

Doelmatigheidsbewijs betreffende een medische technologie wordt verkregen door doelmatigheidsonderzoek, ook wel economische evaluatie genoemd. De kern van dergelijk onderzoek is eenvoudig: keuzealternatieven worden met elkaar vergeleken wat betreft kosten en uitkomsten om zodoende een uitspraak te kunnen doen over de kosten-effectiviteit van een technologie. De meerkosten van iets nieuws zouden moeten kunnen worden gerechtvaardigd uit betere resultaten voor de patiënt, of de zorg zou goedkoper moeten zijn zonder dat daar veel nadelige gevolgen voor de patiënten aan verbonden zijn. In beide gevallen kan worden gesproken van een doelmatige behandeling. Een optie die in vergelijking met een alternatief beter is voor de patiënt én minder kosten met zich meebrengt wordt dominant genoemd en leidt vanzelfsprekend tot een positieve vergoedingsbeslissing.

Het perspectief dat wordt gehanteerd binnen zulke economische evaluaties is belangrijk. Kosten kunnen afhankelijk daarvan betrekking hebben op louter kosten binnen de gezondheidszorg maar ook op de totale maatschappelijke kosten, dus inclusief kosten van eventueel arbeidsverzuim door patiënten, kosten van mantelzorg, maar ook indirecte medische kosten als gevolg van levensverlengende zorg voor de patiënt. Effecten kunnen afhankelijk van het doel van de studie uitgedrukt worden in verschillende eenheden. Ten eerste kunnen puur klinische patiënt-uitkomsten worden gehanteerd zoals aantal succesvol behandelde patiënten (take-home baby rate bij IVF, aantal 'droge' patiënten bij incontinentie), tijd tot recidief ofwel

terugkeer van de aandoening, of een klinische waarde die de ziekte-toestand van de patiënt weergeeft (zoals gezichtsvermogen bij glaucoom behandeling en DAS (disease activity score) bij RA patiënten). Een deel van het nut van economische evaluaties wordt ontleend aan het vergelijkbaar maken van de efficiëntie van interventie in verschillende delen van de gezondheidszorg. Efficiëntie van interventies van bijvoorbeeld diagnostiek bij acute aandoeningen als plotselinge pijn op de borst en toepassing van zogenaamde anti-TNF geneesmiddelen bij reumatoïde artritis kan alleen met elkaar vergeleken worden op basis van een generieke uitkomstmaat. Eventuele invloed van medische zorg op de levensverwachting van patiënten (bij het voorbeeld van diagnostiek bij acute angina pectoris) kan dan in één en dezelfde maat gecombineerd worden met de invloed van medische zorg bij patiënten op hun kwaliteit van leven (zoals bij het voorbeeld van RA-behandeling het geval is). Een veel gebruikte generieke uitkomstmaat bij de evaluatie van medische technologie is daarom de QALY, oftewel de kwaliteit van leven gecorrigeerde levensverwachting.

Besluitvorming

Maar wie zijn nu de potentiële gebruikers van economische evaluaties? In een ongeveer 15 jaar oude beschrijving van het Nederlandse zorgsysteem in termen van toelating van medische technologie en de rol van kosten-effectiviteitsstudies wordt gesproken van een mogelijke rol op macro- en microniveau (Elsinga en Rutten, 1995). Macroniveau beslissingen betreffen beslissingen mede op basis van kosten-effectiviteit ten behoeve van toelating tot een verzekeringspakket of nationaal zorgprogramma. Microniveaubeslissingen betreffen beslissingen betreffende het feitelijk gebruik van een technologie op het niveau van de patiënt-arts relatie. Kosten-effectiviteitsbewijs kan inmiddels echter ook een rol spelen op het mesoniveau, waar bijvoorbeeld beroepsgroepen klinische protocollen definiëren, een ziekenhuis apparatuur aanschaft of een formularium-commissie besluit een geneesmiddel wel of niet voor het eigen ziekenhuis beschikbaar te stellen en dus in te kopen. Bovendien kunnen kosten-effectiviteitsargumenten op mesoniveau een rol spelen bij het actief beïnvloeden van beslissingen op microniveau, middels bijvoorbeeld het ontwikkelen en implementeren van klinische richtlijnen (Hoomans et al., 2009).

Nu is natuurlijk de vraag of en hoe beslissers in een gezondheidszorgsysteem bewijs op basis van economische evaluaties gebruiken, want daaraan ontleent mijn leerstoel toch ook zijn bestaansrecht. In een literatuuroverzicht werden 36 onderzoeksprojecten naar het belang van economische evaluaties bij besluitvorming samengevat (Van Velden et al., 2005). De meeste studies baseerden hun bevindingen op interviews of enquêtes met beslissers en bij slechts enkele studies werden observaties gedaan bij feitelijke vergoedings- of implementatiebesluiten. Uit dit overzicht kwam naar voren dat vooral op het mesoniveau een invloed van kosten-effectiviteitsbewijs wordt gevonden. Op macroniveau leek deze invloed bijna 10 jaar geleden beperkt. In een meer recent overzicht van studies die de invloed van economische evaluaties op de drie besluitvormingsniveaus tegelijkertijd onderzochten werd de sterkste invloed op het macroniveau gezien (Erntoft, 2011). Deze invloed is afhankelijk van het bestaan van een formeel kader voor beoordeling (zoals pharmaco-economische richtlijnen) en het kennisniveau van beslissers wat betreft de methodologie van economische evaluaties. Het bewijs hiervoor blijft echter beperkt tot de Verenigde Staten, Groot-Brittannië en Zweden.

Maar speelt kosten-effectiviteit ook een rol bij beslissingen binnen de individuele arts-patiënt context? Inmiddels zijn er ook studies die rapporteren dat individuele artsen in hun klinische beslissingen kosten-effectiviteitsbewijs meewegen. In een keuze-experiment werd 129 Italiaanse cardiologen steeds twee opties voor behandeling van een hypothetische patiënt voorgelegd (een 65 jaar oude vrouw die 5 jaar geleden een klein hartinfarct had, familiair belast wat betreft cardiovasculaire aandoeningen, cholesterol 230 mg/l, bloeddruk 150/195 mmHg). De hypothetische alternatieven werden beschreven als een behandeling met een geneesmiddel gekenmerkt door de kwaliteit van bewijs (trial bewijs op basis van 3.000 tot 30.000 patiënten), werkzaamheid van het middel (relatief risicoreductie van 5% tot 20%; absoluut risicoreductie van 0.5% tot 2%) en economische impact (variërend van €5.000 / QALY tot €200.000 / QALY). De uitkomsten lieten zien dat naast minimale werkzaamheid en bewijs deze artsen vooral duidelijk leeftijd-effect werd vastgesteld. Jongere Italiaanse cardiologen vinden doelmatigheid een belangrijker keuzecriterium dan oudere (Torbica and Fattore, 2010).

Ik heb niet de illusie dat kosten-effectiviteitsbewijs in alle gevallen een doorslaggevende invloed op de besluitvorming zal hebben. De invloed kan volgens mij beter beschreven worden als richtinggevend bij besluitvorming.

Wetenschappelijk bewijs staat namelijk niet op zich: wetenschappelijk bewijs behoeft interpretatie, al dan niet op basis van eerdere ervaring, observaties en data (Kelly et al., 2010). Ook in het Nederlandse macro-besluitvormingssysteem wordt daarom een duidelijk onderscheid gemaakt tussen op kwalitatief onderzoek gebaseerde 'assessment', vrij vertaald 'schatting' en het meer kwalitatieve oordeel, de zogenaamde 'appraisal' of 'oordeel'. De feitelijke invloed van economische evaluaties op microbesluitvorming lijkt in Nederland voornamelijk beperkt (Banta and Oortwijn, 2009). Hiermee is nog maar eens aangegeven dat hypothetische preferenties van Italiaanse cardiologen niet representatief lijken te zijn voor huisartsen uit Venlo, Mierlo en Hank of medisch specialisten uit Rotterdam, Maastricht of Nijmegen.

Een besluitvormingsniveau dat minder relevant lijkt voor de Nederlandse situatie is de individuele patiënt. In de meeste niet-westerse landen, maar ook in de Verenigde Staten zijn eigen betalingen, ook wel out-of-pocket payments genoemd, gebruikelijk. Er blijkt in de Verenigde Staten een duidelijke relatie te bestaan tussen de out-of-pocket betalingen en het gebruik van de juiste medicatie, volgens de juiste dosering. Bij eigen betalingen zijn patiënten geneigd medicatie helemaal niet te nemen, minder lang of in een lagere dosering te gebruiken (Wagner et al., 2008). In een gezondheidszorgsysteem waar geen sprake is van vergoeding via een verzekering lijken doelmatigheidsargumenten een minder belangrijke rol te spelen op dit micro-beslissingsniveau en spelen absolute kosten in termen van out-of-pocket betalingen een belangrijke rol.

Ook wereldwijd wordt kosten-effectiviteit belangrijk geacht voor zorgsystemen en daarvoor wil ik refereren naar een relatief recent rapport van de wereld gezondheidsorganisatie WHO, getiteld Health Systems Financing, the path to universal coverage (World Health Organisation, 2010). In het voorwoord van de directeur-generaal wordt aangegeven dat, ongeacht het feit of een land arm of rijk is, er economisch gezien sprake is van een spanningsveld: kosten van de gezondheidszorg nemen toe, en tegelijkertijd neemt de levensverwachting van mensen toe, is meer en meer sprake van chronische ziekten, en komen nieuwe en vaak dure diagnostische en therapeutische interventies op de gezondheidszorgmarkt. In het rapport wordt geconstateerd dat de publieke vraag om toegang tot kwalitatief goede, betaalbare zorg toeneemt en dat de politieke druk om tot goed onderbouwde beleidskeuzen te komen daardoor ook stijgt. Universal

coverage, of wel brede, universele dekking c.q. bereik is daarbij het belangrijkste doel. Universele dekking, zoals gesteld in dit rapport, beoogt toegang te geven tot zorg voor iedereen dit dat nodig heeft, zonder dat mensen daarvoor ondraagbaar financieel moeten lijden. De WHO benoemt hierbij drie belangrijke probleemgebieden. Naast het gegeven dat afdoende fondsen voor de zorg gevonden moeten worden en directe betalingssystemen vermeden c.q. afgebouwd moeten worden (die leiden tot financiële barrières en soms armoede van individuen), wordt gesuggereerd bestaande gezondheidszorgsystemen efficiënter te maken. De schatting is dat bestaande gezondheidszorgsystemen wereldwijd 20-40% efficiënter kunnen werken (dus een beter resultaat van het gezondheidszorgsysteem zonder kosten te verhogen). Eén van de aandachtspunten hierbij is de kritische evaluatie van interventies. Daarbij wordt genoemd dat de kosten-effectiviteit van interventies uiteen kan lopen van US\$ 10 tot US\$ 100.000 per gezond levensjaar. Een focus op kosten-effectiviteit en het op basis daarvan kiezen van doelmatige interventies of technologieën heeft hiermee invloed op de efficiëntie van een gezondheidszorgsysteem als geheel.

Maar laten we voor het gemak even uit blijven gaan van een gezondheidszorgsysteem waarvan de financiering geënt is op solidariteitsprincipes. Ook daar is het gebruik van bewijs betreffende de doelmatigheid van medische interventies niet vanzelfsprekend en is een belangrijke waarschuwing op zijn plaats. Er wordt in dit kader gesproken van een paradox: beslissers zouden meer gebruik moeten maken van de resultaten van economische evaluaties teneinde de zorg efficiënter te maken, maar tegelijkertijd worden economische evaluaties steeds complexer, vanwege de steeds verder ontwikkelende methodologie, waardoor beslissers steeds minder in staat lijken de resultaten ervan goed te doorgronden en dus op waarde te schatten (Brousselle and Lessard, 2011).

Modellering

Dit constaterende wil ik met u terug naar de casus van de ATAC-trial. Op macroniveau was er in het kader van het Nederlandse referentieprijssysteem geen noodzaak om kosten-effectiviteitsbewijs in de besluitvorming mee te nemen. Dit vergoedingssysteem is in bijna geen andere Europese en niet-Europese landen gangbaar dus kosten-effectiviteit speelde in het buitenland wel een rol bij deze casus. Maar niet alleen op macroniveau

wordt doelmatigheidsbewijs gebruikt. In klinische richtlijnen die op mesoniveau worden ontwikkeld, wordt aangeraden welke behandeling bij welke indicatie in te zetten en, zoals eerder aangegeven, ook daar spelen doelmatigheidsargumenten steeds vaker een rol. Echter, op basis van de relatief korte termijn werkzaamheid van anastrozole bij de adjuvante behandeling van operabele borstkanker zoals vastgesteld in de ATAC-trial, kon nog geen uitspraak worden gedaan over de lange termijn doelmatigheid ten opzichte van tamoxifen. Het logische gevolg was dat de werkzaamheidsgegevens van de ATAC-trial in economische rekenmodellen werden gebruikt. Eenvoudig gezegd richt een kosten-effectiviteitsmodel zich op het doorekenen van kosten en effecten op basis van beperkte gegevens, uit bijvoorbeeld een relatief kortdurende trial, de literatuur of patiënt-registraties.

Uit een recent literatuuroverzicht bleek dat de gegevens van deze ATAC-trial de basis vormden voor niet minder dan 11 verschillende economische modellen waarin anastrozole met tamoxifen werden vergeleken bij adjuvante behandeling van vroege borstkanker (Frederix et al., in press). Kosten-effectiviteitsmodellen werden ontwikkeld voor Canada, België, Spanje, Brazilië, Engeland en de Verenigde Staten. De resultaten waren zeer divers: doelmatigheidschattingen varieerden van ruim 3.000 euro per gewonnen levensjaar en ook per QALY tot meer dan 100.000 euro per gewonnen levensjaar en bijna 70.000 euro per QALY. U kunt zich voorstellen dat deze spreiding in kosten-effectiviteitsschattingen kan leiden tot verschillende beslissingen over feitelijk gebruik van het nieuwe middel bij een en dezelfde patiëntengroep.

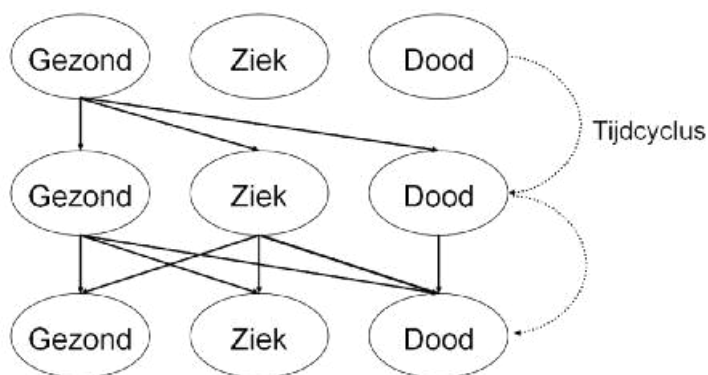
Waar deze opmerkelijke verschillen nu allemaal aan kunnen liggen, wil ik later op terugkomen. Eerst wil ik u iets vertellen over economische rekenmodellen.

Eerder heb ik u verteld dat in een trial patiënten worden gerandomiseerd over verschillende behandelingsopties. U kunt zich voorstellen dat we gedurende de follow-up van patiënten in zo'n trial zowel kosten als patiëntuitkomsten kunnen worden gemeten en zo is sprake van een economische evaluatie. Echter, het nadeel van trials kan onder andere zijn dat de follow-up te kort is om de feitelijke kwaliteit van leven van patiënten tot het moment van overlijden vast te stellen. Daarnaast zijn er wellicht meerdere behandelingsopties en zijn ook complexe strategieën denkbaar om

patiënten te diagnosticeren en te behandelen, welke niet binnen een enkele trial te evalueren zijn. Kortom, er zijn grenzen aan de mogelijkheden om op basis van trials alle relevante kosten en effecten van keuzeropties te bepalen. Economische modellen maken de mogelijkheden van evaluatie schijnbaar grenzeloos.

Een goed kosten-effectiviteitsmodel reflecteert de aandoening, het natuurlijk beloop van de aandoening over de tijd, het effect van een interventie op dit beloop (en dat kan preventie, diagnostiek, behandeling, monitoring, revalidatie en dergelijke zijn), eventuele ongewenste bijwerkingen van de interventie, de gezondheidstoestanden die een patiënt kan ervaren, de waarde die we toekennen aan die gezondheidstoestand en de kosten gerelateerd aan de aandoening, interventies, bijwerkingen, gezondheidstoestanden, zorg voor de patiënt etc. Binnen mijn vakgebied zijn verschillende modelleringsmethoden gangbaar maar ik wil me richten op zogenaamde Markov modellen, ook wel gezondheidstoestand-transitiemodellen genoemd.

In figuur 1 is een eenvoudig Markov model weergegeven. Individuen in een hypothetische populatie kunnen zich bevinden in de gezondheidstoestanden gezond, ziek, of overleden. In het geval we uitgaan van een gezonde populatie kunnen personen na een bepaalde tijd ziek zijn, overleden zijn, of in dezelfde gezonde toestand zijn gebleven. Na weer een periode kunnen patiënten overlijden aan de ziekte, nog steeds ziek zijn of na eventuele behandeling of spontaan herstel weer gezond zijn. Het cohort verdeelt zich dus iedere cyclus opnieuw over de Markov toestanden. En dit kan zo vele cycli doorgaan tot het hele cohort mathematisch gezien overleden is. De verdeling van het cohort over de gezondheidstoestanden, de duur dat de patiënten zich in een toestand bevinden, en de waarde van deze toestanden in termen van kosten en QALY's bepalen de uitkomst van de Markov calculaties. Door het model voor verschillende therapieën door te rekenen met het effect van deze specifieke therapieën op bijvoorbeeld de transitiekans van ziek naar gezond kunnen kosten en effecten geschat worden over langere tijd dan praktisch mogelijk is in een trial.

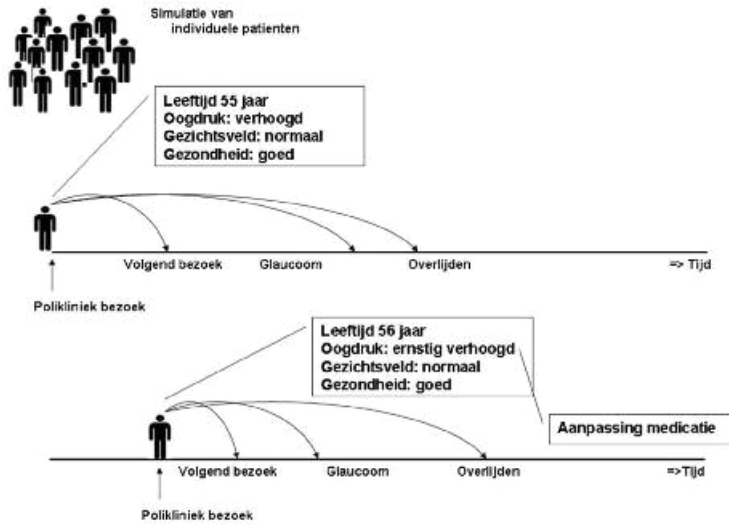


Figuur 1. Schematische weergave van een gezondheidstoestand-transitiemodel, ofwel Markov model.

Wat ik u heb laten zien is een relatief eenvoudig Markov model. U kunt zich voorstellen dat het natuurlijk beloop van een aandoening en alle mogelijke interventies niet altijd recht wordt gedaan in een té sterke vereenvoudiging van de werkelijkheid. Zo bleek een Markov model onhaalbaar bij het modelleren van de kosten-effectiviteit van behandelingen voor glaucoom. Glaucoom is een chronische oogaandoening waarbij, meestal als gevolg van een verhoogde oogdruk, de gezichtsenuw wordt aangetast en als gevolg daarvan het gezichtsvermogen van de patiënt afneemt, eventueel leidend tot blindheid. Het gezichtsveld en oogdruk van patiënten worden door een oogarts levenslang gevolgd en behandeling bestaat vooral uit het verlagen van de oogdruk. Behandelingsopties zijn afwachtend beleid, verschillende geneesmiddelen, lasertherapie, chirurgie en iedere combinatie en opeenvolging hiervan. De keuze van behandeling is afhankelijk van leeftijd van de patiënt, beloop van de ziekte, intolerantie van eerdere behandelingen en de uitkomst van diagnostische tests. Daar komt nog bij dat de progressie van glaucoom niet te definiëren is in ziektestadia, maar dat sprake is van een geleidelijke afname van anatomische en functionele ziektemanifestatie. Zelfs als we deze situatie weergeven in een sterk versimpeld Markov model met 4 gezondheidstoestanden en 10 behandelingsopties zou dit leiden tot 40 Markov toestanden en 1600 transitie mogelijkheden. En dan houdt een dergelijk eenvoudig model nog steeds geen rekening met het feit dat in de klinische praktijk de tijd dat patiënten weer gezien worden variabel is, en dat achtereenvolgende klinische stappen afhankelijk zijn van de

behandelingsgeschiedenis van de individuele patiënten. U begrijpt dan ook dat een Markov model hier theoretisch onvoldoende en praktisch onhaalbaar is.

Een oplossing om dit probleem aan te pakken is een andere modelleringsmethode, discrete-event simulation genoemd, kortweg DES. DES kenmerkt zich door het modelleren van een groot aantal hypothetische patiënten over de tijd waarbij eigenschappen van deze patiënten en gebeurtenissen expliciet gesimuleerd worden. Iedere hypothetische patiënt kenmerkt zich door attributen die vast staan (zoals geslacht of risicofactoren) of variëren over de tijd (zoals leeftijd of in het geval van glaucoom, oogdruk en gezichtsveld). In de loop van de tijd kunnen er gebeurtenissen optreden zoals een bezoek aan de oogarts, het ontstaan van glaucoom of overlijden van de patiënt) en in het model wordt bepaald welke gebeurtenis, op basis van feitelijke data, als eerste verwacht kan worden. Een DES rekenmodel actualiseert dan de attributen van de hypothetische patiënt tot deze naar verwachting van het model is overleden. En om een DES model te gebruiken voor kosten-effectiviteitsberekeningen worden voor alle gebeurtenissen in het model kosten gedefinieerd en wordt de QALY-waarde voor een persoon bepaald aan de hand van de periode in een bepaalde gezondheidstoestand, in dit geval voornamelijk gezichtsvermogen. In figuur 2 wordt schematisch het beloop van een theoretische patiënt weergegeven over de tijd zoals geschat aan de hand van het DES model.



Figuur 2. Schematische weergave van een discrete-event simulation model betreffende verhoogde oogdruk en glaucoom

Het vergelijken van strategieën om glaucoom patiënten te diagnosticeren en te behandelen kan als een DES model eenmaal ontwikkeld is, relatief eenvoudig door modelparameters voor bijvoorbeeld effectiviteit van medicatie en accuraatheid van een diagnostische test aan te passen en de modeluitkomsten van verschillen opties te vergelijken (Van Gestel et al., 2010).

En nu kom ik graag terug op de diversiteit aan bevindingen van de kosten-effectiviteitsmodellen gerelateerd aan de ATAC-trial. De uitkomsten van de 11 modellen, allemaal gebaseerd op de principes van Markov modellering, waren divers waarbij vooral opviel dat de uitkomsten niet alleen varieerden tussen landen maar ook tussen modellen die het beslissingsprobleem binnen één en het zelfde land zouden moeten weergeven. De kosten-effectiviteit uitgedrukt in kosten per gewonnen levensjaar is vooral afhankelijk van de berekende levensverwachting van patiënten, gebaseerd op extrapolatie van de bevindingen van de ATAC-trial. De berekende gewonnen levensverwachting varieerde tussen 0,16 en 0,55 levensjaren, ofwel minder dan 2 maanden tot bijna 7 maanden winst. Deze aanzienlijke verschillen konden uiteindelijk teruggevoerd worden naar de manier waarop de ziekteprogressie en het behandel-effect werd gemodelleerd. Zo verschilde

tussen de modellen het ingeschat risico dat sprake was van recidief van de tumor (effectiviteit van behandeling): de hazard rate, een maat voor kans, was in sommige gevallen gebaseerd op ziektevrije-overlevingskans, die achtergrond sterftekans includeert, soms gebaseerd op ziektevrije overleving zonder achtergrondmortaliteit of werd overleving gerelateerd aan afzonderlijke kansen voor lokaal recidief of metastasen op afstand. Er was ook variatie in het al dan niet tijdafhankelijk schatten van deze kansen. Maar minstens even belangrijk was het feit of in een berekeningsmodel rekening werd gehouden met het zogenaamde carry-over effect, het aanhoudende effect van een geneesmiddel nadat gestopt is met behandeling. Belangrijke bijwerkingen van het product, welke de sterftekans beïnvloeden, (osteoporose met eventueel fracturen, veneuze trombose of baarmoederslijmvlieskanker) werden in sommige modellen zelfs helemaal achterwege gelaten. Daarnaast was er een zeer grote verscheidenheid in lengte van de tijdshorizon tussen de modellen, welke varieerde van 10 jaar tot levenslang. De vraag is of verschillen tussen uitkomsten van modellen veroorzaakt door het op verschillende manieren weergeven van een en dezelfde aandoening, wenselijk of zelfs acceptabel is. Zonder te willen spreken over het hot topic fraude in de wetenschap (Keulemans en Van Raaij, 2011) bereiken we naar mijn idee de grens waar een valide kosten-effectiviteitsmodel minimaal aan zou moeten voldoen: de structuur van een model moet, ongeacht welke interventies geëvalueerd worden het verloop van een aandoening op een valide manier reflecteren.

Methodologische ontwikkelingen

Nu ik u heb proberen duidelijk te maken dat mathematische modellen noodzakelijk zijn om in een nationale, maar ook internationale context de kosten en effecten van een medische technologie te kunnen beoordelen, wil ik met u ingaan op de toenemende complexiteit van onze rekenmodellen.

Een eventuele kosten-effectiviteitsgrens of drempelwaarde kan als een van de factoren dienen bij besluitvorming over een technologie in het kader van vergoedingen, dus een beslissing op macroniveau. Echter, zoals eerder aangegeven, besluitvorming over feitelijk gebruik van een geneesmiddel, diagnostische test, of operatietechniek ligt op het meso-, dan wel micro-niveau. En een individuele dokter heeft te maken met een individuele patiënt die gekenmerkt wordt door individuele kenmerken en individuele

omstandigheden. Nu lijkt het mij praktisch onmogelijk om voor iedere specifieke patiënt die uniek is door een diversiteit aan kenmerken en omstandigheden een volledige kosten-effectiviteitsanalyse te doen, maar de roep om 'individualised medicine' is onmiskenbaar aanwezig. In dit kader wordt gesproken van 'stratified medicine': groepen patiënten kunnen onderscheiden worden binnen een patiëntpopulatie. Zo kan de aanwezigheid van andere aandoeningen (ofwel co-morbiditeit) een rol spelen bij interventiekeuze of kan genetische constitutie van personen mede bepalend zijn voor behandelingsrespons. De doelmatigheid van een pharmacogenetische test en behandeling op geleide van de testuitkomst kan middels een model worden berekend. Zo kan aanwezigheid van een specifiek genetisch polymorfisme bij patiënten met chronische nierziekte voorspellend zijn voor niet reageren op ACE-inhibitor medicatie en blijkt behandelingsstratificatie op basis van genetische testuitslag doelmatig (Costa-Scharplatz et al., 2007).

Economische modellen zullen steeds meer rekening moeten gaan houden met verscheidenheid ofwel heterogeniteit binnen patiëntpopulaties en expliciet op zoek moeten gaan naar doelmatigheid van gestratificeerde geneeskunde (Chalkidou and Rawlins, 2011). Zodoende werd op basis van eerder genoemd model waarmee de kosten en effecten van behandelingsstrategieën voor glaucoom werden doorgerekend, bepaald of het potentieel zinvol is binnen een populatie onderscheid te maken naar subgroepen patiënten. Er bleken tot nu toe geen bekende patiëntkarakteristieken te zijn die voorspellend waren voor welke patiënt wel, en welke patiënt niet doelmatig kon worden behandeld. Differentiatie binnen een glaucoom behandelingsrichtlijn is tot op dit moment dan ook niet aangewezen. Echter, de analyse maakte wel duidelijk dat het doelmatig zou kunnen zijn voor de zorg om te kunnen voorspellen, bijvoorbeeld op basis van een nog te ontwikkelen test, welke patiënten snelle progressie van de ziekte gaan ervaren (Van Gestel et al., in press). Het mag duidelijk zijn dat analyse van doelmatigheid, rekening houdend met heterogeniteit in populaties onze modellen per definitie meer complex maakt.

Uit het voorgaande voorbeeld blijkt dat het zinvol kan zijn potentiële doelmatigheid in te schatten van een technologie die nog niet bestaat of nog in ontwikkeling is. Ik kom hierbij op het terrein van de vroege evaluatie, ook wel vroege HTA of constructieve HTA genoemd. In het recent verschenen Gezondheidsraadrapport 'Waar voor je geld' wordt de grens van

evaluatieonderzoek verschoven (Gezondheidsraad, 2010). Zoals ik tot nu toe beschreef heeft traditioneel economisch evaluatieonderzoek als studieobject een medische technologie die 'klaar' is. Er is sprake van een technologie die uitontwikkeld en stabiel is, die aantoonbaar werkt en die potentieel zinvol is voor een specifieke indicatie. Maar inmiddels wordt evaluatieonderzoek vroeger in de levensloop van een technologie toegepast (IJzermans and Steuten, 2011). De grens van mijn vakgebied is dan ook aan het verschuiven en het feit dat de werkgroep HTA van het CTMM hier expliciet aandacht aan besteedt is tekenend voor deze ontwikkeling. Deze nationale werkgroep bestaat uit economische evaluatieonderzoekers uit academische en niet-academische instituten die binnen het veld van vroege HTA werkzaam zijn. Binnen de werkgroep worden onderzoeksmethoden bediscussieerd en kennis en ervaring uitgewisseld om zo buiten de grenzen van ons vakgebied te kijken.

Een voorbeeld van dergelijke vroege evaluatie is onderzoek naar de potentie van een perfecte technologie voor genetische diagnostiek bij borstkanker. Bijna 10 jaar geleden werd in het Nederlands Kanker Instituut (NKI-AVL) de zogenaamde Mammaprint ontwikkeld (Van 't Veer et al., 2002). Dit betreft een techniek waarmee het mogelijk is op basis van het tumorweefsel van een patiënt een 70-genen expressieprofiel te maken waarmee het individuele risico op afstandsmetastasen kan worden geschat. Op basis van deze inschatting kan worden bepaald welke patiënt wel en welke patiënt naar verwachting geen baat heeft bij aanvullende systemische behandeling zoals chemotherapie. In feite zijn er twee mogelijkheden om deze test uit te voeren: de huidige test op basis van vers ingevroren weefsel of een – verbeterde, nog niet uitontwikkelde – Mammaprint op basis van weefsel geprepareerd met paraffine. De huidige test kent namelijk een zekere kans dat het technisch niet lukt voldoende ingevroren weefsel van voldoende kwaliteit te verkrijgen, terwijl verwacht wordt dat een MammaPrint gebaseerd op paraffine weefsel een grotere kans van slagen heeft omdat de procedure gebruikersvriendelijker is. In een vroege HTA-analyse werd de potentiële kosten-effectiviteit van de nieuwe paraffine-gebaseerde optie geanalyseerd, en werd bovendien met ditzelfde model bepaald dat het economisch rendabel is deze test, die nog niet volledig uitontwikkeld is, te vervolmaken (Retel et al., submitted).

Uit het eerdere voorbeeld van het model van glaucoom blijkt dat het zinvol is op zoek te gaan naar een technologie die de toekomstige progressie van

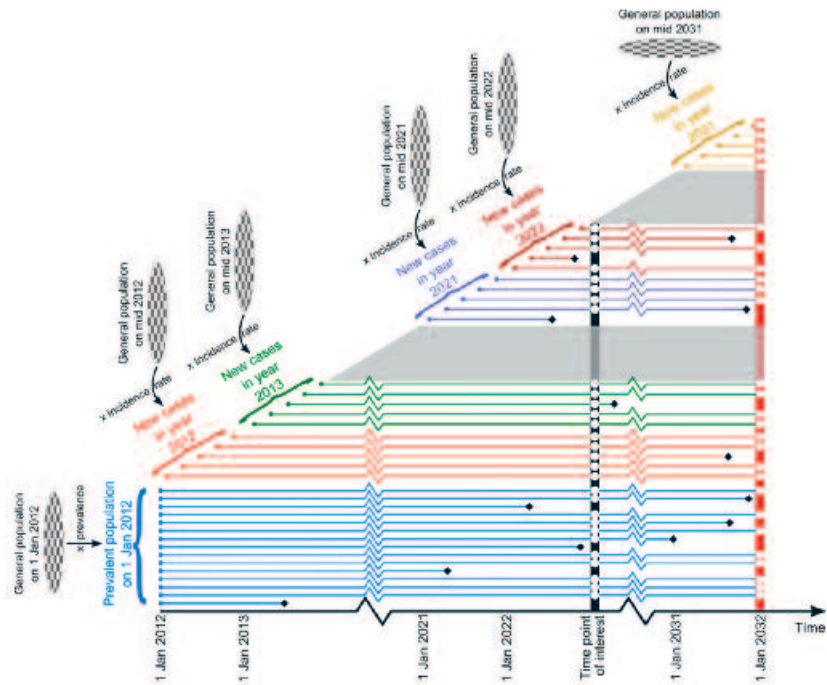
glaucoom kan voorspellen. Hoe deze technologie er uit zou moeten komen te zien is nog volstrekt onduidelijk, maar er lijkt een potentiële doelmatigheidswinst voor de zorg van deze patiënten. Hierbij moet wel de kanttekening geplaatst worden dat de efficiënte investeringsruimte voor deze technologie conform de modelberekeningen slechts 130 euro per patiënt is, hetgeen de maximale prijs van de test zou mogen zijn die op een efficiënte manier relevante patiëntengroepen binnen de totale populatie kan onderscheiden. Uit deze voorbeelden blijkt het nut van economische modellen en gestratificeerde geneeskunde.

Ik raak met deze voorbeelden van de verschuivende grens van economische modellen aan een ander aspect dat de toekomstige gezondheidseconomische modellen meer complex zal maken. Bij het presenteren van kosten-effectiviteitsschatting van een technologie vragen beslissers zich veelal direct af of ze de nieuwe technologie wel kunnen betalen. In mijn vorige oratie refereerde ik daarvoor naar de zogenaamde budgetcurve (Sendi and Briggs, 2001). Deze veronderstelt echter dat er een bestedingsbudget per patiënt bestaat waarbinnen doelmatige interventies kunnen worden ingevoerd. Nu is er in de regel geen expliciet bestedingsbudget per patiënt vastgesteld. Zo weten we ook in Nederland bijvoorbeeld niet wat we als maatschappij bereid zijn te betalen voor het introduceren van een nieuw geneesmiddel voor de behandeling van ankyloserende spondylitis, ook wel de ziekte van Bechterew genoemd. Deze progressieve invaliderende aandoening van de ruggenwervel wordt levenslang behandeld waarbij de onderdrukking van ontsteking een belangrijke component is. Er zijn meerdere geneesmiddelen beschikbaar in het behandelingsarsenaal van de arts en daar zijn de afgelopen jaren zogenaamde anti-TNFalfa middelen bijgekomen. Nu blijkt de kosten-effectiviteit van deze producten (indien gebruikt als 3^e lijnmiddel) goed te zijn, maar is van dominantie, dus betere patiëntuitkomst en tegelijkertijd lagere kosten, geen sprake. Daarentegen heeft een acceptabele kosten-effectiviteit, dus een betere patiëntuitkomst tegen acceptabele hogere kosten wel tot gevolg dat macro gezien de uitgaven voor geneesmiddelen stijgen. En juist in een dergelijke situatie worden ook de budgetconsequenties binnen de gezondheidszorg relevant voor beslissers.

Kortom, de kosten-effectiviteitsmodellen op individueel patiëntniveau zouden idealiter moeten worden opgeschaald naar patiëntpopulatie-niveau. Dit is niet zozeer een nieuw idee (Rutten, 1988) maar toepassing van dergelijke methoden, die eerder wel scenariostudies werden genoemd, is

beperkt. En dat is nu juist wat in een project bij de evaluatie van anti-TNFalfa-middelen bij ankyloserende spondylitis gedaan werd. Een economisch model werd gebouwd op basis van de principes van discrete-event simulation (zoals bij het glaucoom voorbeeld) waarin individuele patiënten met individuele kenmerken en ziektekenmerken werden gemodelleerd wat betreft individuele progressie van hun aandoening over de tijd, en de invloed van behandeling daarop (Tranduy et al., in press). Dit werd gedaan voor de aanwezige ofwel prevalentie populatie in 2011. Behandeling van deze patiënten heeft budgetimplicaties en implicaties betreffende de ziekte-toestand van de populatie als geheel. Echter, in volgende jaren, 2012 en verder, zullen patiënten uit deze prevalentie populatie zijn overleden, maar zullen er ook nieuwe gevallen van ankyloserende spondylitis door artsen worden gediagnosticeerd. Dit is weergegeven in figuur 3. In deze figuur wordt de huidige prevalentie populatie weergegeven als de onderste groep horizontale lijnen die helemaal links start (1 januari 2012) en verder naar rechts kunnen eindigen. Beëindiging van een lijn geeft het overlijden van een patiënt weer. In 2012 en de volgende jaren komen er op basis van bekende incidentie van ankyloserende spondylitis natuurlijk nieuwe patiënten bij die ook over de tijd modelmatig 'gevolgd kunnen worden' wat betreft het beloop van hun aandoening en invloed van behandeling daarbij. Op deze manier ontstaat de dynamiek van een populatie en kan op enig moment in de tijd een voorspelling worden gegeven betreffende de omvang van de patiëntenpopulatie, de ziekte-toestand, de kosten van behandeling en dergelijke (zie bijvoorbeeld de verticale zwarte balk in jaar 2022). Door op basis van dit model wat betreft aanbod en vergoeding van behandelingen voor verschillende beleidskeuzen door te rekenen, kunnen we beslissers de 'realtime' lange termijn kosten en effecten van de opties op populatieniveau voorleggen.

Kortom, een maatschappij heeft te maken met een dynamische patiënt-populatie, hetgeen leidt tot dynamiek in de budgetconsequenties en gezondheidsconsequenties over de jaren heen. Deze boeiende dynamiek en het opschalen van patiënt naar populatie in één enkel, gecombineerd economisch model geeft mijn inziens een belangrijke meerwaarde aan economische evaluaties. Echter, deze uitbreiding van onze methoden maakt eenvoudige kosten-effectiviteitsmodellen al snel complex, niet alleen wat betreft methode maar ook wat betreft databehoeftes.



Figuur 3. Schematische weergave van de constructie van een dynamische populatie ten behoeve van een economisch rekenmodel.

Complexiteit

Ik wil hiermee niet aangegeven dat complexer altijd beter is. Goed modelleren is het vinden van een balans tussen te complex en te eenvoudig. Te eenvoudige modellen doen geen recht aan de complexiteit van een aandoening, aan de veelheid van consequenties van behandelopties en aan lange termijn kosten en patiënt-uitkomsten gerelateerd aan keuzen in de zorg. Het bouwen van te complexe modellen daarentegen vergt veel tijd (vaak jaren), roept om zeer veel gedetailleerde informatie en data en maakt het bijna onmogelijk voor buitenstaanders de precieze werking en validiteit van een model te kunnen beoordelen. Dit laatste ervaren mijn collega's aan zowel de bovenloop en benedenloop van de Maas werkzaam binnen het NICE-project. In dit project is het onze taak de details van een beslissingsmodel dat wordt aangeleverd door

een belanghebbende (meestal de producent van een geneesmiddel) minutieus te ontleden en te beschrijven. Op basis hiervan adviseren wij de commissies binnen het National Institute of Health and Clinical Excellence van Engeland en Wales of een technologie wel of niet kosten-effectief is.

Nu is het bijzondere aan de situatie binnen NICE dat sprake is van een exacte definitie welke technologie wel en welke technologie niet als doelmatig kan worden bestempeld. Hier wordt in principe een drempelwaarde van 20.000 pond per QALY gehanteerd. Eventueel zijn onze westerburen coulant en wordt bij levensverlengende interventies die aan specifieke voorwaarden voldoen een hogere drempelwaarde gehanteerd. Een strikte drempelwaarde voor kosten-effectiviteit kent Nederland niet en deze grens zal inderdaad nog steeds vastgesteld moeten worden, maar tot nu toe heeft de politiek hier zich niet toe laten verleiden. Inmiddels lijkt er wel een consensus te ontstaan dat deze grens niet absoluut is, maar dat deze afhankelijk zal moeten worden van de ziektelast van de aandoening. Ik zou er echter voor willen pleiten dat de budgetconsequenties niet losstaand, maar integraal in deze afwegingen worden gemodelleerd en meegenomen in de beslissing. Een beslissingsalgoritme lijkt me ook hier een politieke keuze, net zoals het vrijstellen van weesgeneesmiddelen voor wat betreft economische evaluaties, maar transparantie en motivatie lijken me van doorslaggevend belang. Ik realiseer me hierbij wel degelijk dat de grenzen van de rigide budgetgrenzen in de zorg expliciet aandacht dienen te krijgen, vooral ook omdat dit een belangrijke beperkende factor is bij gebruik van economisch bewijs op het micro-beslissingsniveau (Erntoft, 2011).

Een toenemende complexiteit van modellen bedreigt echter de transparantie van de methoden en dus het geloof in de resultaten. Het transparant rapporteren van complexe modellen is niet eenvoudig, zeker niet als het werk van jaren in het keurslijf van een wetenschappelijk tijdschrift moet worden geperst (Caro et al., 2010). Maar ook hier kent internet zijn voordelen: technische bijlagen van honderden pagina's worden steeds vaker door tijdschriften geaccepteerd en (zoals bijvoorbeeld gebeurde bij onze studie inzake glaucoom) als achtergrondmateriaal online beschikbaar gemaakt.

Er blijven echter meerdere manieren om de kosten en consequenties van keuzeopties tussen technologieën in maat en getal te voorspellen. Ieder model is een brug tussen onderzoeksvraag en antwoord, op een bepaald moment, gegeven beschikbare data en kennis. Maar wat is nu een goed en

wat is nu een slecht model zult u zich afvragen. Ik kom daarbij op het punt van de validiteit en betrouwbaarheid van economische modellen.

In een overzicht van kosten-effectiviteitsmodellen gebruikt bij het evalueren van geneesmiddelen voor de behandeling van reumatoïde artritis werd geconstateerd dat van de 22 modellen er slechts één studie was waarbij een expliciet valideringonderzoek werd gemaakt (Bansbeck et al., 2008). In die studie gebruikten we een model om o.a. lange termijn effectiviteit te voorspellen en deze voorspelling werd gecontroleerd aan de hand van bestaande lange termijn vervolldata van patiënten (Welsing et al., 2006). In het voorbeeld van het dynamisch populatiemodel betreffende behandeling van ankyloserende spondylitis werd net zolang gesleuteld aan het rekenmodel totdat we in staat waren om op basis van cijfers van 1996 en later, precies die populatie te voorspellen die er in 2011 ook daadwerkelijk bestond zoals bekend was uit een observationele studie. Pas op dat moment hadden we voldoende vertrouwen in ons eigen complexe model om de voorspellingen vanaf 2012 te gaan berekenen. Ik constateer echter dat expliciete validering van zowel structuur als rekenmethoden binnen modellen veel te weinig wordt toegepast en de kennis hieromtrent binnen ons vakgebied meer aandacht dient te krijgen (Chilcott et al., 2010). Het is wat mij betreft dan ook noodzaak expliciet een valideringstoets van het modelontwerp uit te voeren en voordat analyses worden gedaan het model aan een peer-review oordeel te onderwerpen (McCabe and Dixon, 2000).

Ook al is sprake van een gevalideerd rekenmodel, dan nog kan sprake zijn van beperkte beschikbaarheid van de gegevens die nodig zijn voor de berekeningen. Als er slechts beperkte ervaring is opgedaan met gebruik van een technologie dan zijn data statistisch gezien onzeker en kan de inschatting van kosten en consequenties verschillende opties bemoeilijken. Het mag duidelijk zijn dat beslissers dan worden geconfronteerd met onzekerheid.

Er is in de regel sprake van twee beslissingen:

- 1) wordt een nieuwe technologie gegeven de beslissingsonzekerheid al dan niet opgenomen in gezondheidszorgsysteem en
- 2) is het efficiënt om verder onderzoek te initiëren om de statistische onzekerheid van relevante data te reduceren (Sculpher and Claxton, 2005).

Veel westerse landen, waaronder ook Nederland hebben een beleid geïntroduceerd dat neerkomt op een voorlopige vergoeding van geneesmiddelen. Gedurende een periode van enkele jaren wordt een veelbelovend geneesmiddel vergoed waarbij de afspraak wordt gemaakt om gegevens te verzamelen die tot dan toe ontbreken of niet afdoende zijn voor een definitief besluit. Zodoende wordt het risico van vroege toepassing van een technologie gedragen door meerdere partijen (Towse and Garrison, 2010). In de literatuur wordt gesproken van 'coverage with evidence development' of 'only in research' hetgeen we hier 'uitkomstenonderzoek' noemen in het kader van de beleidsregel Dure Geneesmiddelen. Om deze gegevensverzameling in de dagelijkse zorgpraktijk gericht op te kunnen zetten kan wederom gebruik worden gemaakt van rekenmodellen waarbij de bijdrage van statistische gegevens-onzekerheid aan beslissingsonzekerheid wordt vastgesteld. De analyses van de waarde van perfecte informatie, die uiteraard niet voorhanden is, is dan basis voor het inrichten van het uitkostenonderzoek. (Sculpher and Claxton 2005). In Nederland zullen dan de gegevens die relevant zijn in een periode van 4 jaar verzameld moeten worden en de eerste dossiers binnen de beleidsregel zijn nu, op het einde van deze 4-jaars periode, bij de Commissie Farmaceutische Hulp (CFH) in behandeling.

'Transferabiliteit'

Zoals ik eerder aangaf werden de effectiviteitgegevens van de Amerikaanse ATAC-trial gebruikt om in verschillende andere landen de kosten-effectiviteit van anastrozole te berekenen. En daarmee kom ik op het punt van de 'transferabiliteit'. Ik heb dit woord letterlijk vanuit het Engelse 'transferability' vertaald en ik moet erkennen dat ik hiermee geworsteld heb. Van een zeer gewaarde Belgische collega begreep ik dat je in Vlaanderen bijna voor de rechter wordt gesleept als je weigert om het keurige Nederlandstalige woord overdraagbaarheid te gebruiken. Ik denk echter dat het woord overdraagbaarheid niet de lading dekt van 'transferability', vandaar dat ik de Engelse benaming in deze rede zal blijven hanteren.

In de literatuur worden aspecten die de generaliseerbaarheid van kosten-effectiviteitsresultaten van het ene naar het andere land (of jurisdictie) beperken, transferability factoren genoemd (Drummond et al., 2009). Om bijvoorbeeld relevant te zijn in een nationale besluitvormingscontext moet een kosten-effectiviteitsmodel de lokale klinische realiteit zo goed mogelijk

weergeven. Nu zijn er uiteraard verschillen tussen de gezondheidszorgsystemen van landen. Maar vaak zijn er ook verschillen in zorgconstructen, zoals bijvoorbeeld de manier waarop en waar de zorg aan patiënten wordt geboden, en wie de zorg aanbiedt. Zelfs als zorgconstructen generaliseerbaar zijn tussen landen, is de manier waarop kosten en effecten gewaardeerd moeten worden lokaal afhankelijk. Zo kunt u zich voorstellen dat werkelijke kosten van medische zorg verschillen tussen landen, ook al worden die in eenzelfde monetaire eenheid, euro of dollar, uitgedrukt. Zo zijn bijvoorbeeld salarissen van zorgverleners zeker niet gelijk tussen landen.

Meer opvallend is dat waardering van kwaliteit van leven sterk verschilt tussen landen, zelfs als we hierbij een veelgebruikt instrument als de EQ5D gebruiken. Dit instrument beoogt over de grenzen heen bruikbaar te zijn om eenduidig een gezondheidstoestand in een utiliteit, waar ik eerder over sprak, uit te drukken. Maar als we middels de EQ5D vijf dimensies van gezondheid (mobiliteit, dagelijks functioneren, sociale activiteiten, angst / depressie, pijn) proberen te waarderen en vooral ook de invloed van medisch handelen daarbij, dan blijkt de incrementele kosten-effectiviteit landspecifiek. Ongeacht methodologische verschillen in meetmethoden om waardering van gezondheidstoestanden vast te stellen is sprake van cultuurverschillen. Zo lijkt verbetering in angst en depressie een belangrijkere rol te spelen in Spanje dan in Nederland en wordt op verandering in pijn door de Deense populatie als hoogste gescoord van 17 landen waarvoor EQ5D-tarieven bestaan (Knies et al., 2009).

Naast de transferability factoren zijn er ook maatschappelijk gezien verschillende maatstaven waarlangs een medische technologie gelegd wordt. Dit blijkt duidelijk uit het overzicht van de pharmaco-economische richtlijnen die in verschillende landen bestaan (ISPOR, 2011). Dergelijke richtlijnen beogen instructies te geven aan indieners van kosten-effectiviteitsdossiers en kunnen zodoende ook als beoordelingskader worden gezien. Zo wil de Nederlandse overheid bij de besluitvorming over 1B-geneesmiddelen niet dat er rekening wordt gehouden met niet-ziektegerelateerde kosten als gevolg van levensverlengend ingrijpen, terwijl in Zweden deze kosten wel als relevant worden geacht. In het NICE territorium Engeland en Wales vindt men kosten van arbeidsverzuim als gevolg van een aandoening niet meewegen in de besluitvorming terwijl dat in de meeste landen van het Europese vasteland wel het geval is.

In een recent literatuuroverzicht wordt gesteld dat het zinvol zou zijn een instrument te ontwikkelen dat het mogelijk maakt bij de internationale generalisatie van kosten-effectiviteitsgegevens een differentiatie aan te brengen tussen kritische en minder-kritische transferability factoren (Goeree et al., 2011). Een factor die niet of nauwelijks generaliseerbaar is, is dus in potentie een kritische factor en dit impliceert een hoge noodzaak tot aanpassing aan de lokale situatie, zoals bijvoorbeeld de kosten van medisch handelen. Een generaliseerbaar aspect zoals klinische effectiviteit noodzaakt in de regel niet tot een aanpassing. Mijns inziens is het gevalideerd transferability instrument van Welte hier een goede basis voor (Welte et al., 2004). Echter, dit instrument geeft niet aan hoe een factor moet leiden tot mathematische correctie, maar dat expliciete modellering hierbij een rol speelt lijkt mij zonneklaar.

Toekomst

Dames en heren, het voorgaande wil ik omschrijven als een spanningsveld waarbinnen mijn vakgebied opereert: de modellen waar we mee werken worden complexer, inzichten in de kosten en consequenties van nieuwe medische technologie zullen steeds vroeger in de levenscyclus van een technologie gegeven moeten worden, het is belangrijk de validiteit en onzekerheid van modellen expliciet te beschouwen, beslissers willen steeds sneller patiënten toegang kunnen geven tot een technologie, en bestaande buitenlandse modellen en gegevens kunnen niet zonder meer van het ene naar het andere land gengeneraliseerd worden.

Gegeven de grote verschillen tussen landen wat betreft gezondheidszorgsystemen en vooral ook de manier van financiering van deze zorg voorspel ik geen Europese, laat staan wereldwijde besluitvorming welke medische technologieën wel en welke niet voor gebruik en vergoeding in aanmerking komen. Dit geeft aan dat nationale beslissers ten behoeve van goed geïnformeerde beleidskeuzen zullen moeten kunnen beschikken over hun eigen landspecifieke economische modellen. Het ontwerpen en bouwen van meerdere landspecifieke modellen is zoals u kunt begrijpen tijdrovend en leidt internationaal gezien tot dubbel werk en wellicht ongewenste variatie zoals bleek uit de ATAC-casus. Daar komt nog bij dat de mankracht die nodig is om overal steeds weer het wiel uit te vinden, naar mijn ervaring simpelweg momenteel ontoereikend is.

Daarnaast blijven er ook in de toekomst internationale verschillen tussen gezondheidssystemen, de politieke arena, tradities, nationaal inkomen en dergelijke waardoor het dan ook niet vreemd is dat economische evaluaties en de uitkomsten daarvan zullen verschillen. Daarom zullen er ook verschillende beslissingen mogelijk zijn over een nieuwe medische technologie. Echter, er zijn een aantal universele aspecten aan economische studies: er moet sprake zijn van wetenschappelijk valide onderzoeksmethoden met daaraan gerelateerd transparante, efficiënte, voorspelbare en eerlijke besluitvorming (Neumann, 2009).

Hoe kunnen we dit spanningsveld de komende jaren tegemoet treden?

Is het afdoende modellen simpeler te maken zodat sneller en met minder inzet tot resultaten kan worden gekomen? Bijkomend voordeel van simpele modellen zou zijn dat het gevaar van eerdergenoemde paradox dat te complexe modellen niet gebruikt worden door de beslissers, voorkomen kan worden. Ik heb u er hopelijk van overtuigd dat dit niet de weg is om tot valide modellen en goed geïnformeerde beslissingen te komen, alhoewel natuurlijk wel geldt: simpel waar het kan, complex waar het moet. Naar mijn mening kunnen we dat niet alleen overlaten aan de onderzoekers en HTA-organisaties. Daarvoor is het noodzakelijk dat er een expliciete afstemming komt tussen leverancier van bewijs en de gebruiker van dat bewijs. Zelfs binnen het macroniveau van besluitvorming waarbij instanties betrokken zijn die beslissen over toelating enerzijds en vergoeding anderzijds is communicatie en afstemming op dit punt voor verbetering vatbaar (Henshall et al., 2011). Goed informeren en vooral ook opleiden van beslissers zijn hier belangrijke activiteiten.

Kan besluitvorming uitgesteld worden tot ieder land de tijd heeft gehad zijn eigen economisch model te ontwikkelen, analyseren en rapporteren? Volgens mij niet. Het feit dat er in veel westerse landen de mogelijkheid is voorlopige vergoeding in te stellen om zo patiënten vroeg toegang te geven tot potentieel effectieve en doelmatige zorg is hiervoor illustratief. Daar komt nog bij dat voor de industrie niet alleen de tijd tot markttoegang van evident belang is, maar ook de tijd tot vergoedingsbeslissing (Naci, 2011). Het is niet voor niets dat instanties zoals het CVZ en NICE ons als academische partners steeds vaker confronteren met harde deadlines.

Is het mogelijk meer menskracht te mobiliseren? Ja, dat kan en daar wordt ook druk aan gewerkt. Universiteiten in vele Europese en niet-Europese landen hebben inmiddels het vak health technology assessment in hun programma's staan, bijvoorbeeld als vak binnen de gezondheidseconomie of gezondheidswetenschappen. Momenteel is ISPOR, de International Society for Pharmacoeconomics and Outcomes Research binnen de ISPOR-Vision2020-werkgroep waar ik deel van uitmaak, druk doende wereldwijd een inventarisatie te maken op dit terrein. Dit moet leiden tot een overzicht van het aanbod van universitaire opleidingen wat betreft inhoud en capaciteit van HTA opleidingen, maar ook de behoefte van potentiële werkgevers. Ik zie echter naar mijn idee nog te weinig interesse van studenten geneeskunde en de interesse van studenten farmacologie is nog van bescheiden omvang. De studie onder Italiaanse cardiologen waar ik eerder naar refereerde liet echter zien dat kennis van evaluatie het begrip over mogelijkheden en onmogelijkheden van economische evaluatiestudies in relatie tot besluitvorming en doelmatige zorg vergroot. Ik zie hier dan ook nog mogelijkheden tot verdere ontwikkeling van ons onderwijsaanbod aan praktiserende klinici.

Ik constateer wel dat de Erasmus Universiteit actief is om internationale samenwerking wat betreft opleidingen voor mijn vakgebied te realiseren. De samenwerking met universiteiten van Oslo en Innsbruck in een gezamenlijke master-degree en de samenwerkingsinitiatieven richting de universiteiten van Peking en Shanghai zie ik dan ook de komende jaren als onderwerpen op mijn persoonlijke agenda. De samenwerking met collegae binnen ons instituut in de vorm van de master-specialisatie Global Health sluit daar naadloos op aan. Het opleiden van goede onderzoekers vind ik een primaire taak van de universiteit en ik zie in de nabije toekomst kansen om masteropleiding en promotietraject beter op elkaar af te stemmen middels bijvoorbeeld het ontwikkelen van, dan wel aansluiten bij een tweejarige onderzoeksmaster. Hierbij moet opgemerkt worden dat de problematiek betreffende financieringsstructuur van het huidige masteronderwijs nog wel een beperkende factor is, vooral voor buitenlandse studenten. Daarnaast zal verkend moeten worden of er voldoende animo en menskracht te mobiliseren zijn om de lastige horde van accreditatie van een dergelijke opleiding te kunnen nemen. Het mag overduidelijk zijn dat juist een dergelijke topmaster een sterk internationaal karakter dient te krijgen. Maar zorgvuldig geselecteerde en bijzonder gemotiveerde studenten maken met een researchmaster mijns inziens een vliegende start met een aansluitend

promotie-onderzoek. De ontwikkeling van onderwijsprogramma's en het afleveren van studenten is echter een kwestie van lange adem dus op de hele korte termijn geen oplossing voor genoemd spanningsveld.

Een mogelijke oplossing voor het spanningsveld van grote vraag, beperkt aanbod en hoge tijdsdruk zou ook kunnen zijn om modellen op basis van gepubliceerde gegevens na te bouwen en aan te passen. Praktisch gezien is dit zo goed als onmogelijk, gewoonweg omdat inzicht in ieder detail van een model noodzakelijk is. Het beschikbaar zijn van de feitelijke computermodellen en alle achterliggende technische informatie is de enige mogelijkheid anderen modellen te gebruiken (Essers et al., 2010). Ook maakt het beschikbaar zijn van elektronische modellen het mogelijk efficiënt en binnen strikte deadlines modellen aan te passen en ook meerdere modellen, bijvoorbeeld een diagnostiek model en een therapie model, tot een alomvattend model te combineren indien een onderzoeksvraagstelling daartoe leidt. Als voorbeeld wil ik hier een NICE-dossier noemen waarvoor ten behoeve van besluitvorming over een nieuwe generatie CT-apparatuur modellen door collegae van de University of York beschikbaar werden gesteld. Vijf modellen werden gecombineerd tot één: een model diagnostiek van coronair vaatlijden, een behandelingsmodel van patiënten met pijn op de borst, een lange termijn model dat de invloed van stralingsbelasting op levensverwachting weergeeft, een model dat kosten en effecten van beroerte berekende en een gezonde levensverwachting-model berekende, werden zo in een enkel model aan elkaar gekoppeld en doorgerekend (Westwood et al., 2011; Burgers et al., 2011).

Vooralsnog is de beschikbaarheid van elektronische modellen gebaseerd op persoonlijke, internationale relaties van onderzoekers, vooral tussen universitaire instituten. Dit geldt ook voor de meeste databestanden die nodig zijn om gegevens te verkrijgen om de economische modellen mee te vullen. De grote internationale pharmabedrijven zijn ons op dit punt ver vooruit: het hoofdkantoor bouwt een model dat vervolgens gebruikt wordt door de nationale maatschappijen ten behoeve van vergoedingsdossiers. Ik vraag me echter soms wel af of bij deze exercities afdoende rekening wordt gehouden met de transferability factoren.

Hoe kunnen we dit spanningsveld dan wel aanpakken? Mijns inziens is het, zoals ik illustreerde met de ATAC-casus, noodzakelijk internationale coördinatie en afstemming te realiseren met betrekking tot economische modellen, en dat kan eenvoudig beginnen in Europa. Zo is er sprake van Europese

samenwerking betreffende toelatingsbeslissingen geneesmiddelen tot de Europese gezondheidszorgmarkt. Daarnaast bestaat het samenwerkingsverband EUnetHTA waarin sinds 2008 34 landen participeren om betrouwbare, tijdige en internationaal transfereerbare informatie te genereren ten behoeve van evaluatiestudies. En volgens mij valt het ontbreken van coördinatie, afstemming en het beschikbaar maken van economische evaluatiemodellen precies in het gat dat tussen deze twee organisaties ligt. Zeker ook omdat voor geneesmiddelen in veel landen een strikt systeem bestaat voor toegang tot de markt en vergoeding, terwijl dit voor zo iets als medische apparatuur en medische hulpmiddelen en procedures niet of nauwelijks bestaat.

Het verbaast me dan ook niet dat zeer recent het directoraat van Gezondheid en Consumenten van de Europese Unie een oproep gedaan heeft om voor 4 november jongstleden voorstellen in te dienen om te komen tot een permanent secretariaat voor Europese samenwerking op het gebied van Health Technology Assessment. Nu moet me wel van het hart dat de oproep bijzonder vaag is en ik me afvraag of het in te stellen secretariaat wel echt die stappen kan zetten die ik nodig acht, want het beschikbare budget van €100.000 is beperkt terwijl de ambitieuze taken onbegrensd lijken.

Maar wat kunnen we van een internationale samenwerking op het terrein van economische modellen verwachten en hoever zou het mandaat moeten strekken? Ik wil u als voorbeeld van dit idee een recent voorstel voor een Europees project schetsen dat ik onlangs in samenwerking met 9 instituten uit 5 verschillende landen heb ingediend. CHEETA is de projectnaam en dit staat voor Collaboration for Health Economic Evaluation and Transferability Assessment. Binnen CHEETA wordt een ziekteprogressie-model ontwikkeld dat het natuurlijk beloop van hartfalen simuleert op basis van een dynamische populatie, discrete-event simulation model. Binnen de 5 deelnemende landen wordt empirisch onderzoek gedaan naar de dagelijkse praktijk betreffende diagnostiek en behandeling van deze patiëntengroep. De vijf landen zullen het gezamenlijk ontwikkeld en gevalideerd model doorrekenen met landen-specifieke data, zowel wat betreft diagnostische- en behandelingsstrategieën, kosten en kwaliteit van levenwaardering. Hiermee kunnen dan eenduidig de kosten en consequenties van lokale beleidsopties betreffende preventie, diagnostiek en therapie geschat worden. Het fraaie van dit consortium is dat alle deelnemers unaniem van mening waren dat het een taak van Europa is om niet alleen een dergelijk

project te financieren maar ook om de kennis, de vergaarde data, en het elektronische model, in lijn met het E-science principe (Keulemans en Visser, 2011) voor lidstaten beschikbaar te maken. Ik zie hier een rol voor een Europees kennisinstituut weggelegd, dat bestaat uit een netwerk van academische onderzoekscentra. Ik hoop natuurlijk van harte dat het CHEETA-voorstel wordt gehonoreerd om te laten zien dat een dergelijk netwerk van academische centra en beslissers geen fictie is, maar realiteit kan worden. Ik neem mij in dit kader voor, ook los van het wetslagen van dit enkele project, mij ook in te gaan zetten deze ideeën verder te ontwikkelen en ook buiten Europa te brengen. Een belangrijke randvoorwaarde hierbij is het doorbreken van eerdergenoemde paradox. En dat kan! Door beslissers helder te informeren over transparante en internationaal afgestemde economische modellen kunnen we er voor zorgen dat ondanks toenemende complexiteit, beslissers onze studies en modellen juist meer gaan gebruiken bij de efficiënte inrichting van zorgsystemen.

Alleen op die manier kunnen we bereiken dat sprake is van evaluatie zonder grenzen.

Dankwoord

Mijnheer de rector magnificus, dames en heren, het is lastig om alle facetten van de onderwijs- en onderzoekswereld waar ik mij in bevind in deze korte tijd met u te delen. Maar ik hoop u er van te hebben overtuigd dat de universiteit, ook door traditionele gebeurtenissen zoals deze, voor mij een prachtige werkplek is. Daarom wil ik het College van Bestuur van de Erasmus Universiteit en de Raad van Bestuur van het Erasmus Medisch Centrum en in het bijzonder de decaan van de faculteit der Geneeskunde en Gezondheidswetenschappen, de hoogleraar Huib Pols danken voor het in mij gestelde vertrouwen. Ik ben bij instituut Beleid & Management Gezondheidszorg op een rijke voedingsbodem terecht gekomen. Ik dank het instituut en in het bijzonder de sectieleider professor Werner Brouwer en de pro-decaan professor Wilfried Notten dan ook voor de kansen die ik hier krijg.

Het is u ongetwijfeld opgevallen dat ik in mijn rede niemand bij naam heb genoemd. Dat wil niet zeggen dat ik alleen werk, integendeel. In de loop der jaren heb ik met collegae van de universiteiten van Nijmegen en Maastricht en nu inmiddels ook met Rotterdamse collegae gewerkt. Daarbuiten waren en zijn er legio profit en non-profit organisaties waarmee samenwerkingen zijn aangegaan. Gezien het feit dat het vandaag de dag van de duurzaamheid is, wil ik hier u besparen de lange lijst betrokkenen aan te moeten horen. Simpelweg dank ik bij dezen allen voor boeiende projecten, prettige en productieve samenwerking, stimulerende discussies, en intrigerende visies. Deze dank geldt ook voor collegae in verschillende externe gremia als mijn Gezondheidsraad-commissies, wetenschappelijke adviesraden van bedrijven en de Commissie Farmaceutische Hulpmiddelen. Ik vind het motiverend om op deze manier mijn steentje te kunnen bijdragen aan het verbeteren van ons gezondheidszorgsysteem in de brede zin.

Slecht twee collegae wil ik bij naam noemen. Beste Frans Rutten en Werner Brouwer; de Utrechtse gesprekken die we voerden om af te tasten of wisseling van mijn spelerstenu voor beide partijen een meerwaarde kon hebben, waren stimulerend en prettig. In die zin was dit een voorbode voor de samenwerkingen alhier.

Vandaag is het de 11^e van de 11^e van het 11^e jaar van deze eeuw. En wellicht weet u dat vandaag de dag is dat de carnavalsverenigingen in het zuiden van het land, stroomopwaarts aan deze Maas hun nieuwe prins bekend maken.

Ongeacht dat ik van geboorte Limburger ben en mijn zachte 'G' duidelijk opvalt naast de Rotterdamse 'R' ben ik zelf geen echte vierder. Het is een wijdverbreid misverstand dat carnaval met het gekkengetal 11 te maken heeft. Elf november heeft al sinds de Middeleeuwen de betekenis: avond voor het vasten, in het Limburgs heet carnaval dan ook Vastelaovond. Dat zit zo: in die tijd werd er ook voor de kerst gevast. Niet uit religieuze overwegingen, maar uit armoede. Het vlees uit de slachtdaand november moest namelijk bewaard worden tot kerst. Met het vasten werd begonnen op 12 november. De dag ervoor werden nog snel de niet houdbare producten met veel feestvertoon opgegeten. Waarom dan de elfde? Omdat het destijds toch al een feestdag was, namelijk Sint-Maarten, een feest dat vooral in Noord-Holland wordt gevierd. En zo ziet u dus dat Noord-Hollanders en Limburgers, ondanks de afstand van 200km veel gemeen hebben.

Maar ook hier speelt een transferability factor een rol, want voor veel landen is 11 november een feestelijke dag. Onze burens, de Belgen, vieren dan het einde van de Eerste Wereldoorlog, de Wapenstilstandsdag. Deze dag heet in Polen onafhankelijkheidsdag. Dat geldt tevens voor Angola, waar op 11 november 1975 een einde kwam aan het Portugese koloniale bewind. Maar ook in Groot-Brittannië en de Verenigde Staten worden respectievelijk Remembrance Day en Veterans Day gevierd. Dat het voor het overzeese rijkdeel Sint-Maarten een feestdag is, zal iedereen kunnen begrijpen.

Deze editie van de 11^e van de 11^e zal voor mij persoonlijk een fraaie herinnering blijven. Mijn dag was niet mogelijk geweest zonder belangeloze inzet van anderen. Karin van Oorschot, het feit dat je me eigenlijk altijd een stap voor bent, ook wat betreft het organiseren van deze dag, is tekenend voor je kwaliteiten als secretaresse. Judith van den Bosch (Pallas), Werner Brouwer en Frans Rutten, dank voor jullie constructieve commentaar op de concepttekst van mijn oratie. Over de vertaling van 'transferability' moeten we het zeker nog eens hebben. Beste Aukje van Gestel, Geert Frederix en An Tranduy, dank voor het lenen van jullie studies en het voorkomen van medische inhoudelijke fouten in mijn rede.

Lieve familie en vrienden, ik durf jullie toe te zeggen dat dit mijn laatste oratie is. Dank voor jullie aanwezigheid nu, en interesse en steun al die jaren, maar vooral ook voor samen sporten, vrije tijd en uren bijpraten over alles wat het leven brengt.

Lieve pap, wie had ooit gedacht dat ik al rolschaatsend door de gangen van jouw Radboud ziekenhuis zo gegrepen zou worden door de boeiende wereld van de gezondheidszorg. Lieve mam, zonder jou had ik mijn vwo nooit afgemaakt en zo sta jij dus aan de basis van mijn academische levensweg. Jullie namen draag ik ook op déze toga.

Lieve Ellen, 200 kilometer treinen of rijden we om bij elkaar te kunnen zijn. Dit zal nog wel even zo blijven maar deze afstand is niks voor hetgeen jij me geeft: je openheid, liefde en steun, je gezelschap in de bergen, tijdens reizen, op de fiets, op het ijs of gewoon bij jou of mij thuis. Je bent een maatje en ik ben trots op jou. En met jouw lieve dochter en je grote kerels erbij, zijn we samen een prima touwgroep.

Lieve Robbe, groot manneke, ik kan genieten van alle momenten samen, voor de houtkachel, met spek en ei op woensdagmiddag, langs de lijn op zaterdag, samen mountainbiken en hardlopen, of op vakantie in de bergen. Ik ben zo blij dat jij hier bent, vandaag, morgen, altijd.

Dames en heren studenten, bachelors, masters en PhDs, ik richt me ter afsluiting van deze rede tot u. Meer dan 20 jaar geleden gaf een docent hier aanwezig mij de opdracht het boek van Drummond te bestuderen. Ik kon toen niet vermoeden dat dit boek mijn vakgebied zou worden. Na enkele omzwervingen, zowel commercieel als non-profit, bleek de academische wereld het beste passend. Evaluatie van innovatieve medische technologie is continue boeiend, motiverend en zeer relevant voor de gezondheidszorg, nationaal en internationaal. Ik hoop uw enthousiasme te hebben gewekt en ontmoet u graag in academische samenwerking en debat. De wetenschap is grenzeloos!

Ik heb gezegd.

Referenties

Bansback N, Ara R, Karnon J, Anis A. Economic evaluations in rheumatoid arthritis; a critical review of measures used to define health states.

Pharmacoecon 2008; 26 (5): 395-408.

Banta D and Oortwijn WJ. The Netherlands. *Int J Techn Assessm Health Care* 2009; 25 (suppl 1): 143-147.

Baum M, Budzar AU, Cuzick J, Forbes J, Houghton JH, Klijn JG, Sahmoud T; ATAC Trialists' Group. Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. *Lancet* 2002; 359 (9324): 2131-2139. Erratum in: *Lancet* 2002; 360 (9344): 1520.

Brouselle A and Lessard C. Economic evaluation to inform health care decision making: promise, pitfalls and a proposal for an alternative path. *Soc Sci Med* 2011; 72: 832-839.

Burgers LT, Redekop WK, Armstrong N, Lhachimi SK, Severens JL, Westwood M, Sculpher MJ, Walker S, McKenna C, Al M. Using five existing models to comprehensively model the cost-effectiveness of a high definition CT scanner in a coronary artery disease population: a NICE diagnostic guidance project (abstract). *Value Health* 2011; in press. Poster presentation at conference International Society for Pharmacoeconomics and Outcome Research. November 5-8, Madrid, Spain 2011;.

Caro JJ, Möller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluation? *Value Health* 2010; 13 (8): 1056-1060.

Chalkidou K and Rawlins M. Pharmacogenetics and cost-effectiveness analysis: a two way street. *Drug Discov Today* 2011; 16 (19-20):873-877.

Chilcott J, Tappenden P, Rawdin A, Johnson M, Kaltenthaler E, Paisley S, Papaioannou D, Shippam A. Avoiding and identifying errors in health technology assessment models: qualitative study and methodological review. *Health Technol Assess* 2010; 14 (25): iii-iv, ix-xii, 1-107.

Costa-Scharplatz M, Van Asselt ADI, Bachmann LM, Kessels AGH, Severens JL. Cost-effectiveness of pharmacogenetic testing to predict treatment response to ACE inhibitor. *Pharmacogen Genom* 2007; 17: 359-368.

Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, Reed SD, Rutten F, Sculpher M, Severens J. Transferability of economic evaluations across jurisdictions: ISPOR good research practices task force report. *Value Health* 2009; 12(4): 409-418.

Elsinga E en Rutten FFH. Medische technology assessment; toepassing in de Nederlandse gezondheidszorg. *Med Contact* 1995; 50 (1): 13-18.

Erntoft S. Pharmaceutical priority setting and the use of health economic evaluations: a systematic literature review. *Value Health* 2011; 14 (4): 587-599.

Essers BAB, Shanley SC, Tjan-Heijnen VCG, Severens JL, Novak A, Pompen M, Oron UH, Joore MA. Transferability of model-based economic evaluations: the case of trastuzumab for the adjuvant treatment of HER2-positive early breast cancer in the Netherlands. *Value Health* 2010; 13(4): 375-380

Frederix GWJ, Severens JL, Hovels AM, Raaijmakers JAM, Schellens JHM. Reviewing the cost-effectiveness of endocrine early breast cancer therapies. Influence of differences in modelling methods on outcomes. *Value Health* (in press).

Gezondheidsraad. Waar voor ons geld; beslissingen over publieke investeringen in gezondheidszorgonderzoek. Den Haag: *Gezondheidsraad* 2010; publicatienr. 2010/16.

Goeree R, He J, O'Reilly D, Tarride JE, Xie F, Lim M, Burke N. Transferability of health technology assessments and economic evaluations: a systematic review of approaches for assessment and application. *ClinicoEcon Outcom Res* 2011; 3: 89-104.

Henshall C, Mardhani-Bayne L, Frønsdal KB, Klemp M, on behalf of the HTAi Policy Forum. Interactions between health technology assessment, coverage, and regulatory processes; emerging issues, goals, and opportunities. *Int J Techn Assessm Health Care* 2011; 27 (3): 253-260.

Hillner BE. Benefit and cost-effectiveness of anastrozole versus tamoxifen as initial adjuvant therapy for patients with early-stage estrogen receptor-positive breast cancer. *Cancer* 2004; 101 (6): 1131-1322.

Hoomans T, Fenwick EA, Palmer S, Claxton K. Value of information and value of implementation: application of an analytic framework to inform resource allocation decisions in metastatic hormone-refractory prostate cancer. *Value Health* 2009; 12 (2): 315-324.

ISPOR. Pharmacoeconomic guidelines around the world. In: ispor.org. 2011. <http://www.ispor.org/PEguidelines/index.asp> (18 oktober 2011).

Kelly M, Morgan A, Ellis S, Younger T, Huntley J, Swann C. Evidence based public health: a review of the experience of the National Institute of Health and Clinical Excellence (NICE) of developing public health guidance in England. *Soc Sci Med* 2010; 71: 1056-1062.

Keulemans M en De Visser E. Pottekijkers eisen open keuken. *De Volkskrant*, katern Wetenschap; 22 oktober 2011.

Keulemans M en Van Raaij B. Valsspelen in de wetenschap. *De Volkskrant*, katern Wetenschap; 17 september 2011.

Knies S, Evers SMAA, Candel MJJM, Severens JL, Ament AJHA. Utilities of the EQ5D; transferable or not? *PharmacoEcon* 2009; 27 (9): 767-779.

McCabe C and Dixon S. Testing the validity of cost-effectiveness models. *PharmacoEcon* 2000; 17 (5): 501-513.

Naci H. The increasing need to harmonize evidence demands of regulators, payers, and health technology assessment bodies in Europe; opportunities and challenges. *ISPOR connections* 2011; 17 (4): 4-7.

Neumann P. Lessons for health technology assessment; it is not only about the evidence. *Value Health* 2009; 12 (suppl 2): S45-S48.

Retel VP, Grutters JPC, Van Harten WH, Joore MA. Value of research and value of development in early assessments of new technologies. *Submitted for publication*.

Rutten FFH. Medische Technology Assessment; een uitdaging voor wetenschap en beleid. Inaugurale rede, Erasmus Universiteit Rotterdam, 1 december 1988.

Sculpher M and Claxton K. Esthablishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty; when is there sufficient evidence? *Value Health* 2005; 8 (4): 433-446.

Sendi PP and Briggs AH. Affordability and cost-effectiveness: decision making on the cost-effectiveness plane. *Health Econ* 2001; 10: 675-680.

Torbica A and Fattore G. Understanding the impact of economic evidence on clinical decision making: a discrete choice experiment in cardiology. *Soc Sci Med* 2010; 70: 1536-1543.

Towse A and Garrison LP. Can't get no satisfaction? Will pay for performance help?; Towards and economic framework for understanding performance-based risk-sharing arrangement for innovative medical products. *PharmacoEcon* 2010; 28 (2): 93-102.

Tranduy A, Boonen A, Van der Laar MAFJ, Franke AC, Severens JL. A discrete event modelling framework for simulation of long term outcomes of sequential treatment strategies for ankylosing spondylitis. *Ann Rheum Dis* (in press).

Van Gestel A, Grutters J, Schouten J, Webers CA, Beckers H, Joore M, Severens JL. The role of the expected value of individualized care in cost-effectiveness analysis and decision making. *Value Health* (in press).

Van Gestel A, Severens JL, Webers CAB, Beckers HJM, Jansonius NM, Schouten JSAG. Modeling complex treatment strategies: construction and validation of a discrete event simulation model for glaucoma. *Value Health* 2010; 13(4): 358-367.

Van Velden ME, Severens JL, Novak A. Economic evaluations of healthcare programmes and decision-making; the influence of economic evaluations on different healthcare decision-making levels. *PharmacoEcon* 2005; 23: 1075-1082.

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415:530-536.

Wagner TH, Heisler M, Piette JD. Prescription drug co-payments and cost-related medication underuse. *Health Econ Policy Law* 2008; 3: 51-67.

Welsing PMJ, Severens JL, Hartman M, Van Gestel AM, Van Riel PLCM, Laan RFJM. The initial validation of a Markov model for the economic evaluation of (new) treatments for rheumatoid arthritis. *PharmacoEcon* 2006; 24 (10): 1011-1020.

Welte R, Feenstra F, Jager H, Leidl A. A decision chart for assessing and improving the transferability of economic evaluation results between countries. *PharmacoEcon* 2004; 22 (13): 857-876.

Westwood M, Al M, Burgers L, Redekop K, Lhachimi S, Armstrong N, Raatz H, Misso K, Severens J, Kleijnen J. Computed tomography (CT) scanners for cardiac imaging - Somatom Definition Flash, Aquilion One, Brilliance iCT and Discovery CT750: diagnostic assessment report. In: guidance.nice.org.uk 2011. <http://guidance.nice.org.uk/DT/3/DiagnosticAssessmentReport/pdf/English> (20 oktober 2011).

World Health Organisation. The World health report: health systems financing: the path to universal coverage. Geneva: *World Health Organisation* 2010.

IJzermans MJ and Steuten LMG. Early assessment of medical technologies to inform product development and market access: a review of methods and applications / *Appl Health Econ Health Policy* 2011; 9 (5): 331-347.



Fotograaf: Levien Willemse

Prof.dr. J.L. (Hans) Severens (1963) is hoogleraar Evaluatie in de gezondheidszorg bij het instituut Beleid & Management Gezondheidszorg (iBMG) van de Erasmus Universiteit Rotterdam. Hij studeerde Gezondheidswetenschappen, afstudeerrichting Beleid & Beheer aan de universiteit van Maastricht (toenmalig de Rijksuniversiteit Limburg). Hij promoveerde in 1999 aan de universiteit van Nijmegen op het proefschrift 'Some methodological issues in economic evaluation in health care' en was gedurende die periode visiting fellow bij het Center for Health Economics, University of York, UK. In de periode 2001-2010 was hij hoogleraar bij het instituut CAPHRI aan de universiteit van Maastricht. Zijn onderzoek en onderwijs betreffen toepassingen en methoden van economische evaluatie van medische technologie. Betrokkenheid bij de inrichting van de gezondheidszorg blijkt o.a. uit lidmaatschappen van de Gezondheidsraad en de Commissie Farmaceutische Hulp. Internationale betrokkenheid blijkt o.a. uit bestuurslidmaatschap en diverse werkgroepen van ISPOR, consultaties voor MUNDO en incidentele adviseurschappen bij zowel profit als non-profit organisaties.

'Evaluatie zonder grenzen' omschrijft de methoden en toepassingen van economische evaluaties van medische technologie. Er is de laatste decennia internationaal gezien een duidelijke positieve trend in het gebruik van economische evaluatiemodellen. Dergelijke modellen worden in een toenemend aantal westerse en niet-westerse landen expliciet gebruikt om beslissingen over toepassing van medische technologie in de gezondheidszorg te ondersteunen. Het gaat hierbij veelal om beslissingen om een specifieke medische technologie (zoals diagnostische tests, geneesmiddelen of complexe screeningprogramma's) wel of niet te vergoeden. De druk vanuit de maatschappij om zo vroeg mogelijk in de levenscyclus van een technologie te beslissen over toelating tot de zorg neemt toe. Dit kan leiden tot onzekerheid van kosteneffectiviteit van een technologie op het moment dat besluitvorming wenselijk is, maar ook dat buitenlandse gegevens gebruikt worden voor nationale besluitvorming. Daarbij laat de huidige trend zien dat de methoden van economische modellen om meerdere redenen steeds complexer en tijdsintensiever worden. Door deze ontwikkelingen bestaat de kans dat beslissers minder geneigd zijn dergelijke modellen te gebruiken waarvoor ze bedoeld zijn, namelijk het ondersteunen van goed geïnformeerde beslissingen over medische technologie. Hans Severens benadrukt dat internationale samenwerking en coördinatie van economisch evaluatieonderzoek een oplossing is voor deze paradox.



Erasmus Universiteit Rotterdam
instituut Beleid & Management Gezondheidszorg
Burgemeester Oudlaan 50, 3062 PA Rotterdam
Postbus 1738, 3000 DR Rotterdam

Tel. 010 - 408 8555
E-mail info@bmg.eur.nl
Internet www.bmg.eur.nl

ISBN
978-94-90420-17-8