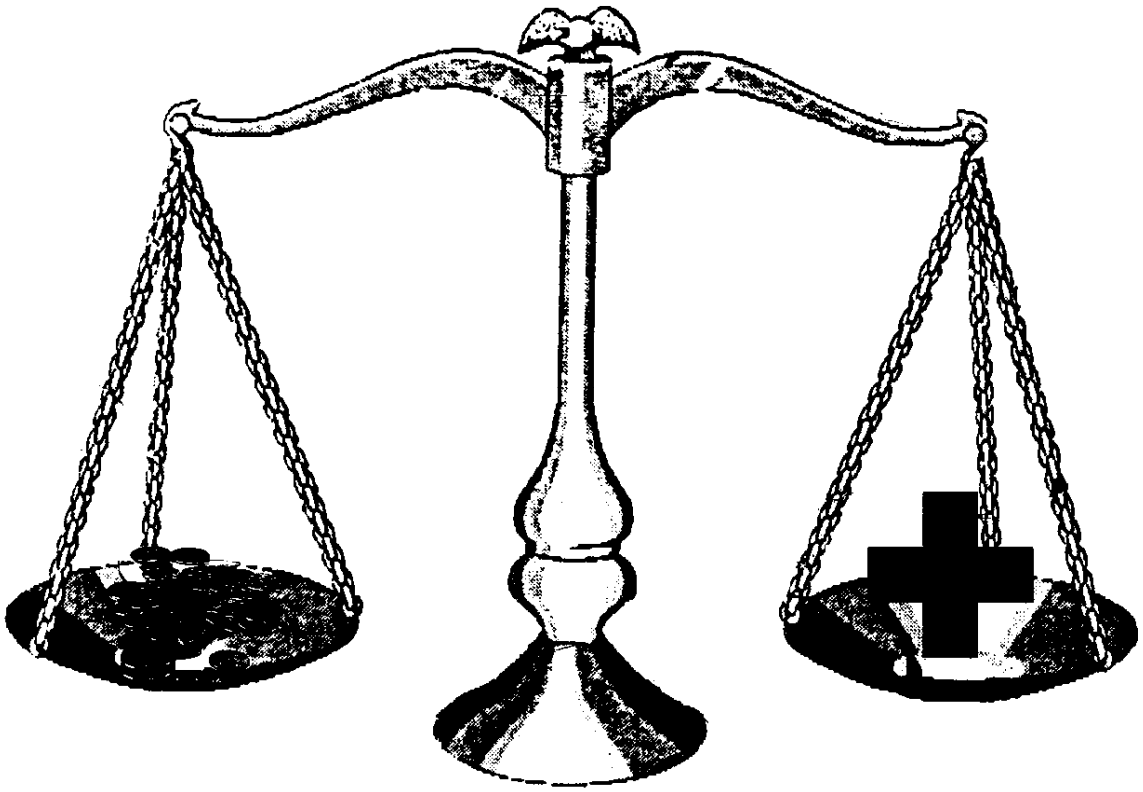
Modelleren van 'kwaliteit' in Data Envelopment Analysis

Benchmarking van de efficiëntie van de personeelsinzet op klinische afdelingen in ziekenhuizen



ORM-seminar

Juli 2008

Michiel Janssen

0488275

Begeleiding:

Ir. J.A.M. Hontelez (*UvA*)

Ir. A.J.M. Humme (*ORTEC*)

Universiteit van Amsterdam

Faculteit voor Economie en Bedrijfskunde

Afdeling Kwantitatieve Economie

Samenvatting

Het consultancy bedrijf *ORTEC* voert jaarlijks een benchmark uit op de klinische afdelingen van een aantal ziekenhuizen. Het doel van deze benchmark is om de onderlinge prestaties van de verschillende ziekenhuizen op het gebied van de efficiëntie van de personeelsinzet te vergelijken en om daarbij aan te geven op welke punten (in de zin van personeelsreductie of productiegroei) deze ziekenhuizen zich nog kunnen verbeteren. Voor een gedeelte van deze benchmark is gebruik gemaakt van de techniek Data Envelopment Analysis (DEA). Een techniek die voor ieder ziekenhuis dat aan de benchmark deelneemt een relatieve (ten opzichte van de andere ziekenhuizen) efficiëntiescore en een verbeterpotentieel levert.

Uit de evaluatie van de resultaten van het DEA-gedeelte van de benchmark uit 2007 is naar voren gekomen dat de dimensie 'kwaliteit van de zorg' nog onvoldoende is meegenomen in het DEA-model. Dit komt doordat in het model van 2007 alleen variabelen zijn opgenomen die betrekking hebben op kostenefficiëntie, waardoor een eventueel verband met de kwaliteit van de geleverde zorg wordt genegeerd in het bepalen van de efficiëntiescores en het verbeterpotentieel van de ziekenhuizen.

Het doel van deze scriptie is dan ook om een manier te vinden om een 'kwaliteitsdimensie' toe te voegen aan het DEA-model. Door deze kwaliteitsdimensie mee te nemen in DEA moet worden voorkomen dat de verschillende ziekenhuizen (DMU's) in het DEA-gedeelte van de benchmark slechts op kostenefficiëntie worden beoordeeld, zonder hierbij de kwaliteit van de zorg van deze ziekenhuizen mee te nemen in deze efficiëntiebeoordeling en de berekening van het verbeterpotentieel.

Door 'kwaliteit' expliciet te modelleren in het DEA-model kan een 'bredere' efficiëntiebeoordeling ontstaan voor de benchmark van 2008. Hiermee kan ook een eventuele trade-off tussen kostenefficiëntie en kwaliteit van zorg gekwantificeerd worden, zodat meer realistische verbeterpotentiëlen kunnen worden berekend voor de verschillende klinische afdelingen van de deelnemende ziekenhuizen.

De uitdaging van deze scriptie is daarom tweeledig:

1. Hoe kan kwaliteit van zorg worden gemeten en
2. hoe kan dit op zinvolle wijze worden geïntegreerd in DEA?

Ten eerste is aan de hand van bestaande literatuur onderzocht hoe 'kwaliteit van zorg' gemeten kan worden. Een aantal veel geciteerde servicemodellen is onderzocht (hoofdstuk 2), zodat de verschillende componenten waar de kwaliteit van een dienst uit bestaat in kaart kunnen worden gebracht. Uiteindelijk

is ervoor gekozen om de verschillende kwaliteitsfacetten te kwantificeren door gebruik te maken van zogenaamde kwaliteitsindicatoren.

Omdat het aantal aan de benchmark deelnemende ziekenhuizen – voor DEA-begrippen – relatief klein is, moet het aantal modelvariabelen ook klein worden gehouden om tot bruikbare resultaten te komen (zie paragraaf 3.1). Om deze reden wordt (in paragraaf 5.2) een relatieve 'kwaliteitsindex' geïntroduceerd die, middels een bepaalde weging, is samengesteld uit de verschillende kwaliteitsindicatoren zodat één samengestelde 'kwaliteitsvariabele' ontstaat. Deze variabele heeft als doel om uiteindelijk op zinvolle wijze in het DEA-model te worden geïntegreerd zodat een eventuele trade-off tussen kwaliteit en kostenefficiëntie gemodelleerd kan worden.

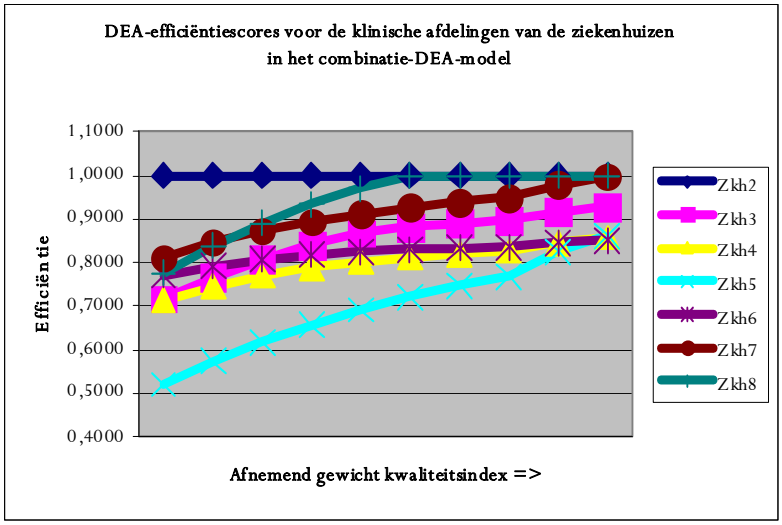
De kwaliteitsindex is 'relatief' in de zin dat het best presterende ziekenhuis op het gebied van een bepaald kwaliteitsaspect voor de bijbehorende kwaliteitsindicator een maximale score van 1 krijgt toebedeeld binnen dit deelgebied van de kwaliteitsindex, en het slechts presterende ziekenhuis een score van 0. Door middel van een weging (die door *ORTEC* in overleg met de deelnemende ziekenhuizen is bepaald) ligt de uiteindelijke kwaliteitsindex (een weging van de verschillende indicatorscores) eveneens tussen 0 (zeer slechte kwaliteit) en 1 (zeer goede kwaliteit) voor ieder ziekenhuis. Gevoeligheidsanalyse toont aan (zie paragraaf 5.4.1.) dat deze kwaliteitsindex tamelijk stabiel is voor de gekozen weging (voor de toepassing op de klinische afdelingen van de deelnemende ziekenhuizen).

Nu, met de introductie van de kwaliteitsindex, 'kwaliteit van zorg' tot op zekere hoogte meetbaar is gemaakt, is onderzocht hoe de kwaliteitsindex op zinvolle wijze geïntegreerd kan worden in DEA. Hiertoe zijn verschillende (uitbereidingen op) DEA-methodes uit de literatuur getoetst op hun bruikbaarheid om de kwaliteitsindex zinvol in DEA te kunnen integreren.

In deze scriptie is (in overleg met *ORTEC*) aangenomen dat de schaalopbrengsten op de klinische afdelingen van ziekenhuizen constant mogen worden geacht, waardoor alleen de toepasbaarheid van uitbreidingen op het CCR-model (dat uitgaat van constante schaalopbrengsten) zijn onderzocht (in hoofdstuk 3 en 4). De onderzochte uitbreidingen bestaan uit: 'absolute gewichtsrestricties', 'absolute virtuele gewichtsrestricties', 'assurance regions', 'virtuele assurance regions', Q-DEA en een grafische methode (al deze technieken worden in hoofdstuk 4 besproken).

Ook zijn er criteria opgesteld waarmee de voor- en nadelen van de verschillende uitbreidingen van het CCR-model met elkaar kunnen worden vergeleken. Aan de hand van deze criteria (zie paragraaf 4.4) is er voor gekozen om het CCR-model uit te breiden met 'virtuele assurance regions' van 'type I' (VAR-I restricties), waarmee het (virtuele) gewicht van de kwaliteitsindex binnen de efficiëntiescore kan worden gekozen en gegarandeerd. Dit betekent dat 'kwaliteit' altijd met voldoende gewicht wordt meegenomen in de benchmark.

Het combinatie-DEA-model, waarin zowel de kwaliteitsindex als VAR-I restricties zijn opgenomen, is toegepast op de data voor de benchmark Kliniek 2008 (zie hoofdstuk 5). Uit de resultaten hiervan kan een aantal zaken worden geconcludeerd:



Figuur 0: Efficiëntiescores voor de klinische afdelingen van de verschillende ziekenhuizen (zkh's) bij verschillende wegingen van de kwaliteitsindex

- Voor de deelnemende ziekenhuizen bestaat er een zekere trade-off tussen kostenefficiëntie en kwaliteit van zorg in bredere zin. Dit is te zien in figuur 0: wanneer het minimale gewicht van de kwaliteitsindex in het model hoger wordt gesteld, neemt de efficiëntiescore af;
- Het bestaan van deze trade-off maakt dat de voorgestelde 'brede' efficiëntiescores, waarin deze trade-off expliciet is gemodelleerd, 'eerlijker' verschillen in (brede) efficiëntie blootleggen tussen de deelnemende ziekenhuizen dan wanneer 'standaard' DEA wordt gebruikt.
- Doordat de kwaliteits-kostenefficiëntie trade-off is meegenomen in het DEA-model kunnen realistischere verbeterpotentiëlen worden berekend dan met 'standaard DEA', Hierdoor kan aan de ziekenhuizen een realistischer beeld worden voorgespiegeld van potentiële efficiëntiewinst;
- De gekozen methode (VAR-I) leidt tot relatief eenvoudig interpreteerbare (radiale) efficiëntiescores en (qua mathematische uitdrukking gemakkelijke) verbeterpotentiëlen. Dit maakt het uitleggen van de gebruikte methode gemakkelijker in de, over het algemeen, minder kwantitatief onderlegde medische wereld. Daarentegen geldt voor de verbeterpotentiëlen niet meer - zoals in 'standaard DEA' wel geldt – dat een zogenaamde 'outputoriëntatie' van het model ertoe leidt dat alle outputtargets gelijk blijven of toenemen, terwijl de inputniveaus gelijk blijven, of, in het geval van een zogenaamde 'inputoriëntatie', dat alle inputtargets gelijk blijven of afnemen, terwijl de outputniveaus gelijk blijven. Dit maakt de verbeterpotentiëlen in dit opzicht weer minder gemakkelijk uit te leggen dan in 'standaard DEA' het geval is;

- Bijkomend voordeel van de gebruikte methode is dat het onderscheidend vermogen van DEA toeneemt, zodat minder ziekenhuizen als relatief efficiënt worden betiteld (zie figuur 0), en het daarom voor een groter aantal ziekenhuizen mogelijk is een verbeterpotentieel te bepalen dan bij het gebruik van een standaard DEA-model. Hierdoor wordt de benchmark voor een aantal ziekenhuizen nog waardevoller;
- De techniek biedt verder het perspectief om te worden toegepast op de individuele (klinische) afdelingen van de deelnemende ziekenhuizen. Per afdeling kan er zo een specifiek op deze afdeling gerichte kwaliteitsindex worden geconstrueerd. Iedere afdeling krijgt zo een voor de kwaliteit van de zorg gecorrigeerde DEA-score, zodat het verbeterpotentieel beter geduid kan worden (gedisaggregeerd). Door de beste DEA-scores van iedere afdeling te combineren kan een theoretisch 'super-ziekenhuis' worden geconstrueerd, zodat voor íeder deelnemend ziekenhuis een verbeterpotentieel is te berekenen (mits niet alle afdelingen van één ziekenhuis de hoogste efficiëntiescores behalen).

Inhoudsopgave

1 Inleiding 1-7
2 Over 'kwaliteit' 2-9
2.1 Het servicemodel van Grönroos 2-9
2.2 Kwaliteit in de zorg en prestatie-indicatoren 2-10
3 Over DEA 3-13
3.1 Het principe 3-13
3.2 Modellen 3-14
4 Over kwaliteit en DEA 4-18
4.1 'Kwaliteit' als output in onaangepaste DEA 4-18
4.1.1 Voorbeelden uit de literatuur 4-18
4.1.2 Beperkingen van de voorbeeldstudies 4-19
4.2 Garanderen van gewicht 'kwaliteit' met gewichtsrestricties 4-22
4.2.1 Restricties op DEA-gewichten 4-23
4.2.2 Restricties op virtuele DEA-gewichten 4-27
4.3 'Kwaliteit' als toetsingscriterium van DEA-resultaten 4-31
4.3.1 Grafische methode 4-31
4.3.2 'Quality-adjusted' DEA 4-33
4.4 Overzicht methoden 4-33
5 Toepassing aangepaste DEA op de klinische afdelingen van de ziekenhuizen 5-36
5.1 Modelleren van kwaliteit op de klinische afdelingen 5-36
5.2 Modelleren van kwaliteit in DEA voor de benchmark Kliniek 5-39
5.3 Beschrijving van de data 5-40
5.4 Resultaten van de benchmark en analyse 5-41
5.4.1 Kwaliteitsindex 5-41
5.4.2 Efficiëntiescores 5-44
5.4.3 Berekening van het verbeterpotentieel 5-47
6 Conclusies en aanbevelingen 6-50
6.1 Conclusies 6-50
6.2 Aanbevelingen 6-51
7 Bibliografie 7-53
8 Lijst met afkortingen 8-55
9 Bijlage 9-56

1 Inleiding

In de afgelopen jaren zijn in Nederland de kosten van de gezondheidszorg sterk gestegen. Voor de toekomst wordt verwacht dat deze kosten alleen maar verder zullen oplopen als gevolg van de vergrijzende Nederlandse bevolking. Mede door deze kostenstijgingen is er politieke en maatschappelijke druk ontstaan om de prestaties van ziekenhuizen onderling beter te kunnen vergelijken. Het is niet alleen voor consumenten van belang om de prestaties van verschillende ziekenhuizen naast elkaar te kunnen leggen, maar zeker ook voor de ziekenhuizen zelf, omdat deze ziekenhuizen steeds meer wettelijk worden verplicht om prestaties openbaar te maken. Het doel van het openbaren van deze cijfers is om de kwaliteit van de zorg te verbeteren door de onderlinge concurrentie tussen verschillende ziekenhuizen te stimuleren.

Een achttal middelgrote Nederlandse ziekenhuizen, verspreid over het land, heeft daarom *ORTEC*, een bedrijf dat is gespecialiseerd in geavanceerde planning- en optimalisatiesoftware en consultancy, gevraagd om een benchmark uit te voeren waarmee de onderlinge prestaties van de verschillende ziekenhuizen kunnen worden vergeleken en waarmee kan worden aangegeven op welke punten deze ziekenhuizen zich nog kunnen verbeteren. De benchmark spitst zich toe op de efficiëntie van de personeelsinzet in relatie tot het aantal behandelingen, de 'productie', van de ziekenhuizen.

De benchmark heeft betrekking op de klinische afdelingen van de ziekenhuizen. Binnen de klinische afdelingen worden de verpleegafdelingen, de Intensive Care en de Cardiac Care onderscheiden. Dit onderdeel van de benchmark is ook in 2007 door *ORTEC* uitgevoerd binnen de acht genoemde ziekenhuizen.

Voor deze benchmark is gebruik gemaakt van de techniek Data Envelopment Analysis (DEA), een non-parametrische techniek waarmee de relatieve efficiëntie en het verbeterpotentieel van verschillende zogenaamde Decision Making Units (DMU's) (e.g. ziekenhuizen) kunnen worden bepaald. Een DMU is gedefinieerd als een entiteit die inputs om kan zetten in outputs. Met DEA kunnen de input- en outputvectoren van de verschillende DMU's met elkaar worden vergeleken. Als resultaat levert DEA (onder andere) voor iedere DMU een relatieve efficiëntiescore (ten opzichte van de andere DMU's) tussen 0 (inefficiënt) en 1 (efficiënt). Hiernaast is het mogelijk om naast deze efficiëntiescores ook het verbeterpotentieel voor de verschillende DMU's te berekenen. Een voordeel van DEA is dat iedere DMU op een voor deze DMU zo optimaal mogelijke manier wordt afgeschilderd, maar een belangrijke beperking van deze techniek is dat er voor voldoende onderscheid tussen de prestaties van de verschillende DMU's maar een beperkt aantal verschillende in- en outputs kan worden meegenomen.

ORTEC heeft, in overleg met de deelnemende ziekenhuizen, ervoor gekozen om voor de benchmark van 2007 als inputvariabele te kiezen voor de "salariskosten van het aantal fulltime-equivalents

(FTE's)" - een goede maat voor de omvang van de personeelsinzet - dat werkzaam is op de klinische afdelingen (inclusief het ondersteunend personeel). Voor de outputvariabelen voor DEA is ervoor gekozen om het "aantal verpleegdagen + dagbehandelingen" en het "aantal klinische opnames" als indicatoren voor de 'productie' van een ziekenhuis te nemen.

ORTEC heeft geconcludeerd dat de dimensie 'kwaliteit' nog onvoldoende is meegenomen in het DEA-onderdeel van de benchmark uit 2007. Dit komt doordat het, voor deze benchmark gebruikte, ('standaard') DEA-model een ziekenhuis als 'het meest efficiënt' beoordeelt, wanneer dit ziekenhuis ten opzichte van de andere ziekenhuizen relatief de laagste salariskosten (personeel) heeft in verhouding tot het aantal behandelingen dat dit ziekenhuis bewerkstelligt. Het model is dus slechts op kostenefficiëntie gefocust, zodat het verband met de kwaliteit van de geleverde zorg wordt genegeerd. Het is dus gewenst om een bredere definitie van het begrip 'efficiëntie' in DEA te integreren.

De uitdaging van deze scriptie ligt in het zo goed mogelijk, binnen de beperkingen die DEA stelt, meetbaar maken van de kwaliteit van de zorg op de klinische afdelingen van ziekenhuizen en het zinvol inpassen van deze outputfactor in een DEA-model, waarmee vergelijkbare efficiëntiescores en een realistisch verbeterpotentieel (op het gebied van personeelsinzet en 'productie') van de klinieken van de verschillende ziekenhuizen in kaart kunnen worden gebracht.

Om te kunnen bepalen hoe de kwaliteit van de zorg kan worden meegenomen in DEA, moet eerst worden bepaald wat er onder 'kwaliteit van dienstverlening' (in de zorg) kan worden verstaan, welke facetten hiervan wenselijk zijn mee te nemen in de analyse en op welke manier deze facetten meetbaar te maken zijn. Deze problematiek wordt behandeld in hoofdstuk 2, op basis van literatuuronderzoek. Vervolgens wordt in hoofdstuk 3 kort uiteengezet wat DEA inhoudt en hoe de methode precies in zijn werk gaat. In hoofdstuk 4 wordt op een rijtje gezet wat in de literatuur bekend is over het meenemen van kwaliteit, en in brede zin expertise, in DEA. Tot slot kan in hoofdstuk 5, op basis van de resultaten uit hoofdstuk 2 en 4, een methode worden gepresenteerd om zorgkwaliteit op bruikbare wijze meetbaar te maken en op een zinvolle manier te integreren in een DEA-model zodat in ditzelfde hoofdstuk betrouwbare resultaten van het DEA-onderdeel van de 'benchmark kliniek 2008' kunnen worden gepresenteerd en geanalyseerd.

2 Over 'kwaliteit'

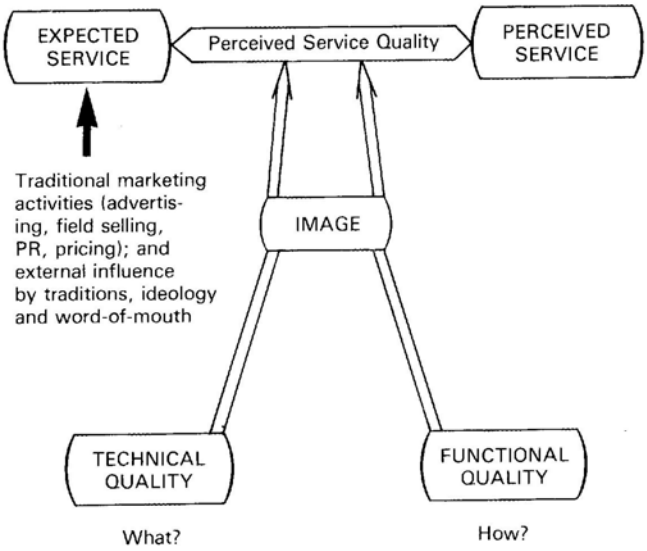
Omdat het begrip 'kwaliteit' centraal staat in deze scriptie, wat een zeer breed begrip is, is het belangrijk om eerst te onderzoeken wat dit begrip allemaal kan omvatten, zodat het voor alle belanghebbenden bij de benchmark helder is welke aspecten van zorgkwaliteit in welke verhoudingen zijn geïntegreerd in het DEA-gedeelte van de benchmark. Dit is belangrijk om te kunnen bepalen welke waarde kan worden toegedicht aan de resultaten van de benchmark (in hoofdstuk 5). Daarom wordt in dit hoofdstuk eerst, in paragraaf 2.1 het in de literatuur veel geciteerd servicemodel van Grönroos (1984) besproken (een ander populair model is dat van Lehtinen en Lehtinen (1991)), waarnaar later (in hoofdstuk 5) kan worden verwezen wanneer wordt vastgesteld welke aspecten van zorgkwaliteit in de benchmark zijn meegenomen. Hierna wordt in paragraaf 2.2 het verband tussen deze modellen en het meetbaar maken van prestaties (op het gebied van kwaliteit) in de zorg toegelicht.

2.1 Het servicemodel van Grönroos

Grönroos (1984) heeft een model bedacht om de kwaliteit van een dienst te kunnen analyseren, i.e. een model dat beschrijft hoe de kwaliteit van een dienst wordt waargenomen door de klanten (zoals patiënten in een ziekenhuis). Grönroos stelt dat de *perceived service quality*, de (achteraf) door de klanten waargenomen kwaliteit van een dienst, het resultaat is van een evaluatie die deze klant maakt, waarin hij de *perceived service*, de (uiteindelijk) ontvangen service, afweegt tegen de zogenaamde *expected service*, het serviceniveau dat de klant bij voorbaat verwachtte.

Grönroos (1984) stelt, naar analogie van Swan en Combs (1976), dat de *perceived service quality* kan worden opgedeeld in twee subprocessen: de *technical quality* en de *functional quality*. Onder *technical quality* (*wat* krijgt een klant) wordt het technische aspect van een product verstaan. Wanneer gekeken wordt naar diensten, dan is dit het resultaat van de productie van een dienst, zoals een medische behandeling in een ziekenhuis. De *functional quality* (*hoe* krijgt een klant de dienst) is daarentegen gerelateerd aan het meer sociaal/psychologische aspect van de servicebeleving van een klant en is gericht op de contacten met personeel en de activiteiten van de dienstverlenende instelling gedurende de duur van de dienst tot de *technical quality* is geleverd. Grönroos (1984) maakt hierbij de belangrijke opmerking dat de *technical quality* van een dienst vaak goed en objectief meetbaar is (wel vaak op langere termijn), terwijl de *functional quality*, de manier waarop deze dienst tot stand is gekomen vaak subjectiever is, en moeilijker te meten.

Grönroos (1984) stelt verder dat er soms nog een derde kwaliteitsdimensie meespeelt in de beleving van kwaliteit door de klanten, namelijk het *image*, het imago, van de instelling. Grönroos (1984) zegt dat het *image* van een instelling voornamelijk bepaald wordt door de diensten die het bedrijf verleent, omdat dit het gedeelte van de instelling is die klanten het meeste zien. Het *image* van de instelling wordt dus eveneens bepaald door de *technical quality* enerzijds, en de *functional quality* anderzijds. Omdat bij een beter imago van een instelling ook hogere verwachtingen van de klanten horen, beïnvloeden de *technical* en de *functional quality* dus zowel de *expected service* als de *perceived service*.



Figuur 1: Het servicemodel van Grönroos (Grönroos, 1984)

Het servicemodel van Grönroos (1984) is schematisch weergegeven in figuur 1. In het figuur staat een aantal factoren vermeld die van invloed kunnen zijn op de verwachtingen van klanten (*expected service*). De belangrijkste factoren zijn mond-op-mondreclame en marketing. Marketing is in het verleden geen kerntaak van de ziekenhuizen in Nederland geweest, maar doordat ziekenhuizen steeds meer wettelijk verplicht zijn om openheid van zaken te geven over hun prestaties, neemt de onderlinge concurrentie toe, en kan marketing op den duur mogelijk een belangrijkere rol gaan spelen.

2.2 Kwaliteit in de zorg en prestatie-indicatoren

Een veel gebruikte manier om de prestaties op zeer diverse gebieden in ziekenhuizen te kwantificeren, is het gebruik van zogenaamde prestatie-indicatoren. Deze prestatie-indicatoren kunnen worden gebruikt om moeilijk meetbare, vaak abstractere, zaken (zoals veiligheid of kwaliteit van dienstverlening/zorg) toch meetbaar te maken door een (grote) verzameling van indicatoren die een indicatie moeten geven over een deelgebiedprestatie van het abstractere geheel te gebruiken. De verzameling van indicatoren kan als geheel een betrouwbaar beeld schetsen van de prestatie van een instelling op een abstracter niveau (zoals 'kwaliteit').

Een dergelijke manier van meten wordt bijvoorbeeld gebruikt door de Inspectie voor de Gezondheidszorg, de Nederlandse Vereniging van Ziekenhuizen, de Nederlandse Federatie van Universitair Medische Centra en de Orde van Medisch Specialisten in de *Basisset Prestatie-indicatoren 2008* (Inspectie voor de Gezondheidszorg, 2008) en o.a. door de Organisatie van Zorgondernemers, Verpleegkundigen & Verzorgenden Nederland,de Landelijke Organisatie Cliëntenraden, de Beroepsvereniging van verpleeghuisartsen en sociaal geriaters, de Landelijke Beroepsvereniging Verzorging, de Inspectie voor de Gezondheidszorg, het Ministerie van Volksgezondheid, Welzijn en Sport en Zorgverzekeraars Nederland, in het *Kwaliteitskader Verantwoorde zorg 2007* (Thuiszorg, 2008).

Het is bij het gebruik van dergelijke prestatie-indicatoren belangrijk, voor de waarde die aan deze data kan worden gehecht (en de resulterende DEA-resultaten), dat de verschillen in de indicatorwaarden tussen de verschillende ziekenhuizen ook daadwerkelijk te wijten zijn aan verschillen in prestaties van deze ziekenhuizen.

Dit betekent dat de indicatoren goed gedefinieerd moeten worden. Wanneer ziekenhuizen bijvoorbeeld verschillende pijngrenzen hanteren bij het meten van het aantal patiënten met een onacceptabele pijnscore, dan zijn deze indicatorwaarden niet vergelijkbaar. En dus ook de resultaten niet.

Naast het belang van een zorgvuldige definitie van de prestatie-indicatoren, moet er voor worden gezorgd dat de zogenaamde 'case mix', de samenstelling van het cliëntenbestand, tussen de deelnemende ziekenhuizen niet verschillend is, of wordt gecorrigeerd wanneer dit wel het geval is. Wanneer de ziekenhuizen die worden vergeleken bijvoorbeeld sterk verschillen in de socio-economische kenmerken van hun patiëntenmix, moet hiervoor worden gecorrigeerd om de indicatorwaarden eerlijk met elkaar te kunnen vergelijken. Verschillende typen patiënten leggen namelijk op verschillende wijze beslag op de middelen in een ziekenhuis (oudere patiënten hebben bijvoorbeeld gemiddeld een langere opnameduur dan jongere patiënten). Wanneer de case mix van de verschillende ziekenhuizen die worden vergeleken wél verschillend is, moet hiervoor worden gecorrigeerd, wat in de literatuur 'risk adjustment' genoemd wordt (zie O'Neill (1998) en Morey (1992) voor toepassingen).

'Medicare', een Amerikaans sociaal verzekeringsprogramma, heeft bijvoorbeeld Diagnosis-related groups (Wikipedia, 2008a) ontwikkeld, waarmee patiënten die op verschillende manier gebruik maken van de middelen van een ziekenhuis in verschillende groepen worden onderverdeeld. Met behulp van de 'Case Mix Index' (Wikipedia, 2008b) kunnen de patiëntkosten nu worden gecorrigeerd voor het type patiënten dat gebruik maakt van een bepaald ziekenhuis, zodat een eerlijke vergelijking tussen de verschillende ziekenhuizen mogelijk is.

Behalve dat de definities van de indicatoren helder dienen te zijn en zo nodig case mix correcties worden toegepast, is het ook van belang dat de meetfrequenties van de verschillende ziekenhuizen vergelijkbaar zijn (en voldoende hoog). Wanneer een ziekenhuis A bijvoorbeeld maar één keer per jaar het aantal decubitusgevallen meet, en toevallig op een dag een laag aantal patiënten met doorligverschijnselen meet terwijl het gemiddelde aantal gevallen in dit ziekenhuis hoger ligt dan in een ziekenhuis B, waar zeer regelmatig wordt gemeten, is de kans groot dat de indicatorwaarden van ziekenhuis A betere prestaties lijken te geven dan deze waarden van ziekenhuis B, wat natuurlijk niet 'eerlijk' is. Bovendien kan een ziekenhuis dat vaker meet, wat een positieve kwaliteitsindicatie is, meer fouten ontdekken dan een ziekenhuis dat minder vaak meet, wat zo een negatieve kwalificatie oplevert.

Door nu de set van, goed vergelijkbare, prestatie-indicatoren precies zó te kiezen dat deze in de juiste aantallen overeenstemt met de facetten van servicekwaliteit (zoals deze geschetst zijn in paragraaf 2.1) die belangrijk worden gevonden om mee te nemen in de benchmark van de klinische afdelingen in de ziekenhuizen, ontstaat een helder afgebakend en meetbaar (en dus bruikbaar) kwaliteitsbegrip.

3 Over DEA

In deze paragraaf wordt een beknopte samenvatting van de techniek Data Envelopment Analysis (DEA) gegeven door eerst in paragraaf 3.1 de principes achter DEA te verhelderen om vervolgens in paragraaf 3.2 de belangrijkste DEA-modellen te benoemen en het voor de benchmark meest relevante model wat uitgebreider te bespreken. In paragraaf 3.3 wordt een techniek gegeven waarmee bij twijfel over de toepasbaarheid van DEA een vorm van gevoeligheidsanalyse uitsluitsel kan bieden.

3.1 Het principe

Data Envelopment Analysis (DEA) is een non-parametrische techniek (i.e. een techniek die niet afhangt van aannames over verdelingsfuncties van data) waarmee de relatieve efficiëntie en het verbeterpotentieel van een zogenaamde Decision Making Unit (DMU) kan worden bepaald in relatie tot de overige DMU's. Een DMU kan worden gedefinieerd als een entiteit die inputs om kan zetten in outputs, in een verhouding die niet nader hoeft te worden gespecificeerd (non-parametrisch) maar door het model wordt vastgesteld op een dusdanige manier, dat iedere DMU op een voor deze DMU zo gunstig mogelijke manier wordt meegnomen in deze benchmark methode DEA.

DEA vergelijkt als het ware de input- en outputvectoren van de verschillende DMU's met elkaar. Een DMU die met minder eenheden van zijn inputs méér eenheden outputs kan produceren dan andere DMU's dat kunnen, wordt in DEA als meer efficiënt dan deze laatste DMU's bestempeld. Voor geen enkele DMU hoeft een productiefunctie te worden gespecificeerd omdat DEA (als doel heeft) om de gewichten (de variabelen in het model) van de in- en outputs van de individuele DMU's zó te stellen dat deze DMU's allen in een voor hen zo gunstig mogelijke input-outputverhouding worden vergeleken met de overige DMU's. Een belangrijke eigenschap van veel DEA-modellen (bijvoorbeeld de modellen die in paragraaf 3.2 worden besproken) is, dat de modellen zogenaamd *unit invariant* zijn (Cooper et al., 2000), wat inhoudt dat de eenheid waarin de in- en outputs zijn gemeten, niet van invloed is op de uiteindelijke efficiëntiescore en targets (verbeterpotentieel).

Als resultaat levert DEA voor iedere DMU een relatieve efficiëntiescore (relatief ten opzichte van de andere DMU's) tussen 0 (inefficiënt) en 1 (efficiënt). Met deze resultaten, in combinatie met de gevonden gewichten voor de in- en outputs voor iedere DMU, kan ook het verbeterpotentieel voor de verschillende DMU's berekend worden. Dit verbeterpotentieel is gebaseerd op de door de andere

DMU's, de zogenaamde *peer* DMU's, gerealiseerde productieverhoudingen (tussen in- en outputs) waardoor dit verbeterpotentieel zeer realistisch is wanneer de DMU's onder vergelijkbare omstandigheden kunnen opereren. Een DMU is een *peer* DMU voor een bepaalde DMUj_0, wanneer deze DMU een efficiëntiescore van 1 behaalt onder de input- en outputgewichten die optimaal zijn voor DMU j_0.

Een DEA-model kan een input- of een outputoriëntatie hebben. De gekozen oriëntatie is, voor de standaard DEA-modellen die in paragraaf 3.2 worden besproken, van invloed op de targets (het verbeterpotentieel) van de DMU's maar niet op de relatieve efficiëntiescores. Een inputgeoriënteerd model tracht in het bepalen van de targets inefficiënties uit de weg te ruimen door met mindere hoeveelheden inputs vergelijkbare outputniveaus te bereiken, terwijl een outputgeoriënteerd model in de targets hogere outputhoeveelheden probeert te bewerkstelligen met een vergelijkbaar niveau van inputs.

Een belangrijk nadeel van de methode is dat de som van het aantal in- en outputvariabelen relatief klein moet zijn ten opzichte van het aantal DMU's dat in het model wordt meegenomen (een verhouding van 1:3 tot 1:4 is raadzaam) omdat het aantal als efficiënt beoordeelde DMU's anders 'te groot' wordt (mogelijk krijgen zelfs álle DMU's een relatieve efficiëntiescore van 1), waardoor het onderscheidingsvermogen van de techniek zeer klein - tot zelfs nul - wordt. Wanneer een complex proces (zoals een kwaliteitstraject in de zorg) onderdeel uit maakt van de 'productie' van een DMU is het wenselijk zo veel mogelijk indicatoren mee te nemen in het model, terwijl het mogelijk kleine aantal DMU's een grens stelt aan de procesindicatoren die kunnen worden meegenomen als in- en outputvariabelen (bijvoorbeeld de kwaliteitsindicatoren uit paragraaf 2.2). Wanneer er relatief weinig DMU's in een onderzoek worden meegenomen, dient het aantal in- en outputvariabelen dusdanig klein zijn dat dit afbreuk doet aan de kwaliteit van de DEA-resultaten.

3.2 Modellen

Het eerste DEA-model is bedacht door Charnes, Cooper en Rhodes, het zogenaamde CCR-model (Charnes et al., 1978). Dit model gaat uit van constante schaalopbrengsten (een verdubbeling van inputs leidt tot een verdubbeling van outputs). Een uitbereiding op dit model is het model van Banker, Charnes en Cooper, het BBC-model (Banker et al., 1984). Dit model gaat uit van variabele schaalopbrengsten. Naast deze twee meest voorkomende DEA-modellen bestaat er nog een aantal andere varianten, zie hiervoor bijvoorbeeld Cook en Seiford (2008).

In overleg met *ORTEC* is besloten om – net als in de benchmark van 2007 – uit te gaan van constante schaalopbrengsten in de ziekenhuizen. Deze aanname is gemaakt omdat de ziekenhuizen qua omvang (aantal behandelingen) ongeveer als 'even groot' kunnen worden beschouwd. Om deze reden zijn alle modellen in deze scriptie gebaseerd op de aanname van constante schaalopbrengsten, en wordt alleen het CCR-model besproken. Bijkomend voordeel hierbij is dat er met het CCR-model altijd minder DMU's als efficiënt zullen worden betiteld dan in het BBC-model. Dit is voor deze benchmark een prettige bijkomstigheid omdat het aantal DMU's relatief klein is (8 ziekenhuizen) en het aantal variabelen relatief hoog (4 stuks, zoals in hoofdstuk 5 wordt besproken) waardoor het aantal ziekenhuizen dat als volledig efficiënt zal worden betiteld relatief hoog zal liggen, wat de resultaten minder aansprekend maakt.

CCR-I-input

$$e_{j_0} = h_{j_0} = \max_{u,v} \frac{\sum_{r=1}^{s} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}}$$

onder de voorwaarden:

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1 \; ; \quad j = 1,\ldots,n$$

$$u_r, v_i \geq 0 \quad r = 1,\ldots,s \quad i = 1,\ldots,m$$

Tabel 0: Overzicht en verklaring van modelnotatie

Indices:		Variabelen:	
j	$j = 1..n$	u_r	outputgewicht van input i
i	$i = 1..m$	v_i	inputgewicht van output r
r	$i = 1..s$		
Data:		Afgeleide variabelen:	
y_{rj}	output-waarde i van DMU$_j$	e_j, h_j, z_j	relatieve efficiëntie DMU$_j$ (gewogen input-outputverhouding van DMU$_j$)
x_{ij}	input-waarde r van DMU$_j$		

Model CCR-I-input (hier gegeven in het geval van een inputoriëntatie, zie Bijlage A voor het outputgeoriënteerde model) geeft de meest intuïtieve vorm van het, op de aanname van constante schaalopbrengsten gebaseerde, CCR-model weer (Charnes et al., 1978). In dit model zijn er n DMU's (j), met elk m inputs (i), en s outputs (r). Het model stelt de inputgewichten, v_i, van de inputvariabelen, x_{ij}, en de outputgewichten, u_r, van de outputvariabelen, y_{rj}, zodanig vast dat deze de gewogen input-outputverhouding van deze DMU j_0, e_{j0},optimaliseert onder de voorwaarde dat geen enkele DMU j met deze gewichten v_i en u_r een gewogen input-outputverhouding realiseert die groter is dan 1. De geoptimaliseerde input-outputverhouding van DMU j_0, e_{j0}, ligt hierdoor tussen 0 en 1, en is te interpreteren als de relatieve efficiëntiescore van DMU j_0. Deze efficiëntiescore, e_{j0}, is een zogenaamde *radiale* efficiëntiescore omdat deze aangeeft welke percentuele inkrimping van alle inputs nodig is als noodzakelijke voorwaarde voor efficiënte productie. De variabelen z_{j0} en h_{j0} worden hier geïntroduceerd omdat deze in Bijlage A gebruikt worden om het outputgeoriënteerde CCR-model af te leiden.

Om het DEA-model gemakkelijk op te kunnen lossen moet het worden getransformeerd naar een lineair programmeringmodel (LP-model). Door gebruik te maken van *lineair fractional programming* (Charnes et al. 1978) ontstaat het gemakkelijk oplosbare CCR-II-input model (equivalent aan het CCR-I model), en het bijbehorende duale model (zie model CCR-II-D-input).

In het CCR-II-D-input model krijgt iedere DMU *j* een gewicht λ_j wat het effect heeft dat er een virtuele DMU ontstaat die een convexe combinatie is van alle DMU's op zodanige wijze dat deze combinatie efficiënt is (Bowling, 1998). De outputslackvariabele s_r^+ is positief wanneer de output van DMU j_0 kleiner is dan de output van de virtuele DMU, wat betekent dat de output van DMU j_0 kan

CCR-II-input

$$e_{j_0} = h_{j_0} = \max_{u,v} \sum_{r=1}^{s} u_r y_{rj_0}$$

onder de voorwaarden

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \quad j = 1,\ldots,n$$

$$\sum_{i=1}^{m} v_i x_{ij_0} = 1$$

$$u_r, v_i \geq 0 \qquad r = 1,\ldots,s \qquad i = 1,\ldots,m$$

CCR-II-D-input

$$e_{j_0} = z_{j_0} = \min_{\lambda_j, s_i^-, s_r^+, \theta} \theta$$

onder de voorwaarden:

$$\theta x_{ij_0} - \sum_{j=1}^{n} \lambda_j x_{ij} - s_i^- = 0 \qquad i = 1,\ldots,m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{rj_0} \qquad r = 1,\ldots,s$$

$$\lambda_j, s_i^-, s_r^+ \geq 0 \quad \theta \in \mathbb{R} \quad i = 1,\ldots,m \quad r = 1,\ldots s \quad j = 1,\ldots,n$$

worden verhoogd onder de gekozen gewichten λ_j terwijl de inputs constant blijven en aan alle voorwaarden voldaan blijft. DMU j_0 is dus inefficiënt wanneer s_r^+ groter is dan 0 (Charnes et al., 1978). Een verhaal van gelijke strekking geldt voor inputslackvariabele s_i^-, die aangeeft in welke mate de inputniveaus kunnen worden verlaagd zonder hiermee het outputniveau te verlagen voor gegeven λ_j. DMU j_0 is dus inefficiënt wanneer s_i^- groter is dan 0.

In een eindig optimum (e_{j0}^*, h_{j0}^*, u_r^*, v_i^*, z_{j0}^*, θ^*, λ_j^*, s_i^{-*}, s_r^{+*}) geldt nu volgens de sterke dualiteitstelling:

$$e_{j_0}^* = h_{j_o}^* = \sum_{r=1}^{s} u_r^* y_{rj_0} = z_{j_o}^* = \theta^*$$

Een DMU moet dus alleen als volledig efficiënt worden aangemerkt wanneer θ^* gelijk is aan 1 en de slackvariabelen gelijk zijn aan 0. Oftewel: een DMU is slechts dan volledig efficiënt wanneer deze DMU een relatieve efficiëntiescore krijgt van 1, en de slackvariabelen gelijk zijn aan 0.

Dit is ook gelijk te zien aan de targets, $\left(\sum_{j=1}^{n} \lambda_j^* x_{ij}, \sum_{j=1}^{n} \lambda_j^* y_{rj} \right) \ \forall j$ (Charnes et al., 1978), die volgen uit de voorwaarden in het duale model:

$$h_{j_0}^* x_{ij_0} - s_i^{-*} = \sum_{j=1}^{n} \lambda_j^* x_{ij} \qquad i = 1, \ldots, m$$

$$y_{rj_0} + s_r^{+*} = \sum_{j=1}^{n} \lambda_j^* y_{rj} \qquad r = 1, \ldots, s$$

Wanneer de slackvariabelen, s_i^{-*}, s_r^{+*}, gelijk zijn aan 0 en de efficiëntiescore, e_{j0}^*, aan 1, dan zijn de targets gelijk aan de uitgangswaarden x_{ij}, en y_{rj}.

4 Over kwaliteit en DEA

In dit hoofdstuk worden de vier belangrijkste bevindingen besproken uit een uitgebreid literatuuronderzoek naar de integratie van een kwaliteitsdimensie in DEA. De eerste methode wordt besproken in paragraaf 4.1. Deze methode behelst het simpelweg opnemen van een kwaliteitsvariabele als outputvariabele in een standaard DEA model (CCR of BCC). In paragraaf 4.2 worden methodes besproken die door het stellen van restricties op de (virtuele) DEA-gewichten het belang van de kwaliteitsvariabele(n) moeten garanderen. In paragraaf 4.3 worden geheel andere methodes besproken die de kwaliteitsfactor niet integreren in het DEA-model, maar nadat DEA is toegepast op de data kunnen worden gebruikt om de DEA-resultaten op de juiste waarde te schatten. In paragraaf 4.4 wordt in tabelvorm een overzicht gegeven van alle positieve en negatieve eigenschappen van de besproken technieken.

4.1 'Kwaliteit' als output in onaangepaste DEA

In deze paragraaf wordt bekeken wat de mogelijkheden zijn om een kwaliteitsvariabele als outputvariabele in een 'standaard' (CCR of BBC) DEA-model op te nemen. In paragraaf 4.1.1 worden de conclusies van een tweetal studies besproken waarin 'kwaliteit' op deze manier is meegenomen in DEA. In paragraaf 4.1.2 wordt vervolgens de belangrijkste beperking van deze studies besproken.

4.1.1 Voorbeelden uit de literatuur

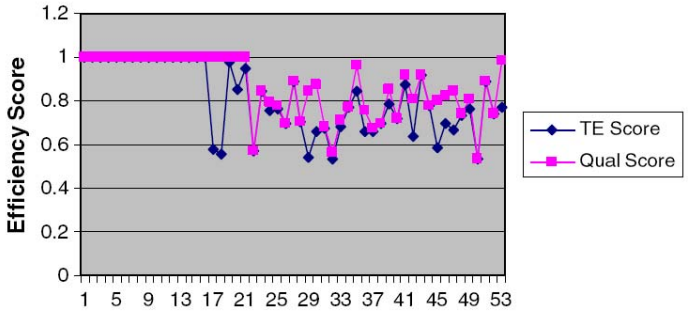
Er zijn verschillende studies uitgevoerd (waaronder in de zorgsector) waarin door het toevoegen van kwaliteitsindicatoren als (output-)variabelen in DEA getracht is om de kwaliteitsdimensie van het productieproces van een DMU te betrekken in de analyse.

In een artikel van Sahin en Ozcan (2000) wordt de reciproque waarde van het sterftecijfer binnen ziekenhuizen (=1/'sterftecijfer') als kwaliteitsindicator in DEA opgenomen. De studie heeft als voornaamste doel om te bepalen hoe efficiënt er gewerkt wordt in ziekenhuizen die eigendom zijn van het

Turkse Ministerie voor Volksgezondheid. Van de 80 onderzochte ziekenhuizen bleek volgens Sahin en Ozcan (2000) 55% inefficiënt te opereren.

Nayar en Ozcan (2008) gebruiken een techniek die gepresenteerd is door Sherman en Zhu (2006a). Volgens deze techniek moeten er twee DEA-modellen worden geschat: een model met daarin alleen technische (niet direct aan kwaliteit gerelateerde) in- en outputvariabelen en hiernaast een DEA-model met kwaliteitoutputs als additionele variabelen. Door de efficiëntiescores van de verschillende instellingen (DMU's) onder beide modellen te vergelijken kan worden bepaald of er een wisselwerking (*trade-off*) is tussen kwaliteit en kostenefficiëntie binnen de DMU's.

In Nayar en Ozcan (2008) wordt geconcludeerd dat ziekenhuizen zowel kwaliteit (van de zorg) als kwantiteit (van het aantal behandelingen) kunnen maximaliseren, zodat kostenefficiëntie niet ten koste gaat van kwaliteit (van de zorg). Dat betekent dat het opnemen van kwaliteit-outputvariabelen in DEA niet hoeft te leiden tot lage gewichten voor de kwaliteitsoutputs. Dit resultaat is te zien in figuur 2, waarin de *TE score*, de technische (kosten-)efficiëntiescore, en de *Qual Score*, de score inclusief de kwaliteitsindicatoren, vrij sterk lijken samen te hangen. Deming (1986) ondersteunt (en verklaart) de conclusie van Nayar en Ozcan (2008) en stelt dat kwaliteit en kosten negatief gecorreleerd zijn (hoe hoger de kwaliteit van zorg, hoe lager de kosten) doordat het afleveren van goede kwaliteit zorgt voor een kleiner aantal cliënten van een DMU die met klachten terugkomen. Bovendien, stelt Deming (1986) dat DMU's die goede kwaliteit leveren minder hoge verzekeringskosten hebben, minder afval overhouden en profiteren van een hogere klantenloyaliteit, zodat hogere kwaliteit van zorg ook via deze wegen kan leiden tot lagere kosten. Uitgaande van de conclusies van Nayar en Ozcan (2008) en Deming (1986) is het simpelweg opnemen van kwaliteitsoutputs in DEA voldoende om DMU's (ziekenhuizen) op prestaties te kunnen vergelijken met DEA.



Figuur 2: TE en Qual-scores (Nayar en Ozcan, 2008)

4.1.2 Beperkingen van de voorbeeldstudies

Hoewel Nayar en Ozcan (2008) en Deming (1986) stellen dat hoge kwaliteit van zorg en hoge kostenefficiëntie samen gaan, kan het tegendeel ook worden beredeneerd: door bijvoorbeeld meer persoonlijke aandacht aan een patiënt te besteden zal de kwaliteit van de zorg toenemen, net als de personeelskosten (hogere personeelsinzet). Hoewel deze extra aandacht de ziekenhuisopname van de

patiënt aangenamer kan maken (en deze dus een hogere kwaliteit van zorg ervaart) hoeft dit niet per definitie te leiden tot medische complicaties of heropnames. Kwalitatief goede zorg kan dus samengaan met een hogere personeelsinzet (en dus meer kosten) dan kwalitatief slechtere zorg zonder dat dit de technische efficiëntie van de zorg in een ziekenhuis beïnvloedt.

Wanneer deze trade-off tussen kwaliteit en kostenefficiëntie bestaat, is er het risico dat DEA een te laag gewicht (mogelijk zelfs 0) toekent aan de 'kwaliteitsvariabelen' wanneer deze simpelweg als outputvariabelen in een standaard (CCR of BCC) DEA-model worden meegenomen. Dit komt doordat het leveren van kwalitatief goede zorg (menselijke) capaciteit kan kosten. Het bestaan van een trade-off tussen kwaliteit en kostenefficiëntie wordt ook in de literatuur verdedigd. Zie bijvoorbeeld Sherman en Zhu, 2006a.

Wanneer er inderdaad een trade-off bestaat tussen kwaliteit en kostenefficiëntie van zorg, dan zal DEA de gewichten van de traditionele, niet direct aan kwaliteit gerelateerde, outputvariabele(n) positief kiezen, terwijl de kwaliteitsoutput(s) een zeer laag (of zelfs gewicht 0) gewicht krijgen toebedeeld. Dit komt doordat de kwaliteit van de productie ten koste gaat van de kostenefficiëntie van deze productie, zodat een DMU (ziekenhuis) met lage gewichten voor de kwaliteitsgerelateerde variabele(n) in enge zin het 'productiefst' kan worden gerepresenteerd (wat het DEA-model nastreeft).

Om deze reden zullen, bij het toevoegen van (kwaliteits-) variabelen aan een DEA-model, de (technische) efficiëntiescores van de DMU's (ziekenhuizen) nooit dalen (de kwaliteitsoutputs krijgen gewicht 0, maar mogelijk stijgen ('kwaliteit' kost dan geen productiecapaciteit). Hierom kan het zo zijn (in het eerste geval) dat de kwaliteitsfactor onvoldoende wordt meegenomen (of feitelijk genegeerd wordt) in de uiteindelijke (brede) efficiëntiebeoordeling van de DMU's (klinische afdelingen van de ziekenhuizen).

Dit effect is duidelijk te zien in figuur 2 in paragraaf 4.1.1., behorende bij het onderzoek van Nayar en Ozcan (2008): de *Qual scores* liggen altijd boven de *TE scores* (geen sterke kwaliteit/kostenefficiëntie trade-off) of 'op' de *TE scores* (de kwaliteitsdimensie wordt feitelijk genegeerd en er kan dus wel degelijk sprake zijn van een trade-off tussen kwaliteit en kostenefficiëntie - wat in tegenspraak is met de conclusie van Nayar en Ozcan).

In het onderzoek van Sahin en Ozcan (2000) worden geen gewichten gepresenteerd. Het is dus niet te beoordelen of de kwaliteitsoutput (het reciproque sterftecijfer) wel voldoende gewicht krijgt in de DEA-analyse. Wanneer dit gewicht 'laag' is, zijn de resultaten dus niet waardevol te interpreteren binnen het veel relevantere 'bredere' kwaliteitskader.

Hierbij moet nog wel worden opgemerkt dat de (LP-) oplossingen van een DEA model vaak niet uniek zijn. De oplossing is bijvoorbeeld afhankelijk van de oplossingsmethode (verschilt per softwarepakket). In de literatuur zijn dan ook methoden ontwikkeld om er voor te zorgen dat alle in- en

output variabelen in de DEA-resultaten zo veel mogelijk gewichten toebedeeld krijgen die 'ver' verwijderd liggen van 0 door op een bepaalde manier op zoek te gaan naar alternatieve optima (met gelijke efficiëntiescores).

Cooper et al. (2007) presenteren een tweestapsmethode (een zogenaamde *facet model*) om zo mogelijk een alternatieve optimale oplossing te vinden die ligt op een *full dimensional efficiënt facet* (FDEF), een gedeelte van de efficiëntiegrens zonder gedegenereerde gewichten (gewichten met waarde 0), zodat de zo gegenereerde gewichten het beste worden ondersteund door de data. Cook et al. (2008) geven een overzicht van een aantal andere *facet models* en van *Contrained Facet Analysis* (CFA). Li en Reeves (1997) presenteren een zogenaamd *multiple objective lineair programming* (MOLP) model waarin verschillende efficiëntiescores gelijktijdig worden geoptimaliseerd, waaronder de gewone DEA efficiëntiescore. Li en Reeves (1997) beschrijven hoe deze methode alternatieve optimale DEA-efficiëntiescores identificeert waarin de gewichten gelijkmatiger zijn verdeeld dan wanneer gewone (enkelvoudige) DEA wordt toegepast.

Het is echter niet heel waarschijnlijk dat een kwaliteitsoutput die in een DEA-model een gewicht van 0 heeft gekregen in een alternatief optimum wél een positief gewicht zal krijgen, omdat er in dit geval een duidelijke trade-off bestaat tussen kwaliteit en kostenefficiëntie. Bovendien is het in een herhaald onderzoek - zoals de benchmark van de klinische afdelingen die centraal staat in deze scriptie - niet wenselijk dat a priori (vóór het toepassen van DEA) niet kan worden bepaald of DEA wel zinvol toepasbaar is op de data. Dit betekent dat de methode van Li en Reeves (1997) dan ook niet altijd voor een oplossing kan zorgen voor (te) lage gewichten voor kwaliteitsvariabelen. Dit maakt de 'standaard DEA' met toegevoegde kwaliteitsoutputvariabelen dan ook onbetrouwbaar.

Wanneer men dus niet alleen geïnteresseerd is in kostenefficiëntie (wat voor de hand ligt in de zorgsector) maar in de gehele prestatie (brede efficiëntie) van een DMU (klinische afdeling van een ziekenhuis) dan kan dus worden geconcludeerd dat de efficiëntiescores van de verschillende ziekenhuizen in een standaard DEA-model waarschijnlijk niet goed te vergelijken zijn (met of zonder kwaliteitsvariabelen in het model).

Dat 'brede efficiëntie' een veel zinvoller criterium is om de prestaties van (afdelingen van) ziekenhuizen te vergelijken dan het criterium van kostenefficiëntie, kan worden beredeneerd door te stellen dat een ziekenhuis een heel hoge kostenefficiëntiescore kan verkrijgen doordat het met zeer weinig personeel zeer veel patiënten kan bedienen, maar hierbij mogelijk een belabberde kwaliteit van zorg aflevert. Om nu te stellen dat dit ziekenhuis beter presteert dan een iets minder kostenefficiënt opererend ziekenhuis dat wel een goede zorgkwaliteit levert is niet redelijk.

Omdat DEA het verbeterpotentieel van de verschillende ziekenhuizen berekent aan de hand van de prestaties van kostenefficiënte (*peer*-) ziekenhuizen, met mogelijk een geheel ander niveau van zorgkwaliteit dan de minder kostenefficiënte ziekenhuizen, is dit verbeterpotentieel (slechts gebaseerd op kostenefficiëntie) mogelijk niet zinvol. Zo kan het lijken dat het personeelsbestand van een ziekenhuis kan worden ingekrompen omdat een ander ziekenhuis dezelfde 'productie' haalt met een lagere personeelsinzet maar hierbij wel een lagere kwaliteit van zorg levert. Deze lagere zorgkwaliteit zou dan ook als target gesteld worden voor het minder kostenefficiënte, maar kwalitatief goede, ziekenhuis.

Samengevat kan dus worden gesteld dat het simpelweg toevoegen van een kwaliteitsgerelateerde output in DEA geen garantie geeft voor goede 'brede efficiëntiescores', waarin mogelijke trade-offs tussen kwaliteit en kostenefficiëntie worden ondervangen, of voor realistische targets (verbeterpotentieel).

4.2 Garanderen van gewicht 'kwaliteit' met gewichtsrestricties

De resultaten van het onderzoek van Nayar en Ozcan (2008) (zie paragraaf 4.1) tonen aan dat het in sommige gevallen mogelijk is dat een kwaliteitsoutput in DEA 'voldoende' gewicht krijgt toegewezen door het model. In andere gevallen is dit echter niet het geval, of in ieder geval een onzekere (bij herhaling niet continue) factor, wat kan worden verklaard door het bestaan van een trade-off tussen kwaliteit en kostenefficiëntie, waarvan het bestaan ook in een aantal gevallen in de literatuur wordt verdedigd (zie bijvoorbeeld Sherman en Zhu, 2006a).

Door het stellen van verschillende typen gewichtsrestricties kunnen beperkingen worden gesteld aan de gewichten die DEA toekent aan de verschillende in- en outputs. Het vaststellen van deze gewichtsgrenzen moet gebeuren op basis van de mening van experts, eventueel in combinatie met kosteninformatie.

Het opleggen van gewichtsrestricties leidt tot efficiëntiescores waarvoor geldt dat de efficiëntiescores van het ongerestricteerde DEA-model een bovengrens vormen. Dit komt doordat de uitkomstruimte van de gewichten wordt beperkt. Wanneer een bepaalde, niet in enge zin productiegerelateerde output, zoals een kwaliteitsoutput, ten koste gaat van de productiecapaciteit van een DMU (ziekenhuis), dan zal DEA een gewicht 0 toekennen aan deze output. Wanneer nu een bepaald minimum gewicht voor deze output wordt afgedwongen door het stellen van restricties aan het model zullen de efficiëntiescores van deze DMU dus dalen. Dit is een intuïtief gemakkelijk te begrijpen resultaat omdat een DMU onder het verwaarlozen van een niet-productie-gerelateerde, capaciteitsbenuttende,

output een hogere productie in enge zin kan behalen en dus efficiënter lijkt te opereren. In feite wordt in dit geval alleen efficiëntie in enge zin beoordeeld, en niet het bredere kwaliteitsbegrip.

In deze paragraaf wordt onderzocht welke typen gewichtsrestricties DEA kan bevatten en of, en zo ja hoe, deze restricties kunnen worden gebruikt om te garanderen dat 'kwaliteit' voldoende zwaar meeweegt in DEA zodat de berekende efficiëntiescores betrouwbaarder zijn (i.e. beter aansluiten bij het doel van de benchmark) dan bij het toepassen van DEA in standaardvorm (zonder restricties) en de targets (verbeterpotentieel) realistischer.

In paragraaf 4.2.1 worden verschillende typen restricties op de DEA-gewichten besproken (u_r, en v_i in de notatie van het CCR-model in paragraaf 3.2). Hierna worden in paragraaf 4.2.2 restricties op de zogenaamde *virtuele* gewichten besproken ($x_{ij*}\ u_r$, en $y_{ij*}\ v_i$ in de notatie van het CCR-model in paragraaf 3.2).

4.2.1 Restricties op DEA-gewichten

Een eerste manier om het aandeel van de kwaliteitsoutputs in DEA te garanderen is het toepassen van restricties op de gewichten die in het DEA-model worden toegedicht aan de verschillende in- en outputvariabelen per DMU. In de notatie van het CCR-I-inputmodel in paragraaf 3.2 (op pagina 3-12) zijn dit restricties op de gewichten u_r, en v_i.

Er zijn twee verschillende manieren om deze restricties op deze gewichten toe te passen: door middel van *absolute gewichtsrestricties* en/of door *assurance regions* (Allen et al., 1997). Deze twee (basis-)typen restricties worden beurtelings besproken.

4.2.1.1 Absolute gewichtsrestricties

Absolute gewichtsrestricties zijn directe beperkingen op de waarden die de in- en output gewichten u_r, en v_i aan kunnen aannemen in het DEA-model. Een inputgewichtsrestrictie krijgt dan de volgende vorm:

$$\alpha_i \leq v_i \leq \beta_i \qquad (4.1)$$

Waarin de waarden voor de constanten α_i en β_i door de gebruiker moeten worden gespecificeerd. En voor een outputgewichtsrestrictie:

$$\gamma_r \leq u_r \leq \delta_r \qquad (4.2)$$

De belangrijkste reden om dergelijke restricties op te nemen is om te voorkomen dat bepaalde in- of outputs genegeerd worden (door lage gewichten aan deze variabelen mee te geven) of om bepaalde in- of outputvariabelen extra te benadrukken (Allen et al., 1997).

Deze vorm van dit type restrictie is niet heel gemakkelijk in het gebruik. De waarde van de gewichtsgrenzen is bijvoorbeeld sterk afhankelijk van de gekozen eenheden van de in- en outputvariabelen van het model, wat het lastig maakt om deze gewichten aan de hand van de gewenste criteria vast te stellen (Allen et al., 1997). Verder stellen Allen et al. (1997) dat de restricties afhankelijk zijn van de keuze van de zogenaamde *normalisatieconstante*, die de maximale theoretische efficiëntie van DMU aangeeft (meestal is deze waarde op 1 gesteld). Bovendien stellen zij dat er een sterke afhankelijkheid tussen de grenzen van de verschillende gewichten bestaat, wat het vaststellen van gewichtsrestricties die het gewenste effect sorteren erg lastig maakt. Hiernaast geldt dat de efficiëntiescores afhankelijk zijn van de oriëntatie van het model (input- of outputgeoriënteerd) wanneer er absolute gewichtsrestricties worden opgenomen in het DEA-model. Bovendien kan dit type restrictie ertoe leiden dat het model infeasible wordt (Allen et al., 1997).

Pas na een modelrun zonder restricties is het duidelijk welke grenzen er kunnen worden aangehouden in relatie tot de gekozen eenheden (Cook en Seiford, 2008). Om een helderder beeld te krijgen van de verschillende gewichten, en zo gemakkelijker de gewichtsrestricties te kunnen bepalen, kan in het standaard (CCR- of BBC-) DEA-model gebruik worden gemaakt van de *unit invariantie* eigenschap van deze modellen. Deze eigenschap houdt in dat lineaire transformaties van de variabelen de efficiëntiescores niet beïnvloeden. Door alle variabelen nu zodanig te normeren dat elke variabele hetzelfde gemiddelde heeft kunnen de gewichtsrestricties van de verschillende in- en outputs gemakkelijker worden vergeleken (Roll en Golany, 1993). Roll en Golany (1993) presenteren een twee-staps-methode en een methode die de gewichtenmatrix na een ongerestricteerde modelrun gebruikt om de gewenste gewichtsrestricties vast te stellen. Bij het analyseren van de resultaten moeten de variabelen dan wel weer terug getransformeerd worden.

4.2.1.2 Assurance regions

Thompson et al. (1990) onderscheiden twee typen assurance regions: assurance regions van type I (AR I) en van type II (AR II). Beide type AR restricties kunnen grafisch worden weergeven in gebieden (*regions*) van toegestane gewichten. Wanneer de gewichten binnen deze gebieden liggen, dan worden de relaties tussen deze gewichten uit het model als het ware gegarandeerd (assured).

AR I restricties leggen relaties tussen ofwel inputgewichten ofwel tussen outputgewichten, maar niet tussen input- en outputgewichten. Met AR II restricties kunnen gewichtsrelaties tussen input- en outputgewichten worden gemodelleerd.

Een voorbeeld van een AR I restrictie die een verband legt tussen de gewichten van verschillende inputgewichten (relatieve ordening) is bijvoorbeeld (Allen et al., 1997):

$$k_i v_i + k_{i+1} v_{i+1} \leq v_{i+2} \qquad (4.3)$$

Waarbij v_i, v_{i+1}, v_{i+2} drie door de gebruiker gespecificeerde inputgewichtsvariabelen zijn, en k_i en k_{i+1} door de gebruiker gespecificeerde constanten. Een ander voorbeeld (Allen et al., 1997) is onderstaande AR I restrictie, die restricties legt op de ratio van twee door de gebruiker gespecificeerde inputgewichtsvariabelen, v_i, en v_{i+1} (relatieve grootte):

$$\eta \leq \frac{v_i}{v_{i+1}} \leq \iota \qquad (4.4)$$

De constanten η en ι kunnen weer door de modelbouwer worden gekozen. In plaats van restricties op de inputgewichten, v_i, kunnen op vergelijkbare wijze restricties worden gesteld aan de ruimte van de outputgewichten, u_i. Roll en Golany (1993) presenteren een methode om de gewichtenvariatie te beperken wanneer weinig informatie beschikbaar is over de relatieve importantie van de verschillende variabelen.

Charnes et al. (1990) hebben een methode ontwikkeld, de zogenaamde *ConeRatio* methode, die door het transformeren van de in- en outputdata gewichtsrestricties 'simuleert' in een standaard DEA-model. Het voordeel van deze manier van modelleren is dat de modellen gewoon zijn op te lossen met een standaard DEA-solver, en dat nulwaardige of negatieve data geen problemen opleveren in het model. Het nadeel van de methode is dat de data moet worden getransformeerd en vervolgens weer terug getransformeerd om de resultaten te kunnen interpreteren. Dit kan soms omslachtiger zijn dan het toepassen van directe gewichtsrestricties in de modelformulering, wanneer de software dit toelaat. Het *Cone Ratio* model levert gelijke efficiëntiescores als onder de AR I restricties van type (4.4) (Allen et al., 1997).

Estellita Lins et al. (2007) stellen dat een model met AR I restricties gemakkelijk infeasible kan worden en presenteren daarom condities waaronder een model met gewichtsratiovoorwaarden (zie 4.4) feasible is. Hiernaast geven Estellita Lins et al (2007) een methode om, wanneer het model infeasible is, een model te specificeren dat wel feasible is en waarbij zo minimaal mogelijk aanpassingen aan de

gewenste gewichten hoeven te worden gedaan. Thompson et al. (1990) bewijzen dat er met AR I restricties altijd minimaal één efficiënte DMU is. Verder is het belangrijk dat de DEA-resultaten, bij het gebruik van AR I restricties in het model, net als bij het toepassen van absolute gewichtsrestricties, gevoelig zijn voor de eenheden die gekozen zijn voor de verschillende in- en outputvariabelen. Een input- of outputoriëntatie is echter níet van invloed op de efficiëntiescore (Allen et al., 1997).

Naast de AR I restricties worden de AR II restricties onderscheiden. Een voorbeeld van een AR II restrictie is:

$$o_i v_i \geq u_r \qquad (4.5)$$

Deze voorwaarde geeft een relatie weer tussen in- en outputgewichten. Dit type restricties is noodzakelijk wanneer niet de individuele variabelen, maar een combinatie van deze in- en outputvariabelen de mate van efficiëntie van een DMU moet weergeven (Allen et al., 1997). Een voorbeeld hiervan is een studie van Thanassoulis et al. (1995), waarin in een DEA-analyse de variabele 'number of babies at risk' als inputvariabele is gekozen en 'number of babies at risk surviving' als outputvariabele, om zo 'survival of babies et risk' te modelleren. Omdat het aantal baby's dat overleeft direct gekoppeld is aan het aantal baby's dat risico loopt om te overlijden, moeten deze in- en outputs worden gekoppeld met AR II restricties.

Allen et al. (1997) stellen dat een DEA-model waarin restricties van type AR II worden gebruikt gemakkelijk *infeasible* (niet oplosbaar) kan worden (net als bij AR I restricties). Ook bij AR II restricties zijn de relatieve efficiëntiescores onafhankelijk van de oriëntatie van het model en zijn de effecten van de restricties afhankelijk van de eenheden van in- en outputvariabelen.

4.2.1.3 Interpretatie DEA resultaten onder gewichtsrestricties

De Production Possibility Set (PPS) bevat in principe alle mogelijke realiseerbare input-outputhoeveelheidscombinaties die zijn afgeleid van de verschillende DMU's uit een bepaalde dataset. Dit zijn dus alle input-outputhoeveelheidscombinaties die liggen in het convexe gebied van alle gerealiseerde input-outputhoeveelheidscombinaties van de deelnemende DMU's.

De zogenaamde efficiëntiegrens (de rand van het convexe gebied) bevat precies die input-outputhoeveelheidscombinaties uit de PPS, die met de informatie van de deelnemende DMU's als (relatief) efficiënt worden betiteld.

Allen et al. (1997) beschrijven dat het opnemen van gewichtsrestricties in een DEA model ertoe leidt dat bepaalde gedeeltes van de efficiëntiegrens niet meer als efficiënt worden beschouwd en dat dit gevolgen heeft voor de interpretatie van de resultaten van deze gerestricteerde DEA-modellen. In het bijzonder onderscheiden Allen et al. (1997) effecten op de drie belangrijkste DEA-resultaten: de efficiëntiescore, de targets (verbeterpotentieel) en de DMU's die door het model als efficiënt worden aangemerkt, de zogenaamde *peer* DMU's.

Allen et al. (1997) tonen aan dat de efficiëntiescores, bij het gebruik van gewichtsrestricties, niet meer zijn te interpreteren als een zogenaamde *radiale efficiëntiemaat*, zoals dat wel het geval is in een DEA-model zonder deze gewichtsrestricties (behalve als alle restricties niet bindend zijn, en feitelijk overbodig zijn). De term *radiaal* houdt in het geval van een inputgeoriënteerd model in dat alle inputvariabelen met hetzelfde percentage ($1- h_{j0}^* = 1-e_j^*$) moeten worden ingekrompen (bij vergelijkbare ouputniveaus) om de DMU efficiënt te laten opereren, terwijl bij een outputgeoriënteerd model alle outputvariabelen met hetzelfde percentage, ($h_{j0}^*-1=1/ e_j^*-1$, zie Bijlage A) zouden moeten kunnen stijgen om deze DMU (j_0) efficiënt te laten produceren (bij vergelijkbare niveaus van de inputvariabelen).

Allen et al. (1997) laten naast het effect van gewichtsrestricties op de efficiëntiescores ook het effect van deze restricties op de targets zien. Het opnemen van deze gewichtsrestricties in het DEA-model kan er voor zorgen dat het behoud van de huidige productmix (de verhouding tussen in- en outputniveaus) van een bepaalde DMU geen voorrang meer krijgt in het DEA-model wanneer de targets worden bepaald, wat in een standaard (CCR-) DEA-model wel het geval is. Hiernaast stellen Allen et al. (1997) dat bepaalde inputtargetsniveaus hoger kunnen liggen dan in de uitgangssituatie onder een inputgeoriënteerd model, en dat bepaalde outputtargetniveau's lager kunnen liggen wanneer een outputgeoriënteerd model wordt gebruikt.

Als laatste effect van gewichtsrestricties op DEA-resultaten stellen Allen et al. (1997) dat deze restricties er toe kunnen leiden dat een inefficiënte DMU niet perse een vergelijkbare input/outputmix levert als zijn *peer* DMU's, zoals bij onaangepaste DEA het geval is. Dit kan er voor zorgen dat een inefficiënte DMU een heroriëntatie moet volgen om als efficiënt te worden betiteld, waarin niet perse de input- en outputvariabelen waarin de DMU het sterkst is worden benadrukt.

4.2.2 Restricties op virtuele DEA-gewichten

Een andere manier van het integreren van expertise, of het stellen van prioriteiten, in DEA is het gebruik van restricties op virtuele input en/of virtuele output, in plaats van het stellen van restricties op de in-

/outputgewichten zoals in paragraaf 4.2.1 is besproken. Virtuele input is gedefinieerd als de 'gewogen inputparameter' $v_i x_{ij}$ en virtuele output als de 'gewogen outputparameter' $u_r y_{rj}$.

Een van de voordelen van restricties op virtuele in- en outputs, in plaats van op de gewichten, is, zo stellen Sarrico en Dyson (2004), dat de resultaten in het laatste geval afhankelijk zijn van de gekozen eenheden, terwijl dit met restricties op virtuele in- en output niet het geval is, omdat de virtuele output dimensieloos is. Wat inhoudt dat de virtuele bijdragen van outputs/inputs onafhankelijk zijn van de gekozen eenheden en dus onderling (direct) vergelijkbaar zijn (ze tellen op tot 1). Hierom zijn de virtuele in/outputfracties te interpreteren als 'procenten van de totale gewogen productie' waardoor virtuele input en/of outputrestricties gemakkelijker zijn te formuleren en te interpreteren dan gewichtsrestricties (Sarrico en Dyson, 2004). Naast het feit dat deze virtuele restrictiemethode het formuleren van restrictiegrenzen gemakkelijker maakt, brengt dit ook met zich mee dat de restricties gemakkelijker uit zijn te leggen aan mensen die onbekend zijn met de DEA-techniek (bijvoorbeeld het management van een ziekenhuis).

Net als bij het gebruik van gewichtsrestricties, zoals besproken in paragraaf 4.2.1, kunnen ook voor het geval van virtuele gewichtsrestricties *absolute restricties* en *assurance regions restricties* worden onderscheiden. Deze verschillende typen virtuele restricties worden beurtelings besproken.

4.2.2.1 Absolute virtuele restricties

Wong en Beasley (1990) stellen voor om niet de variabiliteit van de gewichten te beperken, maar de variabiliteit van de fracties van de totale virtuele output van DMU_j, $\sum_s u_s y_{sj}$, die wordt 'geleverd' door output r. Dat kan worden gelezen als de 'belangrijkheid' van het aandeel van output r in de totale virtuele output, die uiteindelijk de efficiëntiescore van deze DMU_j bepaalt. De virtuele in- en outputs geven dus aan hoeveel de individuele in- en outputs bijdragen aan de efficiëntiescore, e_{j0}.

Een voorbeeld van een virtuele gewichtsrestrictie is bijvoorbeeld:

$$u_r y_{rj} \leq a_r \qquad (4.6)$$

Maar door bijvoorbeeld onderstaande zogenaamde *proportionele restricties* aan het standaard DEA-model toe te voegen wordt de virtuele ('gewogen') outputfractie van output r geforceerd om tussen de, door de gebruiker te specificeren, grenswaarden b_r en c_r te liggen:

$$b_r \leq \frac{u_r y_{rj}}{\sum_s u_s y_{sj}} \leq c_r \qquad (4.7)$$

Een dergelijke modelaanpassing kan natuurlijk ook op de inputvariabelen worden toegepast:

$$d_i \leq \frac{v_i x_{ij}}{\sum_m v_m x_{mj}} \leq e_i \qquad (4.8)$$

Wong en Beasley (1990) en Sarrico en Dyson (2004) geven aan hoe het resulterende niet-lineaire programmeringprobleem toch lineair kan worden opgelost, en beschouwen daarbij verschillende vormen van virtuele input/output restricties.

Sarrico en Dyson (2004) beschrijven waarom het gebruik van restricties van type (4.7) en/of (4.8) op DMU*j* onredelijke restricties kan opleggen aan de overige DMU's en hoe de targets erg complex worden (ze zijn o.a. afhankelijk van de gekozen grenzen (a_r, b_r, c_r, d_i, e_i) en van de input-/outputniveaus van andere DMU's). Een ander nadeel aan dit type restrictie is dat de resultaten van deze techniek afhankelijk zijn van de oriëntatie van het model (Allen et al., 1997). Hiernaast kunnen de restricties ervoor zorgen dat het model onoplosbaar (*infeasible*) wordt (Sarrico en Dyson, 2004). Estellita Lins et al. (2007) beschrijven voor dit type restricties voorwaarden waaronder het model wel oplosbaar is, en presenteren een manier om een onoplosbaar model met minimale wijzigingen in de gewenste gewichten wel oplosbaar te maken.

4.2.2.2 Virtual assurance regions

Naast absolute virtuele gewichtsrestricties worden ook de zogenaamde *virtuele assurarance regions* (VAR) onderscheiden binnen de klasse van virtuele restricties. Wanneer prioriteiten over virtuele inputs en outputs een belangrijke rol in de DEA-analyse spelen (en geen absolute grenzen zoals in paragraaf 4.2.2.1) kan het virtual assurance regions (VAR) concept een oplossing bieden.

Onderstaande restrictie geeft de meest algemene vorm weer van virtuele assurance regions gewichtsbeperkingen:

$$\sum_i a_{ij} v_i x_{ij_0} + \sum_r b_{rj} u_r y_{rj_0} \geq 0 \qquad (4.9)$$

Virtuele assurance regions van type I (VAR I) zijn restricties die slechts verbanden leggen tussen ofwel virtuele inputs, ófwel virtuele outputs, maar niet tussen een combinatie van beiden (vergelijkbaar met AR I). VAR I restricties ontstaan wanneer in (4.9) geldt dat ofwel $a_{ij} = 0 \ \forall i$ of $b_{rj} = 0 \ \forall r$.

Wanneer er minimaal één waarde a_{ij} en één waarde b_{rj} ongelijk is aan 0, dan ontstaan zogenaamde virtuele assurance regions van type II (VAR II) die kruisverbanden leggen tussen virtuele input- en outputvariabelen.

4.2.2.3 Interpretatie DEA resultaten onder virtuele gewichtsrestricties

Sarrico en Dyson (2004) tonen aan dat het gebruik van absolute virtuele restricties (net als het gebruik van niet-virtuele gewichtsrestricties) leidt tot niet radiale efficiëntiematen. Hierdoor zijn de efficiëntiescores en targets vaak moeilijk te interpreteren. Zij (Sarrico en Dyson, 2004) betogen echter dat het gebruik van VAR restricties zonder AR I- of AR II-restricties ($k_w = 0 \ \forall w$ in Bijlage B), die alleen van toepassing zijn op de specifieke DMU_j, die in het enkelvoudige LP-model een efficiëntiescore krijgt, ervoor zorgt dat de efficiëntiescores naar de, in het standaard CCR-model gebruikelijke, radiale scores convergeren (zie bijlage B, met $k_w = 0 \ \forall w$). Ook de targets (zie bijlage B) zijn veel minder complex dan onder gewichtsrestricties (absoluut, AR I, AR II) of absolute virtuele gewichtsrestricties. Het voordeel hiervan is dat het model helderder is, en beter uit te leggen.

Ook voor VAR restricties geldt (Allen et al., 1997), net als in een model met gewichtsrestricties, dat het opnemen van deze gewichtsrestricties in het DEA-model ervoor kan zorgen dat het behoud van de huidige productmix van een bepaalde DMU geen voorrang meer krijgt in het DEA-model wanneer de targets worden bepaald, en dat bepaalde inputtargetsniveaus hoger kunnen liggen dan in de uitgangssituatie onder een inputgeoriënteerd model en bepaalde outputtargetniveau's lager wanneer een outputgeoriënteerd model wordt gebruikt.

Samengevat kan worden gesteld dat AR- en VAR-restricties kunnen worden gebruikt om DEA-gewichten binnen bepaalde kaders te dwingen of om bepaalde verbanden tussen verschillende gewichten te leggen. AR-I- en AR-II-restricties leiden vaak tot *infeasibility* van de DEA-modellen en leveren geen gemakkelijke op 'radiale wijze' interpreteerbare targets op. Ook VAR-restricties leveren moeilijk interpreteerbare (niet radiale) targets op behalve wanneer er slechts VAR-restricties worden opgelegd op (de virtuele gewichten van) de DMU die een efficiëntiescore krijgt toebedeeld in een enkelvoudig LP-model. In dit geval zijn te targets gemakkelijk te interpreteren en blijft het DEA-model *feasible*. Wel leiden deze

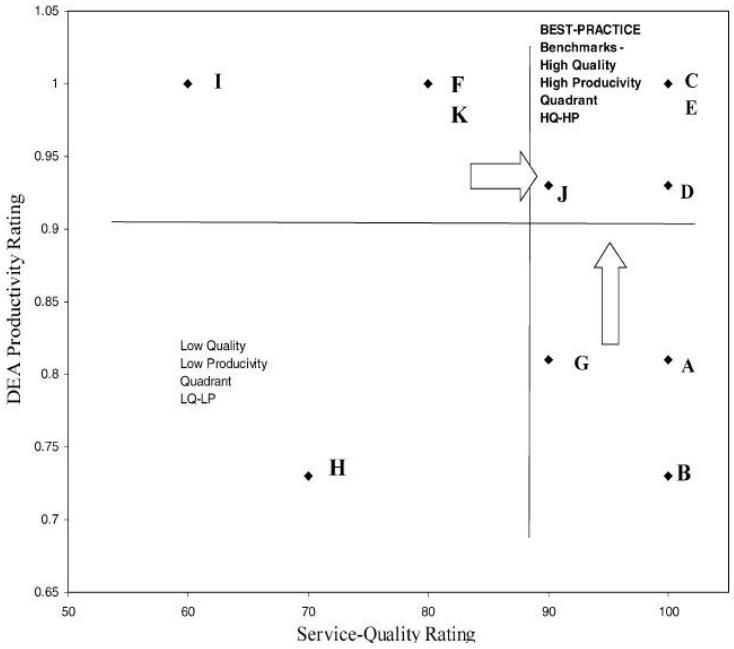
restricties ertoe (net als alle andere gewichtsrestricties) dat het behoud van de 'eigen productmix' geen voorrang meer krijgt en dat de targets niet altijd in de gewenste richting zullen liggen.

4.3 'Kwaliteit' als toetsingscriterium van DEA-resultaten

In plaats van het 'forceren' van een trade-off tussen kwaliteit en productiviteit in een DEA-model, door het opleggen van (virtuele-)gewichtsrestricties (zoals in paragraaf 4.2), kan er ook voor worden gekozen om deze trade-off buiten het DEA-model te houden en minimum kwaliteitseisen te stellen aan de DMU's. In dit hoofdstuk worden twee methodes besproken die op dit idee zijn gebaseerd. De eerste methode die wordt besproken is een methode van Sherman en Zhu (2006a), waarbij de analyse gebaseerd is op een plot met twee dimensies: een kwaliteit- en een kostenefficiëntiedimensie. De tweede methode, Q-DEA genaamd, van Sherman en Zhu (2006b), die wordt behandeld in paragraaf 4.3.2., selecteert DMU's voor de DEA-analyse op basis van kostenefficiëntie en minimale kwaliteitseisen.

4.3.1 Grafische methode

Sherman en Zhu (2006a) presenteren een grafische methode die bestaat uit een plot met twee dimensies: op de horizontale as kan een kwaliteitsvariabele worden uitgezet tegen de DEA-(kosten)efficiëntiescore op de verticale as (zie figuur 3). Door nu een bepaalde 'minimale standaard' voor 'goede kwaliteit' te kiezen en een bepaalde 'minimum standaard' voor een 'goede efficiëntiescore', kan de plot worden opgedeeld in vier kwadranten. In figuur 3 is het minimum voor 'goede kwaliteit' gesteld op een *Service-Quality Rating* van 90 (op een schaal van 1 tot 100), en het minimum voor een 'goede efficiëntiescore' op 0.9 (*DEA Productivity Rating*). Met de typeringen 'lage kwaliteit', ofwel *Low Quality* (*LQ*), 'hoge kwaliteit', ofwel *High Quality* (*HQ*), 'lage



Figuur 3: Q-DEA grafisch (Sherman en Zhu, 2006a)

productie', ofwel *Low Productivity* (*LP*) en 'hoge productie', ofwel *High Productivity* (*HP*), ontstaan zo (tegen de klok in) de kwadranten *LQ-LP*, *HQ-LP*, *HQ-HP* en *LQ-HP*.

Er kan nu worden gesteld dat DMU's in het *LQ-LP*-kwadrant op zowel het gebied van servicekwaliteit, als op het gebied van productiviteit slecht scoren (vb: DMU *H*). DMU's in kwadrant *HQ-LP*-kwadrant presteren goed op het kwaliteitsaspect, maar produceren niet kostenefficiënt (vb: DMU's *A*, *B* en *G*). Deze DMU's zouden zich met behoud van het hoge kwaliteitsniveau in efficiëntie moeten kunnen verbeteren tot DMU's in kwadrant *HQ-HP*, het kwadrant van DMU's die op beide facetten, kwaliteit en kostenefficiëntie, goed presteren (vb: DMU's *C*, *D*, *E* en *J*). In kwadrant *LQ-HP* zitten de DMU's die weliswaar kostenefficiënt produceren, maar waar deze kostenefficiëntie mogelijk ten koste gaat van het kwaliteitsniveau (vb: DMU's *F*, *I* en *K*). Deze DMU's zouden zich eerst in kwaliteitsopzicht moeten verbeteren voor de prestaties van deze DMU's op eerlijke wijze kunnen worden vergeleken met de prestaties van de DMU's in *HQ*-kwadranten.

Deze tweedimensionale plotmethode heeft de beperking dat de DEA-efficiëntiescore met slechts één kwaliteitsvariabele kan worden vergeleken. Hiernaast moet men zich realiseren dat de DEA-efficiëntiescores van DMU's met lage kwaliteitsscores, maar met hoge efficiëntiescores (*LQ-HP*), de productiviteit van DMU's in een *HQ*-kwadrant onderschatten, terwijl de hoeveelheid die *LQ-HP* aan inputs lijken te kunnen besparen wordt overschat in de DEA-efficiëntiescores (Sherman en Zhu, 2006). Voor deze DMU's zijn de met DEA berekende verpeterpotentiëlen en target dus maar van beperkte waarde.

Een tweede nadeel van deze methode is dat 'goede kwaliteit' hierin een statisch begrip vormt: een DMU voldoet óf aan de kwaliteitseis (90 punten in het voorbeeld van figuur 3) óf niet. Let op: figuur 3 geeft slecht aan welke DMU's zich kwalificeren om met elkaar op redelijke wijze vergeleken te worden, vergelijking van efficiëntie met andere DMU's aan de hand va de afbeelding is misleidend! Wanneer een DMU zich bijvoorbeeld in het *LQ-HP*-kwadrant bevindt, dan kan niet onmiddellijk worden geconcludeerd dat deze DMU alleen aan zijn kwaliteit moet werken om in het *HQ-HP*-kwadrant terecht te komen. Wanneer er namelijk een trade-off bestaat tussen kwaliteit en kostenefficiëntie moet er op beide terreinen winst worden geboekt. Of deze trade-off bestaat, en zo ja, hoe dit verband ligt, is niet op te maken uit figuur 3.

Binnen de *HQ*-kwadranten worden de kwaliteitsverschillen niet expliciet meegenomen in de beoordeling van de (brede) efficiëntie van de DMU's (in de efficiëntiescores).

4.3.2 'Quality-adjusted' DEA

Sherman en Zhu (2006b) presenteren een methode, gebaseerd op de ideeën uit paragraaf 4.3.1, die zij *Quality-adjusted DEA* (Q-DEA) noemen. Deze methode heeft ten doel om in DEA-analyses, waarin een kwaliteitsoutput is opgenomen, de efficiëntiescores en het verbeterpotentieel op een dusdanige manier te berekenen dat de onzuiverheden in deze waarden worden vermeden die kunnen worden veroorzaakt door *LQ-HP* DMU's die als referentie-DMU's voor andere DMU's dienen. Dit wordt bewerkstelligd door *LQ-HP* DMU's nu niet meer als referentie-DMU's te gebruiken om de efficiëntiescores en het verbeterpotentieel van andere DMU's te berekenen, maar deze (stapsgewijs) te elimineren uit de dataset.

De door Sherman en Zhu (2006b) voorgestelde methode is een iteratieve tweestapsmethode. In elke iteratie wordt in de eerste stap het standaard DEA-model zonder kwaliteitsoutput doorgerekend. In de tweede stap van de iteratie worden alle *LQ-HP* DMU's met een efficiëntiescore van 1 (en geen slacks) uit het model verwijderd. De iteratie wordt herhaald totdat er geen *LQ* DMU's meer zijn die een efficiëntiescore van 1 krijgen toebedeeld.

Het voordeel van Q-DEA is dat er een minimum aan kwaliteit kan worden geëist door de uitvoerders van de benchmark, en dat de resulterende efficiëntiescores en targets (verbeterpotentieel) voor de resterende DMU's realistische resultaten vormen. In tegenstelling tot de grafische methode uit paragraaf 4.3.1. levert Q-DEA wél (kwantitatieve) realistische targets en efficiëntiescores (voor het gedeelte van de DMU's dat aan de kwaliteitseisen voldoet).

Een nadeel van de methode is dat voor *LQ-HP* DMU's geen realistische efficiëntiescore kunnen worden berekend, en dat voor deze DMU's het verbeterpotentieel niet kan worden bepaald. Verder geldt, net als voor de grafische methode uit paragraaf 4.3.1, het bezwaar dat er maar een enkele kwaliteitsoutput in het model kan worden meegenomen.

4.4 Overzicht methoden

In deze paragraaf wordt een overzicht gegeven, in tabelvorm, van de belangrijkste voor- en nadelen van de in hoofdstuk 4 gepresenteerde technieken om kwaliteitsaspecten mee te nemen in een DEA-analyse. In tabel 1 zijn verticaal de belangrijkste eigenschappen van de horizontaal vermelde technieken te vinden die zijn gepresenteerd in hoofdstuk 4. In tabel 1 is een onderscheid gemaakt tussen drie soorten modelaanpassingen: "geen beperkingen", "absolute beperkingen" en "relatieve beperkingen". Naar gelang

de gewenste toepassing van een DEA-model kan zo overzichtelijk voor een bepaalde methode worden gekozen.

Nu volgt eerst een korte uitleg van de eigenschappen van de verschillende methodes die terug te vinden zijn in tabel 1:

- "Belang 'kwaliteit' gegarandeerd" geeft aan in welke mate een bepaalde techniek zeker stelt dat het kwaliteitsaspect van de 'productie' voldoende (en op zinnige wijze) wordt meegenomen in het model. Wanneer een techniek slecht scoort op dit onderdeel zijn ook de targets en efficiëntiescores niet realistisch/eerlijk;
- "Interpreteerbaarheid modelaanpassingen" geeft aan hoe gemakkelijk een uitleg is te geven aan de bij een bepaalde techniek passende modelaanpassingen (ten opzichte van een standaard DEA-model);
- "Interpreteerbaarheid targets" geeft aan hoe gemakkelijk het is om een uitleg te geven aan de (mathematische) uitdrukking van de targets (verpeterpotentiëlen);
- "Wenselijkheid richting targets" geeft aan of de targets door het model altijd in dezelfde richting worden gesteld zoals is bedoeld in de modeloriëntatie (bijv. in een inputgeoriënteerd model is de hoop dat de targetinputs lager liggen dan de inputs in de uitgangssituatie);
- "Interpreteerbaarheid efficiëntiescores"geeft aan hoe gemakkelijk het is om een praktische uitleg te geven aan de (cijfermatige) efficiëntiescore;
- "Risico's op onoplosbare modellen" geeft aan of het model gemakkelijk onoplosbaar (*infeasible*) wordt onder het gebruik van een bepaalde techniek;
- "Gemak implementatie" geeft aan hoeveel kennis er van het model en het productieproces van de DMU's waarop het model wordt toegepast is vereist voor een succesvolle implementatie;
- "Afhankelijkheid modeloriëntatie" geeft aan of de efficiëntiescores onder verschillende modeloriëntaties (input of output) verschillend zijn.

Aan de hand van tabel 1 kan in het volgende hoofdstuk (5) een bepaalde methode worden gekozen om te gebruiken voor de Benchmark Kliniek 2008. Deze keuze kan worden gemaakt door een afweging te maken tussen het belang van de verschillende criteria voor deze specifieke benchmark en vervolgens de scores van de verschillende methodes op deze criteria met elkaar te vergelijken.

Tabel 1: Overzicht van methodes om 'kwaliteit' te integreren in DEA

Criteria	Belang 'kwaliteit' gegaran-deerd	Interpreteer-baarheid model aanpassingen	Interpreteer-baarheid targets	Wenselijkheid richting targets	Interpreteer-baarheid efficiëntie-scores	Risico's op onoplosbare modellen	Gemak implemen-tatie	Afhankelijk-heid model-oriëntatie
Methodes								
Geen beperkingen:								
Kwaliteit als output in onaangepaste DEA (§ 4.1)	-	n.v.t.	+	+	+	+	+	+
Absolute beperkingen:								
Absolute gewichtsres-tricties (§ 4.2.1.1)	+	-	-	-	-	-	-	-
Absolute virtuele gewichtsres-tricties (§ 4.2.2.1)	+	+	-	-	-	-	+	-
Grafische toetsingsme-thode[1] (§ 4.3.1)	+/-[2]	n.v.t.	-	+	-	+	+	+
Quality-adjusted DEA[1] (§ 4.3.2)	+/-[2]	n.v.t.	+/-[2]	+	+/-[2]	+	+	+
Relatieve beperkingen:								
Assurance regions (type I en II) (§ 4.2.1.2)	+	-	-	-	-	-	-	+
Virtuele assurance regions (type I en II) (§ 4.2.2.2)	+	+	+	-	+	+	+	+

In de tabel zijn positieve scores op een bepaalde eigenschap aangegeven met een "+" en negatieve scores met een "-".

[1] Het is slechts mogelijk één kwaliteitsoutput mee te nemen in het model
[2] Alleen de resultaten van de *HQ*-DMU's zijn waardevol en juist te interpreteren

5 Toepassing aangepaste DEA op de klinische afdelingen van de ziekenhuizen

In dit hoofdstuk wordt, aan de hand van voorlopige data, het DEA-gedeelte van de Benchmark Kliniek 2008 van ORTEC uitgevoerd door gebruik te maken van de informatie uit hoofdstuk 2 (het servicemodel van Grönroos) en hoofdstuk 4 (kwaliteit en DEA). Met behulp van deze informatie kan de benchmark zo worden opgezet dat het 'brede efficiëntie'-begrip (zie hoofdstuk Inleiding) zo goed mogelijk wordt gemodelleerd, waardoor de efficiëntiescores eerlijker en het verbeterpotentieel realistischer kunnen worden berekend dan in de benchmark van 2007.

In paragraaf 5.1 wordt gepresenteerd hoe de kwaliteit van de geleverde zorg op de klinische afdelingen kan worden gemodelleerd. Dit principe wordt vervolgens in paragraaf 5.2 gebruikt om het kwaliteitsaspect van de 'productie' van deze klinische afdelingen te integreren in DEA (gezamenlijk met de meer 'tastbare' productieaspecten). In paragraaf 5.3 wordt beschreven welke data is gebruikt voor de (voorlopige) resultaten van het DEA-gedeelte van de benchmark Kliniek 2008. Deze resultaten zijn tot stand gekomen met behulp van de technieken uit paragraaf 5.1 en 5.2, en worden beschreven in paragraaf 5.4.

5.1 Modelleren van kwaliteit op de klinische afdelingen

Om een 'breed efficiëntie'-begrip te kunnen integreren in DEA, moet het kwaliteitsaspect dat ingesloten zit in dit begrip eerst worden gekwantificeerd. Gezien het relatief kleine aantal deelnemende ziekenhuizen (DMU's) en het belang van niet direct aan kwaliteit gerelateerde outputs (zoals het aantal klinische opnames) in de benchmark, kan er maar een relatief klein aantal (in- en) outputs worden meegenomen in DEA. Het is dan ook de uitdaging om de kwaliteit van de zorg op de klinische afdelingen van de ziekenhuizen op informatieve wijze tot één (kwaliteit-)output (indicator) te verwerken. Het doel van deze kwaliteitsoutput is niet om de kwaliteit van de zorg binnen de klinische afdelingen uitgebreid te analyseren (wat onmogelijk is met één enkele waarde), maar om als corrigerende factor te fungeren voor de meer klassieke outputvariabelen in het DEA-model (paragraaf 5.2).

Voor de DEA-Benchmark Kliniek 2008 zullen prestatie-indicatoren, zoals besproken in paragraaf 2.2, als basis dienen voor het modelleren van 'kwaliteit'. Er is voor gekozen om een aantal, door ORTEC in overleg met de deelnemende ziekenhuizen bepaalde, prestatie-indicatoren te

transformeren tot een 'kwaliteitsindex' op een schaal van 0 tot 1, waarin geldt dat hoe hoger de index, hoe beter de prestaties van een ziekenhuis zijn in het beeld dat de prestatie-indicatoren geven.

Om uiteindelijk tot deze kwaliteitsindex te komen, worden in deze methode eerst alle individuele prestatie-indicatoren geschaald tussen de waarden 0 en 1. Deze schaling is een *relatieve* schaling, waarin het best presterende ziekenhuis op een bepaald indicatiegebied een score van 1 krijgt op dit specifieke indicatiegebied. In plaats van deze relatieve score, is ook overwogen om voor een *absolute* score te kiezen, waarin er voor iedere indicator (op basis van expertise) een vaste onder- en bovengrens moet worden gekozen die aangeeft wat een 'slechte prestatie' respectievelijk een 'goede prestatie' is. Het beoogde voordeel van deze absolute score methode boven de relatieve score methode zou eruit moeten bestaan dat de ontwikkelingen op het gebied van kwaliteit per ziekenhuis door de jaren heen (in de herhaalde benchmark) zichtbaar zouden kunnen worden gemaakt. Dat is niet mogelijk wanneer een relatieve score wordt gebruikt, omdat de onder- en bovengrenzen - de waarden van respectievelijk het slechtst en best presterende ziekenhuis op dit indicatorgebied - met het jaar veranderen doordat er ieder jaar nieuwe data wordt aangeleverd door de ziekenhuizen. Toch is ervoor gekozen de relatieve score toe te passen om drie redenen die nu beurtelings worden besproken:

- De eerste reden is dat een relatieve score geen 'last' heeft van ontbrekende expertise op het gebied van theoretische onder- en bovengrenzen, wat wel het geval is wanneer een absolute score wordt gebruikt. Hierdoor is het 'gokken' van parameters, waarvoor geen expertise beschikbaar is, zo min mogelijk nodig en wordt de kwaliteitsindexscore zo min mogelijk vertroebeld omdat de indexscores direct afhankelijk zijn van de gekozen onder- en bovengrenzen. Wanneer er bijvoorbeeld 'voor de zekerheid' te hoge bovengrenzen worden gekozen zullen de kwaliteitsindexscores worden genivelleerd (vertroebeld), waardoor het onderscheidingsvermogen van de kwaliteitsindex afneemt. De relatieve score is dus datatechnisch de zuiverste en de meest onderscheidende index van de twee indices;
- De tweede reden is dat de relatieve kwaliteitsindexscore gemakkelijk is te interpreteren en goed is uit te leggen aan de deelnemende ziekenhuizen in de DEA-context van de benchmark, omdat, net als in DEA, bij het gebruik van een relatieve score een aantal opzichzelfstaande variabelen worden gecombineerd tot één relatieve score;
- De derde reden is dat de beoogde ontwikkeling van de individuele ziekenhuizen op kwaliteitsgebied over de jaren, ook wanneer er voor een absolute kwaliteitsindexscore wordt gekozen, niet helder in beeld is te brengen. Dit komt doordat het geen zin heeft om dit jaar theoretische maximumwaarden te definiëren voor indicatoren, wanneer de kans groot is dat er een volgend jaar een maximum indicatorwaarde in de data voorkomt die

groter is dan het (theoretische) maximum uit eerdere benchmarks, waardoor het theoretische maximum alsnog moet worden aangepast. Hierdoor zijn de kwaliteitsindexscores alsnog niet vergelijkbaar tussen de verschillende jaren waarin de benchmark is uitgevoerd. Hetzelfde bezwaar geldt wanneer er door de jaren heen een normverschuiving plaatsvindt, waardoor de theoretische onder- en bovengrenzen van 'goede kwaliteit' moeten worden aangepast. Hiernaast geldt nog dat de kwaliteitsindex vanwege zijn hoge aggregatieniveau uiteindelijk niet het beste instrument uit de Benchmark Kliniek is om de ontwikkeling van de kwaliteit van de zorg op de klinische afdelingen door de jaren heen te monitoren (dit geld evenzeer voor de relatieve kwaliteitsindex). De ratio's, die een belangrijk onderdeel van de benchmark uitmaken, zijn hiervoor een stuk beter geschikt.

Het kan voorkomen dat bepaalde ziekenhuizen bepaalde informatie niet voorhanden hebben, waardoor de bijbehorende prestatie-indicatoren niet kunnen worden ingevuld. Het lijkt dan het 'eerlijkste' om aan deze ziekenhuizen, voor de uiteindelijke berekening van de kwaliteitsindex, de gemiddelde indicatorscore van de ziekenhuizen die deze indicator wél hebben ingevuld mee te geven, omdat het toekennen van een gemiddelde score ervoor zorgt dat dit ziekenhuis er in kwaliteitsindexscore niet op voor- of achteruit gaat ten opzichte van het 'gemiddelde' ziekenhuis. Hierdoor springt dit ziekenhuis er gemiddeld niet beter of slechter uit doordat het niets heeft ingevuld. De aanname die hierbij wel moet worden gemaakt is dat een ziekenhuis niet opzettelijk (slechte) informatie achterhoudt. Wanneer een ziekenhuis bijvoorbeeld weet dat het slecht presteert, dan zal de gemiddelde kwaliteitsindicatorscore van alle deelnemende ziekenhuizen waarschijnlijk hoger liggen dan de score van dit individuele ziekenhuis. Hierdoor zou voor het ziekenhuis de prikkel kunnen ontstaan om opzettelijk de indicatorscore niet in te vullen om zo zijn kwaliteitsindex op te hogen.

Omdat ieder ziekenhuis in de Benchmark Kliniek participeert om de bedrijfsvoering te verbeteren, heeft (het is een interne benchmark) een ziekenhuis geen belang bij (integendeel zelfs) het 'oppoetsen' van de kwaliteitscores, en lijkt dit een redelijke aanname. Natuurlijk moet wel worden opgemerkt dat de kwaliteitsindexscores aan kwaliteit inboeten wanneer bepaalde informatie niet is ingevuld (triviaal) en dat hierdoor ook het onderscheidend vermogen van de index kan verminderen omdat de indicatorscores dichter bij elkaar kunnen komen te liggen.

Met behulp van de zojuist besproken genormeerde relatieve indicatorscores kan nu de kwaliteitsindex (met eveneens waarden tussen 0 en 1) worden geconstrueerd door aan de verschillende indicatoren een bepaald gewicht binnen een zogenaamde *pijler* toe te wijzen, en vervolgens deze pijlers ook weer een bepaald gewicht binnen de totale indexscore mee te geven. Al deze gewichten kunnen

worden gebruikt om verschillende facetten van 'kwaliteit van zorg' naar een bepaalde, door de gebruiker (de ziekenhuizen in samenspraak met ORTEC) te bepalen, balans te sturen. In bijlage C1 is een nauwkeurige definitie van de berekening van de kwaliteitsindex terug te vinden.

5.2 Modelleren van kwaliteit in DEA voor de benchmark Kliniek

In deze paragraaf wordt met behulp van hoofdstuk 4 (over kwaliteit en DEA) het kwaliteitsaspect in DEA geïntegreerd, zodat de gewenste 'brede' analyse van de efficiëntie van de klinische afdelingen van de aan de benchmark deelnemende ziekenhuizen ontstaat. Zoals in hoofdstuk 4 is beschreven heeft iedere techniek voor- en nadelen. In paragraaf 4.4 is (in tabel 1) overzichtelijk samengevat hoe deze voor- en nadelen zich tot elkaar verhouden. Afhankelijk van de toepassing kan zo worden gekozen wat de meest geschikte techniek is voor deze specifieke toepassing.

Wanneer nu wordt gekeken naar de Benchmark Kliniek, dan staat het belang van een goede integratie van 'kwaliteit' in DEA voorop, omdat de kwaliteit van de geleverde zorg een zeer essentieel deel is van de prestatie van een ziekenhuis, zodat het 'gewicht van kwaliteit' gegarandeerd moet zijn om de prestaties van de verschillende ziekenhuizen eerlijk te kunnen vergelijken (en voor deze ziekenhuizen realistische targets te kunnen berekenen).

De optie om de kwaliteitsindex als outputvariabele mee te nemen in standaard DEA lijkt niet heel geschikt, omdat de techniek het gewicht van 'kwaliteit' niet garandeert. Zeker in het geval van een herhaalde benchmark is deze techniek niet geschikt, omdat een acceptabel gewicht van 'kwaliteit' in de benchmark van het ene jaar geen enkele garantie biedt dat dit gewicht in een andere jaar opnieuw acceptabel zal zijn. Het zoeken naar alternatieve optima is hierbij zeer omslachtig en biedt ook geen enkele garantie op succes. Deze techniek scoort dan ook een "- " op dit aspect in de tabel in paragraaf 4.4.

Ook de "Grafische toetsingsmethode" en "Q-DEA" vallen af als geschikte kandidaten omdat deze technieken alleen voor DMU's (ziekenhuizen) die een hoge kwaliteit (*HQ*) (van service) afleveren zinnige resultaten kunnen presenteren, hetgeen voor de Benchmark Kliniek ongewenst is.

Van de vier technieken die overblijven, Absolute gewichtsrestricties, Absolute virtuele gewichtsrestricties, AR en VAR, domineert de VAR-techniek de drie andere technieken doordat deze techniek altijd minstens even goed scoort als de andere drie technieken, en op een aantal terreinen beter scoort. Deze Virtual Assurance Regions techniek scoort op alle criteria goed, maar kan niet garanderen, in het geval van een outputoriëntatie, dat de outputtargets altijd strikt hoger liggen dan de waarden in de uitgangssituatie, of dat de inputtargets altijd strikt lager liggen dan de waarden in de uitgangssituatie in het

geval van een inputoriëntatie. Dit resultaat, voor het geval van outputgeoriënteerd model, is niet verwonderlijk, gezien de beperkingen die het kwaliteitsaspect oplegt aan de 'productie'-outputs. Omdat er in de Benchmark Kliniek slechts één inputvariabele wordt meegenomen (zoals wordt besproken in paragraaf 5.3.1) zullen de inputtargets wel degelijk allemaal lager liggen dan de waarden in de uitgangssituatie. Concluderend kan worden gesteld dat de VAR-techniek, van de in hoofdstuk 4 besproken technieken, het beste scoort op de verschillende criteria om gebruikt te worden in de Benchmark Kliniek.

5.3 Beschrijving van de data

In deze paragraaf wordt besproken welke data beschikbaar waren op het moment dat het model is toegepast voor de resultaten van dit onderzoek. De data bestaan uit een set met prestatie-indicatoren van acht ziekenhuizen, die zowel wordt gebruikt voor het berekenen van de inputvariabelen, als voor de niet direct aan de kwaliteit van de zorg gerelateerde outputvariabelen en als inputs voor de kwaliteitsindex, die zelf weer een outputvariabele van het DEA-model vormt.

Om te garanderen dat de informatie die de ziekenhuizen leveren, in de vorm van prestatie-indicatoren (zie ook paragraaf 2.2), ook daadwerkelijk onderling vergelijkbaar is, heeft ORTEC zeer zorgvuldig de meetfrequenties en de definities van deze indicatoren vastgesteld en geëvalueerd. Om een aantal indicatoren te corrigeren voor eventuele verschillen in case mix (zie paragraaf 2.2) is de indicatorvariabele "aantal gewogen patiënteenheden" gebruikt (zie ook bijlage C2). Deze indicator is samengesteld als de gewogen optelling van het "aantal gewogen opnamen" (conform CTG 2001; met werkelijke in plaats van afgesproken aantal verrichtingen), het "aantal gewogen eerste polikliniekbezoeken" (conform CTG 2001; met werkelijke in plaats van afgesproken aantal verrichtingen), het "aantal verpleegdagen" en het "aantal dagverplegingdagen.

Zoals eerder gesteld (in hoofdstuk 1 en paragraaf 5.1) beperkt het relatief kleine aantal aan de benchmark deelnemende ziekenhuizen sterk het aantal in- en outputvariabelen dat in het model kan worden meegenomen. ORTEC heeft om deze reden de in- en outputvariabelen voor het model gekozen die gezamenlijk het productie/serviceproces van de ziekenhuizen voor het belangrijkste deel beschrijven. Het model bevat één inputvariabele, de "salariskosten van de FTE's" die werken op (of voor) de klinische afdelingen, dat een goede indicator is voor de omvang van het personeelsbestand op deze afdelingen,

welke de belangrijkste input vormt voor het productie/serviceproces. Er is verder voor gekozen om vier outputvariabelen op te nemen in het model: het "aantal dagbehandelingen", het "aantal verpleegdagen", het "aantal klinische opnames" en de "kwaliteitsindex". De eerste drie van deze variabelen zijn (gezamenlijk) goede indicatoren voor de fysieke output van de klinische afdelingen, de laatste, de kwaliteitsindex, wordt gebruikt om de (mogelijke) trade-off tussen fysieke productie en de kwaliteit van de geleverde zorg te modelleren (zie de modelbeschrijving in paragraaf 5.4). De in- en outputvariabelen voor het DEA-model zijn weergegeven in tabel 2.

Tabel 2: In- en output variabelen DEA-model

Input variabele (x_{ij}):	Outputvariabelen (y_{rj}):
Salariskosten FTE's	Aantal dagbehandelingen
	Aantal verpleegdagen
	Aantal klinische opnames
	Kwaliteitsindex

De prestatie-indicatoren die voor de kwaliteitsindex zijn gebruikt zijn terug te vinden in bijlage C2. Verder staat in deze bijlage vermeld hoe de prestatie-indicatoren die de ziekenhuizen hebben aangeleverd zijn gebruikt om een twintigtal indicatoren voor de kwaliteitsindex te creëren, binnen welke pijler deze indicatoren vallen (Zorginhoud, Betrouwbaarheid, Snelheid of Flexibiliteit), wat het gewicht is van deze indicatoren binnen deze pijler ("middel" of "hoog") en wat de gewichten van de verschillende pijlers zijn binnen de kwaliteitsindex (zie ook bijlage C1).

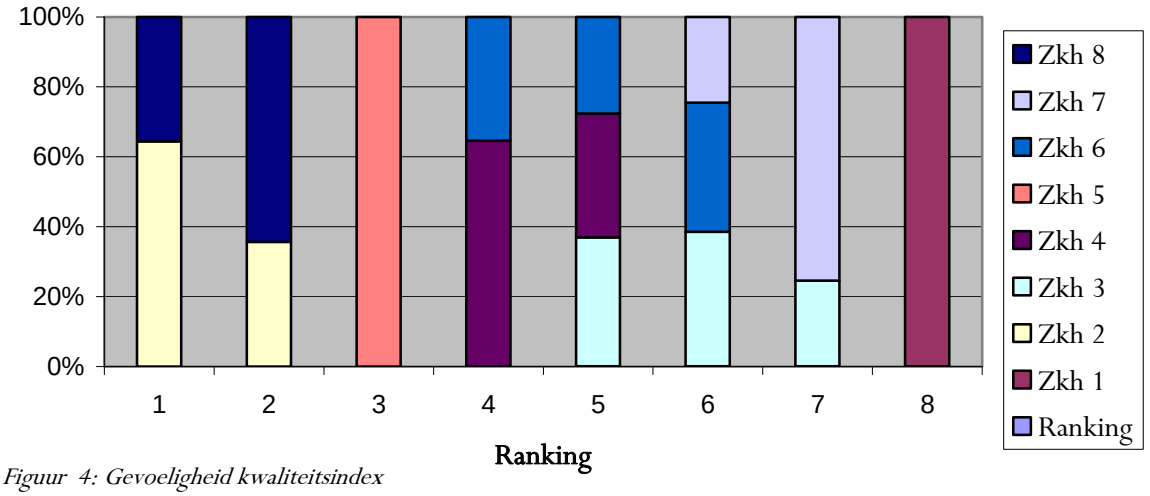
5.4 Resultaten van de benchmark en analyse

In deze paragraaf worden de voorlopige resultaten van de Benchmark Kliniek 2008 gepresenteerd en geanalyseerd. In paragraaf 5.4.1. worden eerst de resultaten rond de kwaliteitsindex besproken, waarna de resultaten van het DEA-model, in de vorm van de efficiëntiescores in paragraaf 5.4.2 en de bijbehorende targets in paragraaf 5.4.3, kunnen worden gepresenteerd en geanalyseerd.

5.4.1 Kwaliteitsindex

ORTEC heeft met ziekenhuizen die aan de benchmark deelnemen afgesproken welke gewichten de verschillende pijlers hebben binnen de kwaliteitsindex. Het gewicht van de verschillende indicatoren binnen deze pijlers is alleen als zijnde "middel" of "hoog" bepaald (zie bijlage C2). Om nu te onderzoeken hoe gevoelig de resultaten van de kwaliteitsindex zijn voor de verhouding tussen de gewichten "middel" en "hoog" kan een gevoeligheidsanalyse worden uitgevoerd op deze verhouding.

Deze gevoeligheidsanalyse is uitgevoerd als Monte Carlo-simulatie in Excel (Microsoft) in combinatie met het programma Crystal Ball (Decisioneering). Voor de simulatie is gebruik gemaakt van Latin Hypercube simulatie met een runlengte van 10000. De verhouding van de gewichten "middel" tot "hoog" is hierbij gevarieerd 1:1.5 tot 1:2.5 (de grenzen rond de, als redelijk veronderstelde verhouding 1:2). De "ranking" van een ziekhuis is de plaats van een ziekenhuis in de volgorde van het ziekenhuis met de hoogste kwaliteitsindexscore (ranking 1) tot en met het ziekenhuis met de laagste kwaliteitsindexscore (ranking 8). In figuur 4 is een stapeldiagram weergegeven met daarin de percentages die aangeven in welk gedeelte van de gesimuleerde kwaliteitsindexscores een bepaal ziekenhuis (Zkh) een bepaalde ranking heeft verkregen bij de wisselende gewichtsverhoudingen tussen "middel" en "hoog".



Figuur 4: Gevoeligheid kwaliteitsindex

Uit figuur 4 valt op te maken dat de kwaliteitsindexranking redelijk ongevoelig is voor de verhouding van de gewichten "middel" en "hoog" van de indicatoren binnen de verschillende pijlers. Er vinden geen "gekke" verschuivingen plaats bij het variëren van deze gewichten (bijvoorbeeld van ranking 1 naar ranking 7). In overleg met ORTEC is besloten om de verhouding "middel" staat tot "hoog" in een verhouding van 1:2 te kiezen.

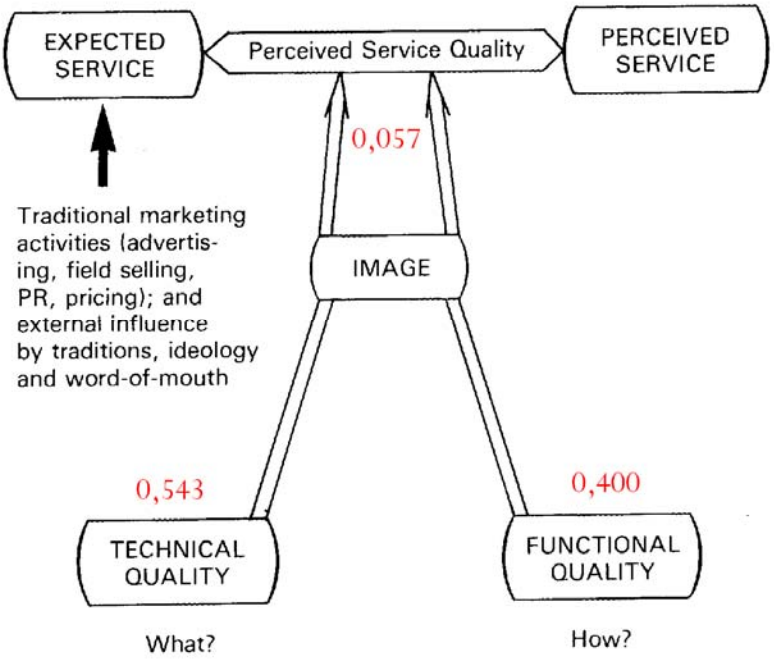
Voor deze verhouding is zowel een relatieve, als een absolute kwaliteitsindex (zie paragraaf 5.1) berekend. Wanneer de resultaten van beide kwaliteitindices worden vergeleken, blijkt dat de relatieve kwaliteitsindex een aanzienlijk hogere standaarddeviatie (35% hoger) heeft dan de absolute kwaliteitsindex. Het onderscheidend vermogen van deze eerste index is dus aanzienlijk groter, en dus, zoals verwacht, geschikter om in DEA in te passen. Voor het model wordt dus gebruik gemaakt van

deze relatieve kwaliteitsindex met de "middel"-"hoog"-verhouding van 1:2. De resulterende kwaliteitsindexscores zijn weergegeven in tabel 3:

Tabel 3: Kwaliteitsindexscores per zkh								
Ziekenhuis (zkh):	1	2	3	4	5	6	7	8
Kwaliteitsindexscore:	0,557	0,752	0,584	0,669	0,676	0,612	0,576	0,812

Om nu een helder beeld te krijgen van het 'soort' kwaliteit dat de kwaliteitsindex bevat, en waarvoor dus gecorrigeerd wordt in DEA (zie paragraaf 5.4.2.), moet gekeken worden naar welke facetten van de kwaliteit van de zorg op de klinische afdelingen door de verschillende indicatoren in de index worden afgedekt, en met welk gewicht deze facetten worden afgedekt.

Een handig kader om dit in beeld te brengen is het servicemodel van Grönroos (1984), zoals dat behandeld is in paragraaf 2.1. Van de indicatoren die zijn opgenomen in de kwaliteitsindex vallen de indicatoren in de pijler "zorginhoud" (zie bijlage C2) onder het door Grönroos gespecificeerde begrip *technical quality*, de technische kwaliteit van de geleverde zorg, behalve misschien het "rapportcijfer patiëntentevredenheid" dat meer direct uitdrukking geeft aan de *percieved quality* van Grönroos (1984), de door de cliënt beleefde kwaliteit van de geleverde zorg. De indicatoren die vallen onder de pijlers "snelheid", "betrouwbaarheid" en "flexibiliteit" (zie bijlage C2) vallen onder Grönroos' (1984) *functional quality*, de manier waarop het zorgtraject door de cliënt wordt beleefd. In figuur 5 is te zien welke facetten van kwaliteit van de zorg met welk gewicht (in rood) meewegen in de uiteindelijke kwaliteitsindex. Te zien is dat de technische kwaliteit het zwaarste mee weegt in de kwaliteitsindex. De gewichten gespecificeerd per indicator zijn terug te vinden in bijlage C3.



Figuur 5: Weging van kwaliteitsaspecten in de kwaliteitsindex

5.4.2 Efficiëntiescores

Door de in- en outputvariabelen te kiezen zoals is weergeven in tabel 2 in paragraaf 5.3, kan het (onaangepaste) CCR-model worden toegepast op de data (dus inclusief kwaliteitsindex, maar zonder VAR-I restricties). De resulterende efficiëntiescores zijn weergegeven in tabel 4, tezamen met de kwaliteitsindexscores van de ziekenhuizen. De kwaliteitsindex krijgt voor iedere DMU een gewicht van 0 of nagenoeg 0 (zoals beschreven in paragraaf 4.1.2). De efficiëntiescores uit tabel 4 zijn dus in feite 'smalle' kostenefficiëntiescores waarin het kwaliteitsaspect van de zorg praktisch wordt verwaarloosd.

Tabel 4: Standaard DEA-efficiëntiescores (zonder VAR-I restricties)

Ziekenhuis (zkh):	1	2	3	4	5	6	7	8
Kwaliteitsindexscore:	0,557	0,752	0,584	0,669	0,676	0,612	0,576	0,812
Efficiëntiescore:	n.b.[3]	1	0,95	0,89	1	0,87	1	1

Door nu de in- en outputvariabelen opnieuw te kiezen zoals af te lezen in tabel 2 in paragraaf 5.3, en het standaard DEA-model uit te bereiden door VAR-restricties (zie paragraaf 4.2.2.2, 5.2 en bijlage B) op te nemen in de specificatie, kunnen er voor de verschillende ziekenhuizen 'brede' efficiëntiescores worden berekend, waarin de kwaliteit van de geleverde zorg expliciet in is meegenomen doordat het gewicht van de kwaliteitsindex gegarandeerd wordt (in tegenstelling tot de resultaten in tabel 4, waarin dit gewicht vrij door het model te bepalen is).

De aan het standaard CCR-DEA-model toegevoegde VAR I-restrictie kan als volgt worden weergegeven:

$$b_1 u_1 \, Aantal\ \ dagbehandelingen + b_2 u_2 \, Aantal\ \ verpleegdagen +$$
$$b_3 u_3 \, Aantal\ \ klinische\ \ opnames + b_4 u_4 \, Kwaliteitsindex \ \ \geq \ \ 0$$

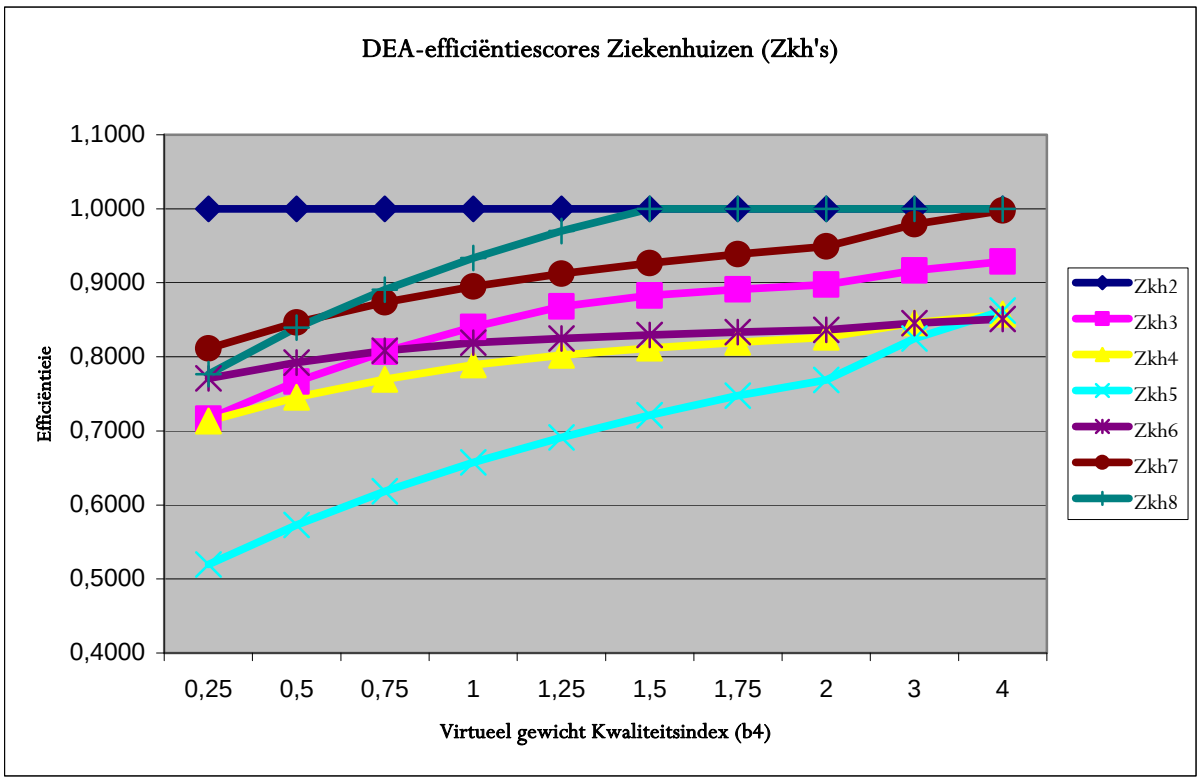
Wanneer b_1,b_2 en b_3 nu gelijk worden gekozen aan 1, kan gemakkelijk de verhouding worden bepaald waarin 'kwaliteit' moet meewegen ten opzichte van de meer fysieke outputs (aantal dagbehandelingen, aantal verpleegdagen en aantal klinische opnames) in de brede efficiëntiescore. De reciproque waarde van parameter b_4 (=1/b_4) geeft dan weer met welke factor 'kwaliteit' moet worden meegenomen ten opzicht van de fysieke outputs. De restrictie wordt dan als volgt:

[3] Voor zkh1 was er op het moment van berekening nog te weinig data beschikbaar om de kwaliteitsindexscore te kunnen berekenen. In het vervolg van de analyse wordt zkh1 dan ook niet meegenomen.

$$b_4 u_4 \, Kwaliteitsindex \;\; \geq \;\; u_1 \, Aantal \;\; dagbehandelingen \; + \; u_2 \, Aantal \;\; verpleegdagen \; + \; u_3 \, Aantal \;\; klinische \;\; opnames$$

Een waarde van de parameter b_4=1/4 geeft bijvoorbeeld aan dat het virtuele gewicht van 'kwaliteit' 4 keer zo groot moet zijn als het virtuele gewicht van de fysieke outputs samen, en een waarde van b_4=3 geeft bijvoorbeeld aan dat het virtuele gewicht van 'kwaliteit' minstens 1/3 keer zo groot moet zijn als het virtuele gewicht van de fysieke outputs samen ('kwaliteit' telt dus voor minstens 1/4 mee in de efficiëntiescore). Wanneer de parameter b_4 ook op 1 wordt gesteld, dan weegt 'kwaliteit' even zwaar in het bepalen van de efficiëntie van een ziekenhuis als de fysieke output samen. Er geldt dus hoe hoger de waarde van b_4 hoe lager de weging van de kwaliteitsindex binnen de DEA-score wordt meegenomen (gegarandeerd).

Voor verschillende gespecificeerde waarden van b_4 zijn de efficiëntiescores van de ziekenhuizen weergegeven in figuur 6 (voor ziekenhuis 1 konden geen efficiëntiescores worden berekend omdat de aangeleverde data niet compleet was). Op de verticale as staan de DEA-efficiëntiescores van de ziekenhuizen uitgezet tegen het reciproque gewicht van kwaliteit ten opzichte van de fysieke outputs (een hogere waarde van b_4 betekent dus dat het kwaliteitsaspect zwaarder meetelt).



Figuur 6: DEA-efficiëntiescores bij verschillende wegingen van de kwaliteitsindex

In figuur 6 valt een aantal zaken op:

- De efficiëntiescores van de ziekenhuizen dalen wanneer de factor 'kwaliteit' zwaarder (b_4 kleiner) wordt meegewogen in de analyse. Er bestaat dus een trade-off tussen de omvang van de productie in enge zin (de fysieke output) en de kwaliteit van de geleverde zorg op de klinische afdelingen. Kwaliteit kost capaciteit (en dus geld).
- Ziekenhuizen kunnen wisselen in de ranking door het veranderen van het gewicht dat aan de kwaliteitsindex wordt toegekend. Een ziekenhuis kan zijn efficiëntiescore verbeteren door een betere kwaliteit van zorg af te leveren.
- Onder de VAR-I restrictie krijgen minder ziekenhuizen een efficiëntiescore van 1 (er zijn minder *peer* DMU's, vergelijk met de efficiëntiescores in tabel 4). Het onderscheidend vermogen van de DEA-techniek neemt hierdoor dus toe.
- Ziekenhuis 2 blijft ook onder de VAR-I restrictie efficiënt (ook de slackvariabelen zijn gelijk aan 0). Ziekenhuis 2 is dus (relatief) efficiënt te noemen in brede zin (bij een kwaliteitsindexscore van 0,752).

Merk verder op dat ziekenhuizen met een hoge kwaliteitsscore niet perse (i.e. bij alle waarden van b_4) beter scoren dan ziekenhuizen met een lagere kwaliteitsindexscore. Ziekenhuis 8 heeft bijvoorbeeld een hogere efficiëntiescore (0,812) dan ziekenhuis 7 (0.576) terwijl de efficiëntiescore van ziekenhuis 8 steiler daalt wanneer 'kwaliteit' zwaarder word meegewogen (kleinere waarde van b_4). Dit komt doordat ook een ziekenhuis wel een hoge kwaliteitsindexscore kan hebben, maar deze kwaliteit inefficiënt 'produceert'. Dit kan bijvoorbeeld het geval zijn wanneer er dusdanig veel personeel per bed beschikbaar is, dat een gedeelte van dit personeel zo nu en dan niets te doen heeft, of simpelweg de tijd niet goed indeelt. Met minder personeel zou dan een gelijke zorgkwaliteit moeten kunnen worden behaald. De ranking (beste ziekenhuis ranking 1, slechtste ranking 8) van de (brede) efficiëntiescores van de verschillende ziekenhuizen convergeert dan ook niet per definitie naar de ranking van kwaliteitsindexscores wanneer de kwaliteitsindex steeds zwaarder meeweegt (b4 => 0) . Dit zou slechts het geval zijn wanneer alle ziekenhuizen precies evenveel personeelkosten zouden hebben (gelijke inputs). Het kan namelijk zo zijn dat een ziekenhuis kwaliteit kostenineefficiënt produceert. Dit is terug te zien in de DEA-scores.

5.4.3 Berekening van het verbeterpotentieel

Met behulp van de formules voor de targets uit bijlage B kan, naast de brede efficiëntiescores, ook het verbeterpotentieel voor de klinische afdelingen van de aan de benchmark deelnemende ziekenhuizen worden berekend. Door de opname van de VAR-I restrictie in het model zijn er minder ziekenhuizen als relatief efficiënt betiteld dan in het standaard DEA-model. Dit betekent dat er voor meer (relatief inefficiënte) ziekenhuizen een verbeterpotentieel is te berekenen,

Voor het geval b_4=1, kwaliteit van de zorg weegt net zo zwaar als de fysieke outputs samen, zijn de targets voor zowel het input- als het outputgeoriënteerde model berekend. Hoewel b_4=1 een behoorlijk zware restrictie is, illustreert deze goed de methode. In tabel 5 is het verbeterpotentieel te vinden wanneer een inputgeoriënteerd model wordt gebruikt, en in tabel 6 wanneer een outputgeoriënteerd model wordt toegepast.

Tabel 5: Targets bij een inputoriëntatie (b_4 =1)

	Salariskosten		# Dagbehandelingen		# Verpleegdagen		# Klinische Opnames		Kwaliteitsindex	
	Target	Huidig	Target	Huidig	Target	Huidig	Target	Huidig	Target	Huidig
Zkh2	€ 7.196.907	€ 7.196.907	9887	9887	50959	50959	10024	10024	0,7517	0,7517
Zkh3	€ 7.210.506	€ 8.579.552	9906	9851	51055	71769	10043	12053	0,7531	0,5844
Zkh4	€ 7.544.848	€ 9.557.692	10365	10928	53423	64973	10509	12206	0,7880	0,6691
Zkh5	€ 9.332.548	€ 14.195.044	12821	14719	66081	118322	12999	21196	0,9747	0,6762
Zkh6	€ 6.476.274	€ 7.908.192	8765	9236	48219	51470	9087	9700	0,6507	0,6120
Zkh7	€ 6.457.124	€ 7.215.009	8871	10709	45721	53972	8994	9359	0,6744	0,5756
Zkh8	€ 10.393.784	€ 11.131.272	14278	13491	73595	111080	14477	16420	1,0855	0,8117

Uit tabel 5 valt af te leiden dat alle ziekenhuizen, op *peer* ziekenhuis 2 na, zouden moeten kunnen besparen op salariskosten (tot 34% voor ziekenhuis 5). Voor sommige ziekenhuizen moeten bepaalde outputs een beetje worden verminderd, terwijl andere outputs iets moeten worden vermeerderd om als efficiënt te worden betiteld (deze eigenschap van de VAR-restrictie is terug te zien in tabel 1 in paragraaf 4.4). Voor niet-efficiënte ziekenhuizen is er verbeterpotentieel op het gebied van de kwaliteit van de zorg. Hierbij moet worden opgemerkt dat de targets op het gebied van kwaliteit niet terug te vertalen zijn naar concrete kwaliteitsaspecten. Een kwaliteitstarget die hoger is dan het uitgangsniveau geeft slechts aan dat er op kwaliteitsgebied winst valt te boeken. De absolute target waarde van deze index heeft geen heldere interpretatie. Ook zijn er targets voor de kwaliteitsindex met een waarde groter dan 1 mogelijk, omdat de normering van de kwaliteitsindex berust op de beste prestaties in de uitgangssituatie die in de targets niet meer van toepassing zijn (zie ook bijlage C1).

De targets in tabel 6, voor het outputgeoriënteerde model, laten zien dat er voor alle ziekenhuizen, behalve het relatief efficiënt opererende ziekenhuis 2, aanzienlijke winst valt te boeken op het gebied van de efficiëntie van de personeelsinzet op de klinische afdelingen, zowel op het gebied van de fysieke outputs als ook zeker op kwaliteitsgebied. Sommige outputtargets liggen lager dan het huidige niveau, dit is een consequentie van de VAR-restrictie (zie tabel 1 in paragraaf 4.4).

Tabel 6: Targets bij een outputoriëntatie (b_4 =1)

	Salariskosten		# Dagbehandelingen		# Verpleegdagen		# Klinische Opnames		Kwaliteitsindex	
	Target	Huidig	Target	Huidig	Target	Huidig	Target	Huidig	Target	Huidig
Zkh2	€ 7.196.907	€ 7.196.907	9887	9887	50959	50959	10024	10024	0,7517	0,7517
Zkh3	€ 8.579.552	€ 8.579.552	11786	9851	60749	71769	11950	12053	0,8961	0,5844
Zkh4	€ 9.557.692	€ 9.557.692	13130	10928	67675	64973	13312	12206	0,9982	0,6691
Zkh5	€ 14.195.044	€ 14.195.044	19501	14719	100511	118322	19771	21196	1,4826	0,6762
Zkh6	€ 7.908.192	€ 7.908.192	10703	9236	58881	51470	11097	9700	0,7946	0,6120
Zkh7	€ 7.215.009	€ 7.215.009	9912	10709	51087	53972	10049	9359	0,7536	0,5756
Zkh8	€ 11.131.272	€ 11.131.272	15292	13491	78817	111080	15504	16420	1,1626	0,8117

Het verschil tussen de resultaten van een input- en outputgeoriënteerd model is, per variabele, te zien in tabel 7. Per variabele zijn er twee kolommen: de linker kolom geeft per in- of outputvariabele steeds de procentuele targetmutaties (het verbeterpotentieel) weer voor een inputoriëntatie, terwijl de rechter kolom deze mutaties geeft voor een outputgeoriënteerd model.

Tabel 7: Procentuele targetmutaties (verbeterpotentieel) t.ov. huidige niveaus (b_4 =1)

	Salariskosten		Dagbehandelingen		Aantal Verpleegdagen		Aantal Klinische Opnames		Kwaliteitsindex	
Oriëntatie:	Input	Output	Input	Output	Input	Output	Input	Output	Input	Output
Zkh2	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Zkh3	-16%	0%	1%	20%	-29%	-15%	-17%	-1%	29%	53%
Zkh4	-21%	0%	-5%	20%	-18%	4%	-14%	9%	18%	49%
Zkh5	-34%	0%	-13%	32%	-44%	-15%	-39%	-7%	44%	119%
Zkh6	-18%	0%	-5%	16%	-6%	14%	-6%	14%	6%	30%
Zkh7	-11%	0%	-17%	-7%	-15%	-5%	-4%	7%	17%	31%
Zkh8	-7%	0%	6%	13%	-34%	-29%	-12%	-6%	34%	43%

Wanneer een ziekenhuis gericht is op kostenbesparingen, dan zijn de inputgeoriënteerde resultaten het meest relevant, waarin de nadruk ligt op het reduceren van personeel. Wanneer een ziekenhuis met het huidige personeelsbestand zijn outputs wil laten stijgen, dan zijn de outputgeoriënteerde resultaten het meest relevant. Alle inefficiënties in personeelsinzet worden dan gebruikt om de output efficiënter te

maken. In tabel 8 zijn dezelfde gegevens weergeven als in tabel 7, maar dan geordend op oriëntatie in plaats van op in- of outputvariabele.

Tabel 8: Procentuele targetmutaties (verbeterpotentieel) t.o.v. huidige niveaus ($b_4 = 1$)

	Inputoriëntatie					Outputoriëntatie				
	Salariskosten	Aantal Dagbehandelingen	Aantal Verpleegdagen	Aantal Klinische Opnames	Kwaliteitsindex	Salariskosten	Aantal Dagbehandelingen	Aantal Verpleegdagen	Aantal Klinische Opnames	Kwaliteitsindex
Zkh2	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Zkh3	-16%	1%	-29%	-17%	29%	0%	20%	-15%	-1%	53%
Zkh4	-21%	-5%	-18%	-14%	18%	0%	20%	4%	9%	49%
Zkh5	-34%	-13%	-44%	-39%	44%	0%	32%	-15%	-7%	119%
Zkh6	-18%	-5%	-6%	-6%	6%	0%	16%	14%	14%	30%
Zkh7	-11%	-17%	-15%	-4%	17%	0%	-7%	-5%	7%	31%
Zkh8	-7%	6%	-34%	-12%	34%	0%	13%	-29%	-6%	43%

6 Conclusies en aanbevelingen

In dit hoofdstuk worden in paragraaf 6.1 de conclusies van deze scriptie op een rijtje gezet en in paragraaf 6.2 worden aanbevelingen gedaan voor verdere toepassing en verfijning van de voorgestelde DEA-modellen in de Benchmark Kliniek.

6.1 Conclusies

Het doel van deze scriptie was het modelleren van een 'kwaliteitsdimensie' in DEA voor de benchmark Kliniek 2008. Door 'kwaliteit' expliciet te modelleren in het DEA-model moest een 'bredere' efficiëntiebeoordeling ontstaan.

Om 'kwaliteit' te integreren in het model was het nodig een eventuele trade-off tussen kostenefficiëntie en kwaliteit van zorg te kwantificeren. Zo kunnen efficiëntiescores van de klinische afdelingen van verschillende ziekenhuizen 'eerlijker' worden vergeleken en zijn er realistischere verbeterpotentiëlen voor de ziekenhuizen te berekenen.

De uitdaging van deze scriptie was daarom tweeledig:

1. Hoe kan kwaliteit van zorg worden gemeten en
2. hoe kan dit op zinvolle wijze worden geïntegreerd in DEA?

Aan de hand van literatuuronderzoek is ervoor gekozen om gebruik te maken van kwaliteitsindicatoren voor verschillende facetten van zorgkwaliteit, die, gecombineerd tot één kwaliteitsindex, de kwaliteit van de zorg tot op zekere hoogte meetbaar maken (zie hoofdstuk 2 en paragraaf 5.1 en 5.4.1). Deze kwaliteitsindex is vervolgens in combinatie met zogenaamde 'virtual assurance regions van type I' (VAR-I restricties) gebruikt om een eventuele trade-off tussen zorgkwaliteit en kostenefficiëntie te modelleren in DEA (zie hoofdstuk 4 en paragraaf 5.2).

De toepassing van deze combinatie van kwaliteitsindex en DEA-model met VAR-I restricties op de data voor de benchmark Kliniek 2008 heeft tot resultaten geleid (uit hoofdstuk 5) waaruit een aantal zaken kan worden geconcludeerd:

- Voor de deelnemende ziekenhuizen bestaat er een zekere trade-off tussen kostenefficiëntie en kwaliteit van zorg in bredere zin;
- Het bestaan van deze trade-off maakt dat de voorgestelde 'brede' efficiëntiescores, waarin deze trade-off expliciet is gemodelleerd, 'eerlijker' verschillen in (brede) efficiëntie blootleggen tussen de deelnemende ziekenhuizen dan wanneer 'standaard' DEA wordt toegepast. Hierdoor wordt de benchmark waardevoller;
- Doordat de kwaliteits-kostenefficiëntie trade-off is meegenomen in het DEA-model kunnen realistischere verbeterpotentiëlen worden berekend dan in 'standaard DEA', Hierdoor kan de ziekenhuizen een realistischer beeld worden voorgespiegeld over de potentiële efficiëntiewinst;
- De gekozen methode (VAR-I) leidt tot relatief eenvoudig interpreteerbare (radiale) efficiëntiescores en (qua mathematische uitdrukking gemakkelijke) verbeterpotentiëlen. Dit maakt het uitleggen van de gebruikte methode gemakkelijker in de, over het algemeen, minder kwantitatief onderlegde medische wereld. Daarentegen geldt voor de verbeterpotentiëlen niet meer - zoals in 'standaard DEA' wel geldt – dat een outputoriëntatie ertoe leidt dat alle outputtargets gelijk blijven of toenemen, terwijl de inputniveaus gelijk blijven, of, in het geval van een inputoriëntatie, dat alle inputtargets gelijk blijven of afnemen, terwijl de outputniveaus gelijk blijven (al blijft het doel wél gelijk per oriëntatie). Dit maakt de verbeterpotentiëlen in dit opzicht weer minder gemakkelijk uit te leggen dan in 'standaard DEA' het geval is;
- Bijkomend voordeel van de gebruikte methode is dat het onderscheidend vermogen van DEA toeneemt, zodat minder ziekenhuizen als relatief efficiënt worden betiteld, en het voor een groter aantal ziekenhuizen mogelijk is een verbeterpotentieel te bepalen dan bij het gebruik van een standaard DEA-model. Hierdoor wordt de benchmark voor een aantal ziekenhuizen nog waardevoller.

6.2 Aanbevelingen

Om de kracht en de kwaliteit van de benchmark verder te vergroten, is het interessant om voor ieder specialisme een speciaal op dit specialisme toegespitste kwaliteitsindex te berekenen. Hiermee kan namelijk per specialisme (aangepaste) DEA worden toegepast, waardoor de kwaliteitsindex beter de specifieke kwaliteit van de zorg binnen een bepaald specialisme weerspiegelt, en waardoor er nu *per*

specialisme een verbeterpotentieel kan worden berekend zodat de DEA-resultaten op een lager aggregatieniveau komen en inefficiënties nauwkeuriger kunnen worden benoemd.

Hiernaast kan er, door alle data van de ziekenhuizen die relatief efficiënt opereren op een bepaald specialisme samen te voegen, een virtueel 'super ziekenhuis' worden gecreëerd, dat op alle specialismen relatief efficiënt opereert. Dit super ziekenhuis kan zo als *peer* ziekenhuis dienst doen om op ziekenhuisniveau voor alle ziekenhuizen een efficiëntiescore kleiner dan 1 te creëren en zo voor al deze ziekenhuizen een verbeterpotentieel te bepalen (voorwaarde hiervoor is dat er géén een ziekenhuis op alle specialismen een efficiëntiescore van 1 scoort). Ook hierdoor wordt de benchmark voor een groter aantal ziekenhuizen aantrekkelijk, omdat voor ieder ziekenhuis nu een perspectief op verbetering kan worden geboden.

Een suggestie voor verder onderzoek is het toetsen van de juistheid van de aanname van constante schaalopbrengsten op de klinische afdelingen van ziekenhuizen. Ook kan dan worden onderzocht hoe de resultaten van dit onderzoek veranderen wanneer van variabele schaalopbrengsten (VRS) en het BBC-model wordt uitgegaan.

7 Bibliografie

Allen, R. et al. (1997). Weights restrictions and value judgements in Data Envelopment Analysis: Evolution, development and future directions. *Annals of Operations Research*, Vol. 73, pp. 13-43.

Banker, R.D. et al. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, No. 30, pp. 1078-1092.

Bowling, W.F. (1998). Measuring Performance: An Introduction to Data Envelopment Analysis (DEA). *Journal of Cost Analysis*,.

Charnes, A. et al. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, No. 2, pp. 429-444.

Charnes, A. et al. (1990). Polyhedral cone-ratio DEA models with an illustrative application to large commercial banks. *Journal of Econometrics*, Vol. 46, pp. 73-91.

Cook, W.D. en L.M. Seiford (2008). Data envelopment analysis (DEA) – Thirty years on. *European Journal of Operational Research*, Vol. 180, pp. 443-458.

Cooper, W.W. et al. (2000). *Data Envelopment Analysis, A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Dordrecht: Kluwer Academic Group.

Cooper, W.W. et al. (2007). Choosing weights from alternative optimal solutions of dual multiplier models in DEA. *European Journal of Operational Research*.

Deming, E.W. (1986). *Out of the crisis: quality, productivity and competitive position*. Cambridge: Cambridge University Press.

Estellita Lins, M.P. et al. (2007). Avoiding infeasibility in DEA models with weight restrictions. *European Journal of Operational Research*, Vol. 181, pp. 956-966.

Grönroos, C. (1984). A service quality model and its marketing implications. *European Journal of Marketing*, No. 4.

Lehtinen, U. en J.R. Lehtinen (1991). Two approaches to service quality dimensions. *The service Industries Journal*, Vol. 11, No. 3, pp. 287-303.

Li, X. en G.R. Reeves (1997). A multiple criteria approach to data envelopment analysis. *European Journal of Operational Research*, Vol. 115, pp. 507-517.

Morey, C. et al (1992). The trade-off between hospital cost and quality of care: an exploratory empirical analysis. *Medical Care*, Vol. 30, No. 8, pp. 677-698.

Nayar, P. en Y.A. Ozcan (2008). Data envelopment analysis comparison of hospital efficiency and quality. *Journal of Medical Systems*, Vol. 32, No. 3, pp. 193-199.

O'Neill, L (1998). Multifactor efficiency in Data Envelopment Analysis with an application to urban hospitals. *Health Care Management Science*, Vol. 1, pp. 19-27.

Roll, Y. en B. Golany (1993). Alternate methods of treating factor weights in DEA. *Omega, International Journal of Management Science*, Vol. 21, No. 1, pp. 99-109.

Sahin, I en Y.A. Ozcan (2000). Public sector efficiency for provincial markets in Turkey. *Journal of Medical Systems*, Vol. 24, No. 6.

Sarrico, C.S. en R.G. Dyson (2004). Restricting virtual weights in data envelopment analysis. *European Journal of Operational Research*, Vol. 159, No.1, pp. 17-34.

Sherman, D.H. en J. Zhu (2006a). *Service productivity management: Improving service performance using data envelopment analysis (DEA)*. New York: Springer.

Sherman, D.H. en J. Zhu (2006b). Benchmarking with quality-adjusted DEA (Q-DEA) to seek lower-cost high-quality service: Evidence from a U.S.bank application. *Annals of Operations Research*, Vol. 145, No. 1, pp. 301-319.

Swan, J.E. en L.J. Combs (1976). Product performance and consumer satisfaction: a new concept. *Journal of Marketing*, april 1976, pp. 26.

Thanassoulis, E. et al. (1995). Exploring output quality targets in de provision of perinatal care in England using data envelopment analysis. *European Journal of Operational Research*, Vol. 80, pp-588-607.

Thompson, R.G. et al. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics*, Vol. 46, pp. 93-108.

Wong, Y.H.B. en J.E. Baesley (1990). Restricting weight flexibility in Data Envelopment Analysis. *The Journal of the Operational Research Society*, Vol. 41, No. 9, pp. 829-835.

Inspectie voor de Gezondheidszorg (2008). Prestatie-indicatoren ziekenhuizen – Basisset 2008 (http://www.snellerbeter.nl/fileadmin/snellerbeter/afbeeldingen/Prestatie-indicatoren/Indicatoren_2007/Basisset_prestatie-indicatoren_ziekenhuizen_2008.pdf), 22 maart.

Thuiszorg (2008). Kwaliteitskader verantwoorde zorg 2007 (http://www.thuiszorg.nl/file?id=107584), 15 maart.

Wikipedia (2008a). Diagnosis-related Group (http://en.wikipedia.org/wiki/Diagnosis-related_group), 4 mei.

Wikipedia (2008b). Case mix index (http://en.wikipedia.org/wiki/Case_mix_index), 4 mei.

8 Lijst met afkortingen

AR I: Assurance Regions van type I

AR II: Assurance Regions van type II

BBC: DEA model van Banker Charnes en Cooper

CCR: DEA model van Charnes Cooper en Rhodes

CFA: Constrained Facet Analysis

CRS: Constant Returns to Scale

DEA: Data Envelopment Analysis

DMU: Decision Making Unit

FTE: FullTime-Equivalents

HP: High Productivity

HQ: High Quality

LP: Low Productivity

LP-model: Lineair Programmeringsmodel

LQ: Low Quality

PPS: Production Possibility Set

Q-DEA: Quality-adjusted DEA

VAR I: Virtual Assurance Regions van type I

VAR II: Virtual Assurance Regions van type II

VRS: Variable Returns to Scale

ZKH: Ziekenhuis

9 Bijlage

Bijlage A : Het outputgeoriënteerde CCR-model

In deze bijlage wordt het outputgeoriënteerde CCR-model gegeven met de bijbehorende sterke dualiteitstelling en targets, zoals deze in paragraaf 3.2 voor het inputgeoriënteerde model is gepresenteerd.

Het outputgeoriënteerde CCR-model (CCR-II-output- en CCR-II-D-output-model) kan worden verkregen uit het duale inputgeoriënteerde CCR-model door de substituties $h_{j_0} = 1/g_{j_o}$ en $\lambda = \mu / g_{j_0}$ te maken:

CCR-II-output

$$1/e_{j_0} = g_{j_0} = \min_{u_r, v_i} \sum_{r=1}^{s} v_i x_{ij_0}$$

onder de voorwaarden

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \quad j = 1,\ldots,n$$

$$\sum_{i=1}^{m} u_r y_{rj_0} = 1$$

$$u_r, v_i \geq 0 \qquad r = 1,\ldots,s \qquad i = 1,\ldots,m$$

CCR-II-D-output

$$1/e_{j_o} = b_{j_0} = \max_{\mu_j, s_i^-, s_r^+, \eta} \eta$$

onder de voorwaarden:

$$-\eta y_{rj_0} + \sum_{j=1}^{n} \mu_j y_{rj} - s_r^+ = 0 \qquad r = 1,\ldots,s$$

$$\sum_{j=1}^{n} \mu_j x_{ij} + s_i^- = x_{ij_0} \qquad i = 1,\ldots,m$$

$$\mu_j, s_i^-, s_r^+ \geq 0 \quad \eta \in \mathbb{R} \quad i = 1,\ldots,m \quad r = 1,\ldots s \quad j = 1,\ldots,n$$

Het bijbehorende 'intuïtieve' outputgeoriënteerde model is het onderstaande CCR-I-output model:

CCR-I

$$1/e_{j_0} = g_{j_0} = \max_{u_r, v_i} \frac{\sum_{i=1}^{m} v_i x_{ij_0}}{\sum_{r=1}^{s} u_r y_{rj_0}}$$

onder de voorwaarden:

$$\frac{\sum_{i=1}^{m} v_i x_{ij_0}}{\sum_{r=1}^{s} u_r y_{rj_0}} \leq 1 \; ; \quad j = 1,\ldots,n$$

$$u_r, v_i \geq 0 \qquad r = 1,\ldots,s \qquad i = 1,\ldots,m$$

In een eindig optimum ($e_{j0}^*, g_{j0}^*, u_r^*, v_i^*, b_{j0}^*, \eta^*, \mu_j, s_i^{-*}, s_r^{+*}$) geldt nu volgens de sterke dualiteitstelling (ook in het geval van een inputoriëntatie geldt de sterke dualiteitstelling):

$$1/e_{j_0}^* = g_{j_0}^* = \sum_{r=1}^{s} u_r^* y_{rj_0} = b_{j_0}^* = \eta^*$$

En voor de targets, $\left(\sum_{j=1}^{n} x_{ij}\lambda_j^*, \sum_{j=1}^{n} y_{rj}\lambda_j^* \right)$, geldt nu:

$$x_{ij_0} - s_i^{-*} = \sum_{j=1}^{n} \mu_j^* x_{ij} \qquad i = 1, \ldots, m$$

$$g_{j_0}^* y_{rj_0} + s_r^{+*} = \sum_{j=1}^{n} \mu_j^* y_{rj} \qquad r = 1, \ldots, s$$

Bijlage B : Efficiëntiescores en targets onder virtuele restricties in het CCR model

Saricco en Dyson (2004) geven een outputgeoriënteerd DEA-model en het bijbehorende duale model waarin virtuele restricties zijn verwerkt:

CCR-II-output met virtuele restricties

$$1/e_{j_0} = g_{j_0} = \min_{u_r, v_i} \sum_{r=1}^{s} v_i x_{ij_0}$$

onder de voorwaarden

$$\begin{aligned} &\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} && \leq 0 \quad \forall j \\ &\sum_{i=1}^{m} u_r y_{rj_0} && = 1 \\ &\sum_{i} a_{ij} v_i x_{ij_0} + \sum_{r} b_{rj} u_r y_{rj_0} && \geq k_w \quad \forall w \end{aligned}$$

$$u_r, v_i \geq 0 \quad \forall r, \quad \forall i$$

CCR-II-D-output met virtuele restricties

$$1/e_{j_0} = b_{j_0} = \max_{\mu_j, \rho_{w_0} s_i^-, s_r^+, \eta} \eta + \sum_{w=1}^{t} k_w \rho_{w_0}$$

onder de voorwaarden:

$$\begin{aligned} &\eta y_{rj_0} - \sum_{j=1}^{n} \mu_j y_{rj} + \sum_{w=1}^{t} b_{rw} \rho_{w_0} r_{r_{j0}} + s_r^+ && = 0 && \forall r \\ &\sum_{j=1}^{n} \mu_j x_{ij} + \sum_{w=1}^{t} a_{iw} \rho_{w_0} x_{i_{j0}} + s_i^- && = x_{ij_0} && \forall i \end{aligned}$$

$$\mu_j, \rho_{w_0}, s_i^-, s_r^+ \geq 0 \quad \eta \in \mathbb{R}, \quad \forall j, \quad \forall w, \quad \forall i, \quad \forall r$$

Wanneer $k_w = 0$ ∀w dan reduceert het model tot een model met alleen VAR I en VAR II restricties (zie paragraaf 4.2.2.2).

Voor de targets volgt nu:

$$\sum_{j=1}^{n} \mu_j^* x_{ij} = x_{ij_0} (1 - \sum_{w=1}^{t} a_{iw} \rho_{w_0}^*) - s_i^{-*} \qquad \forall i$$

$$\sum_{j=1}^{n} \mu_j^* y_{rj} = y_{rj_0} (g_{j_0}^* + \sum_{w=1}^{t} b_{rw} \rho_{w_0}^*) + s_r^{+*} \qquad \forall r$$

Het inputgeoriënteerde CCR-model met virtuele restricties kan nu als volgt worden afgeleid:

CCR-II-input met virtuele restricties

$$e_{j_0} = h_{j_0} = \max_{u_r, v_i} \sum_{r=1}^{s} u_r y_{rj_0}$$

onder de voorwaarden

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \quad \forall j$$

$$\sum_{i=1}^{m} v_i x_{ij_0} = 1$$

$$\sum_{i} a_{ij} v_i x_{ij_0} + \sum_{r} b_{rj} u_r y_{rj_0} \geq k_w \quad \forall w$$

$$u_r, v_i \geq 0 \qquad \forall r, \ \forall i$$

CCR-II-D-input met virtuele restricties

$$e_{j_0} = z_{j_0} = \min_{\lambda_j, s_i^-, s_r^+, \rho_{w_0}, \theta} \theta + \sum_{w=1}^{t} k_w \rho_{w_0}$$

onder de voorwaarden:

$$\theta x_{ij_0} - \sum_{j=1}^{n} \lambda_j x_{ij} + \sum_{w=1}^{t} a_{iw} \rho_{w_0} x_{i_{j0}} - s_i^- = 0 \qquad \forall i$$

$$\sum_{j=1}^{n} y_{rj} \lambda_j + \sum_{w=1}^{t} b_{iw} \rho_{w_0} y_{r_{j0}} - s_r^+ = y_{rj_0} \qquad \forall r$$

$$\lambda_j, s_i^-, s_r^+ \geq 0 \quad \rho_{w_0} \leq 0, \ \theta \in \mathbb{R}, \ \forall j, \ \forall w, \ \forall i, \ \forall r$$

Wanneer $k_w = 0$ ∀w dan reduceert het model weer tot een model met alleen VAR I en VAR II restricties (zie paragraaf 4.2.2.2).

Voor de targets volgt nu:

$$\sum_{j=1}^{n} \mu_j^* x_{ij} = x_{ij_0} (h_{j_0}^* - \sum_{w=1}^{t} a_{iw} \rho_{w_0}^*) - s_i^{-*} \qquad \forall i$$

$$\sum_{j=1}^{n} \mu_j^* y_{rj} = y_{rj_0} (1 + \sum_{w=1}^{t} b_{rw} \rho_{w_0}^*) + s_r^{+*} \qquad \forall r$$

Bijlage C: Prestatie-indicatoren klinische afdelingen en de kwaliteitsindex

In bijlage C1 zijn de prestatie-indicatoren op het gebied van kwaliteit op de klinische afdelingen van de deelnemende ziekenhuizen te vinden, en in bijlage C2 de definitie van de kwaliteitsindex, die wordt berekend op basis van deze prestatie-indicatoren. In bijlage C3 zijn de gewichten van de indicatoren uit de kwaliteitsindex zoals die is toegepast in het model (paragraaf 5.4.1) terug te vinden. De bijlagen zijn afgedrukt op de volgende pagina's.

C1: Definitie van de kwaliteitsindex

z :ziekenhuis van type z
i : pijler van type i
j :indicator van type j

$T = \{'ja','nee'\}$	:verzameling bestaande uit *ja*/*nee*-ruimte
C^p	: verzameling van alle 'positieve'[1] indicatoren
C^n	: verzameling van alle 'negatieve'[2] indicatoren
B	:verzameling van *ja*/*nee* indicatoren
Z	:verzameling van alle ziekenhuizen
I	: verzameling van alle pijlers
J_i	: verzameling van alle genormeerde indicatoren binnen pijler i
$J = \cup J_i \;\; \forall\, i \in I$	:verzameling van alle genormeerde indicatoren
M	: verzameling van alle indicatoren met gewicht 'middel'
H	: verzameling van alle indicatoren met gewicht 'hoog'
$V = C^p \cup C^n \cup B$ $= M \cup H$	:verzameling van alle indicatoren

$k_{ijz} \in \{R^+ \cup T\}$: waarde (bruikbare) indicator j van ziekenhuis z binnen pijler i $\forall\, j \in J \;\; \forall\, i \in I \;\; \forall\, z \in Z$

$w_{ijz}(k_{ijz}) \in [0,1]$: waarde genormeerde indicator in *ja* / *nee* ruimte (T)

$$= \begin{cases} w_{ijz}^{ja} \in [0,1] & \text{als } k_{ijz} = 'ja' \qquad \forall (i,j,z) \in \{\{I,J,Z\} \mid k_{ijz} \in B\} \\ w_{ijz}^{nee} \in [0,1] & \text{als } k_{ijz} = 'nee' \qquad \forall (i,j,z) \in \{\{I,J,Z\} \mid k_{ijz} \in B\} \end{cases}$$

$a_{ijk}(k_{ijk}) \in [0,1]$: waarde genormeerde indicator j van ziekenhuis z binnen pijler i $\forall\, j \in J \;\; \forall\, i \in I \;\; \forall\, z \in Z$

$$= \begin{cases} \dfrac{k_{ijz} - \min\limits_z k_{ijz}}{\max\limits_z k_{ijz} - \min\limits_z k_{ij} z} & \forall\, k_{ijz} \in C^p \\ 1 - \dfrac{k_{ijz} - \min\limits_z k_{ijz}}{\max\limits_z k_{ijz} - \min\limits_z k_{ijz}} & \forall\, k_{ijz} \in C^n \\ w_{ijz}(k_{ijz}) & \forall\, k_{ijz} \in B \end{cases}$$

$g_i \in [0,1]$:gewicht pijler i $\forall i \in I$

$g_{ij} \in R^+$:gewicht indicator j binnen pijler i $\forall j \in J \;\; \forall i \in I$

$$= \begin{cases} g^{middel} & \text{als indicator } j \in M \qquad \forall j \in J \\ g^{hoog} & \text{als indicator j} \in H \qquad \forall j \in J \end{cases}$$

[1] Een indicator is 'positief' wanneer geldt hoe hoger de indicatorwaarde hoe beter de prestatie van dit ziekenhuis

[2] Een indicator is 'negatief' wanneer geldt hoe hoger de indicatorwaarde hoe slechter de prestatie van dit ziekenhuis

C2: Prestatie indicatoren op het gebied van kwaliteit op de klinische afdelingen

Noot: de notatie in de tabel verwijst naar bijlage C1

Lijst van indicatoren (k_{ij})					
nr:	**Indicator type:**	**Indicator (van type *j*):**	**Gewicht indicator (g_{ij}):**	**Pijler (van type *i*):**	**Gewicht pijler *i* (g_i):**
1	'Positieve' indicatoren (C^p):	Gemiddeld rapportcijfer patiententevredenheid klinische patiënten	hoog	Zorginhoud	0.60
2	'Negatieve' indicatoren (C^n):	Met het aantal klinische opnames per afdeling gewogen gemiddelde klinische ligduur	hoog	Snelheid	0.20
3		Met het aantal dagbehandelingen per afdeling gewogen gemiddelde wachttijd dagbehandeling	middel	Snelheid	0,20
4		Met het aantal klinische opnames per afdeling gewogen gemiddelde wachttijd klinische opname	middel	Snelheid	0.20
5		Absoluut aantal klachten, ingediend bij klachtencommissie / Aantal gewogen patiënteenheden	middel	Zorginhoud	0,60
6		Aantal valincidenten / Aantal gewogen patiënteenheden	hoog	Zorginhoud	0,60
7		Aantal medicatiefouten / Aantal gewogen patiënteenheden	hoog	Zorginhoud	0,60
8		Aantal overige fouten / Aantal gewogen patiënteenheden	hoog	Zorginhoud	0,60
9		Aantal patiënten met decubitus op een vast tijdstip / aantal op decubitus onderzochte patiënten (Decubitus prevalentie)	middel	Zorginhoud	0,60
10		Decubitus incidentie: aantal nieuwe gevallen bij totale heupvervanging / Aantal totale heupvervangingen	hoog	Zorginhoud	0,60
11		Aantal patiënten met postoperatieve pijnscore>7 / Aantal patiënten waarbij post operatieve pijnscore is gemeten	middel	Zorginhoud	0,60
12		Aantal door het ziekenhuis afgezegde operaties binnen 24 uur / Totaal aantal electieve operaties	hoog	Betrouwbaarheid	0,10
13		Absoluut verschil tussen de Realisatie klinische opnames en de Interne productieafspraak klinische opnames	middel	Betrouwbaarheid	0,10
14		Absoluut verschil tussen de Realisatie klinische opnames en de Interne productieafspraak klinische opnames	middel	Betrouwbaarheid	0,10
15		Absoluut verschil tussen de Realisatie klinische opnames en de Interne productieafspraak klinische opnames	middel	Betrouwbaarheid	0,10
16	Ja/nee indicatoren (*B*)	Actieve delirium preventie	hoog	Zorginhoud	0,60
17		Actieve valpreventie	hoog	Zorginhoud	0,60
18		Actieve screening ondervoeding	hoog	Zorginhoud	0,60
19		Actieve pijnmeting	hoog	Zorginhoud	0,60
20		Uitwisselbaarheid van bedden	middel	Flexibiliteit	0,10

Bijlage C3: Gewichten van indicatoren binnen de kwaliteitsindex

Gewicht van indicatoren in de kwaliteitsindex		
nr:	**Indicator:**	**Gewicht indicator in de kwaliteitsindex**
1	Gemiddeld rapportcijfer patiententevredenheid klinische patiënten	0,0571
2	Met het aantal klinische opnames per afdeling gewogen gemiddelde klinische ligduur	0,1000
3	Met het aantal dagbehandelingen per afdeling gewogen gemiddelde wachttijd dagbehandeling	0,0500
4	Met het aantal klinische opnames per afdeling gewogen gemiddelde wachttijd klinische opname	0,0500
5	Absoluut aantal klachten, ingediend bij klachtencommissie / Aantal gewogen patiënteenheden	0,0286
6	Aantal valincidenten / Aantal gewogen patiënteenheden	0,0571
7	Aantal medicatiefouten / Aantal gewogen patiënteenheden	0,0571
8	Aantal overige fouten / Aantal gewogen patiënteenheden	0,0571
9	Aantal patiënten met decubitus op een vast tijdstip / aantal op decubitus onderzochte patiënten (Decubitus prevalentie)	0,0286
10	Decubitus incidentie: aantal nieuwe gevallen bij totale heupvervanging / Aantal totale heupvervangingen	0,0571
11	Aantal patiënten met postoperatieve pijnscore>7 / Aantal patiënten waarbij post operatieve pijnscore is gemeten	0,0286
12	Aantal door het ziekenhuis afgezegde operaties binnen 24 uur / Totaal aantal electieve operaties	0,0400
13	Absoluut verschil tussen de Realisatie klinische opnames en de Interne productieafspraak klinische opnames	0,0200
14	Absoluut verschil tussen de Realisatie klinische opnames en de Interne productieafspraak klinische opnames	0,0200
15	Absoluut verschil tussen de Realisatie klinische opnames en de Interne productieafspraak klinische opnames	0,02
16	Actieve delirium preventie	0,0571
17	Actieve valpreventie	0,0571
18	Actieve screening ondervoeding	0,0571
19	Actieve pijnmeting	0,0571
20	Uitwisselbaarheid van bedden	0,1000