

Sturen op patiëntwaarde in de zorgorganisatie: het verbinden van uitkomsten én kosten

Linnean Initiatief Sturen op Uitkomsten én Kosten

Het Linnean Initiatief: een landelijk en multidisciplinair netwerk dat verbindt, inspireert en adviseert over de implementatie van waardegedreven zorg.

Kerngroep

Angelique Weel-Koenders ^a

Matthijs van der Linde ^b

Mariska Hackert ^c

Esmée van der Poort ^d

Jean-Bart Bügel ^e

^a Reumatoloog, Maasstad Ziekenhuis; hoogleraar bewijsgestuurde waardecreatie in de patiëntenzorg, Erasmus Universiteit

^b (Senior) adviseur Linnean Initiatief, Zorginstituut Nederland

^c Postdoc waardegedreven zorg, Maasstad Ziekenhuis; Erasmus Universiteit

^d Promovenda VBHC en gezondheidseconomie, LUMC

^e Manager Finance & Control, Franciscus Gasthuis

Werkgroep

Wilbert van den Hout ^f

Lea Dijkman ^g

Jaco van der Kooij ^h

Gijs van Steenberghe ⁱ

Sjoerd Repping ^j

Welmoed van Deen ^k

Edwin de Beurs ^l

Paul Ulrich ^m

Yoeri Bemelmans ⁿ

Geert van Hoof ^o

Marcel Lagendijk ^p

^f Gezondheidseconoom, LUMC

^g Senior adviseur en onderzoeker WGZ, St. Antonius Ziekenhuis

^h Projectleider VBHC en business controller, Maasstad Ziekenhuis

ⁱ PhD uitkomsten en kosten in de hartzorg en ANIOS cardiologie, Catharina Ziekenhuis

^j Hoogleraar Zinnige Zorg, UvA; voorzitter, ZE&GG

^k Assistent professor HTA, Erasmus Universiteit

^l Hoogleraar klinische psychologie, Universiteit Leiden; senior onderzoeker, Arkin

^m Patiëntvertegenwoordiger en belangenbehartiger GGZ

ⁿ Beleidsadviseur Kwaliteit en Veiligheid, Zuyderland MC.

^o Adviserend Geneeskundige MSZ

^p Product owner waardegedreven zorg, Erasmus MC

Auteurslijst en -volgorde whitepaper worden nader bepaald. Deelname is op persoonlijke titel.

Inhoud

Introductie	3
Doel van dit Linnean whitepaper	6
Praktijkvoorbeelden voor het sturen op uitkomsten én kosten	7
Actiepunten	8
<i>Bijlagen</i>	
I. Kennisvragen voor het meten van waarde	11
II. Overzicht praktijkvoorbeelden voor het sturen op uitkomsten én kosten	16
III. Quality-Cost-Indicator (QCI)-dashboard (Franciscus Gasthuis, Nederland)	18
IV. CVA dashboard (Nordic Healthcare Group, Estland)	20
V. SELFIE (internationaal, coördinator: Erasmus Universiteit, Nederland)	22
VI. Ischemische beroerte dashboard (ziekenhuisgroep, Brazilië)	24
VII. Zorgpaden dashboards (Maasstad Ziekenhuis, Nederland)	25
VIII. Spider chart aortaklepvervangng (Catharina Ziekenhuis, Nederland)	28
IX. Patient-level value analysis (Verenigde Staten)	30
X. GGZ value of efficiency indicator (Nederland)	32
XI. Samen Beter Geboortezorg (Santeon ziekenhuizen, Nederland)	33
<i>Appendix</i> Toelichting begrippen, meetinstrumenten en wetenschappelijke methoden	35
Referenties	43

De beschrijving van de praktijkvoorbeelden is gebaseerd op onze interpretatie van de literatuur. De uiteenzetting van elk praktijkvoorbeeld zal voor publicatie ter goedkeuring worden voorgelegd aan de ontwikkelaars. Wij vragen u vertrouwelijk met de informatie om te gaan.

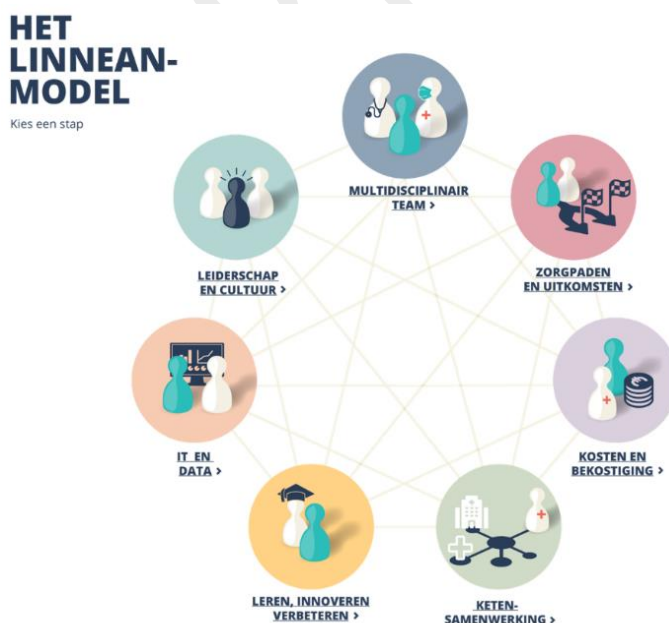
Introductie

Om de kwaliteit, toegankelijkheid en betaalbaarheid van de zorg te waarborgen moet de transitie naar passende zorg worden ingezet (Integraal Zorgakkoord, 2022). Passende zorg is waardegedreven: zorg moet aantoonbaar waarde toevoegen voor de patiënt door zorguitkomsten te optimaliseren passend bij de patiënt zijn of haar doelen. Gezien het eindig budget hebben zorginstellingen conform de Wet Kwaliteit Klachten en Geschillen Zorg (WKKGZ, z.d.) hierbij de taak rekening te houden met zowel de effectiviteit als de kosten van zorg. Porter (2010) definieert waarde om deze reden als een ratio: uitkomsten gedeeld door kosten, ofwel de waardegedreven zorg (WGZ)-breuk (zie Box 1). Hierbij geldt: wie een optimale balans tussen uitkomsten én kosten weet te realiseren levert de hoogste waarde.

Box 1. De WGZ-breuk (Porter, 2010).

$$\text{waarde voor de patiënt} = \frac{\text{uitkomsten}}{\text{kosten}}$$

Het Linnean Initiatief beoogt in dit whitepaper zorgverleners, managers, bestuurders en andere belanghebbenden te stimuleren om te sturen op patiëntwaarde in de medisch-specialistische zorg. We bieden handvatten om aan de slag te gaan met het verbinden van uitkomsten én kosten, zoals in de WGZ-breuk. De whitepaper borduurt voort op eerdere kennis en ervaringen, waaronder de ‘cirkel van gepast gebruik’ geformuleerd in het programma Zorgevaluatie en Gepast Gebruik (ZE&GG, z.d.), en de Linnean werkgroep Versnelling (Hanselaar et al., 2023), waar reeds een handreiking is ontwikkeld voor iedereen die wil starten, of reeds is gestart, met WGZ (zie Figuur 1).



Figuur 1. Linnean-model ‘aan de slag met waardegedreven zorg’ (Hanselaar et al., 2023).

Om uitkomsten én kosten aan elkaar te kunnen verbinden lag de primaire focus, expliciet in Nederland, de afgelopen jaren op het vaststellen en meten van relevante klinische en patiënt gerapporteerde uitkomstmaten (CROMs en PROMs). Deze dragen bij aan het leren en verbeteren in de zorginstelling, het benchmarken tussen zorginstellingen en het samen beslissen in de spreekkamer. Dashboarding wordt voor al deze doeleinden als instrument ingezet, idealiter middels het 7-stappenplan ontwikkeld door de Linnean werkgroep Datavisualisatie / Dashboard (z.d).

Voor het meten van bovengenoemde uitkomsten zijn binnen het programma Uitkomstgerichte Zorg (z.d.) voor 59 aandoeningen, die 50% van de ziektelast in Nederland bedragen, uitkomstensets ontwikkeld waarbij is ingezet op landelijke implementatie. Uitkomstensets ontwikkeld door het Internationaal Consortium for Health Outcomes Measurement (ICHOM, z.d.) vormden daarvoor de basis. De vastgestelde uitkomstensets worden op geaggregeerd niveau gebruikt in instellingen voor het zorginhoudelijk monitoren van veranderingen over tijd, het vergelijken van behandelopties, en het intern en extern benchmarken (Santeon Samen Beter, z.d., SAZ waardegedreven zorg programma, z.d.). Daarnaast worden uitkomstensets gebruikt voor zorgverleners om individuele uitkomsten met patiënten te bespreken, o.a. door middel van samen beslissen (Experiment Uitkomstindicatoren, 2023). Zowel op geaggregeerd als op patiënt niveau is het doel om uitkomsten over de tijd te verbeteren.

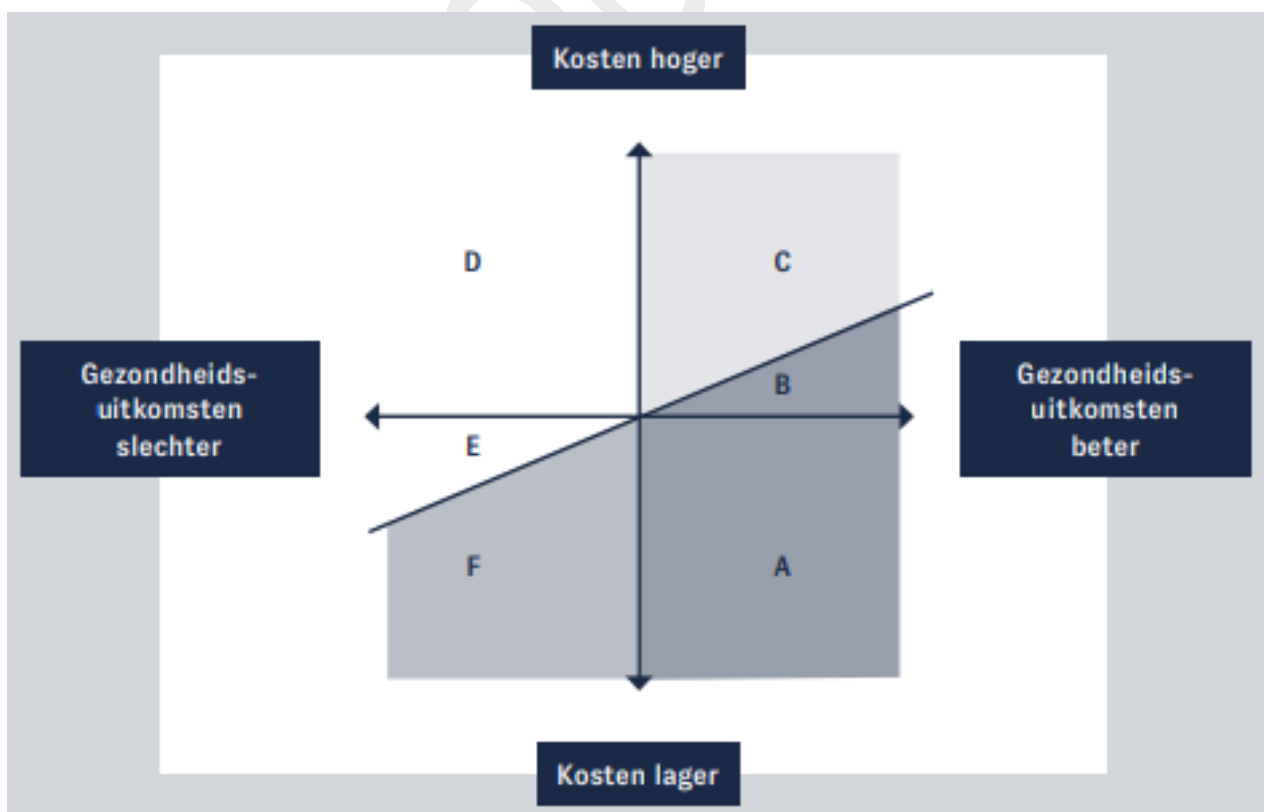
Voor het geven van een impuls aan het meten van kosten over de tijd is de Linnean werkgroep Kosten (Van den Hout et al., 2021) in het leven geroepen. Hun advies is om initieel in de Nederlandse setting de Activity-Based Costing (ABC)-methodiek te gebruiken waarmee reeds veel mogelijk is. Middels de wettelijke verplichting van de Nederlandse Zorgautoriteit (NZa) om kostprijzen te berekenen, bereiken veel ziekenhuizen, zeker met Logex en Performation, al een benaderde ABC kostprijs. Voor gedetailleerder inzicht in kosten over het zorgpad is de Time-Driven Activity-Based Costing (TDABC)-methodiek ontwikkeld (Kaplan en Anderson, 2003). (Inter)nationaal zijn hiermee eerste ervaringen opgedaan (Foster et al., 2023; Keel et al., 2017).

Tot op heden zijn er beperkt initiatieven die ingaan op het weergeven van, en sturen op, de WGZ-breuk: de voor de patiënt relevante uitkomsten in relatie tot de gemaakte kosten. Het kunnen weergeven van de WGZ-breuk kan door middel van evaluaties van zorg(innovaties) en benchmarking bijdragen aan inzicht in de waardetoevoeging van zorg. Een indicator voor waarde biedt inzicht in de kwaliteit en efficiëntie van behandelingen. Wanneer de balans tussen uitkomsten én kosten inzichtelijk kan worden gemaakt, kan als vervolgstap een incrementele kosteneffectiviteitsratio (ICER) worden berekend om de verhouding van het verschil in kosten en het verschil in uitkomsten tussen een interventie en de huidige zorgpraktijk weer te geven. Bijgevoegd raamwerk uit de Linnean Werkgroep Kosten (Van den Hout et al., 2021) kan worden gebruikt om vervolgens te sturen op passende zorg (zie

Figuur 2). Een belangrijke aantekening hierbij is dat de WGZ-breuk in de context moet worden besproken: zorg met betere uitkomsten tegen hogere kosten (C) en zorg met slechtere uitkomsten tegen lagere kosten (F) (zie Figuur 2) kan in bepaalde contexten gerechtvaardigd zijn.

Waarde inzicht kan op micro-, meso- en macroniveau een prikkel zijn tot verbetering, een intrinsieke motivator voor alle stakeholders in de zorg, waaronder patiënt, behandelaar, bestuurder en zorgverzekeraar. Het kan het samen beslissen optimaliseren in de spreekkamer, en de samenwerking en/of afstemming op het leveren van WGZ tussen management en zorgprofessionals versterken. Bovendien kan waarde inzicht informatief zijn voor de transformatie naar waardegedreven bekostiging van de zorg, waarbij betaling per verrichting wordt vervangen door een vergoedingssysteem waarin de gerealiseerde waarde een belangrijke rol speelt. De Linnean werkgroep Bekostiging heeft eerder onderzoek gedaan naar de belemmeringen en adviezen voor bundelbekostiging (Cattel et al., 2021).

Figuur 2. Raamwerk voor het sturen op uitkomsten en kosten op organisatieniveau (Van den Hout et al. 2021). Zorg met betere uitkomsten tegen lagere of beperkt stijgende zorgkosten wordt doorgevoerd (A + B), zorg met slechtere uitkomsten en hogere kosten vindt geen doorgang (D + E), en discussie wordt geïnitieerd bij zorg met betere uitkomsten tegen hogere kosten (C) en zorg met slechtere uitkomsten tegen een flinke kostenbesparingen (F).



Doel van dit Linnean whitepaper

In het kader van Integraal Zorgakkoord (2022) reikt dit Linnean whitepaper zorgorganisaties handvatten om zorguitkomsten te evalueren in relatie tot de gemaakte kosten.

1. *Theorie*

We beogen een eerste stap te zetten om de WGZ-breuk af te stemmen met praktisch bruikbare, evidence-based Health Technology Assessment (HTA)-methodieken. HTA-methodieken worden gebruikt om gezondheidsinterventies te evalueren op macro niveau: dit whitepaper bespreekt wat we hiervan kunnen leren om binnen WGZ op meso niveau te sturen op maximale waarde, ofwel de best mogelijke uitkomsten per euro. We inventariseren HTA-methodieken, en beschrijven middels kennisvragen waar we op moeten letten bij het sturen op de WGZ-breuk op meso niveau (zie bijlage I).

2. *Praktijkvoorbeelden*

We geven een uiteenzetting van praktijkvoorbeelden die uitkomsten én kosten inzichtelijk maken, om hier in de dagelijkse praktijk op te kunnen sturen. De voorbeelden, weergegeven in Bijlage III – XI, tonen hoe de WGZ-breuk kan worden berekend via formules, en hoe uitkomsten én kosten aan elkaar kunnen worden gerelateerd in dashboarding en andere tooling. Hierbij streven we ernaar te **inspireren** zonder oordeel: de whitepaper geeft een uiteenzetting van mogelijke methoden om te kunnen besluiten of de geleverde (innovatieve) zorg waarde toevoegt en dus passend is. Het laat de keuze welke methode het meest passend is, en welke verhouding tussen uitkomsten én kosten wenselijk is, bij sleutelfiguren in de zorgorganisaties zelf.

3. *Advies*

We bieden actiepunten voor zorgorganisaties om zelf aan de slag te gaan met het in kaart brengen van, en sturen op, de WGZ-breuk. We geven aan hoe u dit op eenvoudige wijze kunt doen, en ook waar u het complexer kunt maken om tot een meer precieze inschatting van waarde te komen. Daarnaast presenteren we de limitaties van ons whitepaper en de ontwikkelpunten voor het sturen op uitkomsten én kosten in de zorg.

Praktijkvoorbeelden voor het sturen op uitkomsten én kosten

We bieden handvatten voor het op organisatieniveau sturen op uitkomsten én kosten in de medisch-specialistische zorg. Door een uiteenzetting van praktijkvoorbeelden te geven in Bijlagen III – XI, beogen we managers, zorgverleners, bestuurders en andere belanghebbenden te inspireren hoe kan worden omgegaan met kennisvragen beschreven in Bijlage I, en hoe invulling kan worden gegeven aan de WGZ-breuk, om waardetoevoeging binnen de organisatie inzichtelijk te maken.

We focussen ons op praktijkvoorbeelden die de zorgverlening rondom een patiëntpopulatie in de dagelijkse praktijk beschrijven. Omdat stakeholders met diverse achtergronden moeten kunnen sturen op uitkomsten én kosten, ligt onze nadruk op praktijkvoorbeelden die op een begrijpelijke en overzichtelijke manier transparant zijn op een beperkt aantal eindpunten, met de mogelijkheid voor verdiepende analyses ter verdere exploratie. Eventuele tooling is bij voorkeur gebaseerd op brondata die toegankelijk is en zich leent voor real-time updates, zodat routinematig evaluaties kunnen worden uitgevoerd ter waardebeoordeling.

De werkgroep, bestaande uit een representatieve groep van WGZ-experts in Nederland, hebben vanuit hun netwerk en de literatuur praktijkvoorbeelden aandragen (zie Figuur 3). Wij zeggen niet dat het overzicht compleet is. De praktijkvoorbeelden zijn in tweetallen gescreend op een aantal elementen, die vooraf zijn vastgesteld door de werkgroep (zie Bijlage II). Om consensus te bereiken is de gehele screening ook met de werkgroep besproken. **De uiteenzetting van elk praktijkvoorbeeld zal voor publicatie ter goedkeuring worden voorgelegd aan de ontwikkelaars.**

Figuur 3. Kwadrant type praktijkvoorbeelden.

	<i>uitkomsten én kosten verbonden in WGZ-breuk</i>	<i>uitkomsten én kosten aan elkaar gerelateerd (in grafiek, matrix, o.i.d.)</i>
<i>eenmalige sturing</i>		SELFIE (Erasmus Universiteit, bijlage V)
<i>periodieke sturing</i>	QCI-dashboard (Franciscus Gasthuis, bijlage III)	CVA dashboard (Estland, bijlage IV)
	Patient-level value analysis (VS, bijlage IX)	Ischemische beroerte dashboard (Brazilië, bijlage VI)
	Value of efficiency indicator (GGZ, bijlage X)	Zorgpaden dashboards (Maasstad Ziekenhuis, bijlage VII)
		Spider chart (Catharina Ziekenhuis, bijlage VIII)
		Samen Beter (Santeon, bijlage XI)

Actiepunten

Om te kunnen sturen op patiëntwaarde rekening houdend met de kosten, roepen wij u als zorgverleners, managers, bestuurders en andere belanghebbenden op om samen te werken aan de WGZ bouwstenen (Hanselaar et al., 2023, ZE&GG, z.d.), rekening houdend met onderstaande, aanvullende actiepunten:

1. Inventariseer voor het selecteren van een patiëntengroep voor welke zorgpaden en / of zorginnovaties stuurinformatie wenselijk is.
2. Includeer bij het samenstellen van een multidisciplinair team ook allen die op managementniveau moeten gaan sturen op uitkomsten én kosten. Houdt er rekening mee dat bestaande taken, rollen en verantwoordelijkheden, van managers en andere stakeholders, kunnen veranderen wanneer u aandoeningsgericht gaat sturen op uitkomsten én kosten. Achtergronden en kennis kunnen verschillen, zorg ervoor zodat stuurinformatie voor allen toegankelijk is.
3. Maak ook op managementniveau patiënt relevante uitkomsten én kosten inzichtelijk. Begin met de data die voorhanden is (uit o.a. registraties en audits), en heb de ambitie om uit te bouwen naar een complete set van klinische en patiënt gerapporteerde uitkomsten. Kijk daarbij vooral naar hetgeen binnen het programma Uitkomstgerichte Zorg (z.d.) en ICHOM (z.d.) al is ontwikkeld, en de handvatten geboden in de Linnean werkgroep Kosten (Van den Hout et al., 2021). Dit bevordert benchmarking en bredere kennisdeling.
4. Ontwikkel tooling, zoals een dashboard of spider chart, om uitkomsten én kosten aan elkaar te verbinden. U kunt hierbij gebruik maken van de getoonde praktijkvoorbeelden (zie Bijlagen III - XI), en de handvatten geboden in de werkgroep Dashboard / Datavisualisatie (z.d.).
5. Indien wenselijk, kunt u de WGZ-breuk weergeven. U kunt voor de berekening gebruik maken van de handvatten geboden in de door ons beschreven HTA-methodieken (Bijlage I) en praktijkvoorbeelden (Bijlage III – XI).
6. Start met het sturen op uitkomsten én kosten. Maak afspraken over de aanpak, evalueer uitkomsten én kosten periodiek en formuleer verbeterinitiatieven. Besluitvorming moet tot stand komen door middel van dialoog. Zorg met betere uitkomsten tegen hogere kosten (zie Figuur 2, C) en zorg met slechtere uitkomsten tegen flinke kostenbesparing (F) kunnen in bepaalde contexten gerechtvaardigd zijn. Onze ervaring leert dat het belangrijk is om management te betrekken (o.a. Duaal Management) en korte lijnen te onderhouden met de Raad van Bestuur. Zij dragen verantwoordelijkheid over de strategie, de kwaliteit van

zorg (uitkomsten) en de inzet van middelen. U kunt voor tips kijken naar onze praktijkvoorbeelden.

7. Voor een zinnige vergelijking van waarde tussen zorgaanbieders is het noodzakelijk uitkomsten te corrigeren voor aanwezige casemix verschillen. Daarnaast kunt u stratificeren om variatie in uitkomsten én kosten tussen groepen te onderzoeken.
8. Belangrijk is dat u de aanpak die u gekozen heeft om uitkomsten én kosten te verbinden, en hierop te sturen, evalueert, en geleerde lessen deelt, ook binnen Linnean, zodat we data genereren over welke aanpak leidt tot verbeteringen van zorg, en welke niet. Het delen van deze ervaringen kan de input voor een nieuwe Linnean werkgroep vormen.

Aansluitend hebben we een aantal bredere ontwikkelpunten voor de praktijk, wetenschap en overheid geformuleerd:

1. In onze doelstelling hebben we geformuleerd dat we sleutelfiguren in de zorgorganisaties zelf laten bepalen welke verhouding tussen uitkomsten én kosten wenselijk is. Echter zou kunnen worden onderzocht of één cut-off waarde voor gebruik in de individuele behandelsetting en / of de zorginstelling bepaald kan worden, geïnspireerd door de ICER op macro niveau.
2. Wellicht zijn meer data gewenst, indien het voor sturing noodzakelijk is om uitkomsten én kosten gedetailleerder in kaart te brengen in het zorgpad. Uitkomstendata zouden op verschillende manieren kunnen worden verrijkt: o.a. door koppelingen met andere registraties en / of kwaliteit registers, zoals de Nationale Hart Registratie (NHR) in de cardiologie. Maar ook bijvoorbeeld door afspraken te maken over het op uniforme wijze registreren aan de bron, als mede het stroomlijnen van het datamanagement, zoals het Health Intelligence Platform Santeon (HIPS, z.d.) Kostendata zou met bijvoorbeeld Logex en Performance kunnen worden verrijkt door te verkennen hoe TDABC data kan worden gekoppeld, en wat dit oplevert. Belangrijk hierbij is dat moet worden gewaarborgd dat de registratielast op het niveau van de zorgverlener niet verder stijgt.
3. Daarnaast is het wenselijk om een koppeling met gegevens buiten de zorginstelling te bewerkstelligen voor een completer beeld van uitkomsten én kosten in het zorgpad. WGZ heeft als ultieme benadering de integrale zorg: de uitdaging om uitkomsten én kosten integraal te meten in Nederland. Tot nu toe is er echter een gebrek aan inzicht in het volledige zorgpad. Denk hierbij aan zorgvoorziening die grenzen aan de eerste lijn of daar onderdeel van zijn, zoals de huisarts, GGZ, of mantelzorgbelasting. Echter moeten we ons voorafgaand aan het realiseren van dergelijke koppelingen actief afvragen wat de meerwaarde is, gezien uitdagingen zoals privacy waarborging. Kleinschalige en lokale data

kunnen veel impact hebben, waarbij het toevoegen van koppelingen het stroever en moeilijker navigeren kan maken om te sturen op uitkomsten én kosten.

4. Tot slot willen wij voortbouwen op de werkgroep Bekostiging (Cattel et al., 2021) en roepen wij op te onderzoeken of het verbinden van uitkomsten én kosten informatief kan zijn voor waardegedreven bekostiging van de zorg, waarbij betaling per verrichting wordt vervangen door een vergoedingssysteem waarin de gerealiseerde waarde een belangrijke rol speelt.

VERTROUWELIJK

Bijlage I - Kennisvragen voor het meten van waarde

Voor het inrichten van de WGZ-breuk in de medisch-specialistische zorg kunnen we leren van HTA. We beschrijven eerst kort de HTA-methodiek, waarna we reflecteren op de overeenkomsten en verschillen tussen HTA en WGZ. Vervolgens beschrijven we kennisvragen voor het meten van waarde in WGZ.

Health Technology Assessment

HTA is een multidisciplinair evaluatieproces dat wordt gebruikt om gezondheidsinterventies te beoordelen op aspecten als klinische effectiviteit, doelmatigheid, veiligheid, bruikbaarheid, ethiek en sociale implicaties op populatieniveau (Banta, 2003). HTA heeft als doel beleidsmakers, zorgverleners, patiënten en andere belanghebbenden te ondersteunen bij beslissingen over het gepast gebruik van behandelingen, diagnostische hulpmiddelen en medische apparaten. HTA maakt gebruik van een systematische benadering om beslissingen te nemen op basis van de voordelen en risico's van interventies (WHO, 2023). Een vereiste hiervoor zijn objectieve en betrouwbare gegevens. Voor het beoordelen van doelmatigheid kent HTA verschillende evaluatiemethoden.

De keuze voor welke methode gebruikt wordt, hangt af van de vraagstelling en het doel van de HTA-evaluatie. De belangrijkste voorbeelden van methoden voor het evalueren van doelmatigheid zijn de kosten-effectiviteitsanalyse (KEA) en de kosten-utiliteitsanalyse (KUA). Deze worden vooral gebruikt op macroniveau om beleidsbeslissingen over vergoedingen in het basispakket en het formuleren van klinische richtlijnen te informeren. De KEA biedt een beoordeling van de impact van een (nieuwe) medische interventie op kosten en gezondheidseffecten in vergelijking met een geschikte comparator, meestal standaardzorg. De impact van een interventie kan worden uitgedrukt in de incrementele kosten-effectiviteitsratio (ICER): de incrementele kosten gedeeld door de incrementele gezondheidseffecten van interventie A in vergelijking met B. De KUA is een speciale vorm van KEA waarbij de effectiviteit wordt gemeten in termen van quality-adjusted life years (QALYs), die de kwaliteit van leven combineren met de levensduur (Zorginstituut Nederland, 2016). Het gebruik van QALYs maakt het mogelijk om doelmatigheid van veel verschillende vormen van interventies binnen en tussen indicaties met verschillende effecten op verschillende uitkomstmaten toch op een vergelijkbare manier te beoordelen. Bij het uitvoeren van KEAs kunnen zowel trial-based (op basis van randomized controlled trials), als model-based benaderingen, of een combinatie, worden toegepast om de kosteneffectiviteit van interventies te beoordelen. Verdere uitleg over economische perspectieven en besluitvorming op basis van kosteneffectiviteit staat beschreven in het advies van de werkgroep Kosten (Van den Hout et al. 2021).

Waardegedreven zorg en Health Technology Assessment

Zowel HTA als WGZ gebruiken dus een ratio van uitkomsten en kosten om de waarde van zorg te bepalen (Box 2). Voor het inrichten van de WGZ-breuk kunnen we leren van de duidelijk gedefinieerde ICER in CEA en CUA. O.a. de te maken vergelijking (nieuwe interventie versus oud) en het perspectief voor het meten van kosten zijn duidelijk gedefinieerd en vastgelegd in landelijke richtlijnen (Zorginstituut Nederland, 2016). Toch is de methodiek van HTA niet één-op-één te vertalen naar WGZ. Waar HTA wordt gebruikt voor allocatievraagstukken op maatschappelijk niveau (macro), is WGZ gericht op optimalisatievraagstukken van gezondheidsuitkomsten relatief tot kosten binnen de patiëntenzorg op spreekkamer niveau (micro) en instellingsniveau (meso). HTA beoordeelt de waarde van een interventie op populatieniveau maar houdt geen rekening met de individuele waarden van de patiënt zoals WGZ wel beoogt. De belangrijkste verschillen tussen CEA en VBHC worden weergegeven in Tabel 1 en hieronder verder beschreven.

WGZ	HTA
$Value = \frac{Patient\ relevante\ uitkomsten}{Kosten}$	$ICER = \frac{Kosten(B) - Kosten(A)}{Effectiviteit(B) - Effectiviteit(A)}$

Box 2. Vergelijking value-equation van WGZ met ICER van HTA.

	CEA	VBHC
Goal	Optimize value	
Perspective	Payer	Patient
Level	Macro	Meso/micro
Value definition	Economic	Holistic
Approach	Normative, equitable	Individualistic, patient-centered
Formula	Costs vs Outcomes	
Outcomes	Generic outcomes (QALYs)	Disease specific outcomes
Value assesment	Incremental	Often not incremental
Duration	Lifetime	Care pathway
Challenges	Preferences, resource allocation, operationalization of value	

Tabel 1. Verschillen tussen CEA en VBHC.

Afhankelijk van het doel van de analyse en welke stakeholder in de gezondheidszorg wil sturen op uitkomsten en kosten, kunnen zowel HTA als WGZ gebruikt worden. Op het macroniveau van beleidsmakers vindt, zoals hierboven beschreven, besluitvorming plaats over pakketbeheer en medische richtlijnen. Omdat HTA pakketbeheer en richtlijnen informeert, worden studies vaak éénmalig uitgevoerd bij de introductie van een nieuwe interventie. Ook is in CEA en CUA is een

standaard stramien voor het bepalen van waarde noodzakelijk om interventies voor verschillende aandoeningen op vergelijkbare manier te kunnen beoordelen. Er wordt daarom over het algemeen gebruik gemaakt van generieke vragenlijsten voor het meten van kwaliteit van leven (bijvoorbeeld de EQ-5D) voor het berekenen van QALYs (Versteegh et al., 2016). Kosten worden breed gemeten vanuit maatschappelijk perspectief. Dit omvat o.a. directe en indirecte zorgkosten, productiviteitskosten en informele zorg.

In WGZ is een routinematige werkwijze waarin uitkomsten en kosten continu gemonitord worden juist essentieel. Ook is focus op relevante uitkomsten en kosten vanuit de patiëntreis vereist. Op het microniveau van de spreekkamer wordt er door middel van gedeelde besluitvorming samen beslist tussen arts en patiënt op basis van klinische en patiënt-relevante uitkomsten voor specifieke aandoeningen gemeten met 'patient-reported outcome measures' (PROMs) (Terwee et al., 2019). Binnen samen beslissen ontvangt de patiënt informatie over de toegevoegde waarde van een behandeling op basis van analyses op patiëntgroep niveau, bijvoorbeeld uit effectiviteitsstudies van het macroniveau. Vervolgens is de afweging voor een bepaalde behandeling individueel en afhankelijk van de doelen, normen en waarden van de patiënt. Ook krijgen patiënten op deze manier inzicht in wat er bereikt is met de behandeling, zoals bijvoorbeeld bij Routine Outcome Monitoring (ROM) in de GGZ (De Beurs et al., 2011). Op het microniveau zijn kosten ook relevant voor de patiënt, zeker waar het gaat om bijvoorbeeld de invloed van de ziekte en/of behandeling op arbeidsparticipatie, en voor de maatschappij omdat zorg in het algemeen vanuit de basisverzekering vergoed wordt. Er is nog weinig bekend over hoe er in samen beslissen tussen arts en patiënt gestuurd kan worden op uitkomsten én kosten, maar dit valt buiten de scope van dit whitepaper.

Op het mesoniveau van zorginstellingen en behandelaars worden uitkomstdata al gebruikt om resultaten met anderen te vergelijken (benchmarking) en te monitoren of de zorg binnen een zorgpad over de tijd op peil blijft (leren en verbeteren). Voorbeelden hiervan zijn de kwaliteitsregistraties van the Dutch Institute for Clinical Auditing (DICA) (DICA, 2023; Kuhrij et al., 2018) en het Santeon 'Samen Beter' (z.d.). Daarnaast zijn er verschillende methoden voorhanden voor het berekenen van kostprijzen. Activity-Based Costing (ABC) wordt standard toegepast in veel Nederlandse ziekenhuizen en WGZ raadt Time-Driven Activity-Based Costing (TDABC) voor nauwkeuriger inzicht in de efficiëntie van zorgprocessen. Wel vereist TDABC intensief gebruik van resources en een grote tijdsinvestering, en is daarmee niet altijd geschikt voor routinematige zorgevaluaties (Health, 2020). Er missen tot op heden echter concrete methoden die het sturen op zowel uitkomsten als kosten mogelijk maken op mesoniveau.

Kennisvragen

Hoewel HTA niet geschikt is voor het sturen op uitkomsten en kosten op mesoniveau, biedt de methodiek aanknopingspunten voor kennisvragen voor in het inrichten van de WGZ-breuk. Hieronder worden de drie belangrijkste kennisvragen, naar inziens van de werkgroep, voor het sturen op uitkomsten én kosten op mesoniveau besproken. De praktijkvoorbeelden in volgende bijlagen zullen de verschillende mogelijkheden om deze vragen te beantwoorden laten zien.

1) Welke uitkomsten en kosten dienen gemeten te worden?

Voor veel aandoeningen zijn er al uitkomstensets vastgesteld en wordt veel werk verricht aan generieke PROMs, zowel bij ICHOM (Akpan et al., 2018; Verberne et al., 2019) als het programma Uitkomstgerichte zorg (z.d.) en PROMIS (z.d.). De uitdaging voor het sturen op mesoniveau ligt bij het meten van de kosten. Onduidelijk is de reikwijdte van de kosten die dienen meegenomen te worden. De afbakening van het begin en eind van een zorgpad ontbreekt vaak, vooral bij chronische aandoeningen. Ook kunnen persoonlijke kosten buiten het zorgpad kunnen wel relevant zijn voor de patiënt, zoals reistijd en arbeidsverzuim. Er zijn initiatieven waarbij kosten binnen een zorgpad worden op basis van patiënt-relevantie worden geselecteerd en meegenomen in het leren en verbeteren (Van Steenberghe et al., 2023). De praktijkvoorbeelden zullen bespreken welke uitkomsten worden meegenomen en hoe deze worden gecombineerd, en of kosten worden meegenomen vanuit maatschappelijk perspectief, de instelling en/of de keten.

2) Hoe kan de WGZ-breuk geoperationaliseerd worden?

Doel van de WGZ-breuk is om te komen tot een indicator die aangeeft of uitkomst en kosten in balans zijn op micro-, meso- en macro niveau. Op mesoniveau zijn in principe zijn er meerdere vergelijkingen mogelijk: zorgpaden voor een specifieke aandoening kunnen verbeteren over de tijd of leren van zorgpaden bij een andere aandoening met een vergelijkbare patiëntenpopulatie. De vraag is ook hoe dan de afweging in uitkomsten en kosten gemaakt moet worden: wanneer zijn de extra uitkomsten het geld waard? En hoe wegen meerdere uitkomsten mee? Een voorbeeld van een methodiek die meerdere uitkomsten meeneemt is multi criteria decision analysis (MDCA) (Gongora-Salazar et al., 2023), deze wordt ook gebruikt in het praktijkvoorbeeld over SELFIE.

3) Wat zijn vereisten voor de betrouwbaarheid van data?

Om echt goede vergelijkingen te maken is het noodzakelijk de kans op bias in de data zo klein mogelijk te houden. Vanuit evidence-based medicine is randomisatie de gouden standaard om, bijvoorbeeld interventies, te vergelijken. In de dagelijkse WGZ-praktijk gebruiken we juist observationele data, waar ook veel uit te leren valt. Wel kunnen door patiënt- (case mix) en instellingsverschillen kan er vertekening ontstaan in deze data. Deze vertekening proberen we tegen te gaan met case-mix correctie, een statistische methode om te corrigeren voor deze verschillen. Betrouwbare data zijn nodig om een vinger te leggen op de oorzaak van verschillen in uitkomsten en kosten. Bijvoorbeeld

binnen het benchmarken kan er niet zomaar een (deel van) een protocol worden overgenomen van een andere aanbieder als die aanbieder betere uitkomsten heeft en/of lagere kosten heeft, mede door bijvoorbeeld verschillen in behandelcultuur. Benchmarken en vervolgens transparant publiceren heeft wel de kracht dat instellingen van elkaar kunnen leren. Uniforme definities van gemeten uitkomsten en kosten en zorgvuldige registratie zijn daarbij ook belangrijk. Ook in de praktijkvoorbeelden wordt aandacht besteed aan de betrouwbaarheid van gebruikte data.

VERTROUWELIJK

Bijlage II – Overzicht praktijkvoorbeelden voor sturen op uitkomsten én kosten

Tabel 2. Overzicht praktijkvoorbeelden voor het sturen op uitkomsten én kosten inclusief hun belangrijkste elementen.

	Uitkomsten			Kosten		Uitkomsten én kosten			Betrouwbaarheid
	Klinisch	PROMs	Samengevat in één kwaliteitsindicator	Instelling	Zorgketen (over instellingen heen)	In één visueel overzicht	In WGZ-breuk	Voor sturing in zorgpad	
Quality-Cost-Indicator model (Franciscus Gasthuis, Nederland)	Geïntegreerde zorgenheid selecteert met patiënten landelijke indicatoren en ICHOM set.		Optellen tot dichotome waarde: patiënt voldoet wel of niet aan <u>alle</u> kwaliteitseisen.	Zorgactiviteiten per patiënt x kostprijs. Honorarium via TDABC.	-	Dashboard.	Ratio van patiënten die voldaan hebben aan <u>alle</u> kwaliteitseisen / gemiddelde kosten.	Halfjaarlijkse verbeteracties met Raad van Bestuur.	
SELFIE: integrale zorg bij multi-morbiditeit (Erasmus Universiteit, Nederland)	-	Gevalideerde lijsten voor 5 gezondheid en welzijn indicatoren, geselecteerd via focusgroepen, workshops, lopende dataverzamelingen en literatuurstudie: functioneren, welbevinden, sociale relaties en participatie, levensplezier en veerkracht.	-	-	iMTA MCQ: door patiënten gerapporteerd zorggebruik.	- (wel tool beschikbaar voor MCDA berekening)	8 indicatoren <i>op uitkomsten, ervaringen én kosten</i> zijn gewogen tot één waarde via een Discrete Choice Experiment met patiënten, partners, professionals, zorgverzekeraars en beleidsmakers. Weegfactoren werden bepaald via de keuze die zij herhalend hebben gemaakt tussen 2 integrale zorg-programma's die verschillend scores op de 8 indicatoren. De waarde wordt voor de 5 groepen apart getoond. Deze wordt voor de interventie en controle groep berekend.	In overleg met patiënten, partners, professionals, zorgverzekeraars en beleidsmakers werden resultaten besproken voor 17 integrale zorginnovaties in 8 landen.	Sensitiviteit analyses, deterministisch: andere vorm van wegen / standaardiseren probabilistisch: Monte Carlo simulatie voor onzekerheid.
Ischemische beroerte dashboard (ziekenhuisgroep, Brazilië)	mRS, NIHSS, overleving, opnameduur en infecties.	ICHOM set: mobiliteit, communicatie en zelfzorg.	-	TDABC met: - facturen materialen, middelen, medicatie; - ziekenhuisbezoeken EPD x prijs instelling.	-	Interactief dashboard (o.a. met spider chart).	Enkel weergave kosten per mRS en NIHSS, ook op te splitsen naar 3 niveaus klinische gesteldheid patiënten.	Ontwikkeling zorgstandaard beroerte in Brazilië.	

CVA dashboard (Nordic Healthcare Group, Estland)	Complete ICHOM set	Weging naar perspectief zorgverleners, waarbij PROMs het zwaarst wegen.	Estonian Health Insurance Fund: inzicht in prijs per verrichting, en zorgvolume per patiënt.	-	Dashboard.	Enkel matrix waarin kwaliteitsindicator wordt afgezet tegen de kosten.	Verbeterinitiatieven via workshop ESHF en bestuurders.	Weging uitkomsten niet beschikbaar.
Zorgpaden dashboards (Maasstad Ziekenhuis, Nederland)	(Multidisciplinair) verbeterteam selecteert uitkomsten op basis van KPIs.	-	Zorgactiviteit per patiënt x kostprijs.	-	Dashboard.	-	Periodieke evaluatie vooruitgang met verbeterteam.	
Spider chart aortaklepvervangings (Catharina Ziekenhuis, Nederland)	NHR klinische uitkomst-monitoring.	-	Multidisciplinair team selecteerde kostenparameters uit declaratiedata.	-	Spider chart.	-	Voor implementatie en evaluatie verbeterinitiatieven.	
Value of efficiency indicator (GGZ, Nederland)	-	Eén of meerdere lijsten.	Eén waarde: de pre-to-post test change. Het is aan de gebruiker hoe gewogen wordt.	Gedeclareerde kosten of kosten minus baten.	-	-	Ratio voor behandeluitkomsten versus kosten.	-
Patient-level value analysis (Verenigde Staten)	-	Eén of meerdere lijsten.	Eén waarde: de pre-to-post test change.	TDABC	-	Matrix.	Enkel matrix waarin voor elke patiënt z-score voor totale kosten wordt afgezet tegen z-score uitkomst.	Via regressie wordt onderzocht welke variabelen de waarde van zorg voorspellen.
Samen Beter Geboortezorg (Santeon ziekenhuizen, Nederland)	Verbeterteam definieert scorekaart.	-	Kosten indicatoren per aandoening.	-	Dashboard.	-	Periodiek komt het verbeterteam samen om verbeterinitiatieven te bespreken.	

Bijlage III – Quality-Cost-Indicator (QCI)-dashboard (Franciscus Gasthuis, Nederland)

WGZ is onderdeel van de strategie van het Franciscus Gasthuis. Het is onder meer een instrument om de transitie naar passende zorg te maken. Hierbij is het Quality-Cost-Indicator (QCI)-model ontwikkeld om via een waardeplan WGZ onderdeel te laten worden van de Planning & Control Cyclus.

Het QCI-model bouwt voor op het model van Porter en Lee (2013) door patiëntuitkomsten in kaart te brengen en deze te koppelen aan kosten gemaakt voor de patiënt in de organisatie.

Voor een specifieke diagnosegroep ontwerpt de geïntegreerde zorg eenheid met patiënten en betrokken zorgverleners het zorgpad. Hierin worden klinische uitkomsten én PROMs opgenomen, gebaseerd op landelijke indicatoren en ICHOM. Ook bepalen zij gewenste waarden voor alle uitkomstmaten, en wordt in beeld gebracht welke procesindicatoren de meeste uitkomsten beïnvloeden.

De som van de uitkomsten leidt tot een dichotome waarde, waarbij de patiënt óf voldoet aan alle kwaliteitseisen óf aan één of meerdere kwaliteitseisen niet voldoet. Het percentage patiënten dat voldoet aan alle kwaliteitseisen bepaalt de teller in het QCI-model (zie Box 3). De kosten, die de noemer vormen, worden berekend per patiënt door de gedane zorgactiviteiten te vermenigvuldigen met de bijbehorende kostprijs. Het honorarium wordt door middel van TDABC bepaald, omdat de standaard methode (normtijden vanuit LOGEX) onvoldoende de wijzigingen in het proces binnen het ziekenhuis weergeeft (andere tijden en/of door verpleegkundig- in plaats van medisch specialist). Voor vergelijkbaarheid wordt eenzelfde kostprijs voor verschillende periodes gehanteerd.

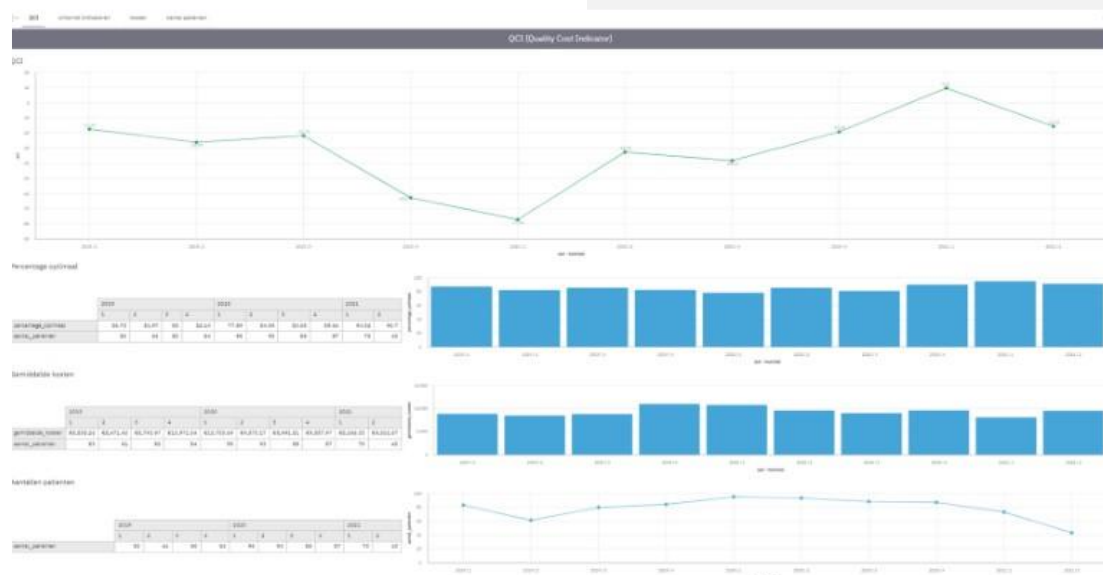
Box 3. Formule behorend bij het Quality-Cost-Indicator (QCI)-model.

$$QCI = \frac{\% \text{ patiënten die voldoen aan alle kwaliteitseisen} * 100}{\text{gemiddelde totale kosten per } \text{€}1000, -}$$

Ontwikkeling op de QCI wordt per kwartaal middels een dashboard gevolgd (zie de bovenste lijn in Figuur 4). Ook worden ontwikkelingen op de teller, noemer, en het aantal patiënten weergegeven (zie de grafieken onder de QCI). Patiënten worden gestratificeerd om vergelijking in de tijd mogelijk te maken. Daarnaast worden procesindicatoren getoond (niet weergegeven in Figuur 4) om handvatten te bieden om te verbeteren op uitkomsten en/of kosten. Halfjaarlijks worden ontwikkelingen en eventuele verbeteracties ook besproken met de Raad van Bestuur.

Figuur 4. Dashboard waarin het Quality-Cost-Indicator (QCI)-model is verwerkt.

QCI Framework



De bespreking van het dashboard is gestart, maar moet nog zijn beslag gaan krijgen. Het Franciscus Gasthuis hoopt in de tweede helft van 2023 het QCI-model te publiceren.

VERTROUWEN

Bijlage IV - CVA dashboard (Nordic Healthcare Group, Estland)

De Nordic Healthcare Group (NHG) heeft in samenwerking met twee academische en twee regionale ziekenhuizen in Estland een CVA dashboard ontwikkeld dat uitkomsten én kosten samenbrengt, en vergelijkt tussen deze ziekenhuizen (ICHOM, 2022). Zij beogen hiermee een meer geïntegreerd en patiëntgericht zorgpad te bewerkstelligen, en als gevolg patiënten hun kwaliteit van leven te verbeteren. Specifiek bij CVA heerst bij veel patiënten onduidelijkheid over het gefragmenteerde zorgpad dat is ondergebracht bij verschillende zorgaanbieders.

Het dashboard bevat de ICHOM uitkomstenset waarin zowel klinische uitkomsten als PROMs zijn geïncorporeerd. Uitkomsten zijn gewogen naar het relatieve belang dat betrokken zorgverleners eraan hechten, om te komen tot één waarde: de 'outcome index'. De weging is niet gepubliceerd, maar bekend is dat PROMs zwaarder zijn gewogen om het belang van het patiëntperspectief te benadrukken. Kostendata zijn aangeleverd door het Estonian Health Insurance Fund (ESHF). Zij leveren globaal inzicht in de prijs per verrichting in het zorgpad, en het zorgvolume per patiënt. Hierbij worden kosten niet gespecificeerd per zorgactiviteit zoals in ABC.

Het dashboard biedt een overzichtswaargave waarin uitkomsten en kosten worden geïntegreerd (zie Figuur 5). Links worden diverse filters voor de data geboden, waaronder het ziekenhuis, de tijdsperiode en case-mix variabelen. Links midden wordt de 'outcome index' weergegeven, evenals de 'average cost of journey': de kosten uitgesplitst per stap in het zorgpad. Rechts boven worden beiden indicatoren tegen elkaar uitgezet. De WGZ-breuk wordt bewust niet verder ingevuld, naar eigen zeggen om een te abstracte waarde, die moeilijk toepasbaar is, te voorkomen. Onderin volgt het benchmarken van de individuele dimensies van de uitkomsten en kosten per ziekenhuis op het gemiddelde van alle ziekenhuizen, zodat gedetailleerd inzichtelijk is op welke terreinen het ziekenhuis verbeteracties kan doorvoeren.

Een meer gedetailleerd beeld van uitkomsten en kosten wordt separaat weergegeven in tabbladen. Uitkomsten kunnen per dimensie worden weergegeven, en ook de absolute waarden op uitkomsten worden getoond, inclusief veranderingen over tijd. Kosten worden ook gespecificeerd per patiënt, en per stap in het zorgpad per ziekenhuis.

Figuur 5. Overzichtswaergave CVA dashboard Estland (ICHOM, 2022).



Het dashboard wordt in een workshop met het bestuur van de vier ziekenhuizen en de ESHF besproken om tot verbeterinitiatieven te komen. Zowel op uitkomsten als kosten bleken bepaalde ziekenhuizen beter te presteren dan anderen, wat leidde tot het bespreken van onderliggende processen om tot verbeterinitiatieven te komen. Procesindicatoren zijn niet opgenomen in het dashboard. Het dashboard wordt momenteel niet verder doorontwikkeld.

Bijlage V – SELFIE (internationaal, coördinator: Erasmus Universiteit, Nederland)

SELFIE (Sustainable intEgrated care modeLS for multi-morbidity delivery, Financing and pErformance, z.d.) is een internationaal project waarin Nederland, Oostenrijk, Kroatië, Duitsland, Hongarije, Noorwegen, Spanje en het Verenigd Koninkrijk middels financiering van Horizon 2020 samenwerken aan persoonsgerichte zorg voor multi-morbide patiënten. Ze hebben een Multi-Criteria Decision Analysis (MCDA)-model ontwikkeld om integrale zorginterventies empirisch te kunnen evalueren.

Als startpunt zijn 17 lopende integrale zorginterventies geselecteerd waarin een besluit moest worden genomen over de duurzaamheid van het programma. De interventies zijn middels documentanalyse en interviews, op micro, meso en macro niveau beschreven op zes zorgsysteem criteria van de Wereldgezondheidsorganisatie (WHO, 2010) (Leijten et al., 2018a).

Daaropvolgend is een kernset van acht uitkomstmaten vastgesteld middels: 1) focusgroepen met multi-morbide patiënten (Leijten et al., 2018b), 2) nationale workshops met vijf stakeholder-groepen, bestaande uit patiënten, mantelzorgers, zorgverleners, zorgverzekeraars en beleidsmakers; 3) uitkomstmaten die reeds werden gemeten in de interventies, en 4) een literatuurstudie op uitkomstmaten in integrale zorg (Looman et al., 2018). Conform de Triple Aim zijn de acht uitkomstmaten verdeeld over drie pijlers: uitkomsten, ervaringen én kosten, allen gemeten middels gevalideerde patiënt-gerapporteerde vragenlijsten (zie Tabel 3). Bij alle 17 zorginterventies is een passende controlegroep gezocht, waarna voor zowel de interventie- als de controlegroep de acht uitkomstmaten gemeten zijn.

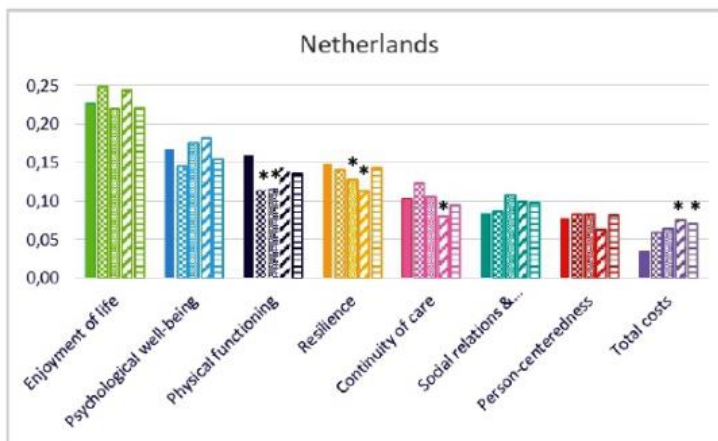
Figuur 6. Voorbeeld van een vraag uit het Discreet Keuze-Experiment (Rutten-van Mólken et al., 2020).

Tabel 3. De kernset van uitkomstmaten.

	Uitkomstmaten	Vragenlijsten
Gezondheid en welzijn	Fysiek functioneren	SF-36, Katz15
	Psychologisch welzijn	MHI-5
	Sociale relaties en participatie	IPA
	Plezier in het leven	ICECAP-O, Q-LES-Q
	Veerkracht	BRS
Ervaring	Persoonsgerichtheid	P3CEQ
	Continuïteit van zorg	NCQ, CPCQ
Kosten	Totale zorg- en welzijnskosten	iMTA_MCQ

	Care programme A	Care programme B
Physical functioning	Severely limited in physical functioning and activities of daily living	Moderately limited in physical functioning and activities of daily living
Psychological well-being	Always or mostly stressed, worried, listless, anxious, and down	Always or mostly stressed, worried, listless, anxious, and down
Social relationships and participation	No or barely any meaningful connections with others	No or barely any meaningful connections with others
Enjoyment of life	Some pleasure and happiness in life	No or barely any pleasure and happiness in life
Resilience	Good ability to recover, adjust, and restore balance	Good ability to recover, adjust, and restore balance
Person-centeredness	Highly person-centred	Highly person-centred
Continuity of care	Poor collaboration, transitions, and timeliness	Fair collaboration, transitions, and timeliness
Total health- and social care costs	5600 pounds per participant per year	7000 pounds per participant per year
Which care programme do you prefer, A or B?		
	A <input type="radio"/>	B <input type="radio"/>

Aansluitend is het relatieve belang van de acht uitkomstmaten bepaald door in ieder land, voor elke stakeholdergroep, een discreet keuze-experiment (DCE) uit te voeren. In de DCE kregen deelnemers keuzevragen voorgelegd waarin ze herhaaldelijk hun voorkeur moesten uitspreken voor één van twee getoonde, hypothetische integrale zorgprogramma's, die telkens verschilden in prestaties op de kernset van uitkomstmaten (zie Figuur 6). Het relatieve belang van elke uitkomstmaat werd gebaseerd op de regressie-coëfficiënt van het beste niveau (niveau 3) gedeeld door de som van alle beste niveaus. Gepoolde analyses tonen dat 'plezier in het leven' het zwaarst wordt gewogen, en de totale zorg- en welzijnskosten het minst zwaar. Verdiepende analyses laten daarnaast een grote variatie in voorkeuren zien, die slechts zwak gerelateerd zijn aan het stakeholdersperspectief van deelnemers. Dit kan worden verklaard doordat stakeholders vaak meerdere rollen op zich nemen, en toont de noodzaak voor een brede uitkomstenset. De weegfactoren voor Nederland staan weergegeven in Figuur 7.



Figuur 7. Weegfactoren in Nederland voor de kernset van uitkomstmaten (Rutten-van Mólken et al., 2020).

De staven tonen opeenvolgend de stakeholdergroepen: patiënten, mantelzorgers, zorgverleners, zorgverzekeraars en beleidsmakers. * weergeeft significante verschillen t.o.v. de patiëntengroep.

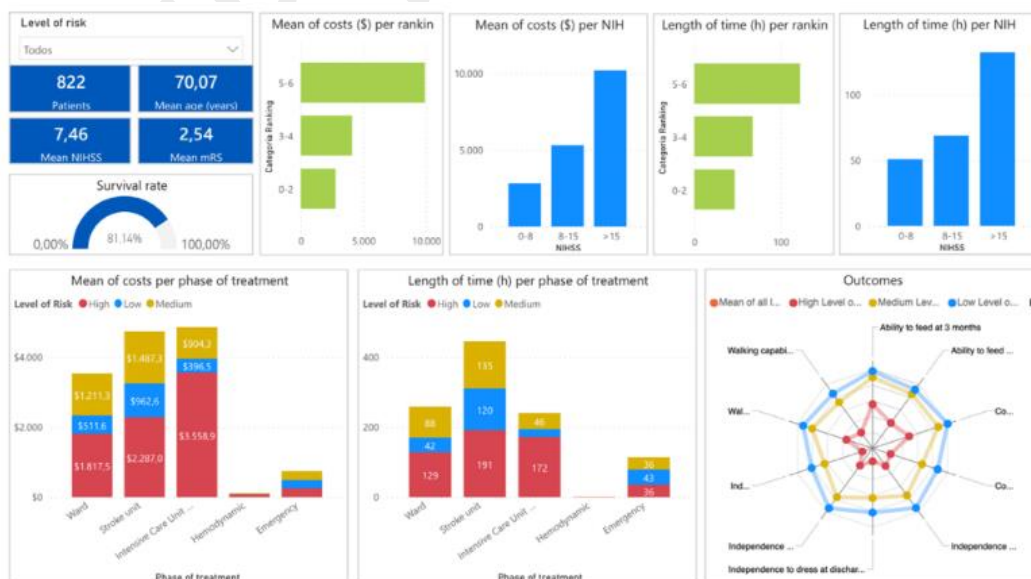
Om te komen tot één waarde score per stakeholdergroep voor zowel de interventie- als de controlegroep in het zorgprogramma, zijn de gemeten uitkomstmaten vermenigvuldigd met bijpassende weegfactoren. Sensitiviteit analyses werden uitgevoerd, zowel deterministisch (andere weeg- en standaardiseringstechniek) als probabilistisch (Monte Carlo simulatie voor onzekerheid). Daarna werden de resultaten met de vijf stakeholdergroepen besproken om te komen tot een besluit betreffende de duurzaamheid van de integrale zorginterventies.

Bijlage VI - Ischemische beroerte dashboard (ziekenhuisgroep, Brazilië)

Om de implementatie van VBHC programma's te faciliteren, standaardiseren en versnellen, hebben één privaat en twee openbaar ziekenhuizen in Brazilië een real-time, ischemische beroerte dashboard ontwikkeld dat breder inzetbaar is voor andere zorgpaden (Beck da Silva Etges et al., 2022). Het dashboard is ontworpen door PowerBI in samenspraak met een multidisciplinair team bestaande uit neurologen, epidemiologen, data analisten en industrieel ingenieurs. Het dashboard bevat zowel klinische uitkomsten als PROMs, in de vorm van de ICHOM uitkomstenset gecombineerd met proces- en uitkomstindicatoren uit het Angel Awards programma waarin het zorgpad voor patiënten is ontwikkeld. TDABC is gebruikt om de kosten per patiënt te schatten: uit facturen en documenten die aanschafkosten beschrijven, zijn de prijzen en hoeveelheden materialen, hulpmiddelen en medicatie afgeleid. Deze zijn aangevuld met informatie over het aantal ziekenhuisbezoeken, afgeleid uit het elektronisch patiëntendossier, vermenigvuldigd met de prijs per instelling.

Het dashboard bevat een viertal interactieve schermen, waarbij data kan worden gespecificeerd naar patiëntgroepen en variabelen. Het eerste scherm toont een beschrijving van de patiëntenpopulatie – ook per ziekenhuis – op overleving, behandeltype, klinische gesteldheid, leeftijd en geslacht. De daaropvolgende schermen tonen een uiteenzetting van proces- uitkomst- en kostenindicatoren naar behandeltype, klinische gesteldheid (zie Figuur 8) en ziekenhuis. Hierbij worden de gemiddelde kosten per uitkomst (modified Ranking Scale, mRS en National Institutes of Health Stroke Scale, NIHSS) weergegeven.

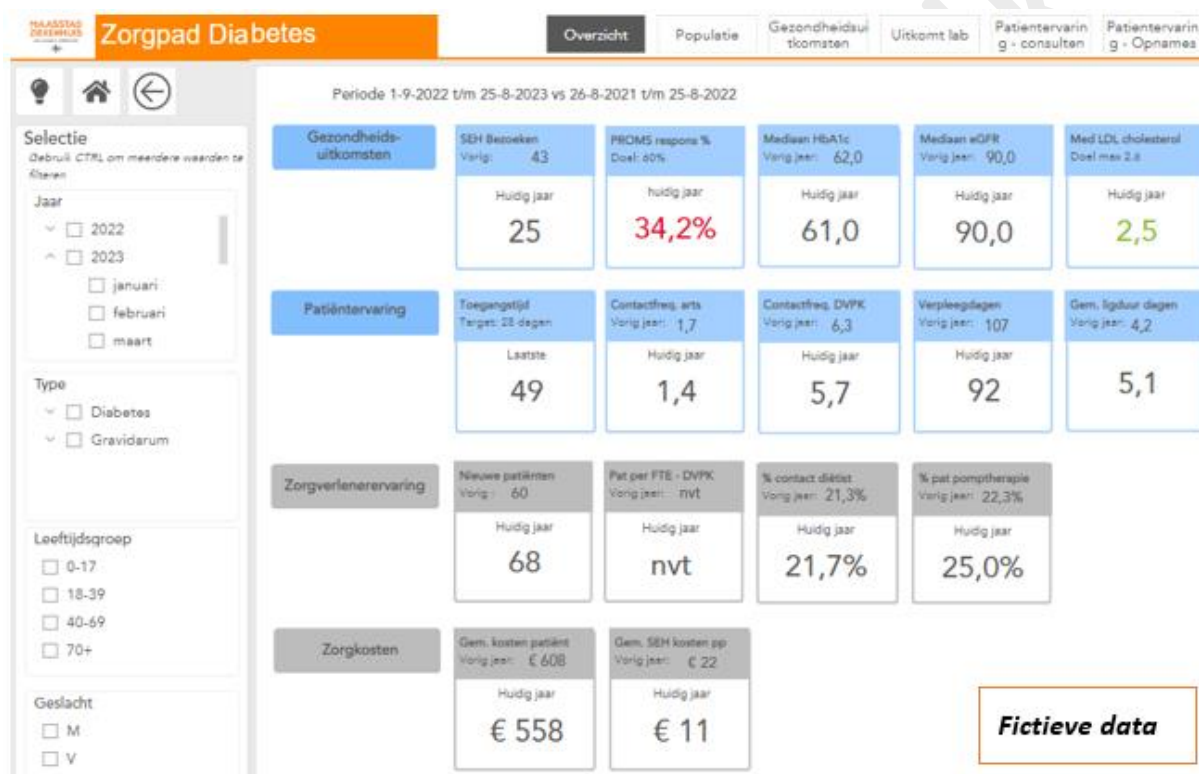
Figuur 8. Ischemische beroerte dashboard Brazilië, te bezichtigen via: <https://webavc.prologica.pt/> (Beck da Silva Etges et al., 2022).



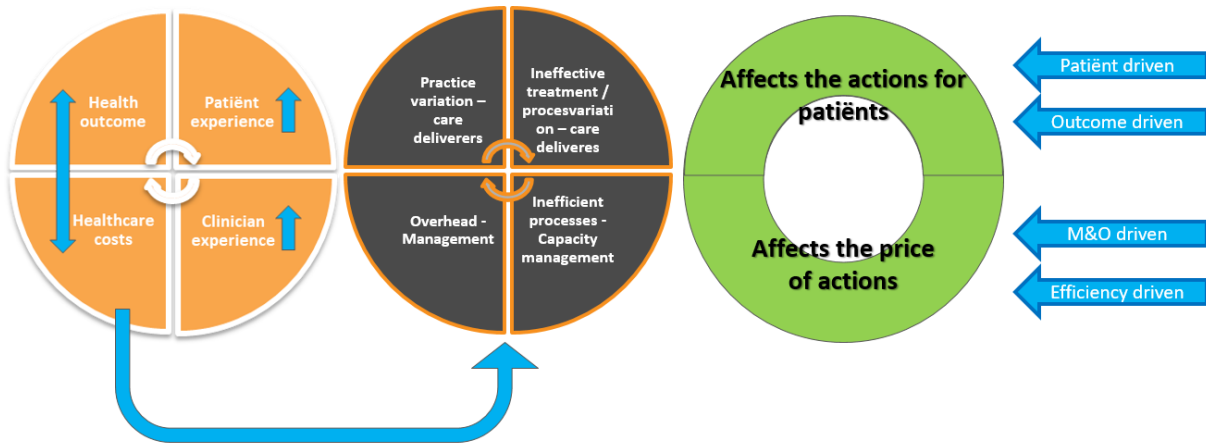
Bijlage VII – Zorgpaden dashboards (Maasstad Ziekenhuis, Nederland)

In het Maasstad Ziekenhuis worden management dashboards ontwikkeld voor een toenemend aantal zorgpaden om sturing op patiëntwaarde en capaciteit mogelijk te maken. Deze worden in de praktijk gebruikt door het behandelteam rondom de aandoening. De landingspagina van de dashboards richten zich op de vier pijlers van de Quadruple Aim: uitkomsten (klinisch en PROMs), patiëntervaringen, zorgverlenerservaring en zorgkosten (zie Figuur 9).

Figuur 9. Landingspagina management dashboard.

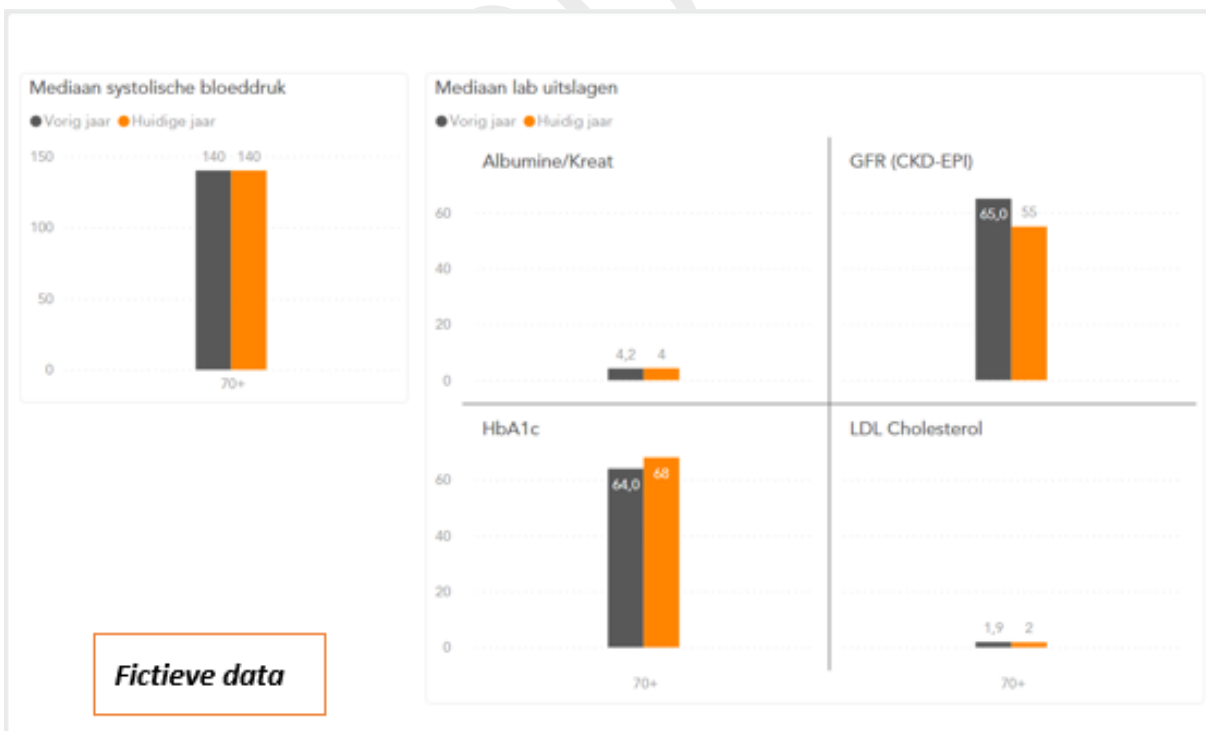


Waarop gestuurd gaat worden binnen de pijlers wordt bepaald middels de prestatie-indicatoren (KPIs) opgenomen in het jaarplan van de afdeling en/of het ziekenhuis. Hieronder valt de inzet van ziekenhuisresources, welke leiden tot kosten ten behoeve van de zorguitkomst. Dit levert de meest directe vorm van sturing op uitkomsten met inzicht in kosten op. Hiermee ondersteunt het dashboard de Planning & Control Cyclus bij de realisatie van de strategie. Ervaringsindicatoren geven inzicht in de efficiëntie van het zorgproces, en bieden handvatten voor verbeteracties. Kosten worden weergegeven in de hoofd KPI: ‘gemiddelde kosten per patiënt’ en, naar gelang verbeteracties, in bijvoorbeeld ‘Gemiddelde kosten SEH per patiënt’ of ‘Gemiddelde kosten ligduur per patiënt’.



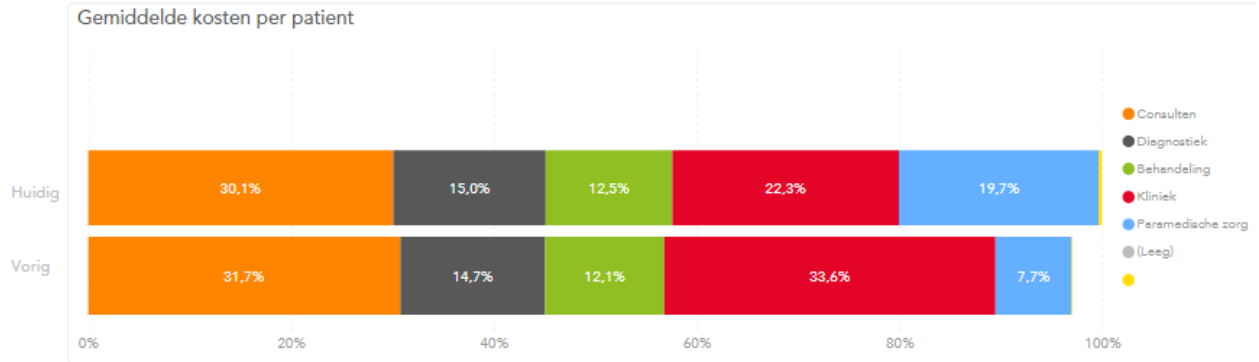
Figuur 10. Sturing via het management dashboard.

Per pijler worden middels sub-tabbladen grafieken getoond die inzicht geven in de populatie en onderliggende waarden. Er is een selectiebar waarbij het verbetersteam populatiestratificaties kan 'bouwen'. Vanuit hier kunnen periodes met elkaar vergeleken worden om de gewenste verbeteringen op te volgen. Tevens kan het stratificeren dienen om niet gewenste praktijkvariatie op te sporen. Indien nodig kan met brondata gericht gezocht worden waar en hoe de variatie tot stand is gekomen.



Figuur 11. Sub-tabblad in het management dashboard (I).

Periode 1-9-2022 t/m 25-8-2023 vs 26-8-2021 t/m 25-8-2022



activiteit hoofdgroep omschrijving	Kosten gem/pat periode	Kosten gem/pat vorige
Consulten	€ 160	€ 193
Diagnostiek	€ 80	€ 89
Behandeling	€ 66	€ 73
Kliniek	€ 119	€ 204
Paramedische zorg	€ 105	€ 47
	€ 0	€ 0
	€ 2	€ 1
Totaal	€ 531	€ 607

Fictieve data

Figuur 12. Sub-tabblad in het management dashboard (II).

Het verbetersteam bespreekt periodiek de voortgang op uitkomsten in relatie tot de ontwikkeling op kosten. Het zorgmanagement monitort van alle zorgpaden de gemiddelde kostenontwikkeling afgezet tegen de budgetverantwoordelijkheden. Hierdoor is er ruimte voor ethische afwegingen, zoals het ondersteunen van zorg met meer kosten dan de benchmark. Indicatoren die gerealiseerd zijn gaan in de follow-up op de achtergrond draaien, waardoor er ruimte komt voor het opnemen van nieuwe KPI verbeterindicatoren in het dashboard.

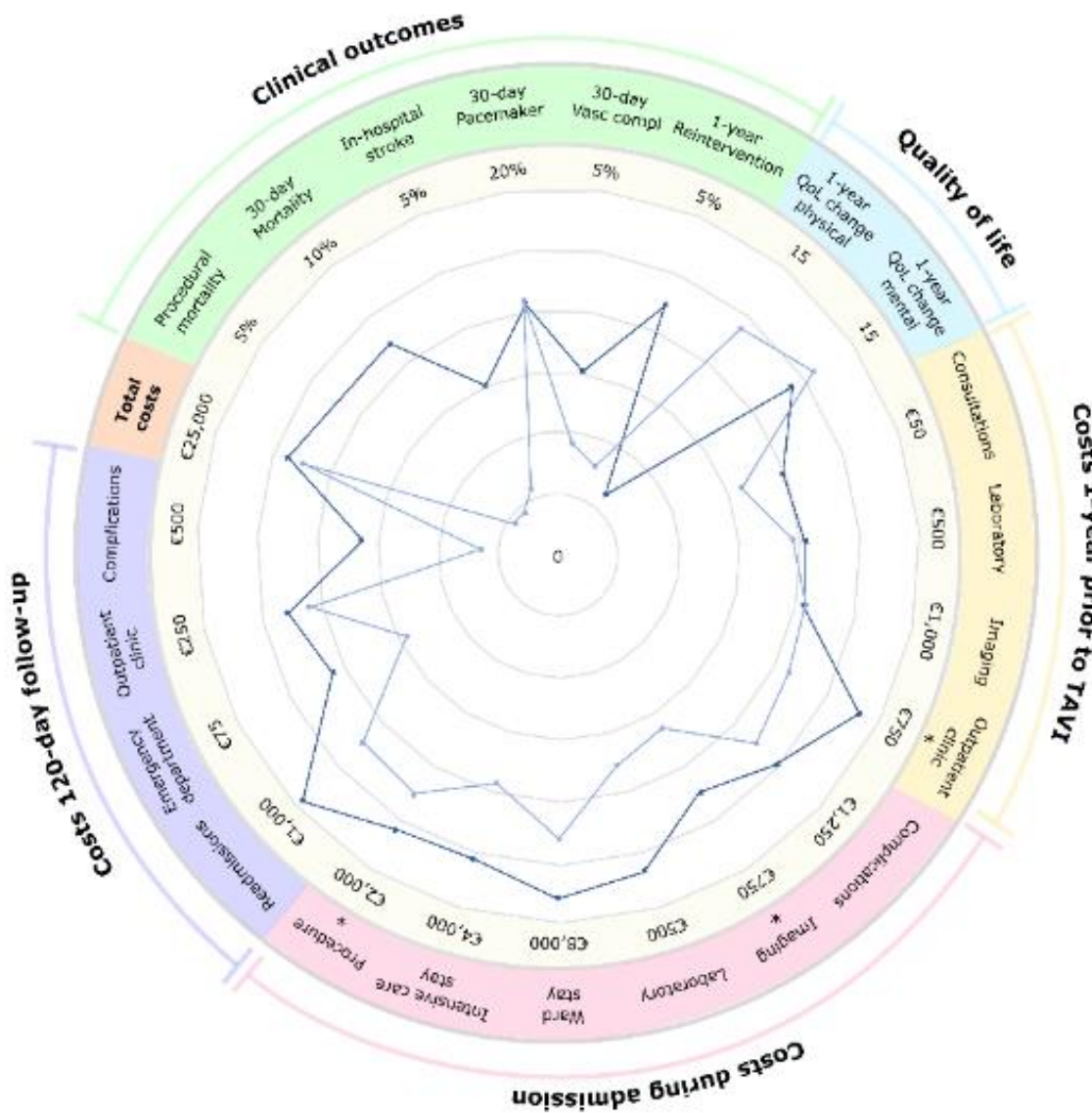
Bijlage VIII - Spider chart aortaklepperving (Catharina Ziekenhuis, Nederland)

Hoewel het monitoren van klinische uitkomsten binnen de cardiologie en -chirurgie op hoog niveau al wordt beoefend, mede dankzij de Nederlandse Hart Registratie (NHR), zijn kostenparameters doorgaans ondervertegenwoordigd. Om deze kloof te dichten is in het Catharina Ziekenhuis een spider chart ontwikkeld (Van Steenberghe et al., 2023) die ondersteunt bij de identificatie en monitoring van kostbepalende factoren, naast de traditionele klinische uitkomsten.

De spider chart is in een casestudie naar aortaklepperving met behulp van een katheter (TAVI) ontwikkeld. Relevante kostenparameters zijn systematisch geselecteerd uit ziekenhuis declaratiedata door een multidisciplinair team van cardiologen, cardio-thoracale chirurgen, onderzoekers, data-analisten, ziekenhuismanagers en patiëntvertegenwoordigers. Aanvullend is uitkomstendata gebruikt uit bestaande klinische uitkomstmonitoring vanuit de NHR.

Om data inzichtelijk en bruikbaar te maken, zijn deze gepresenteerd in een spider chart (zie Figuur 13). Deze visuele tool biedt, naast een tabel met de resultaten, een weergave om verbeteringen in kosten en klinische uitkomsten door de tijd heen te monitoren en waar nodig bij te sturen. De donkerblauwe lijn weergeeft cohort A voor implementatie van een verbeterstrategie, en de lichtblauwe lijn weergeeft cohort B na implementatie. De buitenste ring toont de tijdsperiodes die van toepassing zijn. De uitkomsten en kostendrijvers worden weergegeven in de tweede buitenste ring. De lichtgele, binnenste ring toont voor elke parameter de gebruikte maximale waarde, waarbij het centrum van het spider chart als 'nul' moet worden gezien. Het gebied tussen het maximum en 'nul' voor elke uitkomst en kostendrijver is een schaal waarop de werkelijke waarde per cohort wordt geplotted. De exacte waarden zijn te vinden in de bijbehorende tabellen. De relatie en verschillen tussen cohorten A en B kunnen hierdoor worden bepaald via de positionering van de stip. De geplottede waarden worden met elkaar verbonden met een lijn. De asterisken geven aan of er een statistisch significant verschil is tussen cohort A en cohort B. Een voorbeeld: gemiddelde kosten voor beeldvorming tijdens de opname zijn €565 voor cohort A en €415 voor cohort B. Dit verschil is statistisch significant ($P = .002$).

Figuur 13. Spider chart die klinische uitkomsten, kwaliteit van leven en kosten weergeeft van TAVI patiënten (Van Steenberg et al., 2023).



Het model wordt momenteel toegepast op meerdere zorgpaden zoals hartklep- en omleidingsoperaties. Door het gebruik van het model zijn voor de genoemde zorgpaden al enkele observaties gedaan waar een verbeterinitiatief voor wordt ontwikkeld of reeds is geïmplementeerd en is geëvalueerd.

Bijlage IX - Patient-level value analysis (Verenigde Staten)

Het doel van patient-level value analysis (McCreary et al., 2019) is het ontwikkelen van een waarde evaluatie-model op patiëntniveau. Ervaring is opgedaan binnen de orthopedie bij zowel de zorg voor enkelfracturen (McCreary et al., 2019) als knieervanging (Wise et al., 2020).

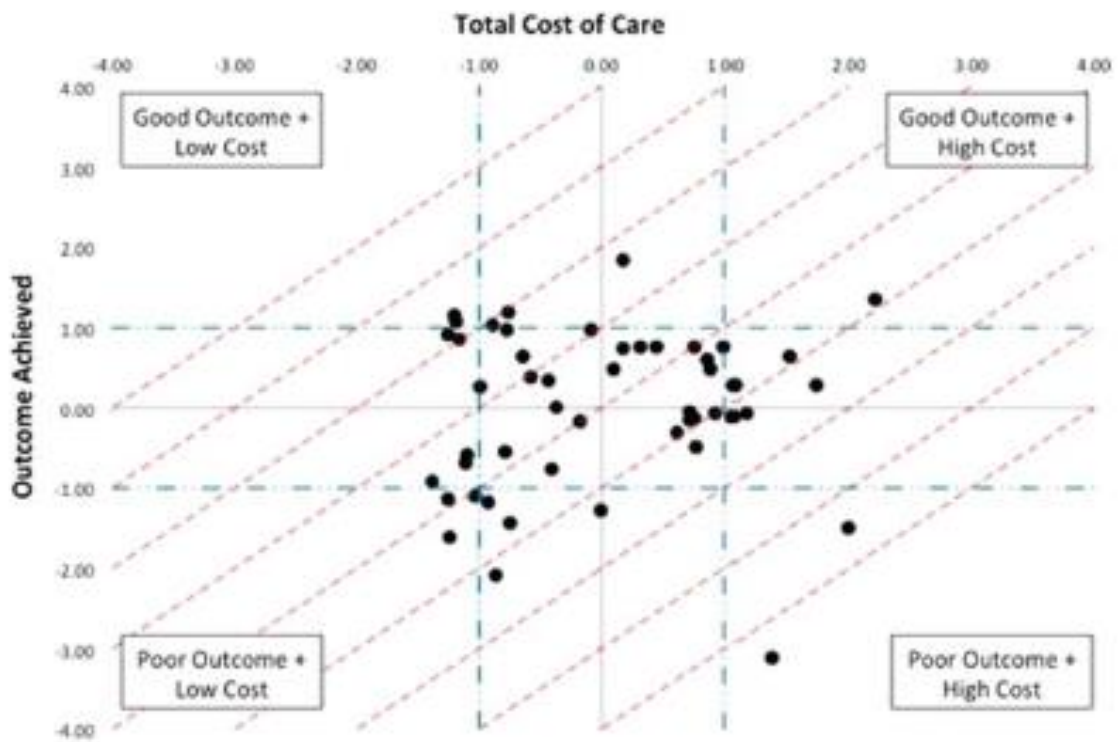
De gebruikte formule voor het bepalen van waardetoevoeging voor elke patiënt staat weergegeven in Box 4. Elke patiënt vult een specifieke PROM (bijv. de short musculoskeletal function assessment, SMFA) voor en na de procedure in, waardoor variatie in waarde binnen een verrichting wordt bepaald. Deze variatie wordt onderzocht om te bepalen welke aspecten van zorg (o.a. patiëntkenmerken, zorgpad procedures en zorgverleners) positief en negatief de geleverde waarde voorspellen. Kosten worden met behulp van TDABC bepaald. De correlatie tussen kosten en resultaat worden bepaald met behulp van de Pearson r-test.

Box 4. Formule voor het bepalen van waardetoevoeging voor de patiënt (McCreary et al., 2019).

$$Waarde = \frac{\text{uitkomst (baseline)} - \text{uitkomst (follow - up)}}{\text{kosten verkregen uit TDABC}} \times 100$$

Om te visualiseren hoe kosten en uitkomsten van elke patiënt variëren ten opzichte van het gemiddelde, wordt de uitkomstscore van elke patiënt en de totale zorgkosten omgezet in z-scores (zie Figuur 14). Een meervoudige lineaire regressie wordt uitgevoerd om te bepalen welke casemixvariabelen de waarde van de verleende zorg significant voorspellen.

Figuur 14. Totale zorgkosten afgezet tegen de patiëntuitkomst voor elke patiënt (z-scores).



Bijlage X - GGZ value of efficiency indicator (Nederland)

In de geestelijke gezondheidszorg (GGZ) is ervaring opgedaan met een formule die een efficiëntie indicator (E) oplevert voor waardetoevoeging per patiënt/behandeling (i) (zie Box 5).

Als outcome kan zowel een post-test score op een zelf-rapportage vragenlijst of een samengestelde score van meerdere zelf-rapportage vragenlijsten worden gebruikt, als een delta voor pre-to-post-test change. De score wordt eerst gestandaardiseerd en vervolgens geëxponentieerd ($z_outcome$) om boven de breukstreep een waarde van 0 tot ∞ te krijgen. Kosten kunnen gedeclareerde kosten zijn of kosten minus baten, en worden uitgedrukt in euro's met een waarde van 0 tot ∞ . GM_{kosten} betreffen de (geometrisch) gemiddelde kosten, een constante.

Een indicator waarde $E_i = 1$ (range 0 - ∞) betekent dat de behandeluitkomst en kosten 'in balans' zijn (bijvoorbeeld de gemiddelde situatie voor beiden). $E_i < 1$ betekent een ongunstige verhouding, en $E_i > 1$ een gunstige verhouding tussen uitkomst en kosten.

Box 5. Formule voor de GGZ value of efficiency indicator.

$$E_i = \frac{2^{z_outcome_i} \times GM_{kosten}}{kosten_i}$$

GM (E_i) betreft de geaggregeerde waarde voor een groep patiënten/behandelingen; zo is de E_i indicator te aggregeren tot ieder gewenst niveau (behandelaar, team, afdeling, locatie, gehele instelling of ziekenhuis). GM (E_i) geeft de verhouding van uitkomsten en kosten weer en is aanvullende informatie bij de gemiddelde uitkomst en de gemiddelde kosten van een groep behandelde patiënten.

Bijlage XI - Samen Beter Geboortezorg (Santeon ziekenhuizen, Nederland)

In het Santeon Samen Beter programma (z.d.) werken voor alle aandoeningen in de zeven betrokken ziekenhuizen verbeterteams aan het verbeteren van zorg. Aan de hand van de scorekaart worden resultaten gemeten en geanalyseerd. Dit is een continu proces van leren en verbeteren. De scorekaart bestaat uit drie onderdelen om kwaliteit van zorg inzichtelijk te maken: uitkomsten, kosten en procesindicatoren.

Het Samen Beter programma Geboortezorg streeft naar het verbeteren van de geboortezorg binnen de Santeon ziekenhuizen aan de hand van de scorekaart (zie Figuur 15). Het programma verzamelt halfjaarlijks data van 16.000 jaarlijkse klinische bevallingen om de zorg rondom zwangerschap en bevalling te analyseren en te verbeteren.

Figuur 15. Scorekaart Samen Beter programma Geboortezorg.

UITKOMST	Sterfte moeder (na bevalling) en kind (voor, tijdens en na bevalling)
	Bloedverlies (fluxus) en bloedtransfusies
	Totaal ruptuur
	Episiotomie (inknippen)
	Heropname moeder
	Opname kind op NICU
	Apgar na 5 minuten
	pH van navelstreng
	Vroeggeboorte, spontaan en niet-spontaan
	Zuurstoftekort kind
	Neonatale infectie
	Borstvoeding, intentie en bij ontslag
KOSTEN	Verblijfsduur moeder en kind
	Aantal keizersneden, binnen en buiten kantooruren
	Operatieduur keizersneden, bruto en netto
	Aantal polikliniekbezoeken
	Zwangerschapsbegeleiding
	Inzet geneesmiddelen: pijnstilling tijdens bevalling
PROCES	Pijnbestrijding, binnen en buiten kantooruren
	Postpartum controles: polikliniekbezoeken en telefonische consulten

Vanaf 2017 is de verbetercyclus vier keer doorlopen. De resultaten van het geboortezorg programma worden transparant gepubliceerd als een interactieve infographic. De belangrijkste verbeterinitiatieven richten zich op het terugdringen van het aantal ongeplande keizersneden, het aantal totaal rupturen, episiotomieën en fluxus tijdens de bevalling, het verbeteren van de uitkomsten

voor het kind, en het efficiënter inrichten van de zorg rondom de bevalling. Resultaten tonen aan dat structurele vergelijking en reflectie op de werkwijzen kan leiden tot een vermindering van ongewenste interventies en complicaties.

De verrassende lessen uit de verbetercycli zijn dat geduld en bewustwording cruciale elementen zijn in de geboortezorg. De keuzes van gynaecologen en verloskundigen moeten na de bevalling gezamenlijk worden besproken voor maximale leerervaring.

Toekomstige verbetercycli zullen zich blijven richten op deze onderwerpen, evenals op het continueren en verbreden van de samenwerking in de gehele geboortezorgketen, meer nationaal en internationaal benchmarken, uitbreiding van de indicatoren, en het opzetten van een gezamenlijke ontwikkelagenda en het gebruik van PROMs.

VERTROUWELIJK

Appendix - Toelichting begrippen, meetinstrumenten en wetenschappelijke methoden

Methoden voor het sturen op uitkomsten én kosten

Samen beslissen

Continu verbeteren

Benchmarken

Het doel van benchmarken is leren en verbeteren door middel van onderlinge vergelijking van zorgprocessen en uitkomsten. Het geeft in het ideale geval een eenduidig antwoord op de vraag: Wat doe jij anders dan ik en levert dat gunstigere of ongunstigere behandelresultaten op? Bij benchmarken vergelijken we dus geaggregeerde behandeluitkomsten van groepen patiënten en proberen we te achterhalen op welke verschillen in aanpak van de behandeling uitkomstverschillen zijn terug te voeren.

De uitkomst van de zorg wordt echter niet alleen bepaald door de interventie of behandeling. Andere zaken zoals de aard en complexiteit van de aandoening, de ernst van de begintoestand van de client, de context van de patiënt of de context van de behandeling zijn ook van invloed op de uitkomst. En - niet te vergeten – welke uitkomst en de manier waarop die wordt vastgesteld, is ook van invloed op het behandelresultaat en op de uitkomst van een vergelijking. Bij gebruik van een heel onnauwkeurig of ongevoelig meetinstrument blijven verschillen tussen instellingen verborgen die er wellicht wel zijn. Voor een leerzame of “eerlijke” vergelijking en bruikbare benchmark resultaten gelden in principe dezelfde uitgangspunten als in wetenschappelijk therapie-onderzoek, zoals RCT's of vergelijkende cohortstudies. We benchmarken bijvoorbeeld bij voorkeur binnen welomschreven aandoeningen (geen appels met peren vergelijken). Er spelen ook dezelfde bedreigingen van de validiteit van de bevindingen als bij ander therapie-effectonderzoek. De twee belangrijkste zijn selectie van patiënten en de invloed van confounding variabelen (Cox et al., 2009). Als niet voor 100% van de behandelde patiënten uitkomstgegevens beschikbaar zijn kan onbedoelde of intentionele selectie bij het meten zijn opgetreden (de Beurs et al., 2019). Van *intentionele* selectie is sprake wanneer voormetingen nagelaten worden bij patiënten waarbij men een ongunstig behandelresultaat verwacht of wanneer nametingen worden vermeden bij patiënten met een ongunstig resultaat. Ongeacht of de behandelresultaten nu intentioneel worden opgepoetst, of dat selectie ongewild optreedt, het gevolg is dat de gemeten patiëntengroep mogelijk niet representatief is voor de gehele behandelde populatie. Confounding treedt op wanneer verschillen in uitkomst tussen groepen patiënten niet alleen toegeschreven kunnen worden aan verschillen in de behandeling, maar bijvoorbeeld veroorzaakt kunnen zijn door achtergrond variabelen of door verschillen in ernst of complexiteit van de te

vergelijken groepen. Het kan dan nodig zijn om statistische correcties op de data toe te passen, zoals casemix-correctie (Iezzoni, 2013). Dat is vooral relevant voor een juiste interpretatie van benchmarkverschillen die worden aangetroffen bij relatief kleine groepen, zoals de caseload van behandelaars. Het is denkbaar dat binnen een behandelteam een net aangekomen stagiair de minder complexe patiënten krijgt toegewezen, waarbij het makkelijker is om gunstige behandelresultaten te behalen. De stagier lijkt dan op het eerste gezicht effectiever dan een ervaren professional die meer complexe problematiek te behandelen kreeg. Casemix-correctie kan deze confounding tegengaan. Bij het onderling benchmarken van behandelresultaten van instellingen of ziekenhuizen met een vergelijkbare casemix is correctie echter in principe niet nodig.

Samenvattend: benchmarken is behandelresultaten vergelijken bij homogene groepen patiënten om te leren wat het best werkt bij wie onder welke omstandigheden. Voor een leerzame vergelijking is het belangrijk een gevoelig meetinstrument te gebruiken, een hoge meetrespons te bereiken en de mogelijke invloed van confounding variabelen mee te nemen bij de analyse van uitkomstgegevens.

Value-based payments

Multi-Criteria Decision Analysis

Kosten-effectiviteitsanalyse

KEA biedt een beoordeling van de impact van een nieuwe medische interventie op kosten en gezondheidseffecten in vergelijking met een geschikte comparator, meestal standaardzorg. De impact van een interventie kan worden uitgedrukt in de incrementele kosten-effectiviteitsratio (ICER) - de incrementele kosten gedeeld door de incrementele gezondheidseffecten van interventie A in vergelijking met B.

Kosten-utiliteitsanalyse

KUA is een speciale vorm van KEA waarbij de effectiviteit wordt gemeten in termen van quality-adjusted life years (QALYs), die de kwaliteit van leven combineren met de levensduur (Nederland, 2015). Het gebruik van QALYs maakt het mogelijk om doelmatigheid van heel verschillende vormen van zorg toch op een vergelijkbare manier te beoordelen.

Meetinstrumenten voor het meten van uitkomsten

Patiënt-gerapporteerde behandeluitkomst (PROMS)

Centraal in Value-based health care staat de uitkomst van de zorg *zoals die wordt beoordeeld door de patiënt*. Patient Reported Outcome Measures (PROMs) zijn gestandaardiseerde zelf-rapportage vragenlijsten om informatie te verzamelen over de gezondheidstoestand, functioneren en de kwaliteit van leven van patiënten vanuit hun eigen perspectief. PROMs worden voor, tijdens en na de behandeling door patiënten ingevuld om veranderingen in de gezondheidstoestand van patiënten te meten over de tijd.

PROMs worden steeds belangrijker bij het bepalen van de kwaliteit van de zorg. Het verzamelen en analyseren van PROMs kan helpen om de gezondheidszorg te verbeteren. Bij individuele patiënten kunnen PROMs worden gebruikt om inzicht te geven in de behoeften en prioriteiten van de patiënt en door te helpen bij het identificeren van levensgebieden van functioneren waar verbetering wenselijk en mogelijk is. Door PROM-gegevens van groepen patiënten samen te nemen kunnen ze gebruikt worden bij het evalueren van de resultaten van interventies en behandelingen. Geaggregeerde gegevens kunnen zo gebruikt worden om de effectiviteit van behandelingen te evalueren.

ICHOM

Door de value-based health care beweging in de VS is al in 2011 een internationale organisatie opgezet om het meten van behandeluitkomsten te ondersteunen en internationaal te harmoniseren: het International Consortium for Health Outcomes Measurement (ICHOM; www.ichom.org). Er zijn werkgroepen georganiseerd om voor tal van aandoeningen standaardsets samen te stellen met (onderdelen van) veelgebruikte meetinstrumenten. In totaal zijn er inmiddels 40 standaardsets beschikbaar, georganiseerd in 13 hoofdgroepen. Zo zijn er standaardsets voor de hoofdgroepen hart- en vaatziekten, ziekten vanwege erfelijke afwijkingen, maag- en darmziekten, besmettelijke ziekten, psychiatrische ziekten, spier- en skelet ziekten, neurologische ziekten, oncologische ziekten, oogziekten, nier- en blaasziekten, moeder-, perinatale- en zuigelingenziekten, en geriatrische ziekten. Voor het GGZ veld zijn er sets voor Depressie en Angst, Dementie en is een set voor verslaving in de maak (zie voor een totaal overzicht van de standard sets de ICHOM website www.ichom.org). Een standaard set bestaat doorgaans uit stoornisspecifieke klachtenlijsten, meetinstrumenten voor functioneren, zoals onderdelen van de WHODAS en vragenlijsten voor gezondheidsgerelateerde kwaliteit van leven (HR-QoL). Er is ook een flinke kruisbestuiving tussen de meetsets. Zo bevatten meetsets voor somatische ziekten ook vaak items of instrumenten voor het meten van symptomen van depressie of angst. De internationale harmonisatie van meetinstrumenten moet het onderling vergelijken van behandeluitkomsten faciliteren, want benchmarken op “value” voor de patient is een expliciet doel van ICHOM.

Generieke versus stoornis specifieke meetinstrumenten

Generieke meetinstrumenten zijn bedoeld om algemene aspecten van de gezondheid en het functioneren van een persoon te meten, onafhankelijk van een specifieke diagnose. Voorbeelden hiervan zijn de SF-36 en de EQ-5D. Stoornisspecifieke meetinstrumenten zijn bedoeld om symptomen en functionele beperkingen te meten die horen bij een specifieke stoornis, zoals een Visual Analog Scale voor pijn bij reuma, de Beck Depression Inventory voor depressie, de PCL5 voor posttraumatische stress stoornis. Sommige instrumenten combineren generiek en stoornisspecifiek meten met een generieke en specifieke module. Een voorbeeld is de European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (QLQ)-C30 (Aaronson et al., 1993) en de EORTC QLQ-BR23 (Sprangers et al., 1996) voor het meten van kwaliteit van leven bij borstkanker.

De belangrijkste voordelen van generieke meetinstrumenten zijn de brede toepasbaarheid en vergelijkbaarheid van meetresultaten tussen diverse patiëntengroepen. Generieke meetinstrumenten meten algemene aspecten van gezondheid en functioneren en kunnen gebruikt worden om de gezondheid en het functioneren van een breed scala aan patiënten te meten, ongeacht de diagnose of aandoening. Generieke meetinstrumenten worden vaak gebruikt in kosten-effectiviteitsanalyses van gezondheidsinterventies, omdat ze een breder perspectief bieden dan stoornis-specifieke meetinstrumenten. De brede toepasbaarheid komt echter met een prijs: generieke instrumenten als algemeen welbevinden zijn vaak minder gevoelig voor veranderingen tijdens de behandeling dan stoornis-specifieke instrumenten, zoals pijn, angst, of depressie. Hieraan gerelateerd is dat generieke instrumenten ook minder gedetailleerde informatie leveren over specifieke symptomen en relevante specifieke aspecten van functioneren.

Hierdoor zijn generiek instrumenten minder geschikt zijn voor het monitoren van de effecten van interventies op een specifieke aandoening bij een individuele patiënt. Generieke instrumenten zijn daardoor ook beperkt bruikbaar voor klinische besluitvorming. Specifieke instrumenten zijn bruikbaar omdat ze juist de specifieke symptomen en functionele beperkingen meten die relevant zijn voor de behandeling van een specifieke aandoening,

Zowel stoornis-specifiek als generiek meten lost dit dilemma op: de stoornis-specifieke meetresultaten zijn goed bruikbaar om verloop en voortgang van de behandeling met de patiënt te bespreken en de efficiëntie van de behandeling in het oog te houden; geaggregeerde specifieke meetresultaten kunnen gebruikt worden voor vergelijken van uitkomsten binnen een specifiek aandoeningsgebied (benchmarking); geaggregeerde generieke meetresultaten kunnen gebruikt worden voor vergelijken over patiëntgroepen heen en leren van elkaar over aandoeningsgebieden heen (ook benchmarking).

EQ-5D

De EQ-5D is een meetinstrument dat de gezondheidstoestand van patiënten meet aan de hand van vijf dimensies van gezondheid – namelijk mobiliteit, zelfzorg, dagelijkse activiteiten, pijn en ongemak en angst en depressie. Het instrument wordt veel gebruikt in kosteneffectiviteitsanalyses en Health Technology Assessments en biedt een nuttige maat voor het evalueren van de effectiviteit van behandelingen. Doordat de EQ-5D de generieke gezondheidstoestand van de patiënt meet, kan de effectiviteit van behandelingen voor verschillende aandoeningen met elkaar vergeleken worden.

PROMIS

Een ander internationaal initiatief dat het waard is om te bespreken is het Patient-Reported Outcomes Measurement Information System (PROMIS) (www.promis.org). PROMIS is meetinstrumentarium dat sinds 2004 wordt ontwikkeld door de National Institutes of Health in de VS om een aantal problemen op het gebied van meten van gezondheid te adresseren. De drie belangrijkste problemen zijn de enorme wildgroei aan meetinstrumenten, de te hoge belasting voor patiënten vanwege te veel en te lange vragenlijsten en de aanzienlijke kosten die zijn verbonden aan commercieel verkrijgbare vragenlijsten.

PROMIS tracht deze problemen op te lossen door de ontwikkeling van een gestandaardiseerde set van vragenlijsten die gebruik maken van moderne meetmethoden, zoals op Item Response Theory (IRT) gebaseerde computer adaptieve testen (CAT), en deze meetinstrument gratis of tegen een beperkte vergoeding beschikbaar te stellen. De Item Respons Theorie gaat uit van het idee dat items van een vragenlijst elk een uniek ernstniveau representeren. De combinaties van antwoorden op items en het ernstniveau van die items bepaalt de uiteindelijke testuitslag. Een voorbeeld: instemmen met de stelling “ik ben in paniek” levert een hogere ernstsore op dan instemmen met de stelling “ik voel me gespannen”. Voor een CAT wordt een itembank met vragen samengesteld, waarmee een unidimensioneel construct wordt gemeten (bijvoorbeeld angst, pijn, vermoeidheid, mobiliteit, welbevinden). Bij afname wordt uit de itembank een eerste item aangeboden en vervolgens wordt, op basis van eerder gegeven antwoorden, de best passende vervolgvraag gekozen en aangeboden, net zo lang tot er een voldoende betrouwbaar meetresultaat is verkregen. Hierdoor kunnen betrouwbare en nauwkeurige resultaten worden verkregen met zo min mogelijk vragen (doorgaans zijn 5 tot 8 vragen voldoende voor een betrouwbaar meetresultaat). Dit leidt tot een meer efficiënte meting van symptomen, functioneren en HRQoL dan met traditionele vragenlijsten mogelijk is. Bovendien zijn de vragenlijsten ontworpen om breed toepasbaar te zijn bij verschillende aandoeningen en onderzoeksgebieden, waardoor de vergelijkbaarheid van resultaten tussen verschillende patiëntengroepen wordt vergroot. CAT vereist digitale aanbidding van een vragenlijst en verbinding

met de “CAT-engine” die vervolgvragen kiest. Dat is niet overal mogelijk of beschikbaar en daarom zijn er naast de CAT versies er ook korte versies van de PROMIS meetinstrumenten gemaakt die zich voor traditionele afname lenen.

Het uiteindelijke doel van PROMIS sluit nauw aan bij de doelen van VBHC: de gezondheidszorg verbeteren door middel van een meer patiëntgerichte benadering, waarbij de nadruk ligt op het meten van de gezondheidsresultaten die belangrijk zijn voor patiënten. Door het gebruik te maken van de gestandaardiseerde en gevalideerde PROMIS vragenlijsten kunnen zorgverleners de gezondheid van patiënten beter volgen, hun zorgplannen aanpassen en de kwaliteit van de zorg verbeteren.

Inmiddels omvat het PROMIS-systeem in de VS meer dan 200 vragenlijsten die zijn ontwikkeld en getest in samenwerking met patiënten en zorgverleners. Een deel van de vragenlijsten (ongeveer 30) is vertaald en psychometrisch onderzocht door de Dutch-Flemish PROMIS group www.dutchflemishpromis.nl en dus beschikbaar in het Nederlands. Een volgende set van 30 vertalingen zijn onderweg en worden nog onderzocht op hun psychometrisch kenmerken. Alle PROMIS meetinstrumenten gebruiken een op de algemene (Amerikaanse) bevolking genormeerde meetschaal van T scores met een gemiddelde van 50 (SD = 10) en een praktische scorering van ongeveer 20 tot 80 (in de algemene bevolking heeft 99.7% een score in de range van -3 tot +3 SD). Uniformiteit van meetresultaten presenteren wordt nagestreefd: een hogere score betekent altijd meer van het gemeten construct, dus meer pijn, vermoeidheid of andere klachten; maar ook meer kwaliteit van leven. Uitslagen op meetinstrumenten worden weergegeven tegen een achtergrond met een kleurverloop van groen via geel naar rood om een duidelijke betekenis aan de score te verlenen. Groen is klachtenvrij, goed functionerend, hoge kwaliteit van leven; rood het tegenovergestelde. Voor de meeste klachtenlijsten geldt dat een score onder de 55 als normaal wordt beschouwd, 55 tot 60 is mild verhoogd, 60 tot 70 matig, en 70 of hoger is ernstig. Voor kwaliteit van leven wordt de score omgedraaid, maar tegen dezelfde kleur achtergrond weergegeven.

PREMs

QALY

VBHC-QALY

Uniforme meetschalen (gestandaardiseerde scores en percentiel scores)

De wildgroei aan meetinstrumenten heeft nog een tweede nadelig bijeffect. Testuitslagen van klachtenlijsten of HR-QoL meetinstrumenten hebben elk hun eigen meetschaal, waardoor de getalswaarde van de testuitslag niet direct duidelijk maakt hoe hoog of hoe laag een score is. Ruwe

scores moeten opgezocht worden in normtabellen om er een betekenis aan te verlenen en dat is bewerkelijk en foutgevoelig. Dit is op te lossen door een gemeenschappelijke meetschaal of common metric) voor klachtenlijsten af te spreken. Om scoring van meetinstrumenten te harmoniseren worden testuitslagen bij voorkeur uitgedrukt in standaardscores (T scores met als referentiegroep de algemene bevolking of klinische percentielscores) (de Beurs et al., 2022). Er zijn talloze publicaties met oversteektabellen om ruwe scores op bestaande meetinstrumenten naar de T score meetschaal om te zetten (zie onder andere www.prosetastone.org) (Choi et al., 2014; Schalet et al., 2021).

Gebruik van T scores is ook door het PROMIS initiatief omarmd en de T score is hierboven beschreven. Klinische percentielscores geven nuttige aanvullende informatie, want ze illustreren voor een patiënt waar hij of zij staat ten opzichte van andere patiënten die in behandeling komen; een percentiel score van 75 geeft aan dat 75% van de andere patiënten een lagere score heeft. T scores en percentielscores helpen bij een juiste interpretatie van de scores en hun gebruik faciliteert de communicatie met andere professionals en met cliënten over de testuitslag. Bijkomend voordeel van genormaliseerde T scores is dat ze de eigenschap van een intervalschaal bezitten en dat verschillen van de voor- en nameting betekenis hebben (wat overigens niet geldt voor percentielscores, die je niet van elkaar af kunt trekken).

Behandeluitkomst als continue en als binaire variabele

Uitkomst van zorg vanuit het gezichtspunt van de patiënt kan gemeten worden met gestandaardiseerde meetinstrumenten, zoals zelf-invul vragenlijsten waarmee de ernst van de aandoening, het niveau van functioneren of gezondheidsgerelateerde kwaliteit van leven (HR-QoL) kan worden vastgesteld. De uitkomst wordt vaak uitgedrukt in een verschilscore tussen een voor- en nameting. Dit is omdat de score op zo'n meetinstrument na behandeling voor een belangrijk deel bepaald wordt door de hoogte van voormeting. De verschilscore drukt dan beter uit welke gezondheidswinst er met de behandeling bereikt is dan alleen de score bij nameting (Gottman & Rushe, 1993; Myers et al., 2010). De verschilscore is een continue variabele die bij de meeste meetinstrumenten een normaalverdeling heeft. <<Ik weet niet hoe het verhaal verder zal gaan maar de keuze van een continue maat voor uitkomst (bijvoorbeeld een gestandaardiseerde verschilscore) heeft consequenties voor een value-indicator.>>

Een alternatief voor de verschilscore dat ook vaak gebruikt wordt is een binaire uitkomstmaat van een klinisch relevante eindstatus bij afsluiting van de behandeling. Voorbeelden zijn “hersteld”, “in remissie”, “vrij van pijn”, of “abstinent” (in de verslavingsszorg). Voor de status van remissie of herstel kunnen grensscores op zelf-rapportage instrumenten gehanteerd worden. Uitkomst kan ook nog wat verfijnder gecategoriseerd worden met meerdere niveaus zoals “hersteld”, “verbeterd”,

“onveranderd”, en “verslechterd”. Voor deze categorieën zijn (statistische) criteria opgesteld (Jacobson et al., 1999; Jacobson & Truax, 1991), zoals de Reliable Change Index (RCI, de mate van verandering die meer is dan de meet-onnauwkeurigheid van het gebruikte meetinstrument) of Minimal Detectable Change en Minimal Important Change (MDC en MIC, (de Vet et al., 2006) en Clinical Significant Change (CSC, klinisch betekenisvolle verandering, doorgaans een score voorbij een grenswaarde die de overgang markeert van ziek naar gezond (Jacobson et al., 1999; Kendall et al., 1999). Beide criteria kunnen los van elkaar worden gebruikt: wel/niet betrouwbare verandering in de richting van gezondheid of wel/niet de grensscore overschreden en CSC bereikt. Maar beiden kunnen ook worden gecombineerd om de vier niveaus te onderscheiden die hierboven zijn beschreven (de Beurs et al., 2018). Bij geaggregeerde gegevens kan men dan bijvoorbeeld het resultaat van de behandeling uitdrukken in de proportie “herstelden” of de proportie “herstelden en betrouwbaar verbeteren”).

VERTROUWELIJK

Referenties

- Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute* 1993;85(5):365-376. <https://doi.org/10.1093/jnci/85.5.365>
- Akpan A, Roberts C, Bandeen-Roche K, et al. Standard set of health outcome measures for older persons. *BMC Geriatrics* 2018;18(1):36. <https://doi.org/10.1186/s12877-017-0701-3>
- Banta D. The development of health technology assessment. *Health Policy* 2003;63(2):121-132. [https://doi.org/10.1016/s0168-8510\(02\)00059-3](https://doi.org/10.1016/s0168-8510(02)00059-3)
- Beck da Silva Etges AP, Nabi J, Geubelle A, Ouriques Martins S, Polanczyk CA. Analytic solutions to support value-based health care: the ischemic stroke care pathway. *NEJM Catalyst* 2022.
- Cattel D, Eijkenaar F, Ahaus K, Van de Laar M. Bundelbekostiging in de zorg is mogelijk, ondanks belemmeringen. *ESB* 2021;106:4794.
- Choi SW, Schalet B, Cook KF, Cella D. Establishing a common metric for depressive symptoms: linking the BDI-II, CES-D, and PHQ-9 to PROMIS Depression. *Psychological assessment* 2014;26(2): 513-527. <https://doi.org/10.1037/a0035768>
- Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML.. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report—Part II. *Value in Health* 2009;12(8):1053-1061. <https://doi.org/10.1111/j.1524-4733.2009.00601.x>
- De Beurs E, Den Hollander-Gijsman ME, Van Rood YR, et al. Routine outcome monitoring in the Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clinical Psychology & Psychotherapy* 2011;18(1):1-12. <https://doi.org/10.1002/cpp.696>
- De Beurs E, Boehnke J, Fried EI. Common measures or common metrics? A plea to harmonize measurement results. *Clinical Psychology and Psychotherapy* 2022;29(5):1755-1767. <https://doi.org/10.1002/cpp.2742>
- De Beurs E, Warmerdam EH, Oudejans SCC, et al. Treatment outcome, duration, and costs: a comparison of performance indicators using data from eight mental health care providers in the Netherlands. *Administration and Policy in Mental Health and Mental Health Services Research* 2018;45(2):212-223. <https://doi.org/10.1007/s10488-017-0818-x>

- De Beurs E, Warmerdam L, Twisk JW. Bias through selective inclusion and attrition: representativeness when comparing providers performance with routine outcome monitoring data. *Clinical Psychology & Psychotherapy* 2018;26(2):430-439. <https://doi.org/10.1002/cpp.2364>
- De Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health and Quality of Life Outcomes* 2006;4(1):54.
- DICA. Dutch Institute for Clinical Auditing. Verkregen 15-08-2023 via: <https://dica.nl/>
- Experiment Uitkomstindicatoren. Samen beslissen met uitkomstinformatie. Santeon 2023. Verkregen via: https://issuu.com/santeon/docs/santeon_magazine_voor_issuu
- Foster F, Kok MR, Van der Kooij J, et al. Dealing with time estimates in hospital cost accounting: integrating fuzzy logic into time-driven activity-based costing. *PharmacoEconomics* 2023;7:593-603. <https://doi.org/10.1007/s41669-023-00413-2>
- Gongora-Salazar P, Rocks S, Fahr P, Rivero-Arias O, Tsiachristas A. The use of Multicriteria Decision Analysis to support decision making in healthcare: an updated systematic literature review. *Value in Health* 2023;26(5):780-790. <https://doi.org/https://doi.org/10.1016/j.jval.2022.11.007>
- Gottman JM, Rushe RH. The analysis of change: issues, fallacies, and new ideas. *Journal of Consulting and Clinical Psychology* 1993;61(6):907.
- Hanselaar T, Van der Kemp M, Wiersma V, Van der Linde M, Bos WJ, Franx A. Aan de slag met waardegedreven zorg. Linnean Initiatief 2023. Verkregen via: <https://www.linnean.nl/inspiratie/bibliotheek/handlerdownloadfiles.ashx?idnv=2331573>
- Health E. Implementing Value-Based Health Care in Europe: Handbook for Pioneers (Director: Gregory Katz). 2020. <https://eithealth.eu/wp-content/uploads/2020/06/Implementing-Value-Based-Healthcare-In-Europe.pdf>
- Health Intelligence Platform Santeon z.d. Verkregen via: <https://santeon.nl/project/health-intelligence-platform/>
- ICHOM. Marking VBHC real – experiences from Estonian stroke patient pathway pilot. 2022. Verkregen via: <https://www.ichom.org/nordic-healthcare-group-blog/making-vbhc-real-experiences-from-estonian-stroke-patient-pathway-pilot-presentation-at-the-ichom-2022-conference/>
- ICHOM. International Consortium for Health Outcomes Measurement z.d. Verkregen via: <https://connect.ichom.org/patient-centered-outcome-measures/>
- Iezzoni L. (Ed.). *Risk adjustment for health care outcomes* (4th ed.). Health Administration Press. 2013. <http://www.ache.org/hap.cfm>
- Infographic geboortezorg. Santeon z.d. Verkregen via: <https://santeon.nl/aandoening/geboortezorg/>

Integraal Zorgakkoord. Samen werken aan gezonde zorg. Den Haag: Ministerie van Volksgezondheid, Welzijn en Sport 2022. Verkregen via:

<https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2022/09/16/integraal-zorgakkoord-samen-werken-aan-gezonde-zorg/integraal-zorg-akkoord.pdf>

Jacobson NS, Roberts LJ, Berns SB, McGlinchey JB. Methods for defining and determining the clinical significance of treatment effects: description, application, and alternatives. *Journal of Consulting and Clinical Psychology* 1999;67(3):300-307. <https://doi.org/10.1037//0022-006x.67.3.300>

Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* 1991;59(1):12-19. <https://doi.org/10.1037//0022-006x.59.1.12>

Kaplan RS, Anderson SR. Time-Driven Activity-Based Costing. SSRN, November 2023. Verkregen via: <http://dx.doi.org/10.2139/ssrn.485443>

Kendall PC, Marrs-Garcia A, Nath SR, Sheldrick RC. Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology* 1999;67(3):285-299. <https://doi.org/10.1037/0022-006X.67.3.285>

Keel G, Savage C, Rafiq M, Mazzocato P. Time-driven activity-based costing in health care: a systematic review of the literature. *Health Policy* 2017;121(7):755-763. <https://doi.org/10.1016/j.healthpol.2017.04.013>

Kuhrij LS, Wouters MW, Van den Berg-Vos RM, De Leeuw FE, Nederkoorn PJ. (2018). The Dutch Acute Stroke Audit: Benchmarking acute stroke care in the Netherlands. *European Stroke Journal* 2018;3(4):361-368. <https://doi.org/10.1177/2396987318787695>

Leijten FRM, Struckmann V, Van Ginneken E, et al. The SELFIE framework for integrated care for multimorbidity: development and description. *Health Policy* 2018a;122(1):12-22. <https://doi.org/10.1016/j.healthpol.2017.06.002>

Leijten FRM, Hoedemakers M, Struckmann V, et al. Defining good health and care from the perspective of persons with multimorbidity: results from a qualitative study of focus groups in eight European countries. *BMJ Open* 2018b;8:e021072. <http://dx.doi.org/10.1136/bmjopen-2017-021072>

Linnean z.d. Verkregen via: <https://www.linnean.nl/inspiratie/kennispaginas/1927360.aspx?t=Waardegedreven-zorg>

Looman WM, Huijsman R, Fabbriotti IN. The (cost-)effectiveness of preventive, integrated care for community-dwelling frail older people: a systematic review. *Health and Social Care in the Community* 2019;27:1-30.

- McCreary DL, Dugarte AJ, Vang SBA, et al. Patient-level value analysis: an innovative approach to optimize care delivery. *Journal of Orthopaedic trauma* 2019;33:S49-S52. <https://doi.org/10.1097/BOT.0000000000001624>
- Myers JL, Well A, Lorch RF. *Research design and statistical analysis* (3th ed.). Routledge. 2010.
- Programma Uitkomstgerichte Zorg. Den Haag: Ministerie van Volksgezondheid, Welzijn en Sport z.d. Verkregen via: <https://www.uitkomstgerichtezorg.nl/>
- Porter ME. What is value in health care? *The New England Journal of Medicine* 2010;363:26. <https://doi.org/10.1056/NEJMp1011024>
- Porter ME, Lee TH. The strategy that will fix health care. *Harvard Business Review* 2013.
- PROMIS z.d. Verkregen via: <https://www.promishealth.org/>
- Rutten-van Mólken M, Karimi M, Leijten F, et al. Comparing patients' and other stakeholders' preferences for outcomes of integrated care for multimorbidity: a discrete choice experiment in eight European countries. *BMJ Open* 2020;10:e037547. <http://dx.doi.org/10.1136/bmjopen-2020-037547>
- Santeon Samen Beter. Santeon z.d. Verkregen via: <https://santeon.nl/onze-aanpak/samen-beter/>
- SAZ waardegedreven zorg programma. Samenwerkende Algemene Ziekenhuizen z.d. Verkregen via: <https://saz-ziekenhuizen.nl/thema/waardegedreven-zorg/>
- Schalet BD, Lim S, Cella D, Choi SW. Linking scores with patient-reported health outcome instruments: a validation study and comparison of three linking methods. *Psychometrika* 2021;86(3):717-746. <https://doi.org/10.1007/s11336-021-09776-z>
- SELFIE, z.d. Verkregen via: <https://www.selfie2020.eu/>
- Sprangers M, Groenvold M, Arraras JI, et al. The European Organization for Research and Treatment of Cancer breast cancer-specific quality-of-life questionnaire module: first results from a three-country field study. *Journal of Clinical Oncology* 1996;14(10):2756-2768.
- Terwee CB, Vonkeman HE, Zuidgeest M, namens de werkgroep PROMs en PREMs. Het menu van generieke PROMs: advies. Linnean Initiatief 2019. Verkregen via: <https://www.linnean.nl/inspiratie/bibliotheek/handlerdownloadfiles.ashx?idnv=1501426>
- Van den Hout WB, Tielen N, Van der Kooij J, et al. Waardegedreven zorg met uitkomsten én kosten. Linnean Initiatief 2021. Verkregen via: <https://www.linnean.nl/inspiratie/bibliotheek/handlerdownloadfiles.ashx?idnv=2026606>
- Van Steenbergen GJ, Schulz DN, Slingerland SR, et al. Introduction of a new method to monitor patient-relevant outcomes and costs: using a quality improvement project in transcatheter aortic valve implantation care as an example. *Quality Management in Health Care* 2023. <https://doi.org/10.1097/QMH.0000000000000401>

- Verberne WR, Das-Gupta Z, Allegretti AS, et al. Development of an international standard set of value-based outcome measures for patients with chronic kidney disease: a report of the International Consortium for Health Outcomes Measurement (ICHOM) CKD working group. *American Journal of Kidney Diseases* 2019;73(3):372-384.
<https://doi.org/10.1053/j.ajkd.2018.10.007>
- Versteegh MM, Vermeulen KM, Evers SMAA, De Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. *Value in Health* 2016;19(4):343-352.
- Werkgroep, C., Nierschade, Uitkomstgerichte, & Zorg. (2021). *Uitkomstenset Chronische Nierschade [onder embargo]*.
- Werkgroep Datavisualisatie/Dashboard. In 7 stappen naar een waardegedreven zorg dashboard. Linnean Initiatief z.d. Verkregen via: <https://www.linnean.nl/werkgroepen/dashboard>
- WHO. Health technology assessment.. Verkregen 08-08-2023 via: https://www.who.int/health-topics/health-technology-assessment#tab=tab_1
- WHO. Key components of a well functioning health system. WHO, Geneva 2010.
- Wise KL, McCreary DL, Parikh HR, et al. Factors associated with a second opioid prescription fill in total knee arthroplasty. *The Journal of Arthroplasty*. 2020;35(6):S163-S167.
<https://doi.org/10.1016/j.arth.2020.03.001>
- WKKGZ z.d. Verkregen via: <https://www.rijksoverheid.nl/onderwerpen/kwaliteit-van-de-zorg/wet-kwaliteit-klachten-en-geschillen-zorg>
- ZE&GG, z.d. Verkregen via: <https://zorgevaluatiegepastgebruik.nl/>
- Zorginstituut Nederland. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. Diemen: Zorginstituut Nederland. 2016. Verkregen via: <https://www.zorginstituutnederland.nl/over-ons/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg>