

# Beoordeling van de klinische effectiviteit in HTA

## Inleiding

Om de impact van een technologie te kunnen beoordelen is uitgebreide informatie nodig die weerspiegelt wat er waarschijnlijk zal gebeuren in een gezondheidszorgstelsel of maatschappij. Een goede analyse vraagt om de bijdrage van deskundig advies en methoden uit de verschillende vakgebieden.

HTA-beoordelingen zijn vergelijkende analyses, die de bestaande standaardzorg vergelijken met de nieuwe technologie om te zien welke waarde de nieuwe technologie zou bieden (de zogeheten 'meerwaarde'). HTA-beoordelingen worden door landen, regio's en ziekenhuizen op verschillende manieren geïmplementeerd. Ze kijken allemaal naar het gezondheidsprobleem in hun lokale context en beoordelen dan de behandeling voor de indicatie die de toezichthoudende organen zijn overeengekomen voor het geneesmiddel.

Binnen die indicatie onderzoeken HTA-instanties de beschikbare gegevens om te evalueren hoe goed de behandeling kan werken in vergelijking met de beste standaardzorg (in termen van veiligheid en klinische effectiviteit). Sommige HTA-instanties beoordelen ook de kosten en kosteneffectiviteit van een geneesmiddel. En sommige instanties voeren formele beoordelingen uit van de ethische, organisatorische, maatschappelijke en juridische aspecten, terwijl andere deze zaken impliciet meenemen in hun beoordeling.

De 'meerwaarde' van een technologie wordt door elke HTA-organisatie op een andere, veelzijdige manier beoordeeld. Conclusies over de 'meerwaarde' van een technologie kunnen per HTA-organisatie verschillen. Het Europese netwerk voor HTA

(EUnetHTA) heeft een structuur ontwikkeld voor het beoordelen van de 'meerwaarde', het zogeheten HTA-kernmodel.<sup>®.1</sup> Het HTA-kernmodel behelst negen domeinen<sup>®</sup>:

1. Gezondheidsprobleem
2. Technische beschrijving van technologie
3. Veiligheid
4. Klinische effectiviteit
5. Kosten en economische evaluatie (kosteneffectiviteit)
6. Ethische analyse
7. Organisatorische aspecten
8. Sociaal-maatschappelijke aspecten
9. Juridische aspecten

EUnetHTA heeft de beoordeling van de eerste vier domeinen gedefinieerd als een 'beoordeling van relatieve effectiviteit' – waarin bijvoorbeeld een nieuwe behandeling wordt vergeleken met de bestaande behandeling(en).

## **Beoordeling van de klinische effectiviteit in HTA**

Bij een beoordeling van klinische effectiviteit wordt het effect onderzocht van een nieuwe technologie op de gezondheid van patiënten in een standaard klinische setting vergeleken met die van de huidige standaardzorg. De impact die een dergelijke technologie heeft op gezondheid wordt meestal geanalyseerd door middel van een nadere bestudering van de gezondheidsuitkomsten. Patiënten willen toegang tot nieuwe geneesmiddelen die:

- de als 'slecht' beschouwde uitkomsten (zoals hartaanvallen, ziekenhuisopnamen en bijwerkingen) verminderen, **en/of**
- de als 'goed' beschouwde uitkomsten (zoals beter functioneren en pijnvrije dagen) vermeerderen.

Gedurende een beoordeling van klinische effectiviteit maken

HTA-instanties gebruik van gevestigde methoden van geassocieerde vakgebieden in de geneeskunde. De evaluatie van klinische effectiviteit wordt vooral uitgevoerd met principes die zijn geleend van de epidemiologie en geneeskunde ('klinische epidemiologie').

Goede beoordeling van klinische effectiviteit is gefundeerd op vier principes:

1. Informatie zoeken,
2. Relevante vragen stellen,
3. Verschillen begrijpen, **en**
4. Verschillen op waarde schatten.

## **Informatie zoeken**

HTA-instanties maken gebruik van klinische informatie om in te schatten welke gezondheidsuitkomsten patiënten zouden kunnen ondervinden wanneer ze een nieuw geneesmiddel krijgen. Eerst moeten ze echter besluiten hoe ze de informatie gaan verzamelen. Er zijn drie verschillende manieren waarop HTA-instanties klinische informatie over nieuwe technologieën kunnen verkrijgen:

1. Bestudering van bestaande informatie over de prestaties van het geneesmiddel,
2. Uitvoering van een nieuw onderzoek om informatie te verzamelen en de prestaties van het geneesmiddel te evalueren in een levensechte situatie, **of**
3. Artsen en patiënten ('ervaringsdeskundigen') vragen wat hun verwachtingen van het geneesmiddel zijn.

HTA-instanties passen vaak een combinatie van deze benaderingen toe. Bijvoorbeeld:

- Ze kunnen voor hun eigen onafhankelijke evaluaties en analyses gebruikmaken van informatie van de vergunninghouder voor de betreffende technologie.
- In geval van ontbrekende informatie kan deskundig advies

een oplossing bieden, bijvoorbeeld om erachter te komen of veranderingen in uitkomsten op de korte termijn (zoals daling van het cholesterolgehalte) zouden kunnen voorspellen wat de uitkomsten op de langere termijn zijn (zoals het vermijden van ziekenhuisopnamen).

HTA-instanties geven zelden opdracht voor nieuwe onderzoeken, omdat het meestal te lang duurt om een onderzoek op te zetten en goed te laten keuren. In sommige gevallen laten de verantwoordelijke instanties vergoeding van een geneesmiddel voorwaardelijk toe, op voorwaarde dat er nog aanvullende informatie zal worden ingewonnen (dit is vergelijkbaar met de verlening van een voorwaardelijke handelsvergunning waar meer informatie voor moet worden verzameld). Het risico dat een nieuw geneesmiddel in de werkelijkheid slechter presteert dan verwacht, kan dan worden gedeeld door de houder van de handelsvergunning en de verantwoordelijke instantie via mechanismen voor prijsonderhandelingen of andere wijzigingen in de voorwaarden voor vergoeding (zoals verdere beperkingen met betrekking tot de patiëntenpopulatie die in aanmerking komt voor vergoeding), terwijl patiënten toch meer directe toegang tot de technologie krijgen.

## **Relevante vragen stellen**

Wanneer de klinische effectiviteit van de nieuwe gezondheidstechnologie wordt beoordeeld, moet de HTA-instantie alle bijbehorende uitkomsten zorgvuldig afwegen. Kennis van deze uitkomsten is van belang om de relevante vragen te kunnen stellen over de effectiviteit van de technologie.

Het wordt steeds duidelijker dat uitkomsten die artsen belangrijk vinden niet altijd als het meest belangrijk worden beschouwd door patiënten. Daarom is het belangrijk dat patiënten betrokken worden in de ontwerpfase van onderzoeken, om er voor te zorgen dat informatie wordt verzameld over de uitkomsten die patiënten belangrijk vinden. Zo is bijvoorbeeld in de afgelopen jaren erkent dat kwaliteit van leven een

belangrijke uitkomst voor patiënten is. Dit heeft geleid tot de ontwikkeling van specifieke methoden voor de creatie van maten voor kwaliteit van leven en zogeheten 'door patiënten gemelde uitkomsten' (patient-reported outcomes, of PRO's) in klinische onderzoeken.

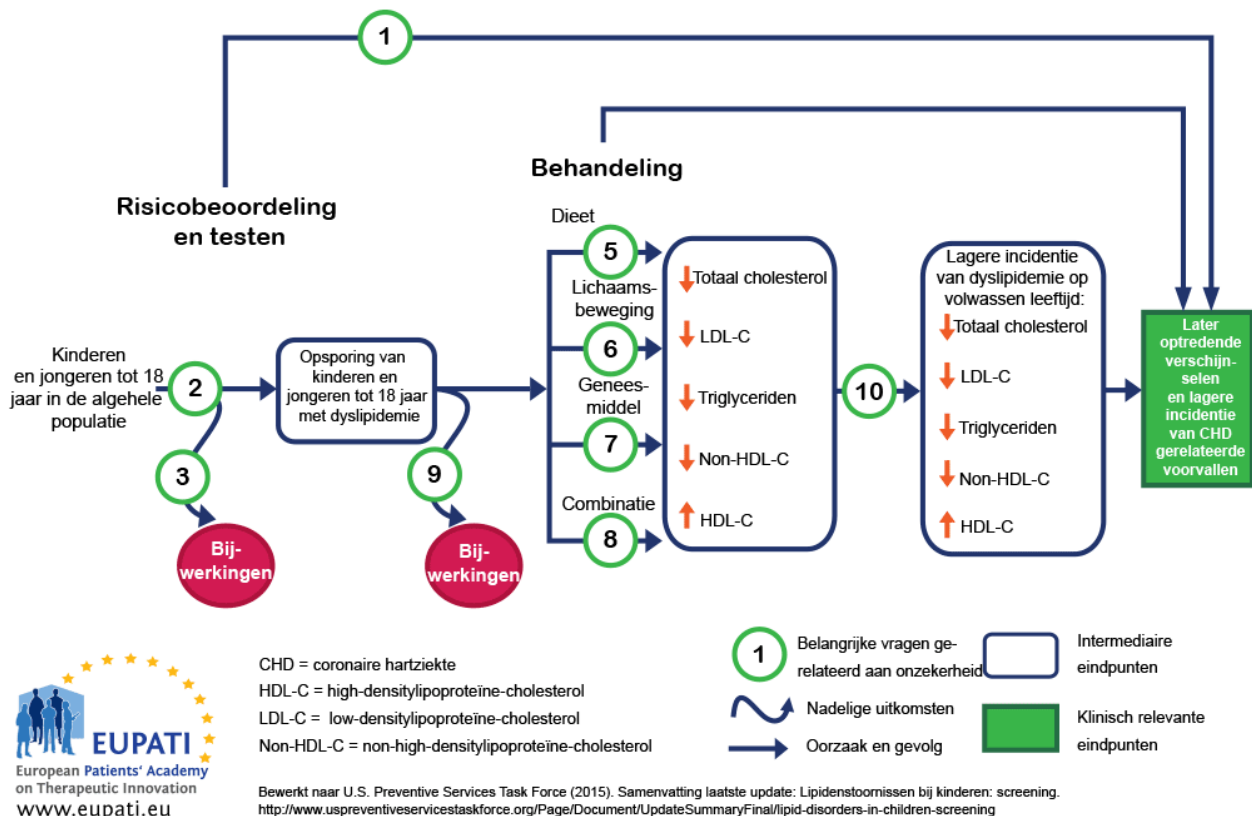
Een manier om er voor te zorgen dat alle belangrijke uitkomsten van een bepaalde technologie worden onderzocht, is door gebruik te maken van een analytische structuur. Bijvoorbeeld een stroomschema zoals in figuur 1.<sup>2</sup> Analytische structuren zijn nuttig om alle uitkomsten met betrekking tot een interventie in beeld te brengen, en om aan te geven waar sprake is van onzekerheden.

In de analytische structuur in figuur 1:

- Oorzaak en gevolg worden met pijlen weergegeven.
  - Gebogen pijlen staan voor schadelijke uitkomsten.
- Uitkomsten met betrekking tot gezondheidsverbetering (zoals daling van het sterftecijfer) zijn met rechthoeken weergegeven.
  - De rechthoeken met scherpe hoeken laten klinisch relevante eindpunten zien (eindpunten die door de patiënt worden opgemerkt, zoals pijn op de borst).
  - De rechthoeken met afgeronde hoeken laten tussentijdse eindpunten zien, zoals surrogaateindpunten (die niet kunnen worden opgemerkt door de patiënt, zoals daling van het cholesterolgehalte in het bloed).
- De sleutelvragen met betrekking tot onzekerheden worden met een getal aangeduid.

# Voorbeeld van een analytisch raamwerk

Ondertitel indien relevant



Deze analytische structuur werd toegepast om de sterke en zwakte punten in het bewijsmateriaal vast te stellen met betrekking tot de effectiviteit van screenen van kinderen en jong volwassenen (adolescenten) op dyslipidemie (stoornissen in de vetstofwisseling) als onderdeel van de standaard eerstelijnszorg. Dyslipidemie is een belangrijke risicofactor voor coronaire hartziekte (CHZ).

De sleutelvragen met betrekking tot deze analytische structuur zijn:

- **Sleutelvraag 1:** Is screenen op dyslipidemie bij kinderen/adolescenten effectief als het gaat om uitstel van de start en vermindering van de incidentie van CHZ (coronaire hartziekten)?
- **Sleutelvraag 2:** Hoe accuraat kan worden gescreend op dyslipidemie met betrekking tot het identificeren van kinderen/adolescenten met verhoogd risico op CHZ en andere uitkomsten?

- **Sleutelvraag 3:** Wat zijn de ongewenste effecten van screening (zoals fout-positief, fout-negatief, bestempelen)?
- **Sleutelvraag 4:** Wat is de effectiviteit van medicatie, eetpatroon, beweging en combinatietherapie bij kinderen en adolescenten met betrekking tot het terugdringen van de incidentie van dyslipidemie op volwassen leeftijd en uitstel van de start en vermindering van de incidentie van CHZ en andere uitkomsten (waaronder optimale leeftijd voor start van de behandeling)?
- **Sleutelvraag 5-8:** Wat is de effectiviteit van medicatie, eetpatroon, beweging en combinatietherapie met betrekking tot de behandeling van dyslipidemie bij kinderen/adolescenten (waaronder het toenemende voordeel van behandeling van dyslipidemie in de kindertijd)?
- **Sleutelvraag 9:** Wat zijn de ongewenste effecten van medicatie, eetpatroon, beweging en combinatietherapie bij kinderen/adolescenten?
- **Sleutelvraag 10:** Zorgt verbetering van dyslipidemie in de kindertijd voor verlaging van het risico op dyslipidemie op volwassen leeftijd?
- **Sleutelvraag 11:** (niet weergegeven): Wat zijn de kosten van screening op dyslipidemie bij kinderen zonder symptomen?

## **Verschillen tussen de uitkomsten doorgronden**

Als alle belangrijke uitkomsten eenmaal zijn geïdentificeerd, kan vergelijking van de effecten van een nieuwe technologie met de standaardzorg en andere behandelingen nog steeds problemen opleveren. Uitkomsten zijn misschien op verschillende manieren gemeten, of de uitkomsten van twee technologieën lijken op het eerste gezicht gelijk totdat bij nader onderzoek verschillen naar boven komen.

Wanneer de belangrijke geïdentificeerde uitkomsten moeilijk

meetbaar zijn, of nooit eerder gemeten zijn, moeten wetenschappers zorgvuldig een maat creëren die in een onderzoek kan worden gereproduceerd. Zo wil een patiënt misschien weten of een geneesmiddel hem/haar weer aan het werk of uit bed kan helpen. Wetenschappers kunnen een numerieke schaal voor pijnbeoordeling maken voor patiënten met pijn in de lage rug. In andere gevallen (bijvoorbeeld waar in een onderzoek wijzigingen in een laboratoriumparameter worden gemeten) moet deze wijziging worden vertaald naar een maatstaf die belangrijker is voor patiënten, zoals weer aan het werk kunnen gaan.

Soms vinden toezichthoudende instanties die geneesmiddelen goedkeuren het voldoende als de fabrikant van het geneesmiddel het effect van een nieuw geneesmiddel aantoont via een uitkomst op de korte termijn (zoals verlaging van de bloeddruk). Een HTA-instantie moet die kortetermijn-uitkomst vertalen naar relevantere uitkomsten voor de patiënt (zoals preventie van vroegtijdig overlijden).

Sommige uitkomsten lijken misschien voor de hand te liggen, maar zijn bij nader inzien moeilijk te interpreteren. Verlaging van het risico op vijfjaarssterfte (overlijden binnen vijf jaar) met 50% wil bijvoorbeeld niet zeggen dat het geneesmiddel vroegtijdig overlijden kan voorkomen. Het kan gewoon betekenen:

- verhoging van de levensverwachting van 4,9 naar 5,1 jaar (of nog minder, van 4,99 jaar naar 5,01 jaar) bij sommige patiënten, **of**
- genezing van de ziekte in een paar gevallen maar geen levensverlenging voor alle anderen.

Zelfs als er verschillen worden waargenomen die belangrijk zijn voor patiënten, kunnen die nog moeilijk te interpreteren zijn. Bijvoorbeeld, uit onderzoek kan blijken dat een nieuw geneesmiddel het risico op ziekenhuisopname door infectie vermindert met 33%. Dit kan echter verschillende dingen



betekenen. Dit kan betekenen dat:

- 33 van de 100 mensen die het geneesmiddel gebruiken die anders in het ziekenhuis zouden zijn opgenomen, nu ziekenhuisopname hebben kunnen vermijden (dit wordt een **absolute** risicovermindering genoemd), **of**
- de kans op ziekenhuisopname is met 33% afgenomen ten opzichte van de kans op ziekenhuisopname zonder geneesmiddel (dit wordt een **relatieve** risicovermindering genoemd). Als de kans op ziekenhuisopname bij afwezigheid van het nieuwe geneesmiddel 3 op de 1000 is, dan wordt deze door een reductie van 33% verlaagd tot 2 op de 1000. Dit wil zeggen dat 1 op de 1000 mensen die het geneesmiddel gebruiken voordeel ondervindt. Dat is heel iets anders dan de 33 op de 100 mensen die voordeel ondervinden in het bovengenoemde voorbeeld.

Een laatste uitdaging voor het doorgronden van de verschillen tussen een nieuwe gezondheidstechnologie en de standaardzorg is het gebruik en misbruik van statistische toetsen. Met behulp van statistische toetsen kunnen onderzoekers weten of de gevonden verschillen waarschijnlijk reëel zijn. Dit wordt vaak vermeld in de vorm van een p-waarde. P-waarden geven echter niet aan wat de omvang van het verschil is, en of dat verschil belangrijk is voor patiënten. Dit betekent dat p-waarden meestal niet veel nut hebben voor patiënten en zorgverleners die beslissingen moeten nemen.

Andere statistische maten zijn onder meer betrouwbaarheidsintervallen. Betrouwbaarheidsintervallen zijn nuttiger omdat ze een idee geven van de omvang van het verschil tussen de nieuwe gezondheidstechnologie en de standaardzorg. Betrouwbaarheidsintervallen weerspiegelen ook eventuele onzekerheden over de berekende omvang van het verschil. Bijvoorbeeld, van een nieuw geneesmiddel is gemeld dat het de kans op een hartaanval in de toekomst verlaagt met 33% (met een 95%-betrouwbaarheidsinterval van 5% tot 45%) ten opzichte van de huidige kans op een hartaanval.

# Verschillen op waarde schatten

Inzicht in hoe de verschillen tussen uitkomsten moeten worden gezien en gewaardeerd vormt de laatste uitdaging. Als een geneesmiddel het leven verlengt met 0,2 jaar moeten HTA-instanties nog steeds weten:

- hoeveel waarde een patiënt hecht aan 0,2 jaar langer leven in het licht van verwachte bijwerkingen en andere problemen,
- of alle patiënten ongeveer dezelfde voordelen ondervinden of dat er grote verschillen zijn tussen patiënten, **en**
- of alle patiënten deze voordelen hetzelfde waarderen.

Een nieuw geneesmiddel dat de levensverwachting met gemiddeld 0,2 jaar verhoogd, zou anders worden beoordeeld als het bij sommige patiënten werkt en niet bij anderen ten opzichte van een middel dat bij alle patiënten, met slechts kleine verschillen tussen patiënten, werkt.

Er zijn verschillende manieren die inzicht verschaffen in de relatieve waarde die patiënten en zorgverleners hechten aan verschillen in gezondheidsuitkomsten. Een daarvan is kwalitatief onderzoek, zoals enquêtes of focusgroepen, bedoeld om inzicht te verschaffen in welke uitkomsten patiënten het belangrijkste vinden. Een ander mechanisme is kwantitatief onderzoek op basis van enquêtes onder patiënten, waarmee precieze waarderingscijfers kunnen worden gegeven voor het belang dat aan verschillende gezondheidstoestanden wordt gehecht.

Kortom, in een beoordeling van klinische effectiviteit moeten de volgende kwesties aan de orde komen:

1. Hoe alomvattend was de informatie?
2. Hoe accuraat is de informatie?
3. Ontbreken er zaken?
4. Hoe begrijpelijk is de informatie?

# Referenties

1. *HTA Core Model*. Geraadpleegd op 8 september 2016, op <http://www.eunethta.eu/hta-core-model>
2. U.S. Preventive Services Task Force (2015). Final Update Summary: Lipid Disorders in Children: Screening. Geraadpleegd op 7 december 2015, op <http://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/lipid-disorders-in-children-screening>

A2-6.03.1-v1.1