



Universiteit  
Leiden

The Netherlands

## Zorgevaluatie is experimenteren voor rechtvaardige zorg

Poolman, R.W.

### Citation

Poolman, R. W. (2020). *Zorgevaluatie is experimenteren voor rechtvaardige zorg*. Leiden: Universiteit Leiden. Retrieved from <https://hdl.handle.net/1887/3458864>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3458864>

**Note:** To cite this publication please use the final published version (if applicable).

Prof. dr. Rudolf W. Poolman

# Zorgevaluatie is experimenteren voor rechtvaardige zorg



Universiteit  
Leiden

Bij ons leer je de wereld kennen

# Zorgevaluatie is experimenteren voor rechtvaardige zorg

Oratie uitgesproken door

**Prof. dr. Rudolf W. Poolman**

bij de aanvaarding van het ambt van bijzonder hoogleraar orthopedie  
in het bijzonder zorgevaluatie  
aan de Universiteit Leiden  
op 21 september 2020



**Universiteit  
Leiden**





doen. Ik vind het belangrijk dat dingen op de juiste manier gebeuren. Vroeger op het schoolplein vocht ik tegen de jongens die in mijn ogen verkeerde dingen deden.. Mensen om mij heen hadden dan commentaar op de vorm en minder op de inhoud. Mijn moeder heeft zich zorgen gemaakt of ik niet in het gevang zou komen. Later, toen ik aan het studeren was voor mijn orthopedie examens stond er veel in mijn studieboeken wat anders was dan hoe mijn leermeesters het deden. Ik heb veel tijd verdaan aan het uitpluizen van de referenties. Bleek het om studies te gaan met 10 patiënten..... en dat stond in een leerboek als de waarheid. Ook heb ik een keer AT5 gebeld als jonge assistent in opleiding omdat ik boos was dat wij patiënten met een gebroken heup naar België over moesten plaatsen. Ik had geadviseerd dat zij een Belgische ambulance in Amsterdam konden komen filmen. Ik moest toen al bij de Raad van Bestuur verschijnen. Daar was ik me over aan het opwinden, ik wilde geen halve waarheden leren. Hier begon mijn zoektocht naar rechtvaardigheid. Dit heb ik daarna, tijdens mijn fellowship in Canada verder uitgewerkt, wat geresulteerd heeft in mijn proefschrift.

Het uitpluizen van dingen, uitzoeken hoe iets werkt, zit al van jongs af aan in mij. Als klein jongetje haalde ik radio's en fietsen uit elkaar om te kijken hoe die werkten. Helaas kon ik ze daarna niet altijd meer goed in elkaar zetten. Maar door mijzelf vragen te stellen, door deze te evalueren met experimenten en daarna te implementeren ben ik gekomen waar ik nu ben. Niet al mijn experimenten lukten, juist daardoor leerde ik.

Zo is ook zorgevaluatie is één groot experiment.

Zorgevaluatie bestaat uit een cyclus van agenderen, evalueren en implementeren.

Ik zal u uitleggen hoe ik denk dat we de gezondheidszorg kunnen verbeteren met behulp van zorgevaluatie. Hiervoor gebruik ik een aantal praktijkvoorbeelden.

Mijn leerstoel dekt de volgende zinnen uit de artseneed. De eed die ik in 1998 aflegde, een eed die iedere dokter in de wereld aflegt.

“Ik zal de geneeskundige kennis van mijzelf en anderen bevorderen.”

“Ik erken de grenzen van mijn mogelijkheden.”

“Ik zal mij open en toetsbaar opstellen en ik ken mijn verantwoordelijkheid voor de samenleving.”

“Ik zal de beschikbaarheid en toegankelijkheid van de gezondheidszorg bevorderen.”

Deze onderwerpen uit de artseneed kunnen we waarmaken met zorgevaluatie, dus agenderen, evalueren en implementeren!

## De zorgevaluatie cyclus

### 1. Agenderen

Beginnen met de vraag te stellen: Wat weten we *nog* niet?

*Praktijkvoorbeeld: 2 patiënten met een gebroken heup:*

Mevrouw Pietersen, 75 jaar, ontvlucht haar huis in Zandvoort voor de sleur en stapt op haar e-bike ten tijde van de Coronacrisis. Ondanks dat het niet druk is op de weg komt zij ten val waarbij zij een gebroken heup oploopt. Een ambulance brengt haar naar het ziekenhuis in Haarlem alwaar zij wordt geopereerd.

Mevrouw Janssen, 76 jaar, is met vakantie in Friesland, geboekt bij Booking.com. Zij rijdt ook op zo'n e-bike, komt ten val met als gevolg ook een gebroken heup en in het ziekenhuis van Leeuwarden volgt een operatie.

De beide dames zijn geopereerd en hebben een kophals prothese gekregen.

Maar hoe zijn zij geopereerd?

De ene dame heeft een operatie gehad via de zijkant van de heup en de andere dame via de achterkant van de heup.

Uit de LROI – Landelijke Registratie Orthopedische Implantaten – komt naar voren dat dit de 2 meest voorkomende benaderingen zijn.

Waarom zijn er verschillende benaderingen?

Deze zijn ontstaan in de loop der jaren. Ontwikkeld in het verleden en regionaal bepaald. Zo komt de direct laterale benadering (benadering van de zijkant) van het vasteland van Europa en vooral uit Midden Europa.

De achterste benadering (posterieur) komt uit de Angelsaksische landen.

Welke benadering een operateur kiest, heeft vooral te maken met zijn leermeester. Welke benadering hij/zij heeft geleerd tijdens de opleiding zal hij /zij ook toepassen.

Beide benaderingen zijn goede benaderingen, maar ze hebben beiden ook voor- en nadelen. Zo maak je bij de benadering van opzij meer spieren los waardoor deze kunnen beschadigen en daardoor minder goed werken. Voordeel is dat de heup goed in de kom blijft zitten en er minder gemakkelijk kan uitschieten. De achterste benadering heeft als voordeel dat deze de spieren minder beschadigt maar dat de heup makkelijker uit de kom kan schieten.

Er is dus twijfel wat de beste benadering is. Terwijl het zo lijkt dat als er een operatie verricht wordt deze techniek wel goed moet zijn uitgezocht. Nogmaals, beide benaderingen zijn goed, alleen is er een onbeantwoorde vraag: welke operatie doet het beter?

Zo zijn er nog veel onbeantwoorde vragen, ook in de zorg die nu standaard geleverd wordt. Het is goede zorg, alleen weten we nog van veel dingen niet wat de beste zorg is.

Deze twijfel over wat de beste zorg is komt tot uiting door de praktijkvariatie die er voor veel behandelingen is. Praktijkvariatie betekent dat er verschil in behandeling zit tussen verschillende dokterspraktijken. Als heel duidelijk is wat een betere behandeling is, verwachten we dat die praktijkvariatie vermindert.

Om meer structuur te brengen in de twijfel, verzamelen we onbeantwoorde vragen en vatten deze samen in een onderzoeksagenda. Bij dit proces van twijfel structureren zijn artsen, patiëntenorganisaties en verzekeraars betrokken. Zij prioriteren samen welke onbeantwoorde vragen het eerste beantwoord moeten worden.

Hierdoor ontstaat een agenda die niet statisch is maar continu verandert omdat vragen worden beantwoord (in Nederland of in het buitenland). Om deze continue ontwikkelingen bij te houden, wordt de agenda in de toekomst dynamisch en niet pas na vijf jaar herzien. Er komen vragen bij en er vallen

vragen af omdat deze beantwoord zijn. Dat is goed omdat de agenda dan de steeds veranderende wereld om ons heen kan volgen. De vragen blijven er zo toedoen.

De volgende stap in de cyclus van zorgevaluatie is evalueren  
**2. Evalueren**

Het uitzoeken van onbeantwoorde vragen.

Elke vraag heeft zijn eigen onderzoeksmethode nodig. Om een causaal verband aan te tonen zal variatie en een vorm van experimenteren nodig zijn. Traditioneel is dit een gerandomiseerde trial (Randomized Controlled Trial RCT) waarbij door loting bepaald wordt wie welke behandeling krijgt omdat hierbij bekende - maar ook onbekende - voorspellende factoren met grote kans gelijk verdeeld worden tussen de groepen. Prognostische factoren zijn voorspellers van hoe mensen het gaan doen, zoals leeftijd, gewicht en de aanwezigheid van chronische ziekten. Die wil je in de groepen, die je met elkaar vergelijkt, in balans hebben. En dat lukt het beste door de randomisatie of loting.

De uitdaging bij gerandomiseerde trials ligt in het werven van deelnemers. Patiënten en dokters vinden het lastig dat het lot beslist welke behandeling er wordt gekozen. Toch bestaan die behandelingen “in het wild” nu naast elkaar. Patiënten willen graag zekerheid. Als het lot bepaalt welke behandeling zij krijgen en de dokter twijfelt over wat de beste behandeling is, maakt dit de patiënt onzeker. Voor een patiënt is het prettiger om een dokter te hebben die zekerheid uitstraalt en zegt te weten wat het beste werkt.

De ethische grondslag om een experiment aan te bieden, is klinische equipoise: de onzekerheid binnen een groep medici (niet slechts 1) welke behandeling het beste werkt. Dit is een belangrijk onderdeel van de ontwikkeling van een onderzoeksagenda. Over alle vragen moet equipoise bestaan.

Hoe komt het dan dat we niet meer experimenteren? Want dan zouden vragen sneller beantwoord worden. Dit kan te maken hebben met het gegeven dat artsen denken te weten wat het beste werkt. Zo zijn zij opgevoed door hun opleider die het ook zo heeft geleerd.

Anekdote: de moeder van een vriend sneed bij het braden van een rollade altijd de kontjes van de rollade af. Had ze zo geleerd van haar moeder. Zijn vader vond dat altijd heel vreemd. Hij vroeg het een keer aan zijn schoonmoeder. Die zei toen: Ja, dat klopt maar dat moest ik doen omdat mijn braadpan niet groot genoeg was.

In het onderwijs aan studenten en artsen is er sinds de introductie van de Evidence Based Medicine meer aandacht voor het uitzoeken wat het beste werkt. Toch bestaan er nog veel onbeantwoorde vragen.

Deze vragen zijn nog niet beantwoord omdat de traditionele RCT minder geschikt is, bijvoorbeeld omdat er grote aantallen patiënten nodig zijn om het antwoord te vinden. Het is dan een uitdaging om zo'n RCT uit te voeren. Gaat dit niet eenvoudiger worden met behulp van Kunstmatige intelligentie (AI)?

Zijn er misschien alternatieve methoden om causaal verband aan te tonen? In het verleden is veel onderzoek verricht waarbij alleen observaties werden gedaan: er werd gekeken hoe mensen het deden na een behandeling. Deze studies hoeven we niet weg te gooien maar kunnen als basis dienen om goede vragen te stellen: hypothesevormende onderzoeken. De vraag kan dan definitief beantwoord worden met een experiment. Wat voor soort experimenten te doen, is een uitdaging.

Laten we even terugkeren naar de 2 ongelukkige dames die gevallen waren met hun fiets. Mevrouw Janssen en Mevrouw Pietersen.

Zij zijn in een soort natuurlijk experiment gekomen. De ene dame viel in de buurt van een ziekenhuis waar de direct laterale benadering wordt uitgevoerd en de andere dame in de buurt van een ziekenhuis waar de posterolaterale – achterste – benadering wordt gedaan op basis van traditie. Het toeval bepaalt welke benadering zij krijgen. Als het toeval bepaalt welke behandeling iemand krijgt is er sprake van een experiment en kunnen er meer causale conclusies uit worden getrokken. Om sneller de onbeantwoorde vragen te kunnen beantwoorden, zullen we moeten zoeken naar dit soort natuurlijke experimen-

ten. Omdat hier alle patiënten aan deelnemen, is de generaliseerbaarheid groter dan van een traditionele RCT waar slechts een kleiner deel van de patiënten aan deelneemt. De kern is dat er een vorm is van een experiment. Dat er door een vorm van toeval bepaald wordt wat de behandeling is. Een experiment om de voorspellers – bekende en onbekende – te balanceren tussen de twee groepen.

Hier kunnen we leren van andere vakgebieden zoals de econometrie en data sciences. De econometrie probeert causale verbanden te leggen tussen economische grootheden op basis van statistische en wiskundige methoden. In dit vakgebied wordt gezocht naar slimme andere manieren om causale verbanden aan te tonen omdat je in de maatschappij niet echt kunt randomiseren.

*Praktijkvoorbeeld 2 patiënten met artrose van de heup:*

Meneer Knoop van 69 jaar en meneer Arlman van 71 jaar met artrose van de heup. Zij willen graag weten wat de beste heupprothese is.

Een heupprothese kunnen we op verschillende manieren vastzetten in het bot. Met een kunsthars - gecementeerd - of met een ingroei oppervlakte -ongecementeerd-.

Laten we kijken wat de econometrie ons te bieden heeft.

Eén van de methoden die we uit de econometrie over kunnen nemen is het Regression Discontinuity Design. Hierbij wordt een afkappunt gebruikt om een behandelingskeuze te sturen. In het implantatenregister LROI zijn alle implantaten geregistreerd, de gecementeerde en de ongecementeerde heupen. Er zijn ziekenhuizen die een leeftijd van 70 jaar gebruiken als afkappunt. Patiënten jonger dan 70 jaar krijgen een ongecementeerde prothese en oudere patiënten krijgen een gecementeerde prothese. Er is dus een afkappunt op basis van een tamelijk arbitraire leeftijd die bepaalt welke behandeling iemand krijgt. Nu is een patiënt van 50 jaar niet te vergelijken met een patiënt van 80 jaar omdat er dan verschil is in de voorspellende factoren.

In de LROI zijn heel veel patiënten geregistreerd (totaal 440.000, 99% van alle prothesen). We hebben dus veel patiën-



ten van 69 jaar en veel patiënten van 71 jaar. Een man van 69 jaar verschilt niet veel van een man van 71 jaar in voorspellende factoren. Je kan ze dus redelijk goed met elkaar vergelijken. Door de toeval van leeftijd wordt bepaald welke behandeling zij hebben gehad. Hierdoor is het mogelijk ze te vergelijken met elkaar.

We hebben nu grote aantallen  $\pm 20.000$  patiënten rond het afkappunt. Dus kunnen we onze vraag, of we moeten cemen-teren of niet, in deze leeftijdsgroep beter beantwoorden. We hoeven daar geen kostbare RCT voor op te zetten en kunnen, maar kunnen, door slim naar bestaande data te kijken, de antwoorden vinden.

Op een andere manier bent u zelf ook onderdeel van verschillende experimenten zonder dat dit aan u is gevraagd.

### **Pak uw telefoon er weer bij.**

**Ziet de site van Booking.com er altijd hetzelfde uit?**

Mentimeter



Websites zoals deze zijn continu aan het experimenteren. Hoe kom je erachter wat op een website het beste werkt? Met AB testing. Hierbij geeft een AB test aan een groep websitebezoekers versie A en aan een ander groep websitebezoekers versie B. Dan wordt gekeken welke versie A of B leidt tot de meeste verkopen. Omdat er veel websitebezoekers zijn, gaat dit heel snel. Snel blijkt of A of B beter verkoopt. Bij A zie je dan bijvoorbeeld de zin: “Op dit moment kijken 8 andere bezoekers ook naar dit hotel”, en bij B krijg je deze melding niet te zien.

Stel A verkoopt beter, dan wordt dit getest tegen de volgende versie van de site. Dus als u ooit een website van Booking.com bezoekt, dan bevindt u zich altijd in een experiment zonder dat u dit weet. De website weet het wel en meet de uitkomst omdat de uitkomst relevant voor hen is, namelijk hoeveel kamers er verhuurd zijn.

Nu is het zo dat er in de zorg ook een experiment plaatsvindt door de praktijkvariatie. Dit gebeurt nu onbewust, zonder vraagstelling en de uitkomst wordt niet gemeten. We worden er dus niet veel wijzer van.

Hier komt zorgevaluatie om de hoek kijken.

Van AB testing kunnen we leren dat we met eenvoudige experimenten meer zekerheid kunnen krijgen of ons handelen werkt. Juist voor bestaande behandelingen is dit ethisch. Het doel van mijn leerstoel is om dit soort AB tests ook in de zorg toe te passen en zo sneller relevante antwoorden te vinden. Voor nieuwe behandelingen is een meer gecontroleerde manier van onderzoek doen nodig. Voor nieuwe behandelingen weten wij immers nog niets over de veiligheid en werkzaamheid. Dit innovatieve onderzoek staat schouder aan schouder - dus ook naast - zorgevaluatie.

Misschien moeten we eerder experimenten doen in het proces wat een patiënt doorloopt als hij bij de dokter komt. Diagnostische tests waaronder beeldvorming zoals röntgenfoto's en MRI scans verdienen extra aandacht.

Veelal zijn diagnostische tests zoals een MRI geïntroduceerd nadat de test eigenschappen waren onderzocht. Zo is er gekeken of de test meet wat hij moet meten en of verschillende beoordelaars hetzelfde meten. Het klinisch nut van een test is veel minder vaak onderzocht. Een test kan heel nauwkeurig zijn maar niet veel klinisch nut hebben. Zo heeft het maken van een MRI van de knie bij een 60-jarige patiënt niet veel klinisch nut omdat de kans al groot is dat er bij artrose ook een afwijkende meniscus wordt gezien. Als er een MRI van de knie bij mensen van die leeftijd zonder knieklachten gemaakt wordt, vinden we ook afwijkingen. De test is immers heel goed in het vinden van afwijkingen. Het nut is minder onderzocht. Dit is van belang

omdat de uitslag van een test vaak leidt tot een interventie. Je kunt eigenlijk zeggen dat de interventie begint bij de diagnostische test. Het diagnose interventie spectrum. Daarom moeten we het nut van de test onderzoeken. Dit kan door patiënten te randomiseren tussen het wel en niet doen van de test en vervolgens te zien hoe het met ze gaat. Dit heeft als voordeel dat direct het hele diagnose interventie spectrum wordt onderzocht.

Een voorbeeld hoe dit slim zou kunnen worden ingezet, komt uit de polsfractuur studies. Er loopt er een studie bij oudere patiënten om te onderzoeken of gips en opereren even goed werken. Nu worden mensen vaak geopereerd maar de vraag is of dit betere klinische resultaten oplevert. De foto wordt wel mooier na de operatie.

De operatie kan echter wel complicaties geven. Bij gips komen de operatie gerelateerde complicaties niet voor maar is de kans dat de stand van de pols afwijkend is, wel groter.

Het blijkt moeilijk om patiënten te vinden die aan deze studie willen meedoen. Hoe komt dit?

Als een arts een suboptimale stand van de pols ziet, wil hij/zij die stand goed hebben. Hoe ziet de arts of de stand goed is of beter kan?

Door naar de röntgenfoto te kijken. Als er twijfel is, dan maakt hij/zij een CT-scan van de pols. Nu is de kans heel groot dat door de nauwkeurige CT de stand veel slechter blijkt te zijn. Het is dan nog moeilijker om te zeggen dat je geen operatie wilt doen. Door het maken van een CT-scan is de vervolgstap dus al gezet.

Als er geen bestaande data zijn waar we dit uit kunnen halen zouden we dus een experiment moeten opzetten dat kijkt naar het klinisch nut van de CT-scan. Patiënten worden dan gerandomiseerd tussen alleen een röntgenfoto en röntgenfoto + CT-scan. Door de randomisatie is de ernst van de breuk vergelijkbaar in beide groepen. We kunnen dan patiënten volgen en

zien hoe het uiteindelijk met ze gaat. Dus hun functie bekijken en vragen wat zij nog kunnen met de pols in beide groepen. Ook kijken we dan hoeveel patiënten er uiteindelijk geopereerd zijn in beide groepen. Zo vangen wij 2 vliegen in 1 klap. We zien wat het effect van de CT-scan op de uiteindelijke uitkomst is en of de CT-scan leidt tot meer operaties. Als er dan uitkomt dat de CT-scan meer operaties oplevert en een vergelijkbare uitkomst geeft, weten we dat het klinisch nut van de CT-scan beperkt is. De kosten voor de gezondheidszorg zullen dalen. Dergelijk onderzoek is nog maar spaarzaam gedaan. Om deze onbeantwoorde vragen te beantwoorden zullen we goed moeten samenwerken met radiologen.

Uitkomsten meten die relevant zijn voor patiënten is een essentieel onderdeel van het evalueren. Dit gebeurt veelal met vragenlijsten om van de patiënt zelf te horen hoe het gaat. Nu is er een - goedbedoelde - wildgroei aan gevalideerde vragenlijsten. Zeker als een patiënt meer aandoeningen heeft kan hij overspoeld worden met vragenlijsten. PROMIS is hier een oplossing voor. Hierbij krijgt de patiënt via de computer een geselecteerde vraag. Door het antwoord wat op deze vraag gegeven is berekent de computer wat de volgende vraag moet zijn. De computer kiest die vraag uit een bestand met 1000 vragen. Zo wordt er nooit meer aan iemand in een rolstoel gevraagd of hij trap kan lopen. Met PROMIS is het mogelijk mensen te vergelijken met de normaalwaarden in de gezonde populatie. Ook is het mogelijk om interventies voor verschillende ziektebeelden met elkaar te vergelijken. Een mens bestaat niet alleen uit een knie, maar kan ook de ziekte van Parkinson hebben. Het gebruik van vragenlijsten zijn momentopnamen. Hierbij moeten we niet vergeten wat de psychische gesteldheid van een mens op het resultaat kan doen. Om dit te evalueren zullen we samen moeten werken met psychologen en zelfs psychiaters.

**Pak uw telefoon:**

## Ik gebruik een stappenteller



M Mentimeter



Een stappenteller is een mooi voorbeeld van een op het lijf gedragen meter. Om te zien hoe patiënten in de dag functioneren kunnen we meer gebruik maken van dit soort wearables die steeds meer kunnen meten naast de vragenlijsten. Op een goede manier uitkomsten meten hoort ook bij zorgevaluatie.

Nu heb ik de eerste twee stappen van zorgevaluatie besproken: agenderen en evalueren.

Hier is duidelijk gemaakt dat we met goed opgezette -liefst experimentele - studies causale verbanden aan kunnen tonen. Daarmee laten we zien welke tests en interventies zinvol zijn en welke we beter niet meer kunnen doen.

De derde stap in de cyclus van zorgevaluatie is implementeren **3. Implementeren**

Zodra de resultaten van een zorgevaluatie onderzoek gepubliceerd zijn, moeten deze bevindingen geïmplementeerd worden in de praktijk. Dit implementeren is een uitdaging. Artsen die al jaren gewend zijn om een behandeling op een bepaalde manier te doen, zijn niet snel overtuigd dit anders te gaan doen. Daarom is het belangrijk het hele proces van zorgevaluatie: agenderen – evalueren – implementeren, in de geneeskundige opleidingen te onderwijzen en het zo onderdeel te maken van de reguliere zorg.

Naast het onderwijs is de financiële prikkel van groot belang. Nu worden specialisten beloond op het aantal verrichtingen

die zij doen. Het kan dus zijn dat zij heel efficiënt niet effectieve behandelingen uitvoeren. Het voorbeeld hiervan is de kijkoperatie van de knie bij mensen boven de 50 jaar met een ouderdomsmeniscus. Op de MRI ziet dit er uit als een scheur. In het begin van mijn carrière werden hiervoor straatjes geregeld en zo vond deze operatie 30.000 keer per jaar plaats. De patiënten werden beter na de operatie. Patiënten die niet geopereerd werden maar een fysiotherapie programma volgden werden echter ook beter. Hierop volgde een aantal RCT's waarbij de kijkoperatie werd vergeleken met fysiotherapie. Ook vonden er placebo gecontroleerde trials plaats. De echte kijkoperatie werd vergeleken met kleine steekgaatjes. Het bleek dat de placebo operatie even goed werkte als de echte operatie. De patiënten werden dus ondanks en niet dankzij de operatie beter.

In verschillende landen zijn deze studies uitgevoerd en zij kwamen tot dezelfde conclusie. Nadat deze studies in de richtlijn zijn gekomen, is het aantal kijkoperaties van de knie drastisch gedaald.

Of zou dit komen doordat de zorgverzekeraars voor code 1801 een kijkoperatie niet meer vergoeden?

Het implementeren van studieresultaten kan dus makkelijk lijken. Maar wat te doen als de resultaten van zorgevaluatie trials minder eenduidig zijn? Juist dan is het belangrijk om aan het gestructureerde systeem van richtlijn ontwikkeling vast te houden. In dit systeem worden de resultaten van een nieuwe studie geduid en vergeleken met de bestaande literatuur met dezelfde vraagstelling. Dit kan snel gedaan worden. Met de Rapid Recommendations, geïntroduceerd door het toonaangevende tijdschrift de BMJ, is een nieuwe trial de vonk die het proces start. Er vindt door een onafhankelijk team een systematische review plaats om de resultaten van de nieuwe trial te duiden in het totaal van bewijs over het onderwerp. Omdat de resultaten niet altijd (vaak niet) zwart/wit zijn, komen er verschillende sterkten van aanbeveling uit naar voren. Er vindt een gradering plaats van het bewijs. Hieruit komt dan een sterkte of zwakte aanbeveling voor de behandeling, of een zwakke of sterke aanbeveling tegen de behandeling. Belangrijk is dat er syste-

matisch gekeken wordt naar *alle* literatuur over het onderwerp. Een trial kan wat sturing geven. Maar als verschillende trials dezelfde richting uitwijzen en goed zijn uitgevoerd dan is het bewijs sterker.

Het implementeren van de nieuwe richtlijn is de volgende uitdaging. De richtlijn is dus gebaseerd op een systematisch review van alle studies die er over het onderwerp gedaan zijn. In deze studies zijn echter niet alle patiënten meegenomen omdat er altijd exclusie criteria zijn waardoor sommige (vaak zwakkere of ziekere) patiënten niet mee hebben gedaan aan de trials. Om te zien of de richtlijn ook opgaat voor die patiënten stel ik een experimentele introductie van de richtlijn voor. Om te zien of iets werkt is een experiment nodig, zoals ik besproken heb bij het agenderen en evalueren. Dus ook om te zien of een richtlijn werkt, is een experiment nodig. Ik stel voor om een richtlijn stapsgewijs in clusters te introduceren en de uitkomsten hiervan te evalueren.

10 Hiervoor is stepped wedge cluster gerandomiseerde methode een uitkomst. Dat is een mondvul, laat me dat uitleggen. Hierbij wordt door loting bepaald welk ziekenhuis begint met de introductie van de richtlijn. Dit gaat op voor alle patiënten in dat ziekenhuis. Dit is veilig omdat we dan kunnen zien of de resultaten uit de studies waaruit patiënten geëxcludeerd werden, ook opgaan voor de grote groepen patiënten uit de dagelijkse praktijk. Als de ziekenhuizen één voor één door loting bepaald, de richtlijn implementeren en de uitkomsten gemeten worden, kan gezien worden of de richtlijn ook echt doet wat we verwachten dat het zou doen. Dit is een verbetering ten opzichte van het huidige systeem waar we zeggen: dit is de nieuwe richtlijn: succes.

Het traditionele passieve implementatieproces wordt zo een experiment waar we causale verbanden uit kunnen trekken. We hebben dan direct een trial die ons informatie geeft over grote groepen patiënten zonder dat hier patiënten worden uitgesloten. De resultaten gaan dus op voor meer mensen, niet alleen de gezondere mensen die vergelijkbaar waren met de mensen uit de studies waar de richtlijn is gebaseerd.

De invoering van dit systeem van experimentele richtlijn in-

troductie lijkt complexer dan het is. Voor veel aandoeningen hebben we routinematige uitkomstmetingen. Helaas gebeurt er met die data tot nu toe niet zo veel. De data worden wel gebruikt voor observationele studies waar helaas vaak geen causale verbanden uit getrokken kunnen worden. Door het randomiseren van welk ziekenhuis als eerste start en welke er volgen met de richtlijn introductie, is er wel sprake van een echt experiment en kunnen er dus causale conclusies uit worden getrokken. De ziekenhuizen die hier uitermate geschikt voor zijn, zijn de Santeon ziekenhuizen, waar het OLVG onderdeel van uitmaakt omdat deze al veel samenwerken op het gebied van uitkomst gedreven zorg.

In het Value Based Health Care programma dat daar loopt, zou deze stepped wedge cluster gerandomiseerde introductie van nieuwe richtlijnen uitermate goed kunnen omdat hier routinematige uitkomsten worden geregistreerd maar ook naar kosten wordt gekeken. Als de richtlijn in de praktijk dan blijkt te werken, weten we ook beter voor wie hij werkt en voor wie niet omdat we meten wat het effect is van de introductie bij alle patiënten. We kunnen hier beginnen met de grote aandoeningen binnen de orthopedie zoals heup en knie artrose. Ook in samenwerking met andere vakken, zoals bij de heelkunde waarbij we het heupfractuur register kennen. Voor deze aandoeningen bestaan registers waarin de uitkomsten van patiënten worden geregistreerd. Er is dus een reeds bestaand uitkomst meetapparaat.

Daar de cyclus van zorgevaluatie continu doordraait, kunnen nieuwe vragen worden geagendeerd en dan weer worden geëvalueerd. Als een nieuwe richtlijn gebaseerd is op zwak bewijs is het extra belangrijk om deze te introduceren als experiment. Zo is de grens tussen evalueren en implementeren minder strikt. In 50% van de aanbevelingen uit richtlijnen is deze aanbeveling gebaseerd op zwak bewijs en consensus tussen de richtlijn opstellers. Om te zien of het klopt wat zij denken, zal dit dus getoetst moeten worden in een experimentele introductie. Tot nu toe is deze stap nog nooit gemaakt; in ieder geval niet in de orthopedie.

De deelname aan een normale RCT vraagt om een informed

consent. Hoe zit dit bij een gerandomiseerde introductie van een richtlijn? Deze vraag kunnen we omdraaien. Waarom mag een richtlijn geïntroduceerd worden zonder dat getoetst is of deze werkt? Dit is de huidige situatie, de richtlijn wordt gewoon geïntroduceerd. Nu kan de richtlijn geïntroduceerd worden en vindt de behandelovereenkomst plaats tussen de arts en de patiënt in de WGBO (Wet Geneeskundig Behandel Overeenkomst). Zorgevaluatie van bestaande zorg is wezenlijk anders dan de introductie van een innovatieve behandeling in een klassieke RCT.

Omdat zorgevaluatie bestaande – reeds eerder geïntroduceerde – behandelingen evalueert, zou de ethische toetsing en het consent proces meer op de praktijk moeten worden toegespitst. Informed consent zou in de WGBO geregeld kunnen worden. Nu zijn we gehouden aan het volgen van een richtlijn zonder dat de richtlijn getoetst is. Door het gestructureerd evalueren van het effect van richtlijn modules lopen we minder achter de feiten aan.

Richtlijnen worden wel geüpdatet, maar de werkzaamheid wordt hooguit observationeel beschreven. Hierdoor zijn conclusies niet goed te trekken en laten we informatie liggen die door een gerandomiseerde stepped wedge introductie wel te verkrijgen zou zijn.

Om het implementeren van richtlijnen te versnellen zouden ziekenhuizen verleid moeten worden voor stepped wedge introducties en ze hiervoor te betalen. Niet doorgaan met betalen per verrichting waar maar zwak bewijs voor is. Er moet een beloningssysteem komen waardoor de ziekenhuizen gestimuleerd worden aan zorgevaluatie bij te dragen zodat het een standaard onderdeel, beter nog de standaard wordt van hoe de zorg georganiseerd is.

De cyclus staat nooit stil.

“Elk nieuw begin komt van het einde van een ander begin” - Seneca.

### **Mijn toekomstvisie**

Zorgevaluatie is dus iets van de patiënten en de dokters. Samen moeten we er voor zorgen dat de zorg rechtvaardig is. Zorg evalueren doen we om beter te worden en hierdoor kunnen artsen enthousiast gemaakt worden. Het gaat dan om het proces van verbetering. Het is niet tegen de dokters. Sommige artsen kunnen zich bedreigd voelen omdat zij een behandeling die zij al jaren doen, niet meer kunnen uitvoeren. Uit zorgevaluatie onderzoek kan blijken dat patiënten beter worden ondanks de behandeling en niet dankzij de behandeling. Het is niet erg dat er behandelingen verdwijnen omdat dit ruimte geeft voor behandelingen die wel werken. Er mag ook uitkomen dat een behandeling én beter én kostbaarder is. Beloond moeten we worden op verbeterprocessen. Niet op het efficiënt uitvoeren van ineffectieve interventies. Hiermee kunnen we tevens de kosten die inmiddels boven de 85 miljard uitkomen in toom houden.

Wie hebben we hierbij nodig?

Patiënten die mee willen doen aan onderzoek en mee willen beslissen welke onderzoeken we moeten doen.

Artsen van verschillende vakgroepen die inzien dat samenwerken buiten de geëigende vakgebieden noodzakelijk is voor de verbetering van de zorg.

Docenten om studenten reeds in hun studie de principes van zorgevaluatie te onderwijzen.

Experts uit andere disciplines zoals de wiskunde, kunstmatige intelligentie en data science hoek, die ik hier in Leiden kan vinden.

Epidemiologen die helpen de studies zo op te zetten dat zij de vragen echt beantwoorden.

Onderzoekscoördinatoren. Zij zijn essentieel om onderzoek vlot te laten lopen.

Tenslotte: zorgverzekeraars die inzicht kunnen geven in welke behandelingen het meeste uit gevoerd worden en het meeste kosten. Zij zijn ook gebaat bij shared savings: het bespaarde geld kan helpen de zorg te blijven evalueren.



## Growth Mindset

Wat hebben we hier voor nodig? Een “growth mindset”. We weten nu *nog* niet hoe het werkt en *nog* niet hoe we het onderzoek het beste kunnen doen. Het woord *nog* is hierbij cruciaal. Met een growth mindset is het mogelijk fouten te maken en daarvan te leren. Dit proces van een lerend systeem kan alle dokters aangehaakt houden. Tijdens de studie is het heel normaal dat je iets *nog* niet weet of *nog* niet kunt. Later als volleerd arts is dit ook *nog* steeds zo. Zorgevaluatie past binnen een leven lang leren. Een continu verbeterproces kan aantoonbaar gemaakt worden. Zorgverzekeraars moeten ziekenhuizen belonen voor deze verbeterprocessen. Niet voor het efficiënt uitvoeren van ineffektieve interventies zoals nu de bekostiging is georganiseerd.

*Om de maximale leeropbrengst uit mijn oratie te halen vraag ik u de laatste keer uw telefoon te pakken en te noteren wat u geleerd heeft vandaag.*

## Wat heeft u geleerd vandaag?



Ik ben ervan overtuigd dat zorgevaluatie de weg is naar rechtvaardige zorg voor alle mensen. Artsen zullen hier meer vol-doening uit halen omdat zij weten dat wat zij doen goed is.

## Dankwoord

Dank aan allen die aan de totstandkoming van mijn benoeming hebben bijgedragen. In het bijzonder dank ik de Nederlandse Orthopaedische Vereniging die deze leerstoel mogelijk maakt. Oud voorzitter Hans André Schuppers en directeur Chris van der Togt speelden hierin een belangrijke rol.

Ook dank ik mijn leermeesters die mij hebben geleerd kritisch te denken. Van de hooggeleerde René Marti leerde ik het belang van het evalueren van ons eigen handelen. Hij was een voorloper op het gebied van zorgevaluatie.

De hooggeleerde Mohit Bhandari leerde mij de methoden om chirurgische trials uit te voeren, maar ook hoe dit te communiceren.

Ik dank mijn collegae uit het OLVG die mij de ruimte geven om een dag per week naar Leiden te gaan, of daar geestelijk aanwezig te zijn.

Dank aan de research coördinatoren Vanessa, Nienke, Maaïke en Sigrid die JointResearch tot een professioneel onderzoeksinstituut hebben gemaakt.

Mijn dank is groot aan de co-assistenten, ANIOS, AIOS en promovendi die mij scherp houden met hun kritische houding en zo mij op het rechte pad houden.

Collegae van de Commissie Zorgevaluatie voor de levendige discussies.

Dank aan hen die geholpen hebben mijn oratie vorm te geven. In het bijzonder Yvette Pasman en Rutger Mollee.

De Raad van Bestuur van het OLVG, in het bijzonder de hooggeleerde Maurice van den Bosch, dank ik om mijn werk in het OLVG te combineren met onderzoek.

De hooggeleerde Nelissen, waarde Rob, dank voor jouw sprankelende enthousiasme waarmee je deze leerstoel naar Leiden hebt gehaald.

Ik dank mijn vrienden voor hun vriendschap.

Neef Coks Bosman heeft mij ooit gevraagd of ik een boekje ging schrijven. Dank voor de zetjes de goede kant op het juiste moment. Nu ga ik veel boekjes begeleiden.

Mijn schoonfamilie

Wijlen mijn schoonvader Adriaan Mazel, die zijn leven lang Leids student is gebleven zou hier graag bij zijn geweest. We missen hem.

Lieve Mama, als ik een slechte jeugd had gehad was ik wel schrijver geworden. Ik bof dat ik opgegroeid ben in een omgeving, door jou en Papa gecreëerd, waar Margreet en ik alle mogelijkheden heb gekregen om het beste uit onszelf te halen. Dank je wel.

Lieve Jet, Wieger en Anna,

Jullie vader houdt dus van experimenteren. Dat merken jullie als ik kook. Ik hoop dat jullie in je leven ook veel experimenten zullen uitvoeren en dat er dan een paar mislukken is alleen maar goed om ervan te kunnen leren. Dus blijf experimenteren en dan zal je nooit een saai leven hebben. Vooral als je vragen blijft stellen helpt dat. Blijf evalueren of wat je doet wel bij je past. Alleen dode vissen zwemmen met de stroom mee.

Lieve Tjitske, “we doen het best goed samen” en dat is zo. Als kinderpsycholoog ben jij de enige die echt met mij kan omgaan. Dankzij jou heb ik een growth mindset.

Ik heb gezegd.

De oratie is terug te kijken op:  
<https://vimeo.com/461124211>







## PROF. DR. RUDOLF W. POOLMAN



Prof.dr. Rudolf W. Poolman is met ingang van 1 december 2019 benoemd tot bijzonder hoogleraar orthopedie - in het bijzonder Zorgevaluatie - aan de Universiteit Leiden. De Nederlandse Orthopaedische Vereniging (NOV) is initiatiefnemer van deze bijzondere leerstoel. Het primaire doel is om wetenschappelijk onderzoek op het gebied van zorgevaluatie te bevorderen, evenals de implementatie van de uitkomsten. Het belang van zorgevaluatie neemt toe. Bovendien wordt het - mede door de toenemende complexiteit van patiënten en behandelmogelijkheden - een steeds grotere uitdaging om de stijgende zorgkosten te rechtvaardigen. Dit maakt dat het voor medisch specialisten een noodzaak is om het eigen handelen te evalueren om zinnige én zuinige zorg te kunnen blijven leveren. Met deze bijzondere leerstoel faciliteert de NOV de systematische evaluatie van de orthopedische zorg.

### Wetenschappelijke carrière

Rudolf Poolman (1970) studeerde Geneeskunde aan de Universiteit van Amsterdam en behaalde in 1998 zijn artsexamen. Van 2000-2006 volgde hij de opleiding tot orthopedisch chirurg. In 2006 volgde hij een fellowship Trauma & Research aan de McMaster University Hamilton Ontario (Canada). Hij promoveerde in 2007 aan de Universiteit van Amsterdam. De titel van zijn proefschrift luidt: *Moving towards Evidence Based Orthopaedic Surgery*. Sinds 2007 werkt hij als orthopedisch chirurg in het OLVG. Hij is tevens oprichter en directeur van Joint Research, het onderzoeksbureau voor het steun en bewegingsapparaat in OLVG.

Poolman heeft een groot orthopedisch netwerk binnen en buiten Nederland. Hij was tot voor kort voorzitter van de NOV-werkgroep Herziening Onderzoeksagenda Orthopedie en is lid van de Wetenschappelijk Adviesraad (WAR) van de Landelijke Registratie Orthopedische Implantaten (LROI).

Hij heeft nationaal en internationaal een voortrekkersrol in het ontwikkelen van methodes voor rationalisatie van de orthopedische zorg volgens de principes van Evidence Based Medicine. In combinatie met een duidelijke visie om het draagvlak voor zorgevaluatie te vergroten, maakt dit hem tot de ideale kandidaat voor deze leeropdracht.



Universiteit  
Leiden