

Effectiviteit in praktijkgericht onderzoek

OP ZOEK NAAR ALTERNATIEVEN VOOR DE RANDOMIZED CONTROLLED TRIAL (RCT)

In het tijdperk van *evidence-based practice* (EBP) lijkt weinig twijfel te bestaan over het belang van het gebruik van onderzoek en onderzoeksgegevens in het uitwerken, de uitvoering en de evaluatie van beleidsmaatregelen, interventies en methodiekontwikkeling. Professionals uit de sectoren zorg, welzijn en cultuur moeten zich de nodige vaardigheden eigen maken om hun acties en interventies op basis van onderzoek te ontwerpen en uit te voeren en daarnaast de nodige kritische zin ontwikkelen voor het interpreteren van (nieuwe) onderzoeksresultaten. Anderzijds moeten vanuit de wetenschappelijke wereld inspanningen worden geleverd om nieuwe inzichten te vertalen naar de praktijk van de sector, antwoorden te formuleren op vragen uit de praktijk, dus *praktijkgericht onderzoek* uit te voeren. Een cruciale vraag is hoe dergelijk onderzoek er moet uitzien en aan welke criteria het moet voldoen.

Door: L. Tavecchio

In de praktijk wordt EBP vaak verengd tot het gebruik van de resultaten van zogenoemd effectiviteitonderzoek en kosteneffectiviteitanalyses – beide overwegend *kwantitatief* georiënteerd – bij de selectie van interventies. De *randomized controlled* (of *clinical*) *trial* (RCT) geldt daarbij als de ultieme vorm van onderzoek, die samen met *meta-analyse* aan de top staat van Geddes¹ hiërarchie van effectonderzoek. Maar dit soort onderzoek is vooral mogelijk in strikt gecontroleerde en afgebakende settings en omstandigheden, waarin de te testen behandeling/interventie wordt uitgevoerd bij een experimentele groep en vergeleken met een controlegroep. Daarbij wordt de toewijzing van de personen aan de verschillende groepen aselekt (*random*) bepaald. Dergelijk onderzoek is echter veel lastiger of zelfs niet uitvoerbaar binnen de praktijk van het dagelijks werk van professionals.

Nadelen van te veel nadruk op RCT's

Howe¹ plaatst daarnaast serieuze kritische kanttekeningen bij het gebruik en de interpretatie van de term randomisering (*randomization*). In de praktijk is 'random selectie' vaak niet mogelijk, in feite gaat het om 'aselecte toewijzing' van personen aan experimentele en controlegecondities. Bovendien is aselecte toewijzing op grond van politieke of juridische overwegingen in de praktijk vaak niet mogelijk. Afgezien van Howe's kritiek blijkt het bewijs voor de *effectiviteit van een interventie* bovendien lang niet altijd ondubbelzinnig – dus met een simpel 'ja' of 'nee' – te leveren (zie onder *Triangulatie*, voor een oplossing van dit probleem).

Een ander behartigenswaardig advies van Howe is om bij praktijkgericht onderzoek de *externe validiteit* (generaliseerbaarheid) van



Louis Tavecchio

Foto: Marian Haringsma

onderzoek en resultaten voorrang te geven boven de interne validiteit (*'trading external validity for internal validity'*). Bij de interne validiteit van een onderzoek(sopzet) gaat het om de vraag of een verschil tussen groepen inderdaad is toe te schrijven aan de interventie/behandeling en niet aan invloeden van buitenaf. Om dergelijke externe invloeden zo veel mogelijk buiten het onderzoek te houden kiest men ervoor om de onderzoeksgroep (populatie) en de behandeling of interventie sterk te begrenzen om die interne validiteit te optimaliseren. Het gevolg is dat de kans toeneemt dat de resultaten *niet* extern valide zijn, met andere woorden dat we gewoon niet weten in welke mate de behaalde resultaten gegeneraliseerd kunnen worden naar andere personen, situaties, condities, metingen en tijdstippen dan die deel uitmaakten van de specifieke onderzoeksopzet. En op de belangrijkste pijlers van interne validiteit – *random* selectie en aselechte *toewijzing* – is ook het nodige aan te merken, zoals bleek uit de reeds genoemde kritiek van Howe.

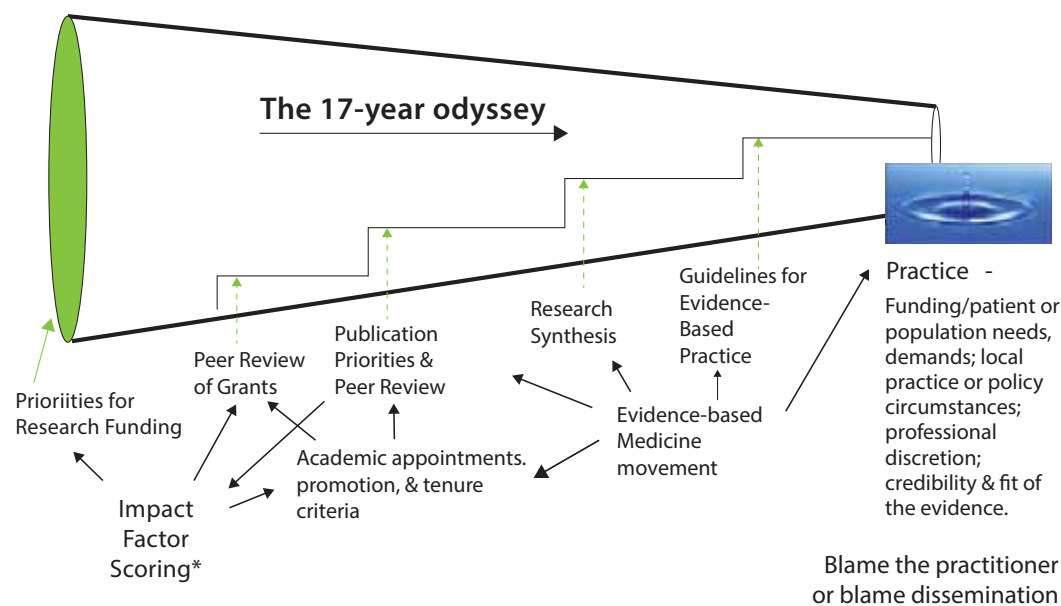
De Amerikaanse onderzoeker en preventiedeskundige Larry Green onderschrijft Howe's visie: 'Wetenschappers zijn obsessief bezig om alles strak te controleren, zodat die interne validiteit maar hoog genoeg is. Als ze zeker weten dat een bepaalde oorzaak in hun studie tot dat effect heeft geleid, zijn ze blij – *al zou het nergens anders geldig zijn*'.¹ Voor het opstellen van richtlijnen voor professioneel handelen in de praktijk zijn resultaten uit dit soort kunstmatig onderzoek vrijwel onbruikbaar, maar dergelijke studies worden wel in groten getale gepubliceerd.

Efficacy en effectiveness en lekkage in de wetenschappelijke pijplijn

Bovenstaande kritische kanttekeningen hebben alle, in meer of mindere mate, te maken met een belangrijk onderscheid. Dit betreft het verschil tussen efficacy en effectiveness. De vertaling van beide termen komt in het Nederlands ongeveer op hetzelfde neer: werkzaamheid, doeltreffendheid. Maar in methodologische literatuur wordt *efficacy* nader gespecificeerd als het getoetste effect van een interventie onder strikt gecontroleerde ('laboratorium') condities, terwijl de term *effectiveness* is voorbehouden aan het getoetste effect van een interventie onder min of meer 'normale' omstandigheden zoals die zich in de dagelijkse praktijk voordoen, dat wil zeggen situaties in *real-time*, met minder controle en in een karakteristieke context, populatie en condities.^{2, 3, 4} Green c.s.¹ beschrijven het proces van kennisoverdracht, verspreiding van onder-

zoeksresultaten en het overnemen van en implementeren van *evidence-based* richtlijnen in het *efficacy* onderzoek als een 'pijplijn' waarin wetenschappelijk bewijsmateriaal wordt geproduceerd dat, na grondige controle en toetsing, wordt verspreid onder beleidsmakers en praktijkwerkers. In figuur 1 zien we de pijplijn. Een verontrustende en frequent geciteerde bewering over de totale 'lekkage' (*leakage*) in de pijplijn en het tijdsverloop, c.q. tijdsverlies tussen wetenschappelijk onderzoek en medische praktijk houdt in dat het in dit vakgebied *17 jaar* duurt om (slechts) *14 procent van het oorspronkelijke onderzoek* toe te passen voor het welzijn van de patiënt.¹ Niet alleen het aanvragen en uitvoeren van onderzoek kost veel tijd, er gaat ook veel tijd verloren vanaf het moment dat het wetenschappelijk onderzoek is afgerond. Tijd die wordt besteed aan het schrijven en aanbieden van een artikel en wachten op publicatie. Daarna volgen indexering in databases en systematische reviews en aanbevelingen voor 'best practices' in leerboeken. Dit alles voor de uiteindelijk tot implementatie van die praktijken in gezondheidszorginstellingen. Voor de praktijkwerker kleven er nog meer specifieke nadelen aan de lekkage die al die jaren in de pijplijn plaatsvindt. 'Negatieve' resultaten worden door onderzoekers vaak niet aan tijdschriften aangeboden, vanuit de veronderstelling dat ze toch niet zullen worden gepubliceerd. Maar voor professionals in de zorg en welzijnsector is het *juist heel relevant* om zulke 'negatieve' interventieresultaten te leren kennen, omdat ze bijvoorbeeld informatie verschaffen over de (on)geschiktheid van een interventie of methodiek voor bepaalde populaties en/of condities (die afwijken van de oorspronkelijke onderzoeksgroep en -condities). Verderop in de pijplijn ontstaat een ander, voor de professionele praktijk mogelijk kostbaar, verlies omdat nogal wat ingediende manuscripten niet worden gepubliceerd vanwege zogeheten 'tekortkomingen' in onderzoeksopzet, statistische power of steekproefomvang. De *grootste verliezen* treden op aan het *eind van de pijplijn*. In het geval van (bio) medisch onderzoek wordt de tijd die verloopt tussen het indexeren van onderzoeksresultaten en het opnemen daarvan in overzichten die zouden kunnen leiden tot richtlijnen voor zogenoemde 'best practices' geschat op zes tot dertien jaar.

Een belangrijke (remmende) factor daarbij is de grote waarde die wordt toegekend aan de RCT, die vaak een doorslaggevende rol speelt bij de uiteindelijke keuze voor opname van resultaten



*J Partic Med 2009;1(1).<http://jopm.org/index.php/jpm/article/view/16/31>.

Figuur 1. The Pipeline Fallacy of Producing & Vetting Research to Get Evidence-Based Practice⁵

in systematische overzichten, richtlijnen, studieboeken en ander lesmateriaal, waardoor veel, voor praktijkwerkers potentieel relevante, informatie daarin niet terechtkomt. En ten slotte komt de kloof tussen onderzoek en praktijk nog eenmaal scherp tot uiting in de constatering dat het – in (bio)medisch onderzoek – gemiddeld *negen jaar* (!) duurt *voordat* de in reviews, richtlijnen, lesmateriaal en dergelijke vermelde *evidence-based* praktijken volledig geïmplementeerd zijn.

Bewijsmateriaal dat op deze wijze wordt toegespitst en gefilterd, gecontroleerd en getoetst werkt prima voor *biomedische interventies*, die *veelal gebaseerd zijn op epidemiologisch onderzoek bij een zeer groot aantal cases*, maar voor veel eerstelijns- en openbare gezondheidszorg is het doel van de interventie veel diverser wat betreft psychologische processen, culturele context en sociaal-economische condities die de relatie tussen interventie en resultaat op allerlei manieren kunnen beïnvloeden en veranderen. Voor dit soort interventies zijn context, flexibiliteit en externe validiteit *van even groot belang* als de strikte experimentele controle, de precisie van de implementatie en de sterke nadruk op interne validiteit die binnen het *efficacy* onderzoek tot norm zijn verheven.

Praktijkgericht onderzoek is in hoge mate gebonden aan methoden die recht doen aan de complexe realiteit van het professionele handelen. Met andere woorden, voor het verbeteren van professioneel handelen is onderzoek naar

de effectiviteit van interventies in de alledaagse praktijk, dus onderzoek naar *effectiveness*, het meest relevant. Reden om een aantal alternatieven voor de RCT kort te introduceren en serieus onder de loep te nemen.

Methodologische oplossingen en alternatieven voor de RCT in praktijkgericht onderzoek

Inmiddels groeit de waardering voor en het gebruik van alternatieve onderzoekstypen en databronnen die de hiervoor besproken nadelen en beperkingen van de RCT ondervangen. Van Yperen¹ zette al eerder vraagtekens bij de relevantie van Geddes' hiërarchie in de context van praktijkgericht onderzoek en de dominante positie van de RCT. Eén van de door hem beschreven alternatieve benaderingen is de hieronder beschreven multi-center opzet.

Multi-center opzet

Een deel van de onderzochte instellingen past de nieuwe interventie, behandeling of methode toe, terwijl de overige instellingen nog enige tijd met de traditionele aanpak/methode (*care as usual*) blijven werken. Na verloop van tijd wordt de groep cliënten die de nieuwe interventie heeft gekregen vergeleken met cliënten uit de andere instellingen. De twee groepen worden gematcht op een aantal relevante kenmerken (seks, leeftijd, ernst en aard van de problematiek enzovoort), zodat ze qua samenstelling zo veel mogelijk op elkaar lijken. Mogelijk zoekt men er nog een derde groep cliënten bij die nog helemaal geen in-

terventie heeft gekregen, maar door de deelnemende instellingen op de wachtlijst is geplaatst. De resultaten van de nieuwe interventie kunnen nu worden vergeleken met die van de traditionele interventie en een conditie zonder behandeling.

Observationele studies

Dit betreft onderzoek waarbij geen interventie of (experimentele) behandeling wordt getoetst, bijvoorbeeld een studie naar het natuurlijk beloop van een ziekte of stoornis. Ook een zogenaamd *case-control* onderzoek, waarbij naar verbanden wordt gezocht tussen een bepaalde uitkomst en 'uitlokkende factoren' is observationeel onderzoek. Er wordt immers niet 'ingegrepen' in de natuurlijke gang van zaken. In een onderzoek naar het effect van het weglaten of toevoegen van observationele studies op RCT-resultaten in meta-analyses constateren Shrier et al.¹ dat 'including information from observational studies may improve the inference based on only randomized trials' en dat 'meta-analyses based on observational studies generally produce estimates of effect similar to those from meta-analyses based on randomized controlled trials'. Observationele studies kunnen dus zeer waardevol zijn en de gegevens die daaruit voortkomen doen niet onder voor RCT's en moeten zeker niet worden opgevat als tweederangs (zie ook Rawlins, 2008).¹

Triangulatie

Triangulatie is het gebruik van meerdere methoden of databronnen bij onderzoek zodat resultaten *kruislings gecontroleerd* (gecross-checkt) kunnen worden. (Denk aan het vergelijken van interviewdata van respondenten met hun antwoorden op vragen in een vragenlijst, of aan een rechercheur die verschillende informanten ondervraagt over de handel en wandel van een verdachte.) Voor het aantonen van de *effectiviteit van een interventie/behandeling* ontwikkelden De Los Reyes en Kazdin (2008) een systematisch uitgewerkte triangulatie variant, het *Range of Possible Changes* (RPC)-model. In dit model worden *evidence-based treatments* (EBT's) onderverdeeld in zes categorieën. Binnen elke categorie worden gegevens die afkomstig zijn van *verschillende* informanten en die zijn verkregen met *verschillende* meetinstrumenten en methoden met elkaar vergeleken ('gecross-checkt'). De hoogst haalbare categorie is de 'beste bewijs voor verandering' (*best evidence for change*) categorie. Een volledige beschrijving van de zes categorieën en bijbehorende criteria van

het RPC-model is te vinden in Tavecchio en Gerrebrands.¹

Mixed methods

Dit is een vorm van *methodische triangulatie*, dat wil zeggen het gebruik c.q. de combinatie van kwalitatieve en kwantitatieve methoden in hetzelfde onderzoek. De groeiende populariteit van mixed methods in praktijkgericht onderzoek is te verklaren vanuit de veronderstelling dat een goed mixed methods onderzoek een vollediger begrip van het bestudeerde verschijnsel oplevert en meer praktijkrelevante kennis dan de kwalitatieve of kwantitatieve onderdelen afzonderlijk zouden kunnen. Een 'echte', dat wil zeggen meest zuivere mixed methods onderzoeksopzet bestaat uit een *integratie* van kwalitatieve en kwantitatieve methoden en onderzoeksresultaten in alle fasen van het onderzoek. Een onderzoeksopzet waarin twee typen data worden verzameld en geanalyseerd om vervolgens een hoofdzakelijk kwalitatieve of kwantitatieve onderzoeksvraag te beantwoorden is in feite een *quasi-mixed design*, omdat er geen sprake is van serieuze integratie.⁶ Een goed voorbeeld van de *integratieve benadering* is een onderzoek van Poortinga et al.⁷ naar de uitbraak van de veeziekte mond- en klauwzeer (MKZ) in Groot-Brittannië. Tijdens het hoogtepunt van de ziekte in 2001 werd in het kader van een survey bij verschillende groepen een vragenlijst afgenomen over onder andere voorlichting, schuldvraag en vertrouwen in over de ziekte verstrekte informatie. Daarnaast werd een kwalitatieve methode – focusgroep interview – toegepast waarin dezelfde thema's en vragen aan de orde kwamen als in de vragenlijst. De focusgroep bestond uit personen die in de vragenlijst hadden aangegeven aan een focusgroep te willen deelnemen. Er vonden drie interviewsessies plaats. Terwijl de surveydata een goed inzicht gaven in het vertrouwen in de verschillende informatiebronnen, leverden de focusgroepresultaten waardevolle extra informatie op, in het bijzonder over de redenen, rationalisaties en argumenten die ten grondslag lagen aan het begrijpen van de ziekte. Op deze manier kregen de onderzoekers een veel vollediger beeld van de MKZ-crisis dan op basis van kwantitatieve of kwalitatieve onderzoeksmethoden afzonderlijk zou zijn verkregen.⁸

Een belangrijk handboek over mixed methods is geschreven door Tashakkori en Teddlie⁶ en in 2007 werd het *Journal of Mixed Methods Research* opgericht.

Pragmatische RCT

Pragmatische RCT of pragmatische trial⁹ is een uitwerking van de klassieke RCT. In een pragmatische trial laat men zich bij het meten van de resultaten op de eerste plaats leiden door de dagelijkse praktijk. *Praktijkrelevante* onderzoeksuitkomsten zullen dan ook de voorkeur verdienen. In het kader van medisch onderzoek kan men er bijvoorbeeld voor kiezen om in een (pragmatische) trial van secundaire preventie als uitkomst alle be-roerten te tellen, ongeacht of deze zijn ontstaan als gevolg van onvoldoende doorbloeding of vanwege een bloeding binnen de schedel. Het *perspectief van de patiënt* telt hier dus zwaar: in beide situaties gaat het immers om een voor de patiënt ongewenste gebeurtenis. Michon en Kroon zeggen er het volgende over: ‘De meerwaarde is dat de pragmatische variant in de alledaagse praktijk wordt uitgevoerd, bij een steekproef van in principe alle mensen die tot de doelgroep van de interventie behoren. Ook de interventie zelf wordt op een natuurlijke wijze aangeboden (...). Doordat aldus de mate van standaardisatie van de interventie, doelgroep en omstandigheden is toegesneden op de werkelijkheid in de praktijk, kunnen de onderzoeksuitkomsten “gegeneraliseerd” worden naar de praktijk van alledag (externe validiteit) – een voordeel ten opzichte van de klassieke verklarende RCT’.¹⁰

Slotsom

De positie c.q. de werksfeer van de fundamentele wetenschapper – evidence-based practice – kan worden getypeerd door de *primaire vraag* of zijn onderzoek(sresultaat) (1) *intern valide* is, vervolgens of het (2) *belangrijk* is en, ten slotte, of het (3) *bruikbaar* is. Vanuit het perspectief van praktijkgericht/praktijkrelevant onderzoek hangt daar een kostbaar prijskaartje aan, immers:

Hoge interne validiteit + Lage externe validiteit = Verminderde relevantie voor praktijk of beleid!

De positie c.q. de werksfeer van de praktijkgerichte onderzoeker – practice-based evidence – kan het best worden getypeerd door de *primaire vraag* of zijn onderzoek(sresultaat) (1) *bruikbaar* is, vervolgens of het (2) *belangrijk* is en, ten slotte, of het (3) *voldoende valide* is. Praktijkrelevant onderzoek heeft hier duidelijk baat bij, want:

Hoge externe validiteit + Soepele interne validiteit = Verhoogde relevantie voor de praktijk!

De ‘verhoogde relevantie voor de praktijk’ vormt de rode draad in dit artikel, waarbij ervoor is gekozen om de evidence-based practice benadering niet terzijde te schuiven of achterwege te laten, maar aan te vullen met en waar wenselijk te vervangen door onderzoeksopzetten en -methoden die dicht bij de praktijk staan, maar tegelijkertijd voldoende bewijskracht garanderen.

Literatuur

De meeste literatuurverwijzingen zijn terug te vinden in hoofdstuk 5 van:

- 1 L. Tavecchio en M. Gerrebrands, Martin (2012). *Bewijsvoering binnen praktijkgericht onderzoek. Methodologische en wetenschapstheoretische reflecties op de onderbouwing van professionele interventies*. Den Haag: Boom Lemma Uitgevers.

Hieronder de overige verwijzingen:

- 2 B.R. Flay (1986). Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine*, 15, 451-474.
- 3 M.H. Pittler en A.R. White (1999). Efficacy and effectiveness. *Focus on Alternative and Complementary Therapies*, 4, 109-110.
- 4 L.W. Green en M.W. Kreuter (2005). *Health program planning. An educational and ecological approach* (4th ed.). New York: McGraw Hill.
- 5 L.W. Green (2008). If it's an evidence based practice, where's the practice based evidence? *J. FamilyMed.*, 25 (suppl_1), 20-24.
- 6 A. Tashakkori en C. Teddlie (2008). Integrating qualitative and quantitative approaches to research. In: L. Bickman en D.J. Rog (eds.), *Handbook of Applied Social Research Methods* (pp. 283-313). Newbury Park, CA: Sage.
- 7 W. Poortinga et al. (2004). The British 2001 Foot and Mouth Crisis: A comparative study of public risk perceptions, trust and beliefs about government policy in two communities. *Journal of Risk Research*, 7, 73-90.
- 8 A. Bryman (2012). *Social Research Methods* (4th edition). New York: Oxford University Press.
- 9 D. Schwartz en J. Lellouch (2009). Explanatory and pragmatic attitudes in therapeutic trials. *Journal of Clinical Epidemiology*, 62, 499-505.
- 10 H. Michon en H. Kroon (2013). En of het werkt!? Over de pragmatische RCT. In Jurriaan Omló, Martijn Bool & Peter Rensen (red.), *Weten wat werkt. Passend evaluatieonderzoek in het sociale domein* (pp. 185-205). Amsterdam: Uitgeverij SWP.

Informatie over de auteur

Louis Tavecchio is emeritus hoogleraar aan de Universiteit van Amsterdam en emeritus lector aan de Hogeschool van Amsterdam.