

“Alternatieven voor RCT bij de evaluatie van effectiviteit van interventies!?”

Eindrapportage



28 oktober 2015

Dr. Anne May, Julius Centrum voor Gezondheidswetenschappen en Eerstelijns Geneeskunde, UMC Utrecht

Dr. Jolanda Mathijssen, TRANZO, Wetenschappelijk Centrum voor Zorg en Welzijn, Tilburg University

INHOUD

Dankwoord		Pagina 5
Samenvatting		Pagina 7
Hoofdstuk 1	Inleiding	Pagina 9
Hoofdstuk 2	De RCT en alternatieve onderzoeksdesigns	Pagina 11
Hoofdstuk 3	RCT en alternatieve onderzoeksdesigns bij eHealth interventies	Pagina 17
Hoofdstuk 4	Resultaten en conclusies uit de interviews	Pagina 21
Hoofdstuk 5	Resultaten en conclusies uit de discussietafels (ZonMw invitationale)	Pagina 37
Hoofdstuk 6	Algemene beschouwing: Stel de vraagstelling centraal	Pagina 45
Bijlage I	Interview vragen	Pagina 53
Bijlage II	Geïnterviewde personen	Pagina 54

Dankwoord

We willen iedereen bedanken, die bijgedragen heeft aan het tot stand komen van deze eindrapportage. Allereerst ZonMw voor het zien van het belang van het onderwerp “Alternatieven voor RCT”. Daarnaast zijn we uiteraard veel dank verschuldigd aan alle geïnterviewden zoals genoemd in Appendix II en aan de deelnemers aan de invitationals voor hun waardevolle bijdrage aan de discussie. We danken Dr. Olaf Dekkers, Prof. Lex Burdorf, Drs. Juanita Heymans, Prof. Patrick Bossuyt, Dr. Rolf Groenwold, Prof. Maroeska Rovers, Prof. Jet Smit, Dr. Lies van Gennip, Prof. Anske van der Bom, Prof. Anton Kunst en Prof. Pim Assendelft voor hun presentaties of de leiding van de discussietafels tijdens de invitationals. Tot slot hebben de volgende personen ons erg geholpen bij het verfijnen van onze rapportages: Prof. Petra Peeters door haar commentaar op het startdocument “Alternatieven voor RCT bij de evaluatie van effectiviteit van interventies!?” en Dr. Rolf Groenwold, Prof. Pim Assendelft, Prof. Patrick Bossuyt, Prof. Henk Garretsen, Dr. Djoeke van Dale en Drs. Juanita Heymans voor hun commentaar op deze eindrapportage.

Samenvatting

Kernboodschap: *Zet bij elke onderzoeksvoorstel de vraagstelling centraal. Kies het passende design, onderbouw waarom dit design in deze specifieke situatie het meest geschikt is en wat deze keuze betekent voor de interne validiteit en toepasbaarheid.*

Diverse opvattingen over de waarde en de (on)mogelijkheden van gerandomiseerd vergelijkend onderzoek (randomized controlled trials – RCTs) voor (kosten)effectiviteitsonderzoek leiden vaak tot discussies binnen de ZonMw-beoordelingscommissies, maar ook tussen de projectindieners, de ZonMw beoordelingscommissies, externe referenten en gebruikers van onderzoeksresultaten.

In opdracht van ZonMw is onderzoek uitgevoerd naar de zienswijzen over de waarde en (on)mogelijkheden van RCT en alternatieve designs van zowel onderzoekers, commissieleden als eindgebruikers uit de werkvelden preventie, care en cure. Het doel was beter inzicht te krijgen in, het beter grip krijgen op en het verder stimuleren van de discussie over de wenselijkheid en (on)mogelijkheid van een RCT en alternatieve designs.

Deze rapportage is het vervolg op de kennissynthese ‘Alternatieven voor randomized controlled trials in onderzoek naar de effectiviteit van interventies’, die, ook in opdracht van ZonMw, is uitgevoerd door het [Dutch Cochrane Centre](#) (Langendam et al., 2013). Voor het schrijven van deze eindrapportage is in aanvulling op het Cochrane rapport een korte literatuurstudie uitgevoerd, hebben er interviews plaatsgevonden met 32 stakeholders uit de verschillende domeinen preventie, care en cure en is er een [invitationeel georganiseerd](#), waar, naast plenaire lezingen, ongeveer 90 stakeholders onder leiding van discussieleiders met elkaar in gesprek gingen over zes verschillende aspecten van het onderwerp ‘alternatieven voor RCTs’.

Hoewel alle geïnterviewden, onafhankelijk van hun achtergrond, het er over eens zijn dat de RCT een valide design is voor effectiviteitsonderzoek, worden er ook beperkingen genoemd. Daarnaast zijn sommige onderzoekers overtuigd dat het bij een effectiviteitsvraag altijd mogelijk is om een RCT uit te voeren, maar anderen zijn het daar helemaal niet mee eens en komen met voorbeelden wanneer het niet lukt of wenselijk is om een RCT uit te voeren. Als het al evident is dat er een effect is of als een goede inschatting gemaakt kan worden van de situatie vóór de interventie, dan is voor het bepalen van de effectiviteit niet nodig om een RCT uit te voeren. Tot slot is een RCT niet mogelijk als het om zeldzame aandoeningen gaat, als het ethisch niet verantwoord is om een groep een bepaalde interventie te onthouden of wanneer random toewijzen niet lukt.

De RCT heeft een hoge bewijskracht en is een goede manier om causaliteit aan te tonen. Echter, er zijn ook nadelen te noemen. Een RCT is vooral geschikt voor enkelvoudige interventies, terwijl veel interventies complex zijn en vaak plaatsvinden in een snel veranderende werkelijkheid waar de onderzoeker geen invloed op heeft. Bovendien heeft de RCT een lage generaliseerbaarheid.

Bij de vraag naar alternatieven voor RCTs, blijken er drie mogelijke interpretaties te zijn.

Ten eerste is er de letterlijke interpretatie, waarbij het gaat om alternatieve designs (of statistische correctietechnieken), die het concept van de RCT zo veel mogelijk benaderen. Hierbij is het de grote uitdaging om confounding en selectie bias te voorkomen en de juiste controlegroep te kiezen. Ten

tweede, wordt er gedacht aan aanvullende designs, die, naast een eventuele RCT, nodig zijn om antwoord te kunnen geven op vragen naar onderliggende (biologische) mechanismen, (zeldzame) bijwerkingen, het evalueren van het proces en de implementatie van de interventie. Ten derde wordt ook genoemd dat men een design moet kiezen dat bij de vraagstelling past en dat, afhankelijk van de vraagstelling, niet persé een RCT hoeft te benaderen. In het algemeen is er behoefte aan consensus over wanneer al dan niet een RCT uitgevoerd moet worden.

Tijdens de invitational bleek dat veel van de discussies, die gevoerd worden over het meest passende design, terug te voeren zijn op het feit dat het 'evalueren' en het 'onderzoeken van de effectiviteit' van een interventie vaak door elkaar gehaald worden. *Evalueren* is een breed begrip: het begint bij de ontwikkeling van de interventie en eindigt bij de uiteindelijke implementatie. Het onderzoeken van de *effectiviteit* van de interventie is een onderdeel van het totale proces. Voor elk onderdeel van de evaluatie van interventies kunnen specifieke onderzoeksvragen geformuleerd worden, waarvoor er dan één of meerdere passende onderzoeksdesigns zijn. Bij de effectiviteitsvraag speelt randomisatie een belangrijke rol (om de uitwisselbaarheid te garanderen), voor veel andere onderzoeksvragen is echter randomisatie niet nodig of te wel niet passend. De onderzoeker moet altijd het gekozen design onderbouwen en aangeven wat deze keuze betekent voor de interne validiteit en toepasbaarheid van de onderzoeksresultaten.

Het is niet mogelijk om een 'kookboek' te schrijven met universele recepten om te komen tot het meest passende onderzoeksdesign. Elk project heeft zijn eigen specifieke vraagstelling(en) en kent zijn eigen randvoorwaarden. Taak van de onderzoeker is het om de vraagstelling(en) scherp te formuleren en de designkeuze te onderbouwen. Om de onderzoeker hierbij te helpen worden in hoofdstuk 6 voorbeelden gegeven van mogelijke onderzoeksvragen omtrent de evaluatie van interventies en het (de) passende onderzoeksdesign(s).

Uit zowel de interviews als de invitational zijn aanbevelingen geformuleerd voor ZonMw. In de subsidieaanvraagformulieren dient aparte ruimte gereserveerd te worden waar de projectaanvrager onderbouwt waarom gekozen is voor een specifiek design. Belangrijk hierbij is dat onderzoekers, reviewers en commissieleden over voldoende kennis van, voor het betreffende programma relevante, onderzoeksmethodologie beschikken.

Tijdens de invitational bleek dat er nog onvoldoende kennis is over hoe alternatieve designs werken en in welke setting alternatieven geschikt of zelfs gewenst zijn. Naast educatie blijkt er ook behoefte aan programmaruimte voor methodologisch onderzoek hiernaar.

Voor complexe interventies is een ruimer programmakader wenselijk waarin de verschillende fasen uit de ontwikkeling en evaluatie van een interventie gefinancierd kunnen worden met gedefinieerde go/no go momenten. Procesevaluaties en haalbaarheidsstudies kunnen dan ook onderdeel zijn van de onderzoeksfinanciering.

Hoofdstuk 1 Inleiding

Diverse opvattingen over de waarde en de (on)mogelijkheden van gerandomiseerd vergelijkend onderzoek (randomized controlled trials – RCTs) voor (kosten)effectiviteitsonderzoek in *preventie, care en cure*, leiden vaak tot discussies binnen de ZonMw-beoordelingscommissies, maar ook tussen de projectindieners, de ZonMw beoordelingscommissies, externe referenten en gebruikers van onderzoeksresultaten. Onderzoekers en beoordelaars redeneren soms vanuit andere perspectieven met als gevolg dat niet alleen een verschil van mening ontstaat over de meest passende dan wel haalbare onderzoeksmethodiek, maar ook onduidelijkheid over de reden waarom de ene methodiek de voorkeur heeft boven de andere.

Echter, kiezen voor een methode passend bij het benodigde bewijs klinkt eenvoudig, maar in werkelijkheid vraagt dit om een complexe afweging. Een afweging waarbij onderzoekers zoeken naar een methode die de vraagstelling beantwoordt en die voldoet aan de validiteitseisen waar het bewijs om vraagt, terwijl zij ook nog eens rekening houden met projectspecifieke omstandigheden.

Om beter inzicht te krijgen in en om beter grip te krijgen op de discussie over de wenselijkheid en mogelijkheid van een RCT en over de verschillende interpretaties van de vraag naar alternatieven voor RCTs heeft ZonMw gevraagd om onderzoek te doen naar de meningen van onderzoekers, commissieleden en eindgebruikers. Hiertoe zijn diverse stakeholders geïnterviewd en zijn tijdens een invitational stakeholders met elkaar in gesprek gegaan over de mogelijkheden en waarde van de RCT, over (de definitie van) alternatieven voor RCTs en over de factoren die belangrijk zijn om te komen tot de juiste designkeuze.

ZonMw heeft in 2012 aan Prof. Patrick Bossuyt (AMC Amsterdam) gevraagd om een kennissynthese uit te voeren over alternatieven voor RCTs. Dit heeft, in samenwerking met het [Dutch Cochrane Centre](#), geleid tot het rapport 'Alternatieven voor randomized controlled trials in onderzoek naar de effectiviteit van interventies' (Langendam et al., Dutch Cochrane Centre, 2013). Dit rapport geeft een beschrijvend overzicht van onderzoeksmethodieken en de bijbehorende consequenties voor de interne en externe validiteit.

In aanvulling op het Cochrane rapport is, een korte literatuurstudie uitgevoerd en hebben er interviews plaatsgevonden met 32 stakeholders uit de verschillende domeinen en is er een invitational georganiseerd, waar ongeveer 90 stakeholders met elkaar in discussie zijn gegaan.

In **hoofdstuk 2** wordt een korte schets gegeven van de discussie, die gevoerd wordt over onderzoeksdesigns om effectiviteit van interventies aan te tonen, en wordt een samenvatting gegeven van de resultaten uit de Cochrane review. Vanwege de dynamiek in de ontwikkeling van eHealth en de breedte van het begrip eHealth wordt in **hoofdstuk 3** specifiek ingegaan op de discussie met betrekking tot evaluatie van eHealth interventies. Een samenvatting van de belangrijkste uitkomsten uit de interviews staat in **hoofdstuk 4**. Dit gebeurt aan de hand van de volgende onderwerpen: mening over de RCT, mening over alternatieven voor de RCT, alternatieven voor *óf* aanvullingen op de RCT, discussiepunten/problemen met de RCT in het veld, situaties wanneer een RCT onmogelijk is, internationale ontwikkelingen en tips voor ZonMw. Drie interviews zijn afgenomen met een focus op de

evaluatie van eHealth interventies, waarbij aanvullend werd gevraagd hoe men in onderzoek recht kan doen aan de dynamiek in de ontwikkeling van eHealth. Het hoofdstuk eindigt met een aantal conclusies. In **hoofdstuk 5** worden de belangrijkste conclusies uit de discussietafels van de invitational gegeven. Deze rapportage eindigt met een algemene beschouwing in **hoofdstuk 6**.

Hoofdstuk 2 De RCT en alternatieve onderzoeksdesigns

De vraag of een interventie effectief is, kan op verschillende manieren beantwoord worden. Vanuit het perspectief van de praktijk kan een interventie als effectief beschouwd worden als de doelen behaald worden, zoals bijvoorbeeld het verminderen van problemen of het tevreden zijn van respondenten. Maar vanuit wetenschappelijk perspectief is een interventie pas effectief als de gevonden resultaten ook daadwerkelijk toegeschreven kunnen worden aan de interventie en niet aan andere factoren (Oomlo, Bool en Rensen, 2013).¹ Dit laatste aspect gaat over de interne validiteit. Bij interne validiteit gaat het over de mate waarin de conclusies van het onderzoek geldig zijn voor de onderzoeksgroep, en er dus geen sprake is van confounding of bias. Om deze interne validiteit te optimaliseren wordt er vaak voor gekozen om de onderzoeksgroep en de interventie sterk te begrenzen (Tavecchio, 2014)². Dit is wat feitelijk gebeurt in een RCT. De te onderzoeken interventie wordt uitgevoerd bij een experimentele groep en die wordt vergeleken met een controlegroep. De controlegroep is een vergelijkbare groep proefpersonen, die of geen of een andere interventie (of standaard zorg) aangeboden krijgt. Om ervoor te zorgen dat bij aanvang geen (prognostische) verschillen tussen de groepen bestaan, die de resultaten kunnen beïnvloeden, wordt toewijzing van personen aan de controle- of experimentele groep random bepaald. Een hogere interne validiteit kan echter tot gevolg hebben dat de generaliseerbaarheid van de resultaten afneemt.

Discussie over designs

De discussie over RCTs en de alternatieven wordt niet alleen bij ZonMw gevoerd, maar ook in de literatuur en bij instituten en organisaties, die op basis van onderzoek beslissingen moeten nemen. De discussie spitst zich toe op de vraag of de RCT nu wel of niet als gouden standaard, als het beste design om effectiviteit van interventies en programma's aan te tonen, beschouwd moet worden. Met andere woorden; is de RCT het meest optimale design om effectiviteit te meten of is het één van de mogelijke designs en zijn andere designs, afhankelijk van de specifieke effectiviteitsvraag, juist beter geschikt?

Is er een hiërarchie in onderzoeksdesigns?

Het 'Centre of Evidence-Based Medicine' (<http://www.cebm.net/>) heeft in 2009 het 'levels of evidence' uitgebracht waarin een hiërarchie van evidentie wordt gepresenteerd. Hierbij worden andere designs dan de RCT als minder waardevol beschouwd. Niet iedereen is het echter eens met deze benadering en sommigen pleiten voor de integratie van meerdere studiedesigns met betrekking tot belangrijke uitkomstmaten. Peineman et al. (2013)³ laten in een systematische review zien dat verschillende onderzoeksdesigns (gerandomiseerde en niet-gerandomiseerde studies), die dezelfde vraag willen beantwoorden in de helft van de voorbeelden andere resultaten opleverden. Zowel niet

¹Oomlo, J., Bool, M., & Rensen, P (2013). Weten wat werkt. Passend evaluatieonderzoek in het sociale domein. Amsterdam, Movisie, SWP.

²Tavecchio, L. (2014). Effectiviteit in praktijkgericht onderzoek. Op zoek naar alternatieven voor de randomized controlled trial (RCT). *Kwaliteit in Zorg*, 2, 16-20.

³Peinemann F, Tushabe DA, Kleijnen J: Using multiple types of studies in systematic reviews of health care interventions--a systematic review. *PLoS One* 2013, 8(12):e85035.

gerandomiseerde als gerandomiseerde studies presenteren onzekere resultaten zonder zeker te weten wat de richting en grootte van het effect is. De auteurs bevelen daarom aan om in systematische reviews niet alleen resultaten van RCTs mee te nemen, maar ook resultaten van niet-gerandomiseerde valide uitgevoerde (prospectieve) observationele studies, waarbij in de statistische analyse gecontroleerd wordt voor versturende factoren. Voor het beoordelen van de validiteit van niet gerandomiseerd onderzoek heeft de Cochrane Collaboration in 2014 een checklist ontworpen: “A Cochrane Risk Of Bias Assessment Tool: for Non- Randomized Studies of Interventions (ACROBAT-NRSI)”⁴.

In hun artikel over alternatieven voor randomisatie bij evaluatie van interventies in de publieke gezondheid concluderen Bonell en collega’s⁵ dat bewijs van niet-gerandomiseerde studies overtuigender is als alle mogelijke versturende factoren goed in kaart gebracht kunnen worden en ervoor gecorrigeerd wordt. Daarnaast is er bewijs nodig over de aanwezigheid van een causaal pad tussen de interventie en de uitkomsten en moet het effect groot zijn. De auteurs merken verder op dat in veel gevallen – als er duidelijk equipoise is, dat wil zeggen als men werkelijk niet weet wat de beste behandeling/interventie is - mogelijke barrières voor RCTs te beslechten zijn en randomisatie wel mogelijk is. Ook wijzen zij op het gevaar dat als men alleen bewijs verkregen in RCTs toelaat, dit tot een uitsluiting van interventies kan leiden die niet (meer) in een RCT onderzocht kunnen worden, bijvoorbeeld op grond van preferenties of ethische gronden.

Vandenbroucke⁶ stelt dat in de medische wetenschap de kracht van randomisatie is dat “bias by indication” wordt voorkomen, wat in observationeel onderzoek maar zeer zelden mogelijk is. Bij het meten van de gewenste (“intended”) effecten van een interventie is daarom randomisatie nodig. Voor het bepalen van nadelige effecten (unintended effects, adverse events), onderzoek naar oorzaken van ziekte of prognose en diagnose kan niet gerandomiseerd onderzoek gebruikt worden.

Beoordelingscriteria voor interventieonderzoek

Verschillende instituten die opdracht geven tot onderzoek, richtlijnen ontwikkelen of beslissingen nemen gebaseerd op onderzoek houden verschillende beoordelingscriteria aan (zie voor voorbeelden Hoofdstuk 6 van het [Cochrane rapport](#)). Het valt op dat steeds meer organisaties gebruik maken van de GRADE (‘Grading of Recommendations Assessment, Development and Evaluation’) systematiek (bijvoorbeeld de Federatie Medisch Specialisten, Zorginstituut Nederland, Integraal Kankercentrum Nederland en de Erkeningscommissie Interventies en internationaal: de Cochrane Collaboration, BMJ en de WHO). Hierbij wordt de kwaliteit van de, voor een specifieke vraag, geselecteerde literatuur beoordeeld aan de hand van: methodologische aspecten, belang van de resultaten, generaliseerbaarheid en het onderzoeksdesign. Er wordt naar drie aspecten gekeken: de grootte van het effect, de wenselijkheid, en de geloofwaardigheid van de te beoordelen interventie. Daarbij kan de

⁴<https://sites.google.com/site/riskofbiastool//home/read-more>

⁵Bonell CP, Hargreaves J, Cousens S, Ross D, Hayes R, Petticrew M, Kirkwood BR. Alternatives to randomisation in the evaluation of public health interventions: design challenges and solutions. *J Epidemiol Community Health*. 2011 Jul;65(7):582-7

⁶Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med*. 2008 ;5(3):e67

kwaliteit van RCTs worden afgewaardeerd, respectievelijk de kwaliteit van observationele studies worden opgewaardeerd, afhankelijk van bepaalde, omschreven beoordelingsgronden.

De Erkenningscommissie Interventies (bestaande uit 9 deelcommissies)⁷ beoordeelt de effectiviteit van interventies op een uniforme manier en kan een erkenning afgeven in één van vier categorieën, die oplopen in mate van bewijskracht. In de nieuwe notitie over de criteria voor gezamenlijke kwaliteitsbeoordeling 2015-2018 wordt expliciet genoemd dat een RCT niet altijd persé noodzakelijk is om effectiviteit aan te tonen. Een goed uitgevoerd observationeel onderzoek kan even hoog gewaardeerd worden of zelfs hoger dan een RCT van slechte kwaliteit. Volgens de criteria van de Erkenningscommissie is een design met een zeer sterke bewijskracht 'een experimentele onderzoeksopzet (d.w.z. er is een aselechte toewijzing van subjecten aan onderzoeksgroepen) of een andere opzet die de causale relatie tussen interventie en effect voldoende overtuigend aantoot'.

Alternatieve designs voor RCTs

De vraag naar alternatieven voor RCTs kan letterlijk geïnterpreteerd worden, dat wil zeggen welke onderzoeksdesigns kan men gebruiken als een alternatief design voor een RCT? Maar het kan ook geïnterpreteerd worden als een vraag naar onderzoeksdesigns die naast een RCT aanvullende informatie verschaffen over de effectiviteit, de implementatie of de bijwerkingen van de interventie. In het in opdracht van ZonMw gemaakte Cochrane rapport "Alternatieven voor Randomized Controlled Trials in onderzoek naar de effectiviteit van interventies" van Langendam et al. (2013) is de eerste interpretatie gehanteerd en werd de literatuur systematisch doorzocht naar alternatieve onderzoeksmethodieken. Vervolgens werden de gevonden methodieken onderverdeeld in vier categorieën:

- 1) Experimenteel onderzoek (toewijzing van de interventie, met parallelle onderzoeksgroepen)
 - RCT
 - Cluster RCT
 - Comprehensive cohort onderzoek
 - Pragmatische trial
 - Patient preference trial
- 2) Quasi-experimenteel onderzoek (toewijzing van de interventie zonder parallelle onderzoeksgroepen)
 - Gecontroleerde voor-na vergelijkingen
 - Stepped wedge design
 - Tijdsreeksonderzoek
 - N=1 opzet
 - Vergelijkende gevalsstudies
 - Regression discontinuity design

⁷Jeugdzorg en psychosociale/pedagogische preventie; Jeugdgezondheidszorg, preventie en gezondheidsbevordering; Ontwikkelingsstimulering, onderwijs-gerelateerd aanbod en jeugdwelzijn; Gezondheidsbevordering en preventie voor volwassenen en ouderen; Sport en Bewegen; Maatschappelijke ondersteuning, participatie en veiligheid; Ouderenzorg; Gehandicaptenzorg; Langdurige Geestelijke Gezondheidszorg

- 3) Observatieel vergelijkend onderzoek (geen toewijzing van de interventie)
 - Cohortonderzoek
 - Case-control onderzoek
- 4) Niet vergelijkend observatieel onderzoek
 - Niet-vergelijkende gevalsstudies
 - Ecologisch onderzoek

Voor meer uitleg over de verschillende designs inclusief voorbeelden en een overzicht van statistische technieken voor het vergroten van vergelijkbaarheid tussen groepen verwijzen wij naar het [rapport](#).

Kwalitatief onderzoek bij aantonen van effectiviteit

Het kwalitatieve onderzoek wordt kort genoemd in het rapport, maar vooral als waardevolle ondersteuning aan effectiviteitsonderzoek, en niet als mogelijk alternatief voor het aantonen van effectiviteit van een interventie. Toch is er steeds meer discussie over de waarde van kwalitatieve onderzoeksmethoden in effectonderzoek en of een RCT alléén kwantitatief van aard dient te zijn of dat er ook kwalitatieve uitkomsten gebruikt kunnen worden. De vraag, die hier onder ligt, is of kwalitatief onderzoek voldoende bewijskracht bezit om een uitspraak te doen over causaliteit. Lub (2014)⁸ beschrijft in zijn boek “Kwalitatief evalueren in het sociale domein” het verschil tussen een causale *beschrijving* en een causale *verklaring*. In sociaal-wetenschappelijk onderzoek is kwantitatief onderzoek vooral gericht op de causale beschrijving; is er een causaal verband tussen twee of meer variabelen? Terwijl in kwalitatief onderzoek meer de waaromvraag centraal staat en dus op zoek gaat naar de verklaring achter de gevonden relatie.

Nieuwe ontwikkelingen

Een recent beschreven nieuw onderzoeksdesign, dat nog niet in het Cochrane review wordt beschreven is het ‘Cohort multiple randomized controlled trial’ (cmRCT) design.⁹ Dit design past in de categorie “experimenteel onderzoek”. In 2010 hebben Relton en collega’s het cmRCT design voorgesteld als een alternatief voor *pragmatische* RCTs. Dit design heeft als doel trials efficiënter en patiëntvriendelijker uit te voeren. Uitgangspunt van een cmRCT is een observatieel cohort van patiënten met eenzelfde aandoening. De patiënten worden over de tijd gevolgd, waarbij regelmatig metingen afgenomen worden. Deelnemers aan het cohort hebben toestemming gegeven om prospectief gevolgd te worden en om gerandomiseerd te worden voor toekomstige interventies. Als men de effectiviteit van een bepaalde interventie wil onderzoeken, kan men binnen het cohort de patiënten identificeren die aan de inclusiecriteria voldoen en per lot bepalen wie de interventie krijgt aangeboden. Alleen de interventiegroep wordt uitgenodigd voor de interventie; de controlegroep krijgt standaard zorg en wordt niet geïnformeerd over de studie. Uitkomsten in de interventiegroep worden vergeleken met de uitkomsten in de controlegroep, verkregen door de metingen die regelmatig in het cohort gedaan worden.

⁸ Lub, V. (2014). Kwalitatief evalueren in het sociale domein. Mogelijkheden en beperkingen. Den Haag; Boom Lemma.

⁹ Relton C, Torgerson D, O’Cathain A, Nicholl J. Rethinking pragmatic randomised controlled trials: introducing the “cohort multiple randomised controlled trial” design. *BMJ*. 2010;340:c1066.;

In de “precision medicine” worden op dit moment nieuwe onderzoeksdesigns toegepast om de effectiviteit van nieuwe medicijnen te onderzoeken, bijvoorbeeld om rekening te kunnen houden met tumor heterogeniteit bij patiënten met kanker: de zogenaamde ‘N-of-1’, ‘basket’ en “umbrella’ trials.^{10,11}

¹⁰Kummar S, Williams PM, Lih CJ, Polley EC, Chen AP, Rubinstein LV, Zhao Y, Simon RM, Conley BA, Doroshow JH. Application of molecular profiling in clinical trials for advanced metastatic cancers. *J Natl Cancer Inst.* 2015;107(4).

¹¹Schork NJ. Personalized medicine: Time for one-person trials. *Nature.* 2015;520(7549):609-11.

Hoofdstuk 3 RCT en alternatieve onderzoeksdesigns bij eHealth interventies

Tijdens dit project is een eerste verkenning uitgevoerd op het terrein van de methodologie voor evaluaties van eHealth interventies. De aanleiding hiervoor was het Consumenten eHealth advies dat de Raad voor de Volksgezondheid en Zorg (RVZ) in het voorjaar 2015 heeft uitgebracht.¹² Eén van de aanbevelingen in het advies luidt: ‘Nictiz en Zorginstituut Nederland ontwikkelen in samenwerking met universiteiten en universitair medische centra en in aansluiting op internationale ontwikkelingen een kader van geschikte methodologie voor de wetenschappelijke evaluatie van de klinische effectiviteit (klinisch nut of meerwaarde) van medische toepassingen van consumenten eHealth’.¹³ Men verwacht dat reguliere zorg en consumenten eHealth steeds meer met elkaar vervlochten zullen raken. Om dit mogelijk te maken is inzicht in het klinische nut van toepassingen van belang.

Definities van eHealth

EHealth is een breed begrip en er worden in het veld verschillende definities gehanteerd. Een definitie werd in 2001 door Eng voorgesteld:

*EHealth is het gebruik van nieuwe informatie- en communicatietechnologieën, en met name internet-technologie, om gezondheid en gezondheidszorg te ondersteunen of te verbeteren.*¹⁴

Deze definitie beperkt zich niet tot één specifieke technologie of tot specifieke gebruikersgroepen (bijvoorbeeld alléén de patiënt of alléén de hulpverlener) en wordt dan ook in Nederland gehanteerd door bijvoorbeeld RVZ, Nictiz en het NIVEL. Voor bepaalde deelgebieden van eHealth kunnen ook aparte definities gehanteerd worden. Zo definieert RVZ bijvoorbeeld consumenten eHealth als *direct op de markt zonder tussenkomst van zorgverleners aangeboden informatie- en communicatietechnologie, die beoogt de gezondheid van gebruikers te ondersteunen of te verbeteren*. Ook kan een nieuwe terminologie ontstaan, zoals mHealth voor mobiele technologie in de gezondheidszorg.

Eysenbach¹⁵ geeft aan dat eHealth méér inhoudt dan alleen een technische ontwikkeling, maar ook een attitude; een manier om denken te beïnvloeden door middel van technologie in een dynamische omgeving:

*“EHealth is an emerging field in the intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies. In a broader sense, the term characterizes not only a technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology.”*¹⁹

¹² http://rvz.net/uploads/docs/Advies_Consumenten_eHealth.pdf

¹³ <http://rvz.net/publicaties/bekijk/consumenten-ehealth#aanbevelingen>

¹⁴ Eng T. The e-Health Landscape – a terrain map of emerging information and communication technologies in health and health care. Princeton NJ: The Robert Wood Johnson Foundation; 2001.

¹⁵ Eysenbach G. What is e-health? J Med Internet Res. 2001;3(2):e20. <http://www.jmir.org/2001/2/e20/>

Discussie omtrent studiedesigns voor de evaluatie van eHealth

Zoals ook in het RVZ advies naar voren komt, is er nog geen eenduidige visie op hoe de klinische effectiviteit van eHealth bewezen zou moeten worden. Liu en Wyatt¹⁶ gaan hierop in in hun publicatie die zich richt op de evaluatie van klinische informatiesystemen (bijvoorbeeld systemen voor decision support, patiënt monitoring of elektronische recepten). De auteurs merken op dat bij sommigen binnen de eHealth gemeenschap de mening heerst dat RCTs maar een beperkte rol spelen bij de evaluatie van klinische effectiviteit. De redenering erachter is dat eHealth (informatie technologie) fundamenteel verschilt van geneesmiddelinterventies en dat daarom een RCT irrelevant is.

De volgende algemene bezwaren voor het gebruik van RCTs in eHealth onderzoek worden beschreven²⁰:

- A. RCTs zijn niet ethisch.
- B. Klinische informatiesystemen zijn te complex om met een RCT geëvalueerd te worden.
- C. Er is gemengd bewijs dat RCTs succesvol uitgevoerd kunnen worden binnen de medische informatica.
- D. Andere studiedesigns kunnen net zou betrouwbaar bewijs leveren als een RCT.
- E. RCTs zijn te duur.
- F. Theorie en case studies kunnen betrouwbaar voorspellen wat werkt.
- G. Klinische informatiesystemen richten geen schade aan.
- H. RCTs duren te lang en de technologie in de medische informatica ontwikkelt zich te snel.
- I. RCTs beantwoorden vragen, die in de medische informatica niet interessant zijn.

Echter, Liu en Wyatt zijn van mening dat ook bij eHealth interventies, net als bij bijvoorbeeld interventies in de publieke gezondheid, het er niet om gaat óf de RCT wel of niet geschikt is, maar of de onderzoeksvraag past bij het onderzoeksdesign. Elk onderzoek moet beginnen met de juiste onderzoeksvraag en vervolgens moet het daarbij passende design gekozen worden (zie Tabel 1).

Een RCT is het geschikte design om te bepalen of en in welke mate een eHealth interventie de klinische praktijk of patiënt-gerelateerde uitkomsten verbetert, maar dit zou vooral relevant kunnen zijn bij een dure eHealth applicatie/interventie of als de applicatie/interventie een hoog risico op schade met zich meebrengt bij de patiënt, hulpverlener of het gezondheidssysteem.²⁰

¹⁶ Liu JL, Wyatt JC. The case for randomized controlled trials to assess the impact of clinical information systems. *Am Med Inform Assoc.* 2011; 18(2):173-80

Tabel 1 EHealth onderzoeksvragen en onderzoeksdesigns (Liu and Wyatt, 2011)

Onderzoeksvraag	Passend onderzoeksdesign
Wat vinden patiënten van het systeem?	Diepte interview
Hoeveel patiënten vinden het een goed systeem?	Beschrijvende survey
Wat bevalt de patiënten aan het systeem?	Focusgroep
Hoe precies is het advies vergeleken met de gouden standaard?	Cross-sectioneel onderzoek
Is het systeem kosteneffectief?	Kosteneffectiviteitsanalyse
Verbeterd het systeem de relevante uitkomsten?	RCT
Hoe kan het systeem verbeterd worden?	Diepte interview, focusgroep, expert mening

Alternatieven voor RCTs

Liu en Wyatt²⁰ noemen ook alternatieven voor de individueel gerandomiseerde RCT, namelijk cluster randomisatie, stepped wedge design, gecontroleerde voor-na vergelijking, (onderbroken) tijdsreeksonderzoek.

Harris en collega's¹⁷ geven aan dat in het eHealth onderzoek veel voor-na vergelijkingen worden gedaan. Zij bevelen aan dat, als men deze gebruikt, er gekozen wordt voor hoog-niveau quasi-experimentele designs, zoals het gebruik maken van een controlegroep met een baselinemeting in beide groepen en van (onderbroken) tijdsreeksen.

De MOST (Multiphase Optimization Strategy) en SMART (the Sequential Multiple Assignment Randomized Trial) strategie

Collins en collega's¹⁸ stellen een methode voor voor de ontwikkeling en evaluatie van eHealth interventies.

De MOST strategie bestaat uit drie fases: (1) de screening fase waarin de werkende interventie componenten geselecteerd worden, (2) de verfijnende fase waarin het nauwkeurig afstellen van de geselecteerde componenten plaats vindt (bijvoorbeeld bepaling van de optimale dosis), en (3) de bevestigingsfase waarin de effectiviteit van de geoptimaliseerde interventie getoetst wordt. In fase 1 en

¹⁷ Harris AD, McGregor JC, Perencevich EN, Furuno JP, Zhu J, Peterson DE, Finkelstein J. The use and interpretation of quasi-experimental studies in medical informatics. *J Am Med Inform Assoc.* 2006;13(1):16-23.

¹⁸ Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. *Am J Prev Med.* 2007;32(5 Suppl):S112-8.

2 vinden gerandomiseerde experimenten plaats, gebruikmakend van factoriele designs (full, fractioneel of 'response surface'). In fase 3 vindt een standaard RCT plaats.

Door middel van de SMART methodiek kunnen in de tijd variërende interventies ontwikkeld worden. Het is een gerandomiseerd experiment waarbij elk individu op verschillende momenten tijdens de interventie gerandomiseerd kan worden voor verschillende componenten (bijvoorbeeld bij een stoppen met roken interventie worden een keer positieve boodschappen en een andere keer negatieve boodschappen gebruikt). Het einddoel is om evidence-based adaptieve interventie strategieën te ontwikkelen. Deze worden uiteindelijk in een RCT getest. SMART kan ook onderdeel zijn van fase 2 van MOST.

Voor meer gedetailleerde informatie verwijzen we naar de publicatie van Harris et al. (2006).¹⁹

¹⁹Harris AD, McGregor JC, Perencevich EN, Furuno JP, Zhu J, Peterson DE, Finkelstein J. The use and interpretation of quasi-experimental studies in medical informatics. J Am Med Inform Assoc. 2006;13(1):16-23.

Hoofdstuk4 Resultaten en conclusies uit de interviews

In totaal zijn er 32 stakeholders uit de domeinen preventie, care en cure geïnterviewd. De belangrijkste conclusies, geïllustreerd met citaten, worden in dit hoofdstuk gepresenteerd. De interviewvragen staan in Bijlage I, de geïnterviewde personen worden in Bijlage II vermeld.

Een RCT is een valide design voor effectiviteitsonderzoek, maar heeft ook zijn beperkingen.

Alle geïnterviewden, onafhankelijk van hun achtergrond (preventie, care en cure; onderzoeker of eindgebruiker), zijn het er over eens dat een RCT uiterst waardevol is; het is een design met een hoge bewijskracht en de beste manier om causaliteit aan te tonen. Het past bij een toetsende vraag naar de mate van effectiviteit van een interventie en hoort bij de laatste fase van het onderzoek, als voorbereidend onderzoek al heeft plaatsgevonden en de interventie helemaal is uitontwikkeld. Belangrijk uitgangspunt is dat de gekozen methode dient te passen bij de vraagstelling.

Er zijn echter ook kritische kanttekeningen te horen. Geïnterviewden geven aan dat een RCT vooral geschikt is voor enkelvoudige interventies, waarbij het mogelijk is om eventueel versturende variabelen onder controle te houden. Het heeft een hoge interne validiteit, maar vaak een lage generaliseerbaarheid. De vraag is steeds of de gevonden resultaten ook gegeneraliseerd kunnen worden naar andere doelgroepen (andere leeftijd, patiënten met comorbiditeit). De vraag naar generalisatie is methodologisch gezien een vraag naar effectmodificatie: is een effect echt anders dan wat je zou vinden als je het onderzoek in de hele, of een iets anders samengestelde populatie had gedaan?

Interventies in de gezondheidszorg en publieke gezondheid vinden vaak plaats in een snel veranderende werkelijkheid, of onder omstandigheden waar wet- en regelgeving de context bepalen en waar de onderzoeker geen invloed op heeft. Dit bemoeilijkt de uitvoer van een RCT. Bovendien blijkt in de praktijk vaak dat als de setting van het onderzoek wordt losgelaten, een interventie vaak anders wordt uitgevoerd, waardoor de effectiviteit (mogelijk) verdwijnt of vermindert.

Ook wordt opgemerkt dat er vaak voor een RCT gekozen wordt, omdat het idee bestaat dat een studie met een RCT sneller subsidie ontvangt en ook beter te publiceren is.

Citaten:

‘RCT is een goed uitgewerkte en goed doordachte onderzoeksmethode.’

‘RCT is het meest valide design voor effectiviteitsonderzoek.’

‘In een bepaalde context is het een mooie aanpak om zonder beïnvloeding van versturende factoren te testen wat wel of niet werkt.’

‘Het is een illusie te veronderstellen dat de RCT de waarheid vertegenwoordigt.’

‘RCT is prachtig, maar wanneer toepasbaar?’

‘RCT geeft een beperkte blik op de werkelijkheid.’

‘In de literatuur blijkt dat er zelden sprake is van effectmodificatie, of dat de effectmodificatie niet zo sterk is dat daardoor beleid zou veranderen. De vraag is: is het er niet of zien we de effectmodificatie niet? Is er echt variatie in effectiviteit tussen subgroepen?’

‘RCT heeft grote maatschappelijke en klinische impact, omdat RCT vaak (makkelijker) in vooraanstaande wetenschappelijke tijdschriften gepubliceerd wordt’.

‘Niet de RCT is het probleem, maar de groep die je selecteert; die is vaak te smal. Men moet goed nadenken over alle onderzoeksvragen die men wil beantwoorden. Een pragmatische trial met ruime in- en exclusiecriteria kan een oplossing zijn.’

Alternatieven?! voor een RCT

Volgens de Van Dale is een alternatief ‘de andere van twee mogelijkheden, die zich voordoen’. Het gaat om een andere gelijke mogelijkheid. De vraag, die opdoemt en ook in de interviews duidelijk naar voren kwam, is echter of het wel juist is om te spreken van alternatieven voor een RCT. Is het niet zo dat ieder design zijn eigen waarde heeft en dus niet als alternatief, maar gewoon als een ander design beschouwd moet worden, dat ingezet wordt om andere onderzoeksvragen te beantwoorden?

De vraag wat de geïnterviewden vinden van alternatieven voor een RCT wordt verschillend geïnterpreteerd. Een deel van de geïnterviewden interpreteert de vraag letterlijk en gaat in zijn / haar antwoord in op mogelijke onderzoeksdesigns of statistische correctietechnieken, die het concept van de RCT zo goed mogelijk benaderen. Anderen spreken niet over alternatieve, gelijkwaardige, designs, maar over andere designs: aanvullend kwantitatief (observationeel), kwalitatief en mogelijk ook

fundamenteel onderzoek is nodig om onderzoeksvragen naar onderliggende (biologische) mechanismen, nadelige effecten, het proces en de implementatie te beantwoorden.

Citaten:

‘Er worden verschillende definities gehanteerd van alternatieven voor RCTs. Vaak wordt het vertaald naar onderzoek dat zo dicht mogelijk in de buurt komt van een RCT, met veel statistische correcties.’

‘Het juiste design is altijd afhankelijk van de onderzoeksvraag. Sommige vragen kan men beter binnen een cohort beantwoorden, voor andere vragen is een RCT nodig.’

‘Er zijn geen echte alternatieven voor RCTs, andere designs geven antwoord op een andere onderzoeksvraag.’

‘Belangrijk is dat kwaliteit nooit opgeofferd wordt. Het gaat om valide waarnemingen. Als die met een alternatief niet verkregen kunnen worden en een RCT niet mogelijk is, dan géén onderzoek doen.’

‘De RCT is een solide oplossing voor confounding, een bias die je in ander onderzoek maar zelden kunt voorkomen. Als dit wel lukt hoeft je geen RCT te doen, dan voldoet niet-experimenteel onderzoek. Bij de vraag naar een RCT gaat het om het feit dat de toewijzing naar de interventie random gebeurt. Hierdoor zijn de groepen vergelijkbaar op prognostische factoren. Als men de vergelijkbaarheid anders kan verkrijgen, is geen RCT nodig. Maar dit is uitermate zeldzaam het geval’.

‘Bij onderzoek naar de effectiviteit van een geneesmiddel moet men altijd een RCT doen, behalve als effecten enorm zijn (bijvoorbeeld penicilline).’

‘Als alle confounders gemeten zijn en het effect plausibel is dan is een cohortstudie een goed alternatief. Als er veel kans is op oncontroleerbare confounders, is randomisatie nodig.’

‘We moeten meer nadenken over alternatieve designs waarin wel wordt gerandomiseerd: bijvoorbeeld het cohort multiple randomized controlled trial (cmRCT) design. Afhankelijk van de uitkomst heb je hiervoor goed opgezette cohorten nodig, maar als de determinant en het eindpunt uit registraties gehaald kunnen worden, dan hoeft je geen cohort op te zetten. Als je een geneesmiddel af wilt zetten tegen care as usual en niet tegen placebo, dan is de cmRCT een goed design’

De geïnterviewden noemden verschillende voorbeelden van designs of statistische correctietechnieken, die als 'echt' letterlijk alternatief van een RCT gezien kunnen worden. Het idee bestaat dat de discussie over alternatieve designs ook veroorzaakt wordt door een tekort aan kennis zowel bij onderzoekers, reviewers, commissieleden als eindgebruikers.

Citaten:

'Er zijn verschillende statistische technieken beschikbaar die gebruikt kunnen worden bij observationeel onderzoek, bijvoorbeeld instrumentele variabele (IV) analyse, inverse probability weighting, Mendelian randomization, propensity scores, marginal structural modelling. Kennis over deze statistische mogelijkheden moet breder bekend zijn.'

'Bij de evaluatie van complexe interventies kan gebruik gemaakt worden van het stepped wedge design of cluster randomisatie.'

'Goede registraties binnen de bestaande klinische praktijk waardoor men goed zicht krijgt op klachten, behandelingen en verbeteringen over de tijd.'

'Goede gestandaardiseerde registraties in meerdere ziekenhuizen maakt natuurlijke vergelijkingen mogelijk.'

'Cohortonderzoek kan gebruikt worden bij de evaluatie van interventies waar de toewijzing is gebonden aan de voorkeur van de artsen en niet aan bepaalde patiëntkarakteristieken.'

'Patient-preference trial – alleen doen als er geen sterke voorkeur voor één van de studiearmen is.'

'Cross-over trial is een goed alternatief bij geneesmiddelen waar effect uitwast.'

'N=1 trial is weer in opkomst (bijvoorbeeld bij personalized medicine): randvoorwaarde stabiele ziekte. Nu wordt meer geaccepteerd dat er verschillen zijn tussen patiënten/patiënt response (bv genetisch).*

'Natuurlijke experimenten' kunnen een goed alternatief zijn; vooral als er een kort tijdsvenster zit tussen de interventie en het effect.'**

*Schork NJ. Personalized medicine: Time for one-person trials. *Nature*. 2015;520(7549):609-11

** Zie ook de Medical Research Council richtlijn voor het gebruik van natuurlijke experimenten in het publieke gezondheid onderzoek: Craig et al. *J Epidemiol Community Health*. 2012; 66(12): 1182–1186.

Alternatieven voor óf aanvullingen op RCT?

Naast de vraag of er al dan niet alternatieven bestaan voor de RCT (de letterlijke interpretatie), wordt door de geïnterviewden ook vaak de vraag gesteld of we niet ook andere designs nodig hebben als aanvulling op de RCT. Een RCT geeft puur informatie of een interventie wel of niet effectief is bij de onderzoeksgroep. Antwoorden over (biologische) mechanismen (mediatie), het proces van de uitvoer van de interventie en de compliantie met de interventie, subgroepen waarbij de interventie meer of minder effectief is (moderatie), ervaringen/meningen van de patiënt en individuele behandel-effecten moeten door een ander type onderzoek of aanvullende analyses verkregen worden. Ook wordt genoemd dat men voor de start van een RCT meer tijd moet besteden aan het ontwikkelen van de interventie samen met de doelgroep, en aan het verkrijgen van commitment van alle stakeholders. Voorafgaand aan een RCT zou ook meer gebruik gemaakt moeten worden van meta-analyses met data van individuele patiënten (IPD-meta-analyse) om bijvoorbeeld de subgroep te identificeren die nog een RCT nodig heeft. Hierdoor kunnen RCTs kosten-effectiever ontworpen en uitgevoerd worden.

Citaten:

‘RCT kan goed aangevuld worden met ander type onderzoek (mixed methods). Bijvoorbeeld kwalitatief onderzoek kan belangrijke informatie geven voor de implementatie van de interventie.’

‘Niet alleen kijken naar effectiviteit, maar ook naar onderliggende mechanismen en bij wie het werkt. Dus niet eindeloos nieuwe RCTs uitvoeren, maar de gegevens uit RCTs gebruiken om te kijken bij wie het werkt, zodat de behandeling in de praktijk beter toegewezen kan worden.’

‘Voor het bewijs van effectiviteit is niet alleen een RCT nodig, ook een biologische onderbouwing is belangrijk. Als bijvoorbeeld homeopathie effectief blijkt in een RCT, is op grond van een ontbrekende biologische onderbouwing de effectiviteit nog steeds onwaarschijnlijk.’

‘RCT is maar één stuk van het bewijs. Na een RCT komen vragen over bijwerkingen van geneesmiddel of opname in het pakket’.

Voor het bepalen van effectiviteit is niet altijd een RCT nodig

Er zijn situaties, die maken dat voor het bepalen van effectiviteit een RCT niet meer nodig is. Dit is het geval als het evident is dat er een effect is (het is of bij iedereen effectief of er is een groot effect), of als

een goede inschatting gemaakt kan worden van de situatie vóór de interventie. Of als bijvoorbeeld blijkt dat een nieuwe (screenings)test beter en makkelijker uitvoerbaar is dan de oorspronkelijke test.

Citaten:

‘Als iedereen ervan overtuigd is dat iets werkt, waarom zou je dan een RCT doen?’

‘Twintig personen met een dwangstoornis, die al 10-15 jaar ernstige klachten hebben. Na invoer van een nieuwe therapie, knappen alle 20 personen op.’

‘Reanimatie: het nut hiervan is evident en hoeft niet onderzocht te worden. Men kan in ieder geval voorspellen dat bij niets doen, iedereen overlijdt. Ander voorbeeld: anesthesie.’

‘Antibiotica werd toegepast op soldaten met infecties, daarna gebruikt bij patiënten met longontsteking. De effecten waren zo groot dat er geen trial meer is gedaan.’

‘Bij sommige middelen kan men vooraf al weten dat ze gelijkwaardig zijn (op grond van biologie). Dan is geen RCT meer nodig.’

Sommige geïnterviewden benadrukken de waarde van modelstudies om effecten van een bepaalde (complexe) interventie te schatten.

Citaten:

‘Effecten van complexe interventies kunnen gemodelleerd worden: complexe interventies grijpen niet aan op één aspect: bijvoorbeeld de veiligheid aanpassen in wijk, heeft ook effecten op toekomstige bewoners en de winkels die er zijn. Effecten hiervan kunnen met modellen onderzocht worden’.

‘Of het modelleren van de effecten van luchtverontreiniging in een stad op de gezondheid.’

Effectiviteitsonderzoek door middel van een RCT vraagt om een terdege voorbereiding en kent specifieke problemen in de uitvoer

Onderzoekers geven aan dat gezien het specifieke karakter van een RCT de voorbereiding over het algemeen lang is en dat zij tegen praktische problemen aanlopen. Dit heeft te maken met de keuze van de controle conditie, met het random toewijzen van respondenten (met als consequentie dat patiënten bij het nemen van de beslissing om deel te nemen niet weten of zij de interventie wel of niet krijgen), de complexiteit van de interventies en de lange duur van het onderzoek. Dit lijkt belangrijke redenen te zijn waarom er vraag is naar alternatieve designs.

Ook wordt aangegeven dat een RCT geen garantie levert voor goed onderzoek, dat altijd tot een valide conclusie leidt. Ieder onderzoek dient goed voorbereid te worden. Immers een slecht opgezet onderzoek, met bijvoorbeeld de verkeerde uitkomstmaten/meetinstrumenten, niet de juiste onderzoeksgroep of te weinig respondenten, levert slechte resultaten – óók al is het een RCT.

Citaten:

‘Het definiëren van de juiste controleconditie bij een RCT is moeilijk.’

‘Vaak is de controlegroep een groep die niets krijgt, terwijl het beter zou zijn als men daarnaast ook arm(en) heeft met andere interventie(s). Zodat men uiteindelijk beter weet wat echt het beste is.’

‘Bij psychologisch onderzoek of onderzoek in het sociale domein kan geen dubbelblind onderzoek uitgevoerd worden. Als men patiënten vraagt om mee te doen aan een RCT, schept men verwachtingen en de patiënten willen de interventie graag. Consequentie hiervan is dat patiënten of niet mee doen, of uitvallen na randomisatie in de controle groep, of als ze in de controlegroep zitten de interventie elders zoeken.’

‘In de public health zijn interventies/maatregelen vaak te complex, factoren zijn ingebed in het dagelijks leven en die kan men er niet zomaar uit tillen en randomiseren. De consequentie hiervan is dat er veel belang gehecht wordt aan wat je wél kunt randomiseren, maar er wordt minder naar relevantie ervan gekeken’.

‘Er wordt vaak het argument gebruikt dat de zorg dynamisch is en een RCT te lang duurt. De maatschappelijke discussie moet dan echter zijn hoe belangrijk het onderzoek is in de besluitvorming. Wij als onderzoeker moeten bedenken hoe het sneller kan’.

‘Door wet- en regelgeving wordt een RCT steeds moeilijker te organiseren (kosten, monitoren).’

Soms is een RCT echt niet mogelijk

Hoewel de meningen over het wel of niet altijd mogelijk zijn van een RCT uiteenlopen, noemen de meeste geïnterviewden wel voorbeelden van situaties waarin het niet mogelijk is om een RCT uit te voeren. Deze hangen ook grotendeels samen met de problemen, die genoemd worden bij de uitvoer van een RCT. Soms is het ethisch gezien niet verantwoord, als al uit eerder vergelijkbaar onderzoek of in de praktijk gebleken is dat er een groot effect is. Bij zeer zeldzame aandoeningen is er een powerprobleem. En tot slot is een RCT niet mogelijk wanneer random toewijzing van de interventie aan individuen of clusters van individuen niet mogelijk is, bijvoorbeeld in het geval van het meten van het effect van landelijke maatregelen, als respondenten sterke voorkeuren hebben voor een specifieke behandeling/interventie (zoals bijvoorbeeld operatie versus observatie bij prostaatkanker) of als er teveel tijd ligt tussen de randomisatie en het kunnen meten van de uitkomst.

Citaten:

‘De relatie tussen roken en longkanker is vooral aangetoond in observationeel (niet experimenteel) onderzoek; en toch gelooft iedereen het. Er is maar één RCT. De onderzoekers hebben echter niet gerandomiseerd naar wel/niet roken (niet ethisch), maar het effect onderzocht van wel/geen anti-rook advies bij rokers.’

‘Je kunt niet randomiseren tussen psychotherapie met opname versus dagbehandeling omdat de patiënten hier sterke preferenties hebben.’

‘Verskil in effecten van borst- versus flesvoeding.’

‘Effect van prijsverhoging van sigaretten op het rookgedrag.’

‘Nut van screenen van familieleden van patiënten met een hersenbloeding na een geruptureerd aneurysma. In dit geval kan bijvoorbeeld observationeel onderzoek gecombineerd worden met modellering’.

Ook internationaal zijn er interessante ontwikkelingen

Geïnterviewden zijn ook gevraagd naar internationale ontwikkelingen op het gebied van alternatieven voor RCTs. Hieronder de initiatieven die zijn genoemd, waarbij blijkt dat in het buitenland een vergelijkbare discussie wordt gevoerd. Vooral in Engeland zijn er veel initiatieven en zijn ook al richtlijnen beschikbaar.

Het Medical Research Council (UK) heeft in 2008 een update van hun richtlijn “Developing and evaluating complex interventions” uitgebracht.²⁰ Hierin beschrijven zij de stappen die onderzoekers moeten nemen vanaf de ontwikkeling van een interventie naar de evaluatie en uiteindelijke implementatie. Opvallend is dat zij ingaan op zowel de letterlijke vertaling van alternatieven voor de evaluatie middels een RCT (Heeft een RCT de voorkeur en is een RCT dan haalbaar? Als conventionele individuele parallelle randomisatie niet haalbaar is, is cluster randomisatie of een stepped-wedge design dan een mogelijkheid? Een observationeel design zou mogelijk overwogen kunnen worden bij een te verwachten groot effect, een effect dat op korte termijn te verwachten valt, en bij een zeer klein risico op selectie bias.) als ook op het aspect dat naast een RCT aanvullend onderzoek nodig is. Zij stellen dat bij de ontwikkeling van een interventie een pilotstudie nodig is, een proces- en economische evaluatie naast de RCT plaats dient te vinden en er ook implementatieonderzoek van de interventie in de praktijk nodig is. Ook geven zij adviezen over de rapportage van de onderzoeksgegevens, bij voorkeur bevelen zij aan om de resultaten op te nemen in een update van een systematische review.

EU/IMI project GETREAL (Incorporating real-life clinical data into drug development, www.imi.europa.eu/content/getreal): Doel van dit project, waarbij universiteiten en farmaceutische bedrijven zijn betrokken, is het ontwikkelen van nieuwe technieken om ‘real-life’ klinische data te incorporeren bij de ontwikkeling van geneesmiddelen. Ook wordt er gewerkt aan studieopzetten die iets kunnen zeggen over de ‘real world’ (relatieve) effectiviteit van geneesmiddelen.

Een initiatief in de VS om ‘real world’ data te verzamelen is “PCORnet: The National Patient-Centered Clinical Research Network” (<http://www.pcori.org/content/pcornet-national-patient-centered-clinical-research-network>).

Binnen het EU project Developing methodologies to reduce inequalities in the determinants of health (DEMETRIQ) is een handleiding opgesteld om te komen tot een onderzoeksdesign om de effecten van natuurlijke beleidsexperimenten (bedoeld om socio-economische verschillen te verkleinen) te meten. Voor meer informatie zie <http://www.demetriq.eu/>.

²⁰<http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/>

Tot slot nog enkele tips van de geïnterviewden voor ZonMw

Tijdens de interviews kwamen soms spontaan gedachten over de subsidieprocedure en tips voor ZonMw naar voren. Deze staan hieronder op een rij.

Citaten:

‘Bij subsidieaanvragen maakt een RCT meer kans. Men krijgt vaak veel meer vragen bij een ander design.’

‘Als een onderzoeker straks zijn/haar designkeuze moet onderbouwen, dan ook de referenten dit laten doen.’

‘Bij uitleg zou de nadruk dan moeten liggen op waarom niet op patiëntniveau gerandomiseerd wordt.’

‘Het zou goed zijn als er een lijst voor onderzoekers komt, zodat ze keuzen langs kunnen lopen ter onderbouwing waarom ze een bepaald design wel of niet hebben gekozen.’

‘Het is belangrijk dat er een balans aan expertise in de commissies aanwezig is.’

‘Bij de beoordeling van onderzoeksvoorstellen is het bepalend wie de reviewer is. De subsidiegever zou goed naar de expertise van de reviewers moeten kijken, en deze bewust in moeten zetten. Reviewers kunnen in verschillende scholen zitten.’

‘Commissieleden hebben niet allemaal evenveel kennis van onderzoeksdesigns en data analyse. Alle commissieleden zouden geschoold moeten worden in onderzoeksdesigns en mogelijkheden voor data analyse.’

‘Door de enorme waarde die wordt gehecht aan RCTs krijgen soms RCT projecten met minder interessante vraagstellingen de voorrang boven niet-RCT projecten (die soms wel excellent beoordeeld zijn).’

‘Projectduur van 3-4 jaar is te kort voor studieopzet, werving, etc. De looptijd van onderzoeksprojecten (bepaald door subsidiegevers) past niet bij RCTs.’

Resultaten uit de interviews met betrekking tot eHealth interventies

Er zijn drie interviews afgenomen bij in totaal vijf personen met een focus op de evaluatie van eHealth interventies, waarbij aanvullend werd gevraagd hoe men in onderzoek recht kan doen aan de dynamiek in de ontwikkeling van eHealth.

Een RCT is een valide design voor effectiviteitsonderzoek, maar heeft ook zijn beperkingen

Alle geïnterviewden gaven aan moeite te hebben met het gebruik van een klassieke RCT bij de evaluatie van eHealth interventies. Redenen waren de snelheid van de ontwikkeling van technologie, zodat de resultaten van een RCT bij publicatie al niet meer van toepassing zijn op de stand van de huidige technologie. Ook vindt men het moeilijk om een passende controlegroep te definiëren en wordt aangegeven dat de controleconditie niet constant is over de tijd. Een geïnterviewde gaf aan dat men zich met name moet richten op de schadelijke bijeffecten; dat men dus uit moet sluiten dat de interventie/toepassing schade oplevert.

Uit de interviews kwam een ander aspect naar voren waarmee men op het gebied van eHealth worstelt, namelijk de vraag of men voor het evalueren van eHealth interventies dezelfde eisen moet hanteren als bij geneesmiddelonderzoek. Gezien de brede definitie van eHealth en de grote variëteit aan interventies/toepassingen die onder dit begrip vallen, is er behoefte aan consensus voor welke aspecten van eHealth wel of geen hard bewijs nodig is. Hierbij kan men denken aan interventies/toepassingen met een laag risico en lage kosten versus hoog-risico en dure interventies/toepassingen. Volgens een geïnterviewde houdt men innovatie tegen als voor eHealth interventies/toepassingen dezelfde eisen worden gehanteerd als voor geneesmiddelonderzoek.

Alternatieven ?! voor een RCT

Als mogelijke alternatieven werden genoemd: pragmatische trials volgens het MOST principe gebruikmakend van factoriele designs, quasi-experimenteel onderzoek gebruikmakend van natuurlijke (bijvoorbeeld gegevens uit de NIVEL database) of historische controles, onderbroken tijdsreeksen, observationeel onderzoek, N=1 studies, en linked evidence.

Om te achterhalen welke componenten van een interventie bijdragen aan het klinische effect kan men gebruik maken van pragmatische, factoriele designs (MOST strategie).

Alternatieven voor óf aanvullingen op de RCT

Echter, net als bij andere interventies in de gezondheidszorg en publieke gezondheid, gaat het om meer dan het evalueren van de effectiviteit van een eHealth interventie. Er is ook een onderzoeksdesign nodig voor het ontwikkelen van eHealth interventies. Dit zou kunnen gebeuren volgens de sociotechnische benadering, gebruikmakend van user-centered designs. De implementatie van een interventie start niet aan het einde van het traject, maar al aan het begin van het ontwerp (contextueel ontwerpen). Gedurende het hele traject kan men gebruikmaken van het mixed methods design.

Aan de Universiteit van Twente is een roadmap (CeHRes-roadmap, 2011)²¹ ontwikkeld voor onderzoek naar de ontwikkeling van gezondheidstechnologie, die ook internationaal wordt gebruikt. Bij elk iteratief moment wordt getest of technologie bruikbaar is. Dit vraagt om andere meetmethodes dan voor- en nameting: formatieve evaluatie (behoefte vertalen in prototypes, testrondes met eindgebruikers); verder business modellering met stakeholders, die macht hebben om technologie te verbeteren (waardetoewijzing ('value mapping')), zodat het implementatieplan meeloopt bij het ontwerpen van de technologie.

Als de technologie is geïmplementeerd, dan meet men real time wat mensen ermee doen. Zo kan men voorspellen wat de effecten van (onderdelen van) de technologie zijn op gedragsverandering (=procesevaluatie).

Binnen KSYOS TeleMedisch Centrum, een instelling voor medisch specialistische zorg die zich richt op ontwikkeling, onderzoek en invoering van TeleConsultatie in de reguliere gezondheidszorg, wordt de volgende route gehanteerd (Health Management Research model²²):

Fase 1: Intern gebruiksvriendelijkheid testen; fase 2: extern testen door toekomstige gebruikers (gebruiksvriendelijkheid); fase 3: aan hand van vooraf vastgelegde indicatoren evalueren of het leidt tot snellere, betere zorg, dichterbij de patiënt voor lagere kosten (50-100 toekomstige gebruikers); fase 4: zoveel mogelijk gebruikers testen de consequenties van het grootschalig inzetten van de TeleMedicine dienst. Ook de logistiek en financiële haalbaarheid worden dan onderzocht.

Dynamiek in de ontwikkeling van eHealth interventies

Op de vraag hoe bij evaluatieonderzoek recht gedaan kan worden aan de dynamiek in de ontwikkeling van eHealth interventies gaf een geïnterviewde aan dat een methode nodig is waarbij men in onderzoek elementen in kan bouwen die het mogelijk maken om verandering in technologie mee te nemen. Een andere geïnterviewde vindt dat het bij eHealth interventies niet om een tool/ding gaat (bijvoorbeeld een slim horloge of virtual reality devices), maar om een concept. Het gaat om het testen van principes, en niet van apparatuur. Daarom is het niet erg als de technologie verandert, het concept/de principes achter het apparaat blijven object van studie. De vraag naar hoe om te gaan met dynamiek komt voort uit het feit dat technologie vaak wordt gezien als een ding en niet als een concept.

Methodologische vraagstukken

De volgende methodologische vraagstukken werden genoemd waaraan op dit moment gewerkt wordt (bijvoorbeeld aan de Universiteit van Twente):

- Hoe kunnen wij de effectiviteit meten van medische informatie op het internet? Medische informatie over bepaalde ziekten die vrij beschikbaar is op internet kan er bijvoorbeeld toe leiden dat een patiënt wel of niet naar de huisarts gaat. Hoe bepaal je de effectiviteit hiervan?
Een verdere belangrijke ontwikkeling in de gezondheidstechnologie is dat er niet altijd meer wordt

²¹<http://www.utwente.nl/igs/ehealth/cehres/>

²²<http://www.ksyos.nl/onderzoekers/>

gemeten door zorgverleners, maar door de “Googles” etc. Hiervoor zijn dus andere kwaliteitscriteria nodig: hoe meet je gebruik en effectiviteit van technologie buiten de zorg om?

- Er bestaan online interventies op twee niveaus: Behandelinterventies met een vaste duur, bijvoorbeeld negen weken en een verwacht gebruik (getunnelde interventie). Hierbij kan adherentie goed vastgesteld worden. Echter, bij de zogenaamde niet-getunnelde interventies, vaak leefstijlinterventies bij chronisch zieken, hebben de interventies geen vaste duur en hangt gebruik af van de klachten van de patiënt (bijvoorbeeld intermitterende pijn bij chronische pijnklachten). Niet-getunnelde interventies moeten dus anders geëvalueerd worden en er moet een methodiek ontwikkeld worden hoe men in dit geval adherentie meet.

Ook internationaal zijn er interessante ontwikkelingen

Bij de vraag naar internationale ontwikkelingen op eHealth gebied wordt de naam Jeremy Wyatt (Leadership Chair in eHealth Research (Health Informatics), Faculty of Medicine and Health, University of Leeds) genoemd; een wetenschapper die veel doet op het gebied van (methodologie van) eHealth. Zie bijvoorbeeld voetnoot 20.

Verder worden er met name in Canada, Noorwegen en Finland heel andere designs dan de klassieke RCT gebruikt. Nederland loopt achter op dit gebied. In genoemde landen, en ook in Duitsland, Oostenrijk, wordt steeds meer onderzoek gedaan naar hoe men de gebruiker kan betrekken, en wat implementatie van technologie bewerkstelligt. Dit onderzoek vindt veel plaats in de innovatiewereld. Dit heeft te maken met bepaalde scholen waarin multidisciplinair werken meer geïntegreerd is in onderzoek.

In September 2014 werd er in Londen een seminar gehouden over “Evaluation of digital health interventions and medical devices”, georganiseerd door het eHealth Unit, UCL en UCL Interaction Centre in samenwerking met het “Science, Medicine, and Society Network of UCL” (Dr. Sonali Wayal, Dr. Rosie Webster, Dr. Chris Vincent). Tijdens het seminar werden verschillende presentaties gegeven die in een rapport zijn samengevat.²³ De mogelijkheden van online RCTs werden besproken om eHealth interventies te evalueren. Een mogelijk voordeel zou kunnen zijn dat men in korte tijd een grote groep deelnemers kan includeren. De volgende methodologische uitdagingen werden genoemd: de definitie van de meest geschikte dosering en het meten ervan; het bereiken van een hoge interventie adherentie en retentie in de studie; het langdurige evaluatieproces; het risico op contaminatie door andere vrij beschikbare eHealth interventies. Het belang van een procesanalyse werd onderstreept en het gebruik van ‘mixed methods’ om de context en psychologische factoren die de studie kunnen beïnvloeden te begrijpen. Ook werd voorgesteld dat onderliggende theorieën belangrijk zijn bij de ontwikkeling en evaluatie van eHealth interventies zodat men deze goed aan kan passen aan technologische ontwikkelingen (“De technologie verandert, maar de mens niet.”). Toekomstige doelen zijn het ontwikkelen van een theorie voor generaliseerbaarheid en om te komen tot alternatieven van de traditionele RCT, zoals een fractionele factoriële proefopzet. Voor meer informatie verwijzen we naar het rapport.²⁶

²³<http://www.ucl.ac.uk/pcph/research-groups-themes/ehealth/news/techsharingseminarreport>

Tot slot, was er ook een tip voor ZonMw:

Het ontbreekt vaak bij evaluaties van eHealth aan een overkoepelend model. Bij veel evaluaties worden uitkomstmaten gekozen, zonder dat er goed is nagedacht over wat het uiteindelijke doel van de interventie is (bijvoorbeeld bepaalde mate gezondheidswinst, mortaliteit verlaging, kostenbesparing) en via welke subdoelen men daar komt. Het zou goed zijn als er in Nederland door bijvoorbeeld ZonMw een evaluatiemodel wordt gepropageerd dat landelijk geaccepteerd wordt.

Citaten (eHealth):

‘Het gaat bij eHealth om een attitude, een manier om denken te beïnvloeden door middel van technologie. EHealth is geen ding, tool, maar het is een manier om gezondheidszorg te organiseren. EHealth als device is een gamma benadering. Men gaat veel meer uit van verandering, vernieuwing en het kijken hoe eHealth dat kan versnellen of vertragen.’

‘Bij eHealth gaat het om het testen van principes, en niet van devices.’

‘Kwalitatief onderzoek wordt nog weinig op waarde geschat en toegepast, maar kan wel degelijk goed bewijs leveren.’

‘Men moet voor eHealth niet dezelfde eisen hanteren als voor geneesmiddelonderzoek. Er is nog nooit onderzoek gedaan naar de klinische effectiviteit van de ene dokter vergeleken met de andere dokter. Variaties per arts, patiëntcommunicatie, opleiding, geslacht of leeftijd in reguliere zorg zijn nauwelijks onderzocht. EHealth is eenzelfde variant die nu onder een vergrootglas lijkt te worden gelegd terwijl andere variaties in reguliere zorg nooit zo onderzocht zijn.’

‘Wanneer heb je echt harde evidence nodig? Moet je niet gewoon met iets starten en kijken of het werkt? Als het om leven en dood gaat, hoge kosten, etc., dan is wel een RCT nodig.’

‘Het begrip eHealth is ook niet nauw omljnd, er is een grijs gebied tussen medisch technologie en eHealth.’

Belangrijkste conclusies

Er lijkt weinig discussie te bestaan over het feit dat de keuze voor een onderzoeksdesign moet passen bij de onderzoeksvraag. Maar dat betekent niet dat er geen discussie is over wat dan het meest geschikte onderzoeksdesign is. Is er bijvoorbeeld sprake van een hiërarchie in designs? Daar waar sommigen de RCT bovenaan zetten, geven anderen aan af te willen van de RCT als “heilige graal”. Het is in ieder geval duidelijk dat er behoefte bestaat aan handvaten om een designkeuze te kunnen maken en te onderbouwen.

Een RCT kent beperkingen

Hoewel alle geïnterviewden, onafhankelijk van hun achtergrond, het er over eens zijn dat de RCT een valide design is voor effectiviteitsonderzoek, worden er toch ook wel beperkingen genoemd. Daar waar een RCT vooral geschikt wordt geacht voor enkelvoudige interventies, wordt de uitvoer van een RCT bemoeilijkt bij complexe interventies, bij interventies, die plaatsvinden in een snel veranderende werkelijkheid (zoals eHealth interventies) en bij interventies waar wet- en regelgeving de context bepalen. Daarnaast lopen onderzoekers in de voorbereiding van een RCT vaak tegen praktische problemen aan, zoals het kiezen van een controle conditie, het random toewijzen van respondenten en de lange duur van het onderzoek (dat niet altijd overeenstemt met de duur van gesubsidieerde projecten).

Een RCT is soms niet mogelijk of nodig

Daar waar sommige onderzoekers ervan overtuigd zijn dat het bij een effectiviteitsvraag altijd mogelijk is om een RCT uit te voeren, zijn anderen het daar helemaal niet mee eens en komen met voorbeelden wanneer het niet lukt om een RCT uit te voeren. Er worden ook situaties genoemd waarin het niet nodig is om een RCT uit te voeren om effectiviteit aan te tonen. Bijvoorbeeld als het evident is dat er een (groot) effect is, als de situatie vóór de interventie goed in beeld gebracht kan worden of als een nieuwe diagnostische test makkelijker uitvoerbaar is en dezelfde diagnostische accuratesse heeft als de oude effectief gebleken test.

Door de brede definitie van eHealth blijkt er behoefte te zijn om te definiëren voor welk type eHealth interventies of toepassingen wel of niet een klinische effectiviteitsevaluatie (door middel van een RCT) nodig is.

Drie interpretaties van de term ‘Alternatieven voor RCTs’

Bij de vraag naar alternatieven voor RCTs blijken er drie mogelijke interpretaties te zijn.

Ten eerste is er de letterlijke interpretatie, waarbij het gaat om alternatieve designs (of statistische correctietechnieken) die het concept van de RCT zo veel mogelijk benaderen. Hierbij is het de grote uitdaging om confounding en selectie bias te voorkomen en de juiste controlegroep te kiezen. Ten tweede wordt er gedacht aan aanvullende designs, die nodig zijn om antwoord kunnen geven op vragen naar onderliggende (biologische) mechanismen, (zeldzame) bijwerkingen, het evalueren van het proces en de implementatie van de interventie. Ten derde wordt genoemd dat men een design moet kiezen dat bij de vraagstelling past en dat niet persé een RCT hoeft te benaderen. Alle drie interpretaties zijn

legitiem en relevant, maar zullen tot andere antwoorden leiden, die meegenomen dienen te worden in handvaten voor het opstellen en beoordelen van subsidieaanvragen.

(Inter)nationale ontwikkelingen

Bij de vraag naar interessante (inter)nationale ontwikkelingen valt op dat deze wel genoemd worden, maar dat de kennis hierover nogal uiteenloopt. Dit hangt ook samen met één van de opmerkingen, die door meerdere geïnterviewden is gemaakt, namelijk dat zowel onderzoekers als commissieleden meer kennis zouden moeten hebben van onderzoeksdesigns en mogelijkheden voor data-analyse.

Een interessant buitenlands product is de handleiding van de Medical Research Council (MRC, UK) voor het ontwikkelen en evalueren van complexe interventies.²⁴

²⁴<http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/>

Hoofdstuk 5 Resultaten en conclusies uit de discussietafels

Tijdens de ZonMw invitationale 'Alternatieven voor RCT' van 7 september 2015 werd aan zes tafels, in drie rondes, gediscussieerd over verschillende aspecten van 'alternatieven' voor RCTs. Hieronder worden de belangrijkste conclusies van de discussie weergegeven. Allereerst volgt een aantal algemene conclusies en daarna volgen de conclusies per discussietafel.

Deelnemers aan de discussietafels zien een belangrijke rol voor ZonMw, maar ook onderzoekers moeten 'opgevoed' worden

Er wordt gepleit voor methodologische scholing van zowel commissieleden als reviewers. In ieder geval is het belangrijk dat ze goede instructies mee krijgen bij het beoordelen van onderzoeksvorstellen. Maar het is ook nodig dat de onderzoekers het gekozen design motiveren.

Bij aanvang van een programma zou een bijeenkomst georganiseerd kunnen worden waarin een verheldering van de vragen, die beantwoord dienen te worden met het programma, plaats moet vinden. Daarna kan een discussie gevoerd worden over het meest passende design voor de specifieke vraagstelling(en).

Voor complexe interventies is een ruimer kader in een onderzoeksprogramma nodig met financiering in verschillende fasen (go/no go momenten). Ook procesevaluaties, ontwikkelstudies en haalbaarheidsstudies zouden onderdeel moeten zijn van de onderzoeksfinanciering.

Niet alleen bij het begin, maar ook aan het eind van een programma zou door ZonMw evaluatie van de gekozen designs plaats moeten vinden. Bijvoorbeeld: Zijn de studies goed uitgevoerd? Is het gelukt om vast te houden aan het design dat beschreven stond in de aanvragen? En als dat niet gelukt is, hoe kwam dat dan?

Discussietafel 1. Alternatieven voor RCTs: voor- en nadelen van de verschillende designs

Het evalueren van een interventie is iets anders dan het bepalen van de effectiviteit

'Evalueren' is een breed begrip met veel dimensies, het beoordelen van effectiviteit is er daar één van. Om 'uitwisselbaarheid' (het onderling kunnen verwisselen van de experimentele groep en controlegroep) te garanderen ontkom je bijna niet aan een vorm van randomiseren, bijvoorbeeld op individueel of cluster niveau. Alle andere designs waarbij niet gerandomiseerd wordt, is second-best. Dat wil niet zeggen dat randomiseren altijd nodig of mogelijk is (bijvoorbeeld de beschermende effecten van een motorhelm kunnen niet gerandomiseerd onderzocht worden), maar als een onderzoeker randomisatie niet nodig of onhaalbaar vindt, dient hij/zij hiervoor wel met argumenten te komen. Het ondervinden van praktische problemen om te randomiseren betekent nog niet dat een alternatieve oplossing (observationeel onderzoek) plotseling meer valide wordt.

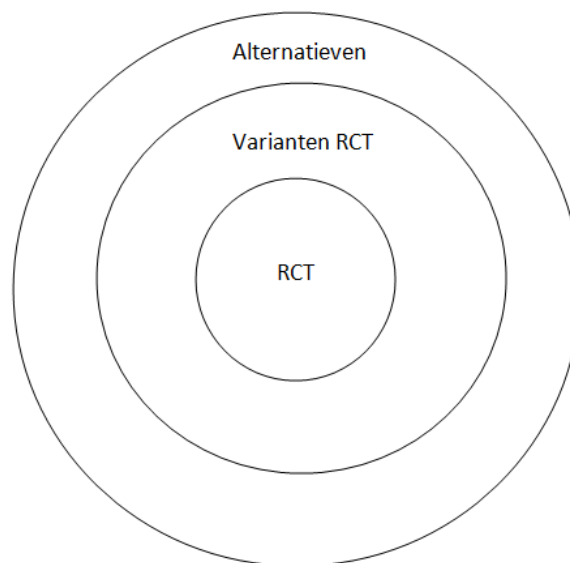
Als je andere aspecten dan effectiviteit wilt evalueren, bijvoorbeeld gebruiksgemak of acceptatie van een interventie, dan ligt het niet voor de hand om te randomiseren. Belangrijk hierbij is dat de

vraagstelling heel scherp gedefinieerd wordt, dus dat die heel specifiek is. Er kan niet worden volstaan met aan te geven dat een interventie wordt geëvalueerd, maar duidelijk moet zijn welk aspect van die interventie geëvalueerd wordt. Bij een vraag naar de werkzaamheid liggen designs waarin gerandomiseerd wordt voor de hand. Voor andere aspecten, zoals de vraag wat werkt bij wie en waarom werkt iets soms niet, zijn er alternatieve designs beschikbaar, die ook vooral gebruikt moeten worden.

Naast 'alternatieven' wordt een nieuwe term geïntroduceerd: 'de variant'

Een belangrijke vraag in de discussie over methoden voor effectiviteitsonderzoek is of een gerandomiseerd onderzoek als de gouden standaard gezien moet worden? Velen (zowel onderzoekers als beoordelaars) gebruiken de parallelle groepen RCT als 'benchmark' in het onderzoeks- of beoordelingskader. Dit gebeurt, omdat het relatief eenvoudig te schetsen is hoe zo'n trial er grofweg uit zou komen te zien. Op deze trial zijn dan variaties mogelijk (denk aan cluster RCT, gerandomiseerde cross-over trial, cohort multiple randomized controlled trial, stepped wedge design, n= 1 trial, etc). De term 'variant' is daar meer op zijn plek dan alternatief, omdat er nog steeds gerandomiseerd wordt en uitwisselbaarheid dan aannemelijk is.

In een plaatje ziet er dat als volgt uit:



Figuur 1 Alternatieven en varianten van een RCT

Voor het beantwoorden van effectiviteitsvragen lijkt observationeel onderzoek niet zo geschikt, maar vertrouwen in observationeel onderzoek verschilt tussen onderzoekers

De rol van observationeel onderzoek om effectiviteitsvragen te beantwoorden leek vooral die van 'last option'. Als om verschillende redenen randomisatie niet kan, is er een rol voor observationeel onderzoek. De mate waarin de resultaten uit dergelijk onderzoek vertrouwd worden, verschilt sterk tussen onderzoekers. Ook het vertrouwen dat onderzoekers hebben in de mate waarin statistische methoden kunnen corrigeren voor mogelijke bias (bijvoorbeeld ongemeten confounding) verschilt en de aanwezige kennis hierover is niet altijd up-to-date en accuraat. Het lijkt alsof met name onderzoekers met een methodologische achtergrond scepsis hebben en terughoudender zijn in het vertrouwen op observationeel onderzoek en statistische methoden om bijvoorbeeld te corrigeren voor ongemeten confounding. In het [Cochrane rapport](#) staan de alternatieven voor de RCT bij effectiviteitsonderzoek beschreven. Echter, er is nog onvoldoende methodologische kennis hoe en in welke setting deze alternatieven werken bij effectiviteitsonderzoek.

Discussietafel 2. Wat is nodig voor valide en efficiënt onderzoek in Nederland?

De vraag naar valide onderzoek kan zowel op uitvoerend als op structuurniveau beantwoord worden.

Uitvoerend

- Vraagstelling moet passen bij het design, RCT reflex onderdrukken (ook alternatieven toelaten).
- Er zijn voldoende respondenten nodig.
- Duidelijke rapportage van de onderzoeksopzet, methodologie en resultaten is nodig.
- Bewijsvoering van de te onderzoeken interventie(s) opbouwen: start met evidence-based ontwikkelen van de interventie, dan een haalbaarheidsstudie doen, dan (gerandomiseerde) effectstudie. Door gebruik te maken van 'value of information analysis'/vroeg HTA modellen (d.w.z het schatten van de potentiële benefit van een interventie of componenten ervan gebruikmakend van statistische modellen) kun je trials ook specifiek ontwerp.
- Er is methodologisch onderzoek nodig naar de voor- en nadelen van de verschillende designs (zodat onderzoekers op basis van deze kennis een weloverwogen keuze kunnen maken), maar ook naar nieuwe technieken en naar het combineren van observationele studies en trials.

Structuur

- Alle onderzoeken moeten geregistreerd worden, zowel RCTs als observationeel onderzoek.
- Er zou ook een vrij toegankelijke website moeten zijn, waar zowel data, als de codeboeken, syntaxen, analyse strategieën en studieprotocollen beschikbaar zijn voor iedereen.
- Educatie van studenten, onderzoekers, reviewers en commissieleden gericht op de vraagstelling versus design. Wat is het best passende design bij welke vraagstelling? Welke alternatieven zijn er? Wanneer werkt welke methodologie (suggestie: een handleiding met de voors- en tegens van methoden/designs).

- Keer op keer ook kritische blik werpen op de bezetting binnen de diverse ZonMw commissies. En misschien moet er wel bij iedere aanvraag, die in een commissie besproken wordt, de vraag gesteld worden of er ook een alternatief design mogelijk is en wat daar de voor- en nadelen van zijn met betrekking tot de validiteit. Wellicht dat dit ook aan de referenten gevraagd dient te worden (om zo de RCT reflex te onderdrukken).

Ook aan efficiënt onderzoek kunnen zowel uitvoerende als structuuraspecten onderscheiden worden.

Uitvoerend

- Maatschappelijk draagvlak creëren (bv. patiënt en/of eindgebruiker betrekken van het begin tot en met het eind).
- De vraag stellen of het onderzoek efficiënter kan.
- Goede planning.
- Leren van gemaakte fouten uit eerder onderzoek. Daarnaast dient de onderzoeker goed na te denken over de haalbaarheid en implementatie van het onderzoek.

Structuur

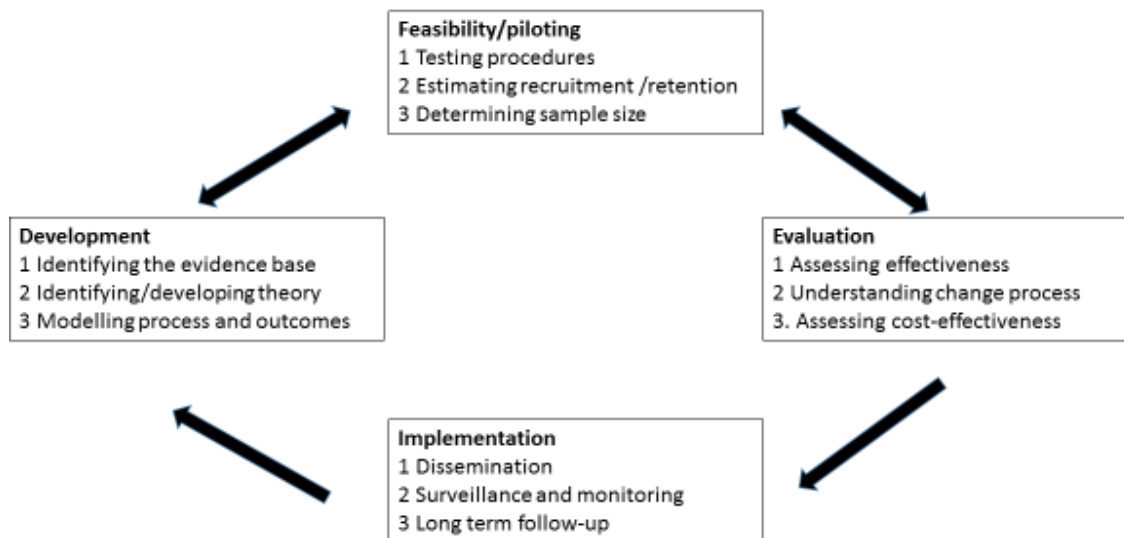
- Financiering voor onderzoek efficiënt besteden, bijvoorbeeld prioritering van onderwerpen, ontwikkeling van een wetenschapsagenda, financiering uitzetten op één onderwerp (grootschaliger onderzoek) en structurele financiering.
- Beter organiseren van onderzoek in Nederland (weten welk onderzoek we doen in Nederland).
- Een ander honoreringssysteem opzetten om zo echte samenwerking te bevorderen. Goede samenwerking kan de kwaliteit van het onderzoek in Nederland ten goede komen.

Discussietafel 3. Toepassing MRC handleiding 'Developing and evaluating complex interventions: new guidance'²⁵ in Nederland

De handleiding heeft meerwaarde en is bruikbaar

In de handleiding worden op een logische wijze de stappen gepresenteerd, die je moet volgen bij het ontwikkelen van een interventie (zie figuur 2). Daarnaast wordt er op gewezen dat voorafgaande aan een effectiviteitsstudie een pilotstudie uitgevoerd dient te worden om de uitvoerbaarheid van de interventie te toetsen (details interventie, hoe meet je het effect etc.). Door middel van de juiste vragen/handvaten in de handleiding wordt ruimte gecreëerd voor het overwegen van andere designs dan de RCT. En tot slot is een sterk punt van de handleiding dat de nadruk gelegd wordt op het ook uitvoeren van een procesevaluatie parallel aan een effectiviteitsstudie. Een procesevaluatie kan inzicht verschaffen in het waarom een interventie niet werkt of onverwachte uitkomsten laat zien. Maar het verschaft ook inzicht in waarom het wel werkt en hoe de interventie eventueel verbeterd kan worden.

²⁵ <http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/>



Figuur 2 Kernelementen van het ontwikkelings- en evaluatieproces
 Afkomstig van <http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/> (pagina 8).

Maar....., er zijn ook wel een paar minpunten

Deze minpunten hebben vooral te maken met het gebrek aan concretisering. Bij sommige deelnemers was er behoefte aan meer uitwerking van de alternatieve designs. Daarnaast is in de handleiding onvoldoende duidelijk hoe complexe interventies geëvalueerd dienen te worden indien er sprake is van meerdere relevante uitkomstmaten.

De handleiding zou verder geconcretiseerd kunnen worden met bijvoorbeeld de effectladder, die gebruikt wordt bij de Erkeningscommissie Interventies²⁶.

Discussietafel 4. Casus ‘Welke alternatieven zijn er voor eHealth, waar RCT’s de snelheid van ontwikkelingen niet bij kunnen houden?’

²⁶Zwikker, M., van Dale, D., Dunnink, T., Willemse, G., van Rooijen, S. Heeringa, N., Rensen, P. (2015). Erkenning van interventies Criteria voor gezamenlijke kwaliteitsbeoordeling 2015-2018

Definitie van eHealth nodig om te bepalen of andere onderzoeksmethoden nodig zijn

EHealth is een breed begrip en omvat veel soorten producten en toepassingen. Volgens de Raad voor de Volksgezondheid en Zorg (2002) is eHealth 'het gebruik van nieuwe informatie- en communicatietechnologieën, en met name internettechnologie om gezondheid en gezondheidszorg te ondersteunen of te verbeteren'²⁷. Afhankelijk van de manier waarop men eHealth definieert (bv. is het een medical device, heb je het over een patiëntendossier, domotica, over een app of een gadget bedoeld voor zelfmanagement, of gaat het over een interventie), kun je beoordelen of er wel of geen specifieke onderzoeksmethoden nodig zijn.

Veel eHealth toepassingen kunnen beschouwd worden als 'medical devices'. De manier waarop daar onderzoek naar gedaan wordt, kan gebruikt worden voor eHealth (veiligheid, performance en de berekenen van de benefit-to-risk-ratio).

RCT ook mogelijk bij eHealth interventies, maar niet altijd nodig

Bij onderzoek naar de effectiviteit van een eHealth interventie, kan een RCT uitgevoerd worden. Het is belangrijk dat indien eHealth onderdeel is van een interventiepakket, er integraal geëvalueerd wordt en niet naar losstaande interventies gekeken wordt. Het is dan geen probleem dat gedurende de looptijd de eHealth interventie doorontwikkeld wordt. Dat kan worden toegestaan, omdat het interventiepakket wordt geëvalueerd, en niet een bepaalde versie van de eHealth. Ter vergelijking werd genoemd dat de software van MRI's ook ieder half jaar wordt aangepast. Dat wordt binnen een trial ook geaccepteerd. Als eHealth gebaseerd is op een reeds erkend werkingsmechanisme, dan is effectonderzoek naar dit mechanisme niet interessant, maar is ontwikkel- en implementatieonderzoek veel meer aan de orde.

Discussietafel 5. Casus 'Alternatieven voor RCT: Bijwerkingen een geval apart'

Onderstaande casus is een voorbeeld van medische kennis vergaard uit observationeel onderzoek met potentieel belangrijke consequenties voor patiënt en maatschappij. Om te komen tot 'wijze' besluitvorming aan de hand van dergelijke resultaten is diep inzicht vereist in de structuren en potentiële invloed van bias.

In een cohortonderzoek onder nieuwe gebruikers van een recombinant factor VIII product was de kans op het ontwikkelen van remmende antistoffen tegen factor VIII 60% groter bij één van de producten, namelijk Kogenate Bayer. Bias en confounding leken geen aannemelijke oorzaken voor deze bevinding en ook de kans op een toevalsbevinding was klein. Gegeven gelijke effectiviteit en kosten zou een 'wijze' dokter dus bij voorkeur niet kiezen voor Kogenate Bayer. Op symposia en in de literatuur werden veelvuldig "levels of evidence" en "non-randomised" als argumenten gebruikt om de resultaten van de cohortstudie te bagatelliseren en te negeren. Andere factoren van invloed op de keuze voor een product, zoals financiële belangen, werden niet openlijk besproken.

²⁷ Raad voor de Volksgezondheid en Zorg (2002) EHealth in zicht.

Replicatiestudies zijn belangrijk

Data kunnen op verschillende manieren geanalyseerd worden, dus het is wel belangrijk dat er een check is (hebben we het wel goed gedaan?).

Belangrijk om belangen inzichtelijk te maken

De “hiërarchie van bewijs” wordt vaak misbruikt om resultaten die ongunstig zijn voor de belanghebbende te ontcrachten. “Het is observationeel en retrospectief dus resultaten zijn onbetrouwbaar.” Maar als de financiële belangen groot zijn, dan is er bij negatieve uitkomsten van een RCT een zelfde weerstand.

Studieresultaten verkregen uit routinematig verzamelde klinische gegevens hebben veel beperkingen

Beperkingen bestaan uit vertekening door confounding, door informatie- of selectiebias en vertekening door inadequate analysemethoden. Er is behoefte aan betere en/of nieuwe onderzoeks- en analysemethoden, die de mate van vertekening kunnen kwantificeren of die vertekening kunnen tegengaan. Deze onderzoeks- en analysemethoden moeten onderzoek naar effectiviteit en bijwerkingen van medische interventies naar een hoger plan brengen.

Discussietafel 6. Casus ‘Alternatieven voor RCT: een casus uit integraal gezondheidsbeleid

De casus, die aan deze discussietafel werd besproken ging over de effecten van de verruiming van de openingstijden van de horeca rondom twee pleinen in Amsterdam. De vraag was of deze verruiming het aantal ongevallen en gewelddelicten waar alcohol in het spel was beïnvloedde. Bij dit natuurlijke experiment werden via registratiegegevens de ‘interventie- en controlegebieden’ met elkaar vergeleken.

Met observationeel onderzoek kun je effecten van beleid aannemelijk maken.

Het is niet mogelijk om 100% zekerheid over causaliteit te krijgen, maar met de uitkomsten van observationeel onderzoek is er voldoende zekerheid om het beleid te kunnen informeren. In beleidsvorming wordt gevraagd om evidentie van enige vorm, waarbij relevantie en toepasbaarheid minstens zo belangrijk zijn als bewijskracht. Dit geldt vooral voor beleidsevaluatie.

Replicatie van observationele studies in verschillende situaties is belangrijk

Deze replicatie is nodig ten behoeve van de geloofwaardigheid van de resultaten, maar ook om te onderzoeken of resultaten uit één setting ook in een andere setting van toepassing zijn. Voor replicatie is het belangrijk dat er wordt geïnvesteerd in registraties, zodat observationeel onderzoek snel uitgevoerd kan worden, vooral als het gaat om evaluatie van ‘natuurlijke experimenten’.

Niet alleen aantonen van effect is belangrijk, maar ook het begrijpen van het effect

Het is belangrijk dat wij niet alleen streven naar het aantonen van effecten, maar vooral ook ervoor zorgen dat we deze begrijpen. Kennis over mechanismen is van belang voor geloofwaardigheid, maar in geval van complexe interventies kun je pas lessen voor beleid trekken als de mechanismen zijn ontrafeld.

Juist het begrijpen van effecten, en ontrafelen van mechanismen, vraagt om innovatieve onderzoeksdesigns.

Hoofdstuk 6 Algemene beschouwing: Stel de vraagstelling centraal!

De discussie is gestart met de vraag naar alternatieven voor een RCT bij het evalueren van de effectiviteit van een interventie. Tijdens de interviews en de ZonMw invitationale ‘Alternatieven voor RCT’ bleek dat veel onderzoekers deze vraag breder opvatten en de vraag betrokken op het hele proces van evalueren van interventies. Discussies, die gevoerd worden over het meest passende design, zijn mede terug te voeren op het feit dat het ‘evalueren’ en het ‘onderzoeken van de effectiviteit’ van een interventie vaak door elkaar gehaald worden. De evaluatie van een interventie begint bij het opzetten van de interventie tot de uiteindelijke implementatie. Het onderzoeken van de *effectiviteit* van de interventie is een – belangrijk – onderdeel van het totale proces. Voor elk onderdeel van de evaluatie van interventies kunnen specifieke onderzoeksvragen geformuleerd worden, waarvoor er dan één of meerdere passende onderzoeksdesigns zijn. Bij de effectiviteitsvraag speelt randomisatie een belangrijke rol (om de uitwisselbaarheid te garanderen), voor veel andere onderzoeksvragen is echter randomisatie niet nodig of te wel niet passend.

Onderzoekers die tijdens de interviews of de invitationale vroegen naar alternatieven of zich uitspraken tegen de RCT deden dit vaak bij onderzoeksvragen waarbij het *niet* om toetsen van effectiviteit ging. Ze noemden dan bijvoorbeeld de vraag *hoe* een interventie werkt of *wat* de gebruiker van de interventie vindt. Bij dit soort vragen is de RCT niet persé het startpunt en juist vaak niet het meest geschikte design.

Het is dus van groot belang om elk onderzoek te starten met de juist geformuleerde vraagstelling, zodat duidelijk is welk aspect van de interventie men wil evalueren: bijvoorbeeld de werkzaamheid (effectiviteit), het werkingsmechanisme of mogelijke bijwerkingen. Ook is een actieve afweging van interne en externe validiteit nodig.

Het Nederlands Tijdschrift voor Geneeskunde besteedde recent aandacht aan het belang van de vraagstelling²⁸. In een methodologie serie wordt ingegaan op verschillende onderzoeksdesigns²⁹.

Tijdens de ZonMw invitationale werd duidelijk dat het niet mogelijk is om een ‘kookboek’ te schrijven met universele recepten om te komen tot het meest passende onderzoeksdesign. Elk programma en elk project hebben hun eigen specifieke vraagstelling(en) en kent zijn eigen randvoorwaarden. Taak van de onderzoeker is de vraagstelling(en) scherp te formuleren en de designkeuze te onderbouwen.

In Tabel 2 geven we, als hulpmiddel, een overzicht van mogelijke onderzoeksvragen omtrent de evaluatie van interventies en mogelijk(e) onderzoeksdesign(s), mede gebaseerd op elders gepubliceerde onderzoeksvragen^{30,31}. De tabel illustreert dat er verschillende vragen geformuleerd kunnen worden bij het ‘evalueren’ van een interventie resulterend in verschillende keuzes voor het onderzoeksdesign. Elke specifieke vraagstelling en daaraan gekoppeld het design moet keer-op-keer opnieuw gedefinieerd worden binnen de daarvoor geldende context.

²⁸ Van der Graaf Y and Zaat J. Onderzoeksartikelen lezen. NTvG 2015;169:A9004

²⁹ <https://www.ntvg.nl/artikelen/dossiers/methodologie-van-onderzoek>; <https://www.ntvg.nl/methodologie>

³⁰ Liu JL, Wyatt JC. The case for randomized controlled trials to assess the impact of clinical information systems. Am Med Inform Assoc. 2011; 18(2):173-80

³¹ <http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/>

De onderzoeksvragen in Tabel 2 zijn ingedeeld volgens de verschillende fasen uit het evaluatieproces van de interventie, te weten (zie ook Figuur 2):

- Ontwikkeling
- Uitvoerbaarheid/pilot
- Effectevaluatie: Beoordelen effectiviteit
- Implementatie

De discussie over de RCT en de alternatieven hiervoor valt met name onder het kopje 'Effectevaluatie'.

De keuze voor een onderzoeksopzet (RCT of alternatief design) hangt af van de context van de specifieke interventie als ook van antwoorden op vragen gesteld in de haalbaarheidsfase zoals:

Is randomisatie haalbaar/mogelijk?

Is er gevaar voor contaminatie in de controlegroep?

Zijn er (sterke) preferenties van patiënten/participanten (bijvoorbeeld als iedereen uiteindelijk de interventie wil krijgen, dan kan men kiezen voor een wachtlijstcontrolegroep)?

Antwoorden op deze vragen kunnen, wanneer relevant, in een projectaanvraag opgenomen worden ter onderbouwing van de designkeuze.

Zie ook de checklist voor het kiezen en beoordelen van de beste onderzoeksopzet in het [Cochrane rapport](#) op pagina's 50-51.

Tabel 2 Voorbeelden van mogelijke onderzoeksvragen, mogelijke onderzoeksmethodiek(en) en meetinstrumenten

Onderzoeksvraag	Mogelijke onderzoeksmethodiek(en) en meetinstrumenten*
<i>Ontwikkelen van de interventie</i>	
Hoe moet een interventie eruit zien om in potentie effectief te zijn? Weke onderdelen zijn hiervoor nodig?	Literatuur review Kennis van experts, gebruikers/patiënten, eindgebruikers d.m.v. focusgroepen, diepte interviews, expertmeetings, Delphi methode, discrete choice experimenten Modelleringsstudies ('value of information analyses' (vroeg HTA modellen))
Hoe werkt de interventie in potentie (theoretische onderbouwing)?	Literatuur review Expertmeeting Sociaalpsychologisch onderzoek Fundamenteel onderzoek (bijvoorbeeld bij geneesmiddel onderzoek)
<i>Haalbaarheid van de interventie en het beoogde studieopzet (pilot fase)**</i>	
Is de ontwikkelde interventie geschikt voor de doelgroep? Is de uitvoering van de ontwikkelde interventie haalbaar (in de specifieke context)?	Pilotstudie met: Registratie (bijvoorbeeld aan- en afwezigheid bij interventie sessies, protocol compliantie, percentage rekrutering en retentie) Directe observatie Survey (enquête), focusgroepen, diepte interviews
Wat vindt de doelgroep van de interventie? Sluit de interventie aan bij de wensen/ verwachtingen van de doelgroep? Complexe interventie [†] : Wat vindt de doelgroep van de onderdelen van de interventie? Welke onderdelen bevallen wel/niet?	Survey (enquête), focusgroepen, diepte interviews Patient preference trial

Onderzoeksvraag	Mogelijke onderzoeksmethodiek(en) en meetinstrumenten *
Complexe interventies/snel evoluerende interventies (bijvoorbeeld eHealth interventies): Welke onderdelen werken wel/niet? Wat is de optimale volgorde/frequentie van aanbieden?	MOST (Multiphase Optimization Strategy) en SMART (the Sequential Multiple Assignment Randomized Trial) strategie (zie hoofdstuk 3 (pagina 18-19) en Collins et al. 2007 ³²).
Hoeveel proefpersonen zijn nodig voor toekomstige effectiviteitsstudie (steekproefgrootte berekening)? ⁺⁺	Literatuuronderzoek Observationeel onderzoek, bestaande registraties Fase 2 studies Voor-na vergelijkingen (als men in de controle groep een stabiele situatie verwacht) Pilot RCT (bij verwachte veranderingen op de uitkomstmaat in de controle groep)
Is het mogelijk om in de toekomstige effectiviteitsstudie de interventie gerandomiseerd aan te bieden? Is er kans op contaminatie (d.w.z. dat controlegroep (een deel van) de onderzochte interventie krijgt)?	Pilotstudie waarbij men beoogde randomisatie toepast (bijvoorbeeld individueel randomisatie (parallele groepen) of cluster randomisatie)
Is de test procedure van de toekomstige effectiviteitsstudie haalbaar?	Pilotstudie met: Registratie (worden alle metingen gedaan, aantal missende waarden), Survey (enquête) Interviews

³²Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med. 2007;32(5 Suppl):S112-8.

Onderzoeksvraag	Mogelijke onderzoeksmethodiek(en) en meetinstrumenten *
<i>(Effect)evaluatie van de uiteindelijke interventie</i>	
Hoe effectief is de interventie? Heeft de interventie een positief effect op de relevante uitkomst(en)?	RCT (individueel gerandomiseerd (parallele groepen) of varianten ervan: cluster randomisatie, pragmatische RCT, stepped wedge design) of een alternatief (zie Cochrane rapport en MRC handleiding ³³ voor mogelijke alternatieve designs).
Is de interventie kosteneffectief?	Kosteneffectiviteitsanalyse (empirische data verzameld in effectiviteitsstudie, modellering of een combinatie).
Hoe wordt de interventie uitgevoerd, hoe worden onderdelen ervan uitgevoerd? Wat zijn de knelpunten? Wat zijn de sterke punten van de interventie? (Procesevaluatie)	Survey, compliantregistraties, directe observatie, interviews, focusgroepen
Hoe komt de effectiviteit (de verandering van de relevante uitkomsten) van de interventie tot stand (via welk mechanisme, welke factoren? (Mediatie analyse) [§]	Metten van potentiële mediërende factoren (bijvoorbeeld biologische, psychologische, demografische, procesmatige factoren) d.m.v., afhankelijk van de setting, biologische samples, vragenlijsten, diepte interviews, focusgroepen Fundamenteel onderzoek
Bij wie werkt de interventie (het beste)? (Moderatie (subgroep) analyse (enkele modererende factor of een combinatie van verschillende modererende factoren (eventueel gecombineerd in een risicoscore))) [§]	Metten van potentiële modererende factoren (bijvoorbeeld biologische, psychologische, demografische) d.m.v. biologische samples, vragenlijsten, diepte interview, focusgroep

Implementatie van de interventie

³³<http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/>

Onderzoeksvraag	Mogelijke onderzoeksmethodiek(en) en meetinstrumenten *
Wat is de 'uptake'/disseminatie van de interventie?	Observationeel cohort onderzoek ('real world uptake), (routine)registratie
Werkt de effectief bevonden interventie in de praktijk? Wordt de interventie in de praktijk aangepast? Worden, bij een complexe interventie, alle componenten in de praktijk uit-/ingevoerd? Wat is de compliantie van een geneesmiddel interventie na implementatie in de zorg? Wat is de adherentie in de praktijk?	Observationeel cohort onderzoek, (routine)registratie Survey (enquête), focusgroepen, diepte interviews
Heeft de interventie bijwerkingen ('unintended effects')?	Observationeel cohort onderzoek (mogelijk al binnen RCT onderzocht, mits haalbaar, maar voor zeldzame bijwerkingen niet mogelijk)

* In deze kolom worden passende onderzoeksmethodieken gegeven. De onderzoeker kiest de methodiek(en) die bij zijn specifieke vraagstelling pas(sen)t.

** In de haalbaarheidsfase wordt de interventie aangeboden aan een steekproef uit de doelgroep en – afhankelijk van de onderzoeksvraag – ook een controlegroep geïnccludeerd.
De relevantie van de hier genoemde onderzoeksvragen zijn afhankelijk van het doel van het onderzoek.

† Een complexe interventie is een interventie die uit verschillende met elkaar interacterende componenten bestaat.

†† Een pilotstudie wordt niet alleen uitgevoerd om de sample size berekening te onderbouwen, maar het kan één van de subvragen zijn.

§ Bij deze vraagstelling gaat het niet om een te kiezen onderzoeksdesign, maar om factoren die in de effectiviteitsstudie gemeten moeten worden om het gevonden effect te kunnen onderbouwen of om subgroepen te identificeren bij waarbij de interventie het grootste effect heeft. Met geplande analyses naar effect modificatie moet de onderzoeker bij de werving van deelnemers rekening houden: het vereist meer deelnemers en bij voorkeur een ongeveer gelijke verdeling van degenen met en zonder de modererende factor. Bijvoorbeeld als geslacht een potentieel modererende factor is, dan wil men in het onderzoek ongeveer 50% mannen en 50% vrouwen includeren (onderzoeksvraag: verschilt het effect tussen mannen en vrouwen?)

Aanbevelingen aan ZonMw

Aanleiding voor dit project 'Alternatieven voor RCT' was de discussie, die regelmatig gevoerd wordt over het gekozen onderzoeksdesign binnen de ZonMw-beoordelingscommissies, maar ook tussen de projectindieners, de ZonMw beoordelingscommissies, externe referenten en gebruikers van onderzoeksresultaten. Om deze discussie te stroomlijnen wordt ZonMw aanbevolen om in de ZonMw subsidieaanvraagformulieren specifieke ruimte op te nemen waarin de projectaanvrager aangeeft wat het meest geschikte design is om binnen de programmakaders de betreffende onderzoeksvraag te beantwoorden. De onderzoeker beargumenteert de keuze voor dit design en beschrijft waarom dit in de specifieke situatie het meest geschikte design is. Als een ander alternatief design gangbaar is, maar niet gekozen wordt, dient de onderzoeker uit te leggen waarom een andere keuze is gemaakt. Hierbij gaat de aanvrager verder dan "het kan niet", maar beschrijft de redenen voor de keuze in de specifieke context van het programmakader. Vervolgens onderbouwt hij/zij de designkeuze en geeft aan waarom de te verkrijgen resultaten voldoende valide zijn. De aanvrager maakt hierbij een afweging van de interne validiteit en de toepasbaarheid van de te verwachten resultaten.

Hierbij kunnen de in Tabel 2 genoemde onderzoeksvragen hulpbieden. Verder biedt ook het [Cochrane rapport](#) en de [MRC handleiding](#) (p. 10-12) handvaten om de keuze te onderbouwen.

Om een onderbouwing van de designkeuze te kunnen geven en om de juiste discussie in de commissies te kunnen voeren, is educatie in onderzoeksmethodologie van onderzoekers en commissieleden nodig. Bij de keuze voor reviewers moet erop gelet worden of expertise over de verschillende mogelijke designs aanwezig is. Ook moet er bij de samenstelling van de commissies nagegaan worden of voldoende methodologische expertise aanwezig is.

Daarnaast is methodologisch onderzoek naar de voor- en nadelen en de toepasbaarheid van de verschillende (alternatieve) designs bij effectiviteitsonderzoek nodig. Tijdens de invitation bleek dat er nog onvoldoende kennis is over hoe goed alternatieve designs werken en in welke setting alternatieven geschikt zijn. Ook is er methodologisch onderzoek nodig voor een optimale statistische analyse bij de evaluatie van complexe interventies indien er sprake is van meerdere relevante uitkomstmaten.

Gezien de verschillende stappen in de evaluatie van een interventie wordt aanbevolen om een ruimer kader in onderzoeksprogramma's mogelijk te maken waarbij de verschillende fasen in het evaluatieproces gefinancierd worden met gedefinieerde go/no go momenten. Ook procesevaluaties en haalbaarheidsstudies zijn dan onderdeel van de onderzoeksfinanciering.

Bijlage I Interview vragen

- Wat is uw mening over een RCT?
- Wat is uw mening over alternatieven voor een RCT?
- Is een RCT altijd mogelijk? Zijn er ook situaties te bedenken waarin een RCT niet mogelijk / optimaal is?
- Kunt u een concreet voorbeeld in uw eigen werkveld benoemen waarbij een RCT niet geschikt is of te wel niet het passend bewijs levert?
- Heeft u zelf ooit wel eens een RCT uitgevoerd?
- Waarom was dit de methode om antwoord te krijgen op de onderzoeksvragen?
- Heeft u andere designs overwogen? Waarom wel / niet ?
- Is er een probleem m.b.t. designkeuze in uw werkveld. Zo ja, welke discussiepunten zijn er?
- Heeft u zelf wel eens op basis van effectonderzoek beslissingen moeten nemen (als eindgebruiker, of richtlijncommissie)? Waar let u dan op? Aan welke eisen moet het onderzoek volgens u voldoen?
- Gaat u wel eens met onderzoekers in gesprek over onderzoeksdesign? Waarover gaan die discussies dan? En zo nee, waarom gebeurt dat niet?
- Ziet u voor uw werkveld belangrijke (internationale) ontwikkelingen m.b.t. studie design?

Vragen specifiek voor eHealth interventies:

- Wat is volgens u de definitie van eHealth?
- Welke designs zijn nodig/ geschikt voor de evaluatie van de klinische effectiviteit van eHealth interventies?
- Hoe kan daarbij recht gedaan worden aan de dynamiek in de ontwikkeling van eHealth interventies?

Bijlage II Geïnterviewde personen

Prof. dr. W.J.J. (Pim) Assendelft, Radboud UMC Nijmegen
Dhr. T.A. (Teus) van Barneveld, Kennisinstituut van Medisch Specialisten Utrecht
Prof. dr. N.M.A. (Nicole) Blijlevens, Radboud UMC Nijmegen
Dr. H.R. (Hennie) Boeije, Universiteit Utrecht
Dr. M.J.H. (Marja) van Bon-Martens, Trimbos Instituut Utrecht
Drs. M. (Martijn) Bool, Movisie Utrecht
Prof. dr. P.M.M. (Patrick) Bossuyt, AMC, Amsterdam
Prof. dr. A. (Lex) Burdorf, Erasmus MC Rotterdam
Prof. dr. J.J.V. (Jan) Busschbach, Erasmus MC Rotterdam
Dr. D. (Djoeke) van Dale, RIVM Bilthoven
Prof. dr. L. (Lisette) van Gemert-Pijnen, Universiteit Twente, Enschede
Prof. dr. Y. (Yolanda) van der Graaf, UMC Utrecht
Prof. dr. D.E. (Rick) Grobbee, UMC Utrecht
Dr. R.H.H. (Rolf) Groenwold, UMC Utrecht
Dr. J. (Job) van der Heijden, KSYOS TeleMedisch Centrum, Amstelveen
Drs. J. (Juanita) Heymans, Zorginstituut Nederland, Diemen
Dr. N.F. (Nicolette) de Keizer, AMC, Amsterdam
Dhr. J. (Johan) Krijgsman, Nationaal ICT Instituut in de Zorg (Nictiz), Den Haag
Dr. F.J. (Frank) van Lenthe, Erasmus MC Rotterdam
Dr. M. (Mirella) Minkman, Vilans Utrecht
Prof. dr. K.G.M. (Carl) Moons, UMC Utrecht
Drs. V. (Veronica) van Nederveen, Patient Advocate
Dr. M.C.G. (Madelon) Pijls, CZ zorgverzekeringen, Sittard
Drs. P. (Peter) Rensen, Movisie Utrecht
Prof. Dr. D. (Dirk) Ruwaard, Universiteit Maastricht
Prof. dr. R. (Robbert) Sanderman, Rijksuniversiteit Groningen
Dr. W.J.M. (Wilma) Scholte op Reimer, Hogeschool van Amsterdam
Prof. dr. G.J.J.M. (Geert-Jan) Stams, Universiteit van Amsterdam
Dr. G.A. (Ardine) de Wit, RIVM, Bilthoven
Prof. Dr. L. (Leonard) Witkamp, KSYOS TeleMedisch Centrum, Amstelveen en AMC, Amsterdam
Prof. dr. G.A. (Gerhard) Zielhuis, Radboud UMC, Nijmegen
Drs. M. (Machteld) Zwikker, Nederlands Jeugd Instituut, Utrecht