

STAND VAN ZAKEN

Het 'stepped wedge'-design

METHODOLOGIE VAN ONDERZOEK 1

Olaf M. Dekkers

- Standaard trials lenen zich niet voor alle therapeutische vraagstukken. Het 'stepped wedge'-clusterdesign is een speciale vorm van gerandomiseerd onderzoek waarbij een interventie op groepsniveau getrapt wordt ingevoerd. Het design kan worden gebruikt wanneer randomiseren op individueel niveau niet goed mogelijk is en wanneer het om ethische, logistieke of financiële redenen wenselijk is dat de interventie getrapt wordt ingevoerd.
- In dit design wordt een controleperiode gebruikt: de tijd voordat de experimentele interventie in een cluster is ingevoerd.
- Belangrijk is dat bij de statistische analyse altijd rekening moet worden gehouden met het gegeven dat deelnemers binnen een cluster veel overeenkomsten kunnen vertonen.
- Een zorgvuldige afweging van de ethische aspecten van het stepped-wedgedesign is van belang, vooral wanneer informed consent niet op individueel niveau kan worden verkregen.

In de rubriek Stand van zaken verschijnen de komende periode regelmatig bijdragen over methoden die gebruikt worden bij het opzetten van wetenschappelijk onderzoek. De artikelen in deze serie illustreren op begrijpelijke wijze wat een bepaalde methode behelst, zonder dat hier uitvoerige methodologische kennis voor nodig is. Zowel oude als nieuwe methodologische principes worden zo inzichtelijk gemaakt voor artsen die klinische onderzoeken goed willen interpreteren.

Gerandomiseerde studies worden gezien als de gouden standaard voor het aantonen van therapeutische effecten. Omdat standaardtrials zich niet voor alle therapeutische vraagstukken lenen, wordt toenemend gezocht naar, en gebruik gemaakt van andere onderzoeksvormen.^{1,2} Het 'stepped wedge' clusterdesign is een speciale vorm van gerandomiseerd onderzoek waarbij een interventie op groepsniveau getrapt wordt ingevoerd. Het design kan worden gebruikt wanneer randomiseren op individueel niveau niet goed mogelijk is en wanneer het om ethische, logistieke of financiële redenen wenselijk is dat de interventie getrapt wordt ingevoerd.

RANDOMISATIE: INDIVIDUEN VERSUS CLUSTERS

In de meeste gerandomiseerde studies worden individuele deelnemers gerandomiseerd tussen de experimentele interventie en een andere interventie (placebo of standaardbehandeling). Het is echter voor sommige interventies niet wenselijk of mogelijk om individuen te randomiseren. Wanneer bijvoorbeeld het effect van een campagne voor stoppen met roken middels posters en folders in huisartsenpraktijken wordt onderzocht, is randomisatie op het niveau van individuen niet mogelijk. Alle patiënten die een

*Leids Universitair Medisch Centrum,
afd. Klinische epidemiologie, Leiden.*

*Dr. O.M Dekkers, internist-endocrinoloog en
klinisch epidemioloog (tevens: afd. Endocrinologie
en Metabole ziekten) (o.m.dekkers@lumc.nl).*

huisartspraktijk bezoeken zullen de informatie uit de campagne immers meekrijgen. Ook voor een studie naar het effect van telemonitoring op de Intensive Care (IC), is individuele randomisatie niet mogelijk.³

Daarnaast zijn er studies te bedenken waarbij weliswaar individuen gerandomiseerd zouden kunnen worden, maar waarbij (a) de effecten van een interventie niet tot een individu beperkt blijven en (b) de individuen die indirect meebehandeld worden in de controlegroep van de studie zitten. Dit kan het geval zijn wanneer bijvoorbeeld het effect van educatieve programma's op therapietrouw wordt onderzocht bij patiënten met een hiv-besmetting. Educatie zal mogelijk een indirect effect hebben op vrienden of familieleden van de patiënten die de educatiebijeenkomst hebben bezocht. Ook kan dit het geval zijn wanneer vaccinatieprogramma's worden onderzocht; vaccinatie van een individu verlaagt ook de kans op infectie van verwante, niet-gevaccineerde personen. Wordt in dergelijke situaties gerandomiseerd op het niveau van individuen dan wordt het therapie-effect onderschat.

Wanneer randomisatie op individueel niveau niet mogelijk of wenselijk is, kan er voor worden gekozen om groepen individuen (een cluster) te randomiseren. Ziekenhuizen, afdelingen, huisartsenpraktijken, dorpen, geografische regio's en behandelaars zijn voorbeelden van clusters.

Randomisatie op het niveau van clusters betekent dat een of meerdere clusters de experimentele interventie toegewezen krijgen en een of meerdere cluster de standaardinterventie. In het voorbeeld van de studie naar het effect van een campagne voor stoppen met roken middels posters en folders worden de deelnemende huisartsenpraktijken gerandomiseerd tussen praktijken die actief campagne voeren en praktijken met een standaard huisartsenbeleid. In principe worden alle deelnemers binnen een cluster 'behandeld' met de interventie zoals toegewezen door randomisatie. In het voorbeeld van de studie naar het effect van telemonitoring op de IC kregen patiënten van een IC-unit allemaal wel of allemaal niet de experimentele interventie.

STUDIEOPZET

In het stepped-wedgedesign wordt een experimentele interventie getrapt geïmplementeerd binnen de verschillende clusters. Een belangrijk kenmerk van deze studieopzet is dat uiteindelijk alle clusters de experimentele interventie krijgen. Wanneer de studie begint wordt in alle deelnemende clusters aan alle patiënten de standaardbehandeling gegeven; dit geldt als de controleperiode. Na een bepaalde tijd wordt eerst in 1 cluster, en vervolgens sequentieel in de andere clusters, de experimentele interventie geïntroduceerd. De volgorde van implementatie is door randomisatie bepaald. Hoewel in een stepped-wedgedesign meestal clusters worden gerandomiseerd, kan

het ook worden toegepast na randomisatie op individueel niveau.⁴ Een belangrijke karakteristiek van het stepped-wedgedesign is dat de interventie na introductie in een bepaald cluster voortgezet wordt, zodat aan het eind van de studie in alle clusters de experimentele interventie wordt gegeven. In een studie naar het effect van telemonitoring op de IC werd de interventie sequentieel ingevoerd op 7 verschillende IC-units over een tijdbestek van 1 jaar. Aan het eind van de studie werden patiënten op alle 7 IC's behandeld met behulp van telemonitoring.³

Hoewel in het stepped-wedgedesign van randomisatie gebruik wordt gemaakt, zal het voor de onderzoekers en deelnemers duidelijk zijn wanneer in een cluster de experimentele interventie wordt ingevoerd. In praktijk is een clustertrial dus veelal niet geblindeerd.¹ Onderzoekers en deelnemers weten in principe dat op een IC-unit telemonitoring wel of niet is ingevoerd. In dat opzicht is een clustergerandomiseerde trial vergelijkbaar met een pragmatische trial, die veelal ook ongeblindeerd is, maar die het effect van een interventie in de dagelijkse praktijk beter zou weergeven.⁵ Gerandomiseerde studies zonder blinding gaan gepaard met een hoger risico op bias.⁶ Het is echter in principe mogelijk om in een stepped-wedgedesign de uitkomstmaten te blinderen voor de beoordelaar, zodat deze niet weet uit welk cluster de resultaten afkomstig zijn. Dit kan een bias in de beoordeling van de uitkomstmaten tegengaan. Dit blijkt echter in de praktijk slechts zelden te gebeuren.¹

De analyse van een stepped-wedgedesign biedt 2 mogelijkheden. Ten eerste kan een vergelijking worden gemaakt tussen de effecten van alle clusters ten tijde van de experimentele interventie en alle clusters ten tijde van de standaardbehandeling. Ten tweede kan ook het effect van de periode vóór de experimentele interventie worden vergeleken met het effect binnen hetzelfde cluster tijdens de interventie.

Belangrijk is dat bij de statistische analyse altijd rekening moet worden gehouden met het gegeven dat deelnemers binnen een cluster veel overeenkomsten kunnen vertonen door ongelijke distributie van risicofactoren tussen clusters. Dit wordt het clustereffect genoemd. Niet alleen de distributie van risicofactoren kan verschillen tussen clusters (bijvoorbeeld de proportie rokers), maar ook de zorg op het niveau van de clusters kan verschillen. Zo zal de organisatie van huisartsenpraktijken kunnen verschillen, of zal de beschikbaarheid van medicatie kunnen verschillen tussen clusters (bijvoorbeeld dorpen of regio's) in ontwikkelingslanden waar vaccinatieprogramma's worden ingevoerd. Wordt in de analyse geen rekening gehouden met het clustereffect, dan zijn de betrouwbaarheidsintervallen van de effectschatter te smal en wordt de significantie overschat (te kleine p-waardes).⁷ Het clustereffect betekent eveneens dat voor gelijke

power, het aantal deelnemers in een clusterstudie groter moet zijn dan voor een studie met randomisatie op individueel niveau.⁶ Omdat voor het stepped-wedgedesign in de analyse ook gebruik kan worden gemaakt van de informatie verkregen uit de individuele clusters ('vóór-tijdens behandeling'-vergelijking) kan de benodigde groepsgrootte kleiner zijn dan voor een gebruikelijke clustertrial, waarbij de interventie niet getrapt is ingevoerd.

CLUSTERGERANDOMISEERDE STUDIES EN CLUSTER-CROSS-OVERSTUDIES

Evenals het stepped-wedgedesign maken ook het clustergerandomiseerde onderzoek en het clustercross-overonderzoek gebruik van clusters. In een clustergerandomiseerd design worden clusters gerandomiseerd tussen de experimentele interventie en de standaardbehandeling. Anders dan in het stepped-wedgedesign wordt niet in alle clusters na een bepaalde tijd de experimentele interventie geïmplementeerd. Net als in een standaardtrial met randomisatie op individueel niveau blijven de clusters gedurende de hele studie doorgaan met de initieel toegewezen interventie.

In een clustercross-overtrial krijgen alle clusters gedurende een bepaalde tijd zowel de experimentele als de standaardinterventie. Maar anders dan een stepped-wedgedesign blijven de clusters die beginnen met de experimentele interventie deze niet gebruiken. Of een cluster begint met de standaardinterventie of de experimentele interventie wordt bepaald door randomisatie. Een stepped-wedgedesign kan dus worden gezien als een clustercross-overtrial waarbij de cross-over maar in één richting gaat: van standaardinterventie naar experimentele interventie. Dat kan van belang zijn wanneer het effect van de experimentele interventie lang aanhoudt (zoals bij vaccinaties, of

aanpassingen in huizen) of wanneer de interventie blijvend is (zoals telemonitoring op de IC).

ETHISCHE ASPECTEN

Equipoise en informed consent van deelnemers zijn 2 ethische basisbeginselen van gerandomiseerd onderzoek. De ethische rationale voor een gerandomiseerde studie is dat er onzekerheid bestaat over de vraag of de nieuwe interventie werkzaam zal zijn. Op voorhand kan dus niet voorspeld worden of de nieuwe interventie beter zal zijn dan de standaardinterventie of beter zal zijn dan niets doen. Dit wordt equipoise genoemd. Zou wel bekend zijn dat de nieuwe interventie werkt, dan zou gerandomiseerd onderzoek onethisch zijn omdat aan patiënten een effectieve behandeling wordt onthouden. Het stepped-wedgedesign kan in sommige gevallen uitkomst bieden zoals wanneer een interventie op individueel niveau effectief is gebleken maar nog onderzocht moet worden of implementatie op populatieniveau werkzaam is. Binnen dit design krijgen uiteindelijk immers alle deelnemers de experimentele interventie. Empirische studies tonen aan dat in het merendeel van de studies met een stepped-wedgedesign de onderzochte interventie op individueel niveau al effectief is gebleken.¹ De ethische rationale voor de randomisatie is dat het om financiële of logistieke redenen niet mogelijk is om de interventie in alle clusters gelijktijdig in te voeren. Randomisatie bepaalt in het stepped-wedgedesign niet of een cluster de nieuwe interventie krijgt, maar wanneer het wordt ingevoerd. Financiële en logistieke redenen blijken in de praktijk vaak een belangrijke overweging voor de keus van een stepped-wedgedesign.¹ Een voorbeeld van een studie waarbij gebruik gemaakt werd van het stepped-wedgedesign en de ethische overwegingen daarbij staat in de tabel.

TABEL Voorbeeld van een studie waarbij gebruik gemaakt werd van het stepped-wedgedesign.

Onderzoek naar het effect van verbeterde woonomstandigheden

Slechte woonomstandigheden (zoals oude verwarmingssystemen, slechte isolatie en hoge vochtigheid) zijn gerelateerd aan een verhoogde kans op verschillende ziekten. Zo'n 15 jaar geleden werd in Engeland een studie opgezet om het effect van aanpassingen in woonomstandigheden op ziekte te meten.⁹ De interventies hadden tot doel om de vochtigheid en schimmel in huizen te verminderen en de isolatie te verbeteren. Er werden 119 huizen geïncludeerd in de studie, met in totaal 580 inwoners. Er was subsidie voor 50 huizen in het 1e jaar. Huizen die in aanmerking kwamen voor de studie werden gerandomiseerd tussen aanpassingen in jaar 1 en aanpassingen in jaar 2. De onderzoekers keken vooral naar het effect van de interventie op gezondheid en kwaliteit van leven. Van alle deelnemers werden gegevens verzameld vóór start van het onderzoek, na 1 en na 2 jaar.

In de studie werd gerandomiseerd op het niveau van clusters (huizen), terwijl de uitkomsten voornamelijk op individueel niveau werden gemeten. De onderzoekers wilden om ethische redenen geen gebruikelijk clustergerandomiseerd onderzoek uitvoeren, waarbij de helft van de huishoudens niet voor de interventie in aanmerking zou komen. De onderzoekers verwachtten immers dat een interventie waarbij woningen worden geïsoleerd en boilers worden vervangen op een heel aantal punten effectief zou zijn. De keuze voor een stepped-wedgedesign bracht het voordeel met zich mee dat alle deelnemende woningen uiteindelijk werden verbeterd. Omdat het budget van de studie slechts kon voorzien in 50 woningen per jaar was er een goede rationale voor het getrapt invoeren van de interventie. Randomiseren tussen direct een interventie aan de woning en de wachtlijst was, gegeven de beperkte economische middelen, wel ethisch te verdedigen.

LEERPUNTEN

- **Standaardtrials, waarbij individuen worden gerandomiseerd tussen experimentele interventie en standaardbehandeling of placebo, lenen zich niet voor alle therapeutische vraagstukken.**
- **Het 'stepped wedge'-clusterdesign kan worden gebruikt wanneer randomiseren op individueel niveau niet goed mogelijk is en wanneer het om ethische, logistieke of financiële redenen wenselijk is dat de interventie getraptd wordt ingevoerd.**
- **In de analyse moet rekening worden gehouden met het clustereffect; wordt dit niet gedaan dan zijn de betrouwbaarheidsintervallen van de effectschatter te smal en wordt de significantie overschat (te kleine p-waardes).**
- **Een zorgvuldige afweging van de ethische aspecten van het stepped-wedgedesign is van belang, vooral wanneer informed consent niet op individueel niveau kan worden verkregen.**

Wanneer voor een clusterstudie een interventie onderzocht wordt die van invloed is op het gehele cluster, dan is informed consent op individueel niveau niet mogelijk. Een voorbeeld hiervan is een clusterstudie naar het effect van waterontsmetting. Waterontsmetting gebeurt centraal en iedereen in het betreffende gebied zal ontsmet water krijgen, waarbij er voor een individu praktisch geen keus is om niet deel te nemen. Hetzelfde geldt voor een patiënt die wordt opgenomen op een IC waar de implementatie van telemonitoring wordt onderzocht. Ook deze patiënt heeft geen mogelijkheid om niet deel te nemen. Niet deelnemen aan de studie zou immers neerkomen op het overplaatsen van de patiënt naar een ander ziekenhuis. In zulke situaties moeten door medisch ethi-

sche commissies en vertegenwoordigers van de clusters de ethische aspecten van de studie zorgvuldig worden afgewogen.⁸ In de studie naar het effect van telemonitoring heeft de ethische commissie dan ook geoordeeld dat toestemming op individueel niveau niet nodig was.³ Wanneer een interventie op clusterniveau wordt geïmplementeerd, maar op individueel niveau wordt toegevend, is het wel een overweging om informed consent van deelnemers te verkrijgen.⁸ Bij de invoering van een vaccin binnen een clusterdesign kan bijvoorbeeld informed consent van de individuele deelnemers worden verkregen.

CONCLUSIE

Het stepped-wedgedesign kan worden ingezet voor onderzoeksvragen waarbij randomisatie op individueel niveau niet mogelijk of wenselijk is en waarbij ethische, logistieke of financiële redenen een getrapte invoering vereisen. De studievorm is daarmee aantrekkelijk voor het beantwoorden van vragen waarbij interventies op groepsniveau onderzocht worden.

Prof.dr. Jan P. Vandenbroucke leverde commentaar op een eerdere versie van dit artikel.

Belangenconflict: geen gemeld. Financiële ondersteuning: geen gemeld.

Aanvaard op 7 november 2011

Citeer als: Ned Tijdschr Geneeskd. 2012;156:A4069

[➤ Meer op www.ntvg.nl/klinischepraktijk](http://www.ntvg.nl/klinischepraktijk)

LITERATUUR

- 1 Mdege ND, Man M-S, Taylor Nee Brown CA, Torgerson DJ. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol*. 2011;64:936-48.
- 2 Hordijk M. Klinische trial krijgt concurrentie. *Med Contact*. 2011;66:2045-7.
- 3 Lilly CM, Cody S, Zhao H, et al. Hospital mortality, length of stay, and preventable complications among critically ill patients before and after tele-ICU reengineering of critical care processes. *JAMA*. 2011;305:2175-83.
- 4 Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol*. 2006;6:54.
- 5 Algra A, van der Graaf Y. Pragmatische en pathofysiologische trials: een kwestie van vraagstelling. *Ned Tijdschr Geneeskd*. 1999;143:514-7.
- 6 Jüni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ*. 2001;323:42-6.
- 7 Van Houwelingen JC. Randomisatie op het niveau van behandelaars. *Ned Tijdschr Geneeskd*. 1998;142:1662-5.
- 8 Edwards SJ, Brauholtz DA, Lilford RJ, Stevens AJ. Ethical issues in the design and conduct of cluster randomised controlled trials. *BMJ*. 1999;318:1407-9.
- 9 Somerville M, Basham M, Foy C, et al. From local concern to randomized trial: the Watcombe Housing Project. *Health Expect*. 2002;5:127-35.