



## Effect meten bij kleine aantallen

### Het N=1 design toegepast in een effectonderzoek naar de justitiële gedragsinterventie MTFC

*Miranda Witvliet en Maartje Timmermans*

#### Aanbevolen citeerwijze bij dit artikel

Miranda Witvliet en Maartje Timmermans, 'Effect meten bij kleine aantallen', *Beleidsonderzoek Online* december 2015, DOI: 10.5553/BO/221335502015000021001

## Het N=1 design als alternatief voor het experimentele trial

Onderzoek naar de werkzaamheid van een interventie, zoals een behandeling, therapie of training, kan op verschillende manieren worden uitgevoerd. Om de effectiviteit van een interventie vast te kunnen stellen, wordt gerandomiseerd onderzoek met een controlegroep (het experimentele trial, oftewel RCT) veelal beschouwd als de gouden standaard (Shadish, Cook & Campbell, 2002). Deze manier van onderzoek doen houdt in dat een vergelijking wordt gemaakt tussen een onderzoeksgroep waarbij een interventie wordt uitgevoerd en een controlegroep die de interventie niet ontvangt. Om de vergelijkbaarheid van de beide groepen te kunnen waarborgen is het van belang dat men het toeval laat bepalen of een persoon tot de onderzoeksgroep behoort of in de controlegroep valt. Vervolgens wordt onderzocht of de onderzoeksgroep zich beter ontwikkelt dan de controlegroep.

In de praktijk is deze manier van onderzoek doen echter lang niet altijd haalbaar. Het is bijvoorbeeld niet altijd mogelijk om personen willekeurig toe te wijzen aan een interventie- of controlegroep. Soms is het ethisch niet te verantwoorden om personen bewust een interventie te onthouden waarvan het aannemelijk is dat het van nut is. In andere gevallen staat van tevoren al vast wie wel in aanmerking komt voor een interventie en wie niet. Willekeurige toewijzing is dan niet meer

aan de orde. Een andere reden waarom effectonderzoek met een RCT niet altijd mogelijk is, is dat hiervoor een relatief grote onderzoeksgroep nodig is. Sommige interventies zijn bedoeld voor een kleine doelgroep. Het RCT is dan geen geschikte onderzoeksmethode om de effectiviteit te kunnen vaststellen (Van Yperen & Veerman, 2008). In geval van interventies met een kleine doelgroep, waarbij tevens willekeurige toewijzing aan een onderzoeks- en controlegroep niet mogelijk is, zal dus moeten worden uitgeweken naar een ander onderzoeksdesign. Een voorbeeld van zo'n alternatieve opzet is het N=1 design.

Het N=1 design is een bruikbare onderzoeksopzet voor het identificeren van het effect van een interventie bij individuen, in de opmaat naar een RCT-studie en in onderzoek naar de effectiviteit voor bepaalde doelgroepen of voor bepaalde varianten van een interventie. Het N=1 design wordt gezien als een gedegen manier om de resultaten van een interventie, gericht op een klein aantal personen, in kaart te brengen. Volgens sommigen wordt dit design gezien als volwaardig alternatief voor de RCT (Task Force APA, 1995; Van Yperen & Bijl, 2006) en volgens anderen als goede opmaat voor een RCT (Bartels, Spreen, Schuringa & Teeken, 2008). Volgens Robey (2004) speelt N=1 onderzoek een belangrijke rol in verschillende stappen van klinisch (effect)onderzoek. N=1 onderzoek komt in Robey's hiërarchisch model vooral aan bod in het identificeren van therapeutisch effect bij individuen (stap 1), in de voorbereiding voor RCT-onderzoek (stap 2) en om effectiviteit in speciale (sub)populaties en in varianten van de interventie te onderzoeken (stap 4). Stap 3 in het model bestaat uit het uitvoeren van klinische trials (RCT-onderzoek).

Centraal bij de N=1 aanpak staat dat bij verschillende observanten op meerdere tijdstippen informatie wordt verzameld over de ontwikkeling van de deelnemers. Een klein aantal deelnemers is typisch voor N=1 onderzoek, wat niet wil zeggen dat het onderzoek letterlijk één onderzoeksobject kan bevatten; dat zijn er veelal meerdere. Door meerdere informanten per deelnemer te betrekken in het onderzoek kan 'triangulatie' worden gerealiseerd. Triangulatie houdt in dat vanuit verschillende invalshoeken metingen worden verricht, bijvoorbeeld via verschillende informanten en het combineren van kwantitatieve en kwalitatieve onderzoeksmethoden binnen een onderzoeksopzet (Oates, 2005). Dit heeft positieve gevolgen voor de betrouwbaarheid van de onderzoeksresultaten. Als er sprake is van vergelijkbare gewenste patronen in de ontwikkeling van de deelnemers volgens verschillende informanten, onderbouwt dit de bewijskracht voor de werkzaamheid van de interventie (Van Yperen et al., 2008). Essentieel in N=1 onderzoek is het gebruik van instrumenten die sensitief zijn voor verandering en waarmee nauwkeurig gemeten kan worden (dat wil zeggen, dat er een goede aansluiting met de doelen van de interventie is). Dit is juist van belang

om ondanks kleine aantallen respondenten toch voldoende power te realiseren om het eventuele effect van de interventie aan te kunnen tonen. In de N=1 aanpak kunnen zowel kwalitatieve als kwantitatieve maten worden gebruikt. Een meer uitgebreide beschrijving van het N=1 design en de bijbehorende analyses is beschreven in Bartels et al. (2008).

## Het N=1 design toegepast

Recent hebben Timmermans, Witvliet en Homburg (2015), in opdracht van het WODC van het ministerie van Veiligheid en Justitie, het N=1 design toegepast in onderzoek naar de doeltreffendheid van de justitiële gedragsinterventie Multidimensional Treatment Foster Care (MTFC). Deze interventie is ontwikkeld voor jongens en meisjes in de leeftijd van 12 t/m 17 jaar met ernstig antisociaal gedrag, meestal gepaard gaand met delinquent gedrag. MTFC wordt ingezet ter vervanging of bekorting van een gesloten behandeling in een jeugdgevangenis (Justitiële Jeugdinstelling, JJI). De interventie is in 1983 in de Verenigde Staten ontwikkeld door Patricia Chamberlain en collega's en houdt een intensieve begeleiding in binnen een opvoedgezin (dus niet het oorspronkelijke gezin), waarbij jongeren worden getraind in sociaal gedrag. De plaatsing in het opvoedgezin voorziet tevens in een (tijdelijke) nieuwe woonomgeving, buiten de eigen sociale context. Op de lange termijn is het doel van MTFC recidivevermindering van delinquent gedrag, het behalen van een startkwalificatie/schooldiploma en het hebben van een stabiele en permanente verblijfplaats. De verwachting is dat deze langetermijndoelen worden bereikt door tijdens MTFC te werken aan een aantal kortetermijndoelen: het verminderen van antisociale gedragsproblemen, een toename van sociale- en probleemoplossende vaardigheden, het aangaan van nieuwe vriendschappen met prosociale leeftijdgenoten en tot slot het volgen van onderwijs of het hebben van een andere zinvolle dagbesteding (MTFC-programmahandleiding, 2010). Het totale programma duurt circa een halfjaar tot een jaar.

In de periode waarop de doeltreffendheidsstudie van MTFC betrekking had, april 2012 t/m maart 2014, zijn in totaal dertien jongeren van start gegaan met het programma. Met dit aantal wordt direct duidelijk dat zicht op de effectiviteit van MTFC niet verkregen kan worden via het doen van een grootschalige experimentele trial. In het onderzoek is daarom uitgeweken naar het N=1 design. De primaire doelstelling was om te onderzoeken in hoeverre MTFC doeltreffend is. Daarnaast is onderzocht in hoeverre MTFC toegevoegde waarde biedt ten opzichte van de gebruikelijke zorg bij de doelgroep, namelijk verblijf in een Justitiële Jeugdinstelling.

Van de dertien jongeren die in de onderzoeksperiode zijn gestart met MTFC, wilden er twee geen medewerking verlenen aan het onderzoek

en zijn er drie gedurende het MTFC-traject uitgevallen. De ontwikkeling van de acht MTFC-deelnemers die uiteindelijk in het onderzoek zijn betrokken, is via vragenlijstonderzoek, interviews en overig materiaal dat tijdens het MTFC-traject is verzameld, in kaart gebracht voor de kortetermijndoelen. Per deelnemer zijn verschillende informanten – opvoedouders, leerkrachten op school, vrienden, familie en/of eerdere begeleiders, medewerkers van het MTFC-team en de jongere zelf – geraadpleegd. De analyse richtte zich op de vraag in hoeverre er sprake is van een significante daling of stijging in de korte termijndoelen van MTFC bij de deelnemers en hoe sterk deze veranderingen zijn. Daarnaast is de concrete situatie van de MTFC-deelnemers bij uitstroom uit het programma inzichtelijk gemaakt voor wat betreft school, werk, wonen en contacten met justitie.

Voor het vaststellen van de toegevoegde waarde zijn de acht MTFC-deelnemers vergeleken met een vergelijkingsgroep van acht jongens in een Justitiële Jeugdinstelling, om zo het verschil in ontwikkeling tussen MTFC en de gebruikelijke zorg bij deze doelgroep te meten. De beperkte omvang van beide groepen maakt het onmogelijk om verschillen in de ontwikkeling op de korte termijndoelen van MTFC op de gebruikelijke wijze, via het vergelijken van gemiddelde scores, te toetsen. Het aantal personen per groep (acht) is immers te laag om eventuele verschillen in groepsgegevens te kunnen detecteren. Om de toegevoegde waarde van MTFC toch zinvol te kunnen beoordelen is van een alternatieve methode gebruikgemaakt, namelijk het statistisch genereren van een voldoende grote vergelijkingsgroep op basis van de verzamelde gegevens van de acht jongens in de JJI. Het aantal cases in de gegenereerde vergelijkingsgroep is gebaseerd op de (geschatte) doelpopulatie in Nederland en bestaat uit 800 cases. Deze omvang van de vergelijkingsgroep is gebaseerd op gegevens over deze populatie, zoals het aantal jongeren in een JJI (*JJI in getal 2008-2012*, Den Haag: DJI) en het aantal jongeren met ernstige antisociale problematiek (Vreugdenhil et al., 2004).

Voor iedere MTFC-deelnemer is vervolgens op basis van rangschikking vastgesteld in hoeverre zijn of haar ontwikkeling beduidend gunstiger uitpakt dan verwacht mag worden op basis van de ontwikkeling in de gegenereerde vergelijkingsgroep. Op basis van internationaal onderzoek naar MTFC (Eddy, Whaley & Chamberlain, 2004; Leve & Chamberlain, 2005; Leve, Chamberlain & Reid, 2005) werd van tevoren verwacht dat de MTFC-deelnemers zich bij de 'beste' twintig procent zou bevinden voor wat betreft hun ontwikkeling op de kortetermijndoelen. Zo verwachtte men bijvoorbeeld dat de MTFC-deelnemers gedurende het programma een toename zouden laten zien in pro sociaal gedrag die vergelijkbaar was met de toename van de twintig procent 'sterkste' stijgers in de gegenereerde vergelijkingsgroep. Deze norm sluit aan bij de relatief grote effecten van MTFC die in de internationale empirische literatuur zijn

gevonden.

Op basis van de analyses van de doeltreffendheidstudie kan worden geconstateerd dat de acht MTFC-deelnemers op iets meer dan de helft van de gemeten gedragingen gedurende MTFC geen (significante) veranderingen lieten zien. De overige gemeten gedragingen ontwikkelden zich vaker in lijn met de tussentijdse doelen (dit gold voor 32% van alle gemeten gedragingen) dan in tegengestelde richting (dit gold voor 12% van de gemeten gedragingen). Van de zeven jongeren waarbij de situatie na uitstroom uit MTFC bekend was, volgden er zes na MTFC een opleiding, hadden twee een (bij)baan en waren er vier niet meer in aanraking gekomen met justitie. Twee jongeren zijn na MTFC onder begeleiding zelfstandig gaan wonen. Eén jongere bevond zich na MTFC in detentie. Vier jongeren zijn weer teruggekeerd naar hun oude thuissituatie. Alles bij elkaar genomen duiden deze resultaten op een beperkte mate van doeltreffendheid van MTFC. Ten aanzien van de analyse van de toegevoegde waarde van MTFC kon, alles bij elkaar, voor een minderheid van de kortetermijndoelen worden geconstateerd dat de MTFC-deelnemers zich ten opzichte van de gegenereerde vergelijkingsgroep, volgens verwachting, bij de beste twintig procent bevonden. Dit duidt op een beperkte mate van toegevoegde waarde van MTFC ten opzichte van de gebruikelijke zorg binnen de Justitiële Jeugdinstellingen.

## Bruikbaarheid van het N=1 design

De beschreven illustratie van de toepassing van het N=1 design voor het onderzoeken van de effectiviteit van de justitiële gedragsinterventie MTFC laat zien dat dit design bruikbaar is als alternatief voor de gecontroleerde experimentele trial (RCT). Er zijn echter ook (mogelijke) beperkingen en risico's waarmee in de toepassing van het N=1 design rekening gehouden moet worden. Uit het MTFC-onderzoek bleek bijvoorbeeld dat het moeilijk was om alle beoogde informanten te bevragen, omdat niet alle deelnemers hier toestemming voor gaven of omdat informanten niet op alle meetmomenten bereikt konden worden of medewerking wilden verlenen. Dit geeft aan dat, hoewel dit type effectonderzoek draait om een klein aantal respondenten, niet de aanname gemaakt moet worden dat de dataverzameling eenvoudig en snel uit te voeren is. Om alle belangrijke informanten rond een kleine doelgroep te bevragen, is veel tijd, energie en doorzettingsvermogen nodig. Ook moet met de toepassing van het N=1 design goed worden gelet op een vergelijkbare systematiek van metingen tussen de interventiegroep en de controlegroep waarmee vergeleken wordt. In het MTFC-onderzoek bleek bijvoorbeeld dat er een aanzienlijk kortere tijd zat tussen de voor- en de nameting in de JJI-groep dan in de MTFC-groep. Bovendien had er geen follow-upmeting plaatsgevonden bij de JJI-groep, terwijl dit wel is gebeurd bij de MTFC-groep. Hoewel er dus ook

zeker mogelijke beperkingen en risico's ten aanzien van het N=1 design te identificeren zijn, en bovendien de meningen verdeeld zijn over de status van het design als volwaardig alternatief voor het RCT of als aanvulling/opmaat hiervoor (Bartels et al., 2008; Robey, 2004; Task Force APA, 1995; Van Yperen & Bijl, 2006), geeft de studie naar de doeltreffendheid van MTFC wel aan dat het een kansrijk design kan zijn voor het onderzoeken van de effecten voor interventies waarbij het aantal deelnemers laag is.

## Literatuur

Bartels, A., Spreen, M., Schuringa, E., & Teeken, V. (2008). *N=1: nauwkeurige en sensitieve behandel-evaluatie op individueel niveau*. Utrecht/Groningen/Doorwerth: Expertisecentrum Forensische Psychiatrie, FPC Dr. S. van Mesdag, Dr. Leo Kannerhuis.

Eddy, J.M., Whaley, R.B., & Chamberlain, P. (2004). The prevention of violent behavior by chronic and serious male juvenile offenders: A 2-year follow-up of a randomized clinical trial. *Journal of Emotional and Behavioral Disorders, 12*, 2-8.

*JJI in getal 2008-2012*. (2013). Den Haag: Dienst Justitiële Inrichtingen, Ministerie van Veiligheid en Justitie.

Leve, L.D., & Chamberlain, P. (2005). Association with delinquent peers: Intervention effects for youth in the juvenile justice system. *Journal of Abnormal Child Psychology, 33*, 339-347.

Leve, L. D, Chamberlain, P., & Reid, J. B. (2005). Intervention outcomes for girls referred from juvenile justice: Effects on delinquency. *Journal of Consulting and Clinical Psychology, 73*, 1181-1185.

Oates, B.J. (2005). *Researching information systems and computing*, SAGE Publications, Middlesborough.

Robey, R.R. (2004). A five-phase model for clinical-outcome research. *Journal of Communication Disorders, 37*, 401-411.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Task Force American Psychological Association (APA), 1995.

Timmermans, M., Witvliet, M., & Homburg, G. (2015). *Evaluatie proces en doeltreffendheid Multidimensional Treatment Foster Care (MTFC)*. Amsterdam: Regioplan.

Van Yperen, T.A., & Bijl, B. (2006). De opzet van praktijkgestuurd onderzoek. In T. van Yperen & J.W. Veerman (red.), *Zicht op effectiviteit. Bronnenboek voor praktijkgestuurd onderzoek in de jeugdzorg* (deel 3, pp. 7-30), NIZW, Praktikon en Ministerie van VWS.

Van Yperen, T.A., & Veerman, J.W. (red.). (2008). *Zicht op effectiviteit. Handboek voor praktijkgestuurd effectonderzoek onderzoek in de jeugdzorg*. Delft: Eburon.

Vreugdenhil, C., Doreleijers, T.A.H., Vermeiren, R., Wouters, L.F.J.M., & Van den Brink, W. (2004). Psychiatric disorders in a representative sample of incarcerated boys in the Netherlands. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 97-104.