



[www.DirectieveTherapie.nl](http://www.DirectieveTherapie.nl)

## Dank voor uw download

U kunt er natuurlijk uit citeren, graag zelfs, maar dan wel met bronvermelding. U mag dit artikel ook ruimhartig verspreiden mits het niet voor commerciële doeleinden is. In die gevallen pas na onze schriftelijke toestemming.

Opname in bloemlezingen en readers moedigen wij aan, maar wel graag eerst even overleggen.

Alle rechten van de artikelen liggen bij  
**de Stichting Cognitie en Psychose.**

Voor alle vragen:  
[info@gedachtenuitpluizen.nl](mailto:info@gedachtenuitpluizen.nl)

# De rol van ‘informed consent’ in placebo-gecontroleerde, dubbelblinde, gerandomiseerde onderzoeken naar de effectiviteit van psychofarmaca

Willem van den Burg\*

---

## Samenvatting

*Gecontrasteerd worden twee mogelijke effecten van de informed consent-procedure bij dubbelblinde, placebo-gecontroleerde en gerandomiseerde onderzoeken naar de effectiviteit van psychofarmaca: 1 onderschatting van de effectiviteit van het psychofarmakon door onzekerheid of men misschien geen placebo gekregen heeft en 2 overschatting van die effectiviteit door gespitsheid op de vraag wát men gekregen heeft en als gevolg daarvan onblinding op grond van de bijwerkingen. Met betrekking tot het gepleegde onderzoek naar de effectiviteit van antidepressiva blijken er nauwelijks of geen aanwijzingen te zijn voor onderschatting en sterke aanwijzingen voor overschatting. De problematiek laat zien dat de onderzoeken minder ‘objectieve’ resultaten opleveren dan vaak wordt aangenomen.*

## Inleiding

Ingewikkelder varianten daargelaten, is de opzet van een *placebo-controlled, double-blind randomized trial* (PDRT) zeer eenvoudig en, in theorie, onovertroffen om de effectiviteit van een psychofarmakon te onderzoeken: *at random* geven we de helft van een groep patiënten het betreffende middel en de andere helft een identiek ogend placebo, terwijl we noch de patiënt noch de clinicus/beoordelaar laten weten wie wat gekregen heeft. Als dan de patiënten die het psychofarmakon kregen na verloop van tijd méér verbeterd zijn dan de patiënten die de placebo kregen, kan redelijkerwijs worden aangenomen dat *selection bias* – de twee groepen verschilden van tevoren op belangrijke kenmerken – geen plausible, alternatieve verklaring voor dit verschil vormt. Daarnaast kan, volgens de opzet, de dubbelblindheid garanderen dat het verschil niet vertekend wordt door allerlei *biases* in de beoordeling. Het is dus alleszins begrijpelijk dat PDRT's als de basis worden gezien voor het ideaal van *evidence based medicine* en, in het verlengde daarvan, voor het uitvaardigen van behandelingsvoorschriften ('richtlijnen').

Er kunnen echter nogal wat haken en ogen zitten aan dit type onderzoek zoals dat gewoonlijk plaatsvindt. Zo kan men betwijfelen of de bevindingen naar de dagelijkse behandelingspraktijk gegeneraliseerd kunnen worden, omdat zowel mild depressieve als zeer ernstig depressieve patiënten zelden in de onderzoeken opgenomen worden, evenmin als patiënten met comorbiditeit (die veelvuldig voorkomen). Of omdat de onderzoeken meestal slechts kort duren terwijl men de effectiviteit op langere termijn zou willen weten. De meeste van die problemen zijn, althans in theorie, door ander onderzoek oplosbaar: zij zijn niet inherent verbonden aan de proefopzet. Zo zou men meer 'pragmatische onderzoeken' kunnen uitvoeren, waarbij de deelnemers niet

---

\* DR. W. VAN DEN BURG, psycholoog, is werkzaam als universitair docent en onderzoeker bij de afdeling Neuropsychologie van de Rijksuniversiteit Groningen. Correspondentieadres: Postbus 30 001, 9700 RB Groningen.

geselecteerd worden op een 'net' door het DSM gesanctioneerd syndroom, maar op patiëntengroepen zoals men ze in de praktijk tegenkomt (Hotopf, Churchill, & Lewis, 1999). Of men zou langer durende onderzoeken kunnen uitvoeren.

Maar er zijn twee complicerende aspecten die wél als inherent aan PDRT's kunnen worden beschouwd. Het eerste betreft de moeilijk vermijdbare omstandigheid dat de clinicus/beoordelaar weet dat onderzoek plaatsvindt en daarvan de proefopzet kent, hetgeen zijn beoordelingen kan beïnvloeden. In het verleden is een enkele keer onderzoek uitgevoerd waarbij dat niet het geval was, zoals in het fraaie, *triple blind*-onderzoek van Paul, Tobias en Holly (1972). (Dit liet geen duurzame, aanvullende effectiviteit van antipsychotica zien toen psychosociale behandelingsprogramma's geïntroduceerd werden voor in een inrichting verblijvende, schizofreenachtige patiënten.)

Het tweede is eigenlijk een complement van het eerste en van betrekkelijk recente datum. Het betreft de eis dat de patiënten hun toestemming dienen te geven voor deelname aan het onderzoek, na precies te zijn geïnformeerd over de aard en opzet van het onderzoek; zij dienen hun informed consent te geven. Dit houdt onder andere in dat ook hun verteld wordt dat zij een placebo zouden kunnen krijgen. Vroeger, tot zo'n 35 jaar geleden, was het leven eenvoudiger. Zoals Bradford Hill, de architect van het beroemde onderzoek in 1948 naar de effectiviteit van streptomycine als tuberculosebestrijder, in een terugblik schreef (Hill, 1990, p. 78):

'Of course, there were no ethical problems in those days: we did not ask for the patient's permission or anybody's permission. We did not tell them they were in a trial – we just did it.'

Dat de patiënten ook wat te zeggen zouden kunnen hebben over hun behandeling was een non-issue (Faden & Beauchamp, 1986) en artsen die bij een onderzoek betrokken waren, werden gewoon geacht (praktisch) alle patiënten te leveren die in aanmerking kwamen (Silverman & Altman, 1996). De bijbehorende denkwereld waarin de arts als allesbepaler figureert, is overigens slechts langzamerhand veranderd. Zo doet Joyce, in 1982, nog steeds of patiëntendeelname aan een PDRT eigenlijk alleen een zaak is van de arts en meent hij dat als naar diens oordeel het effect van de behandeling verkleind zou kunnen worden als de patiënt verteld wordt dat hij een placebo kan krijgen, dit niet gezegd moet worden.

Weinigen zullen, op ethische gronden, nu nog de klok willen terugdraaien. Dit neemt echter niet weg dat de informed consent-procedure onderzoekstechnisch en methodologisch aanzienlijke problemen met zich heeft meegebracht. Na de invoering van deze verplichting werden onderzoekers geconfronteerd met sterk teruglopende aantallen deelnemers (Silverman, 1995; Silverman & Altman, 1996). Mede daardoor bestaat nu veel literatuur met als onderwerp hoe voldoende mensen te recruter (Spilker & Cramer, 1992). 'Recruitment problems are undoubtedly the single most important reason for the failure of clinical trials' (Spilker & Cramer, 1992, p. XVII). Tegen deze achtergrond is dan de vraag in hoeverre de resultaten van PDRT's, verkregen bij selecte groepen van vrijwilligers, ook gelden voor 'gewone' patiënten. Met betrekking tot psychofarmacologisch onderzoek is deze kwestie onder andere aan de orde gesteld

door Anseau (1992), Edlund, Craig en Richardson (1985) en Licht, Gouliaev, Vestergaard en Frydenberg (1997) .

Daarnaast is er het probleem in hoeverre de informed consent-procedure op zich kan leiden tot een onrealistische taxatie van de effectiviteit van een psychofarmacon (hetgeen de zorg van Joyce was). Dit is het onderwerp van het navolgende. Het gaat in meerdere opzichten om zowel een spiegelbeeld als verdubbeling van de altijd al aanwezige problematiek die ontstaat doordat de clinicus/beoordelaar op de hoogte is van de proefopzet. Mogelijkheden van vertekening zullen eerst in het algemeen besproken worden en vervolgens in het kader van de vele placebo-gecontroleerde, dubbelblinde, gerandomizeerde trials die zijn uitgevoerd om de effectiviteit vast te stellen van antidepressiva bij depressieve patiënten (afgekort PDRT-AD's).

### **Leidt informed consent tot onderschatting van de effectiviteit van een psychofarmacon?**

Het ligt voor de hand dat een patiënt zich tijdens een PDRT regelmatig afvraagt: wat heb ik nu gekregen, het echte middel of het pilletje-met-niks? Het is dan goed denkbaar dat hij zich in die context van onzekerheid bij zijn zelfbeoordeling en rapportage aan de clinicus/beoordelaar enigszins terughoudend en sceptisch opstelt, met als gevolg een onderschatting, in absolute zin, van de effectiviteit van het verum en een te klein verum-placeboverschil. Misschien ligt, zoals Kienle (1996) veronderstelt, aan zo'n terughoudendheid ten grondslag dat de patiënt 'might easily feel that he himself and his ability to give correct answers are on trial, rather than the drug being tested' (p. 79), en geen grote 'fouten' wil maken. Op analoge wijze zou de clinicus/beoordelaar de effectiviteit van een farmacon kunnen onderschatten. Veatch (1987, p. 62) daarentegen meent dat de informed consent-procedure de accuraatheid van de rapportage ten goede komt.

Een experiment van Kirsch en Rosadino (1993) naar de invloed van cafeïne op de ervaren gespannenheid van gezonde psychologiestudenten zou kunnen wijzen op onderschatting. Tegen één groep studenten werd gezegd dat zij, in het kader van een vergelijkend onderzoek naar het effect van cafeïnevrije en gewone koffie, in de 'gewone-koffie'-conditie zaten. Dit was echter bedrog: de helft kreeg cafeïnevrije koffie (de placebo). Een andere groep kreeg de gebruikelijke informed consent-informatie: u krijgt hetzij gewone koffie, hetzij cafeïnevrije koffie, en wat u krijgt vertellen wij niet. Wie in de gewone-koffieconditie zat en ook daadwerkelijk gewone koffie kreeg, bleek gemiddeld meer gespannenheid te ervaren dan wie gewone koffie kreeg in de informed consent-conditie. En het verschil met cafeïnevrije koffie was groter in de gewone-koffieconditie dan in de informed consent-conditie. Hoewel Kirsch en Rosadino hun bevindingen anders verklaren dan Kienle zou doen, is de boodschap van de auteurs dezelfde: 'Current double-blind methods may fail to detect some genuine drug effects' (Kirsch & Rosadino, 1993, p. 442). Bergmann, Chassany, Gandiol, Deblois, Kanis, Segrestaa, Caulin en Dahan (1994) kwamen tot dezelfde conclusie op grond van een dubbelblind, cross-over-onderzoek naar de effectiviteit van een pijnstiller bij kankerpatiënten. At random werd eerst de pijnstiller gegeven en een dag later de placebo, of andersom. Eén groep patiënten kreeg zonder enige informatie de pillen. De tweede groep werd netjes om toestemming gevraagd voor deelname aan het

onderzoek, na geïnformeerd te zijn over de opzet. In deze laatste groep was het verum-placeboverschil kleiner dan in de eerste groep.

De resultaten van deze twee onderzoeken kunnen echter niet automatisch gegeneraliseerd worden: eenzelfde onderzoeksopzet als van Kirsch en Rosadino (1993), waarbij het ging om de effectiviteit van nicotine bevattend kauwgum om het roken te laten, liet geen spoor van resultaten zien die op onderschatting van die effectiviteit wijzen door de informed consent-procedure (Hughes, Gulliver, Amori, Mireault, & Fenwick, 1989). Vergelijkbaar ander onderzoek is schrijver dezes onbekend en de uitvoering daarvan zal tegenwoordig op grote bezwaren stuiten, zeker als de onderzoeksgroep uit patiënten bestaat. De onderzoeksopzetten impliceren bedrog (Kirsch & Rosadino, 1993) of een inbreuk op het recht op informatie (Bergmann et al., 1994) en zullen waarschijnlijk geen medisch-ethische commissie passeren. (Bergmann et al. schrijven dat toen hun onderzoek uitgevoerd werd, in 1988, de Franse voorschriften nog niet zo streng waren.)

Een andere benadering is te bezien of de effectiviteit van een farmacon blijktens placebo-gecontroleerd onderzoek kleiner is dan de effectiviteit van datzelfde farmacon blijktens vergelijkend onderzoek met andere farmaca. Dat zou men verwachten als de informed consent-procedure in PDRT's tot onderschatting leidt. Aldus werden op een Noorse kraamafdeling vlak na elkaar en zonder dat het protocol of de omstandigheden aldaar veranderden, twee onderzoeken gedaan naar de effectiviteit van paracetamol als postpartale pijnstillers (Skovlund, 1991). In het eerste onderzoek werd paracetamol vergeleken met een placebo, in het tweede met een andere pijnstiller. In beide gevallen werd informed consent verkregen. De pijnstillende werking van paracetamol werd door de patiënten in het placebo-gecontroleerde onderzoek als duidelijk geringer beoordeeld. Skovlund becommentarieert:

'When patients participating in a clinical trial are informed that they may receive placebo, the effect of the active drug in the trial may be reduced, and the trial will tend to underestimate the effect of the same drug in ordinary clinical practice.'

Maar een mogelijke interpretatie is natuurlijk ook, andersom en in de geest van Veatch (1987), dat positieve verwachtingen en beleefdheid van de patiënten de effectiviteit in het vergelijkende onderzoek chargeerden.

In de vorm van meta-analyses is dit type onderzoek ook uitgevoerd met betrekking tot de effectiviteit van antidepressiva, zij het dat de beoordelingen niet die van de patiënten zelf betroffen, maar van de onderzoekers (die natuurlijk voor een belangrijk deel oordelen op geleide van wat de patiënten hun vertellen). Wechsler, Grosser en Greenblatt (1965) onderzochten de gerapporteerde effectiviteit bij opgenomen patiënten in 79, in de jaren 1958-1963 gepubliceerde onderzoeken. Conform de bevindingen van Skovlund, bleek die effectiviteit als aanmerkelijk groter beoordeeld te worden in onderzoeken waarin twee of meer antidepressiva vergeleken werden dan in simpele PDRT-AD's. De destijds nog ongebruikelijke informed consent-procedure kan daarmee weinig te maken hebben. Een nieuwe analyse van Trivedi en Rush (1994) wijst daarentegen niet op grote verschillen in de effectiviteit van antidepressiva volgens PDRT-AD's en volgens vergelijkende onderzoeken. De geanalyseerde onderzoeken,

met in totaal 229 *treatment cells*, werden gepubliceerd in de periode 1975-1990, toen de informed consent-procedure redelijk ingeburgerd was.

Alles bij elkaar genomen moet de conclusie zijn dat de informed consent-procedure bij de gangbare placebo-gecontroleerde onderzoeken misschien soms een neerwaartse invloed zou kunnen hebben op de getaxeerde effectiviteit van antidepressiva, maar dat daar vooralsnog weinig evidentie voor is.

## Onblindering

In het voorgaande is er stilzwijgend van uitgegaan dat patiënten en beoordelaars in het ongewisse blijven over wie wat gekregen heeft. Technisch-methodologisch gezien was dan de vraag in hoeverre de bevindingen, verkregen in die situatie, generaliseerd kunnen worden naar de gewone situatie waarin men wél weet wat men krijgt; de externe validiteit van PDRT's is in het geding. Maar dezelfde informed consent-procedure die de patiënt inprent dat hij een placebo kan krijgen en die tot een onderschatting van de effectiviteit zou kunnen leiden, zal, kan men verwachten, ook gespitstheid op de vraag met zich meebrengen wát men eigenlijk gekregen heeft. Een psychofarmakon zal doorgaans sterkere lichamelijke gewaarwordingen veroorzaken dan een placebo – en vaak zeer sterke (misselijkheid, droge mond, andere 'bijwerkingen'), zodat een patiënt wel degelijk een idee krijgt in welke groep hij zit. A fortiori kan dat gelden voor de clinicus/beoordelaar die de typerende effecten van het psychofarmakon gewoonlijk goed kent. Als de groep waarin de patiënt zit aldus vaak goed gegokt wordt, is de feitelijke dubbelblindheid doorbroken en is de integriteit van de PDRT in het geding, dat wil zeggen de interne validiteit. Dan ligt de weg open naar een overschatting van de effectiviteit.

Het probleem is sinds de intrede van de PDRT's regelmatig onder de aandacht gebracht (Fisher & Greenberg, 1993; Oxtoby, Jones, & Robinson, 1989). Desondanks wordt de interne validiteit zelden onderzocht door navraag bij de patiënten naar wat zij denken gekregen te hebben, of bij de beoordelaars naar wat de patiënten volgens hen kregen. Dat was vroeger zo (Ney, Collins, & Spensor, 1986) en is nog steeds zo (Mulrow et al., 1999). Het begrip *pact of ignorance* van Orne (1969) lijkt hier van toepassing: men doet of zijn neus bloedt, want niemand heeft er belang bij te weten of de proefopzet misschien eigenlijk mislukt of half gelukt is. Oxtoby et al. (1989) noemen voorbeelden die suggereren dat tijdschriftredacties artikelen weigerden omdat daarin het probleem aan de orde werd gesteld.

Niettemin is uit het onderzoek dát is gedaan langzamerhand duidelijk geworden dat 'the breaching of the double-blind' eerder regel dan uitzondering is (Fisher & Greenberg, 1997), zij het dat middelen als lithium, haloperidol en imipramine ongetwijfeld eerder tot doorbreking leiden dan bijvoorbeeld de nieuwe non-benzodiazepine anxiolytica. Schweizer en Rickels (1997, p. 34.) vermoeden dat de effectiviteit van deze laatste middelen zo weinig duidelijk uit PDRT's blijkt, omdat deze middelen, in vergelijking met benzodiazepinen met hun onmiddellijk waarneembare sederende effecten, zo weinig bijwerkingen hebben

'(...) which leaves both the patient and the clinician with no way of telling whether random assignment has been to active drug or placebo. This is particularly a problem for those drugs that need weeks to show any effect.'

Alsof de effectiviteit van de middelen pas goed kan blijken als de dubbelblindheid doorbroken is! De overtuiging dat gewoonlijk geen sprake is van feitelijke dubbelblindheid heeft geleid tot krachtige uitspraken:

'Physicians who administer active medication and placebo may be "blind" in terms of official knowledge of the patient's assignment, but I have often wondered if their vision is 20/20 when it comes to "knowing" which medication a specific patient has received' ( Heimberg, 1997).

'There is very little evidence that the double blind trial is blind for anybody, except those who read the report' ( Ney et al., 1986).

Onblinding hoeft niet desastreus te zijn. In de eerste plaats kan, ook al wordt de groep waarin een patiënt zich bevindt goed geraden, dit vermoeden of deze overtuiging door onzekerheid omgeven zijn, hetgeen 'gewenste reacties' kan temperen. Ook zal men het zeker niet altijd bij het rechte eind hebben; sterke staaltjes van bijkans absolute zekerheid dat de patiënt het psychofarmakon kreeg, terwijl het toch een placebo bleek, vindt men overal in de literatuur (bijvoorbeeld Rickels, 1986). Verder hoeft het besef van een patiënt dat hij in de controlegroep zit niet per se tot teleurstelling of pessimisme te leiden. Kramer en Shapiro (1984) bespreken een onderzoek naar de mogelijkheid via een complexe set interventies dood door hartfalen te voorkomen, waarbij de controlegroep niets kreeg. Geen verschil werd gevonden, maar beide groepen lieten een lager aantal sterfgevallen zien dan men normaliter kon verwachten. De auteurs geven als verklaring dat de wetenschap tot de controlegroep te horen ertoe leidde dat de deelnemers door hun gedrag in gezonde richting bij te stellen, die wetenschap compenseerden. Misschien kan dergelijk compensatiegedrag soms ook optreden in psychofarmacologische trials. De mededeling van sommige patiënten dat zij de voorkeur gaven aan de controleconditie omdat zij hoopten te 'bewijzen' geen antidepressivum nodig te hebben ( Thase, 1999), wijst op die mogelijkheid. Een belangrijk punt is ten slotte dat de onblinding eerder een irrelevante bijkomstigheid kan zijn dan een probleem, namelijk als het gewenste effect sterk is en de onblinding geheel het gevolg van de ervaring van dat gewenste effect. Voor dat geval hebben Ney et al. (1986). Philip's paradox bedacht (naar de voornaam van de eerste auteur): hoe sterker het gewenste effect, des te slechter bewijsbaar in dubbelblinde onderzoeken. In principe kunnen juiste *medication guesses* zowel met de gewenste effecten als met irrelevante of ongewenste effecten samenhangen. Samenhangen met de gewenste effecten van psychofarmaca zijn een enkele keer gerapporteerd ( Schweizer & Rickels, 1997).

## Onblinding in PDRT-AD's

Ook in PDRT-AD's kan onblinding in principe natuurlijk soms voortkomen uit de gewenste effecten van de antidepressiva, maar gegevens dat dit een belangrijke route is, zijn er niet. Die route ligt ook niet voor de hand. In PDRT-AD's plegen zowel de antidepressivum- als de placebogroepen sterk te verbeteren, terwijl het uiteindelijke verschil in verbetering gemiddeld gering is ( Bollini, Pampallona, Tibaldi, Kupelnick, & Munizza, 1999; Khan, Warner, & Brown, 2000; Mulrow et al., 1999). Bovendien treden de bijwerkingen van antidepressiva doorgaans veel eerder op dan de gewenste effecten. Dit patroon suggereert niet dat juiste medication guesses vooral zijn toe te schrijven aan de ervaren vooruitgang.

Vier soorten aanwijzingen daarentegen suggereren dat de bijwerkingen een belangrijker route vormen en tot overschatting van de effectiviteit van antidepressiva geleid hebben of kunnen leiden. Zoals besproken zal worden is geen van deze aanwijzingen spijkerhard, maar tezamen genomen maken zij overschatting wel waarschijnlijk.

1. Een aantal onderzoeken is uitgevoerd waarin 'actieve placebo's' gebruikt werden teneinde bijwerkingen te veroorzaken die vergelijkbaar zijn met die van de antidepressiva. De gemiddelde effectgrootte van de antidepressiva volgens die onderzoeken valt als duidelijk geringer te berekenen dan wanneer gewone, 'inerte' placebo's worden gebruikt (omstreeks 0,20; Moncrieff, Wessely & Hardy, 1998). Een probleem met deze onderzoeken is dat de meeste wel erg oud en klein zijn, mede daardoor problemen opleveren bij de berekeningen, en dat atropine, de meest gebruikte actieve placebo, misschien geen echte placebo is, maar zelf ook enige antidepressieve activiteit bezit ( Cooney, 1998; Even, Siobud-Dorocant, & Dardennes, 2000; Healy, 1998). Het is zeer te betreuren dat dit type onderzoek niet vaker gedaan werd en wordt.

2. Greenberg, Bornstein, Zborowski, Fisher en Greenberg (1994) berekenden uit PDRT-AD's, waarin fluoxetine het antidepressivum was, een forse correlatie tussen de effectgrootte en de gerapporteerde hoeveelheid bijwerkingen. Dit kan betekenen dat onblinding samenhangt met de effectgrootte, maar misschien ook, volgens de auteurs, dat antidepressiva via de bijwerkingen hun effectiviteit ontlenen 'by, among other things, mobilizing body awareness or serving as a distractor' (p. 550). Een derde mogelijkheid is dat de mensen die baat hebben bij antidepressiva, ook gevoelig zijn voor bijwerkingen. Greenberg et al. (1994) vonden daarnaast een forse negatieve correlatie tussen de grootte van de proefgroepen en de effectgrootte. Zij suggereren dat dit misschien ook met blindheidsproblemen te maken heeft. Bij grote groepen kent de clinicus de patiënten minder goed, of voelt zich, als het onderzoek uitgevoerd wordt in meerdere centra en door vele beoordelaars, wellicht minder bij het onderzoek betrokken. Het gevolg zou dan een geringere gespitsheid zijn op wie wat gekregen heeft, en een kleinere bias. De correlatie kan echter ook te wijten zijn aan publicatie-bias: kleine onderzoeken worden makkelijker gepubliceerd als er positieve resultaten te melden vallen. Het bezien of kleine gepubliceerde onderzoeken een grotere effectiviteit laten zien dan grote gepubliceerde onderzoeken vormt in feite een standaardmethode om publicatie-bias op het spoor te komen. Op grond daarvan rapporteerden Mulrow et al. (1999) een publicatie-bias met betrekking tot, in het algemeen, de nieuwere antidepressiva. Wat ook de reden is van de relatie, een overschatting van de effectiviteit is het gevolg.



3. Veel patiënten die aan PDRT-AD's deelnemen, maken niet voor de eerste keer een depressie door. Zij zullen vaak al eerder behandeld zijn met antidepressiva, hetgeen de herkenning van de groep waarin men zit ten goede kan komen door de mogelijkheid van vergelijking (Even et al., 2000). Uit de meta-analyse van Mulrow et al. (1999) bleek dat in 75 procent van alle 315 trials die in de berekeningen waren opgenomen (waaronder 81 placebo-gecontroleerde trials) meer dan 50 procent van de deelnemers eerdere depressies had. Swartzman en Burkell (1998) bespreken twee onderzoeken die fraai illustreren hoe de mogelijkheid van vergelijking, in samenhang met de gegeven voorlichting, de 'effectiviteit' kan beïnvloeden. In beide onderzoeken ging het om de vraag of clonidine kan helpen bij de bestrijding van opvliegers in de menopauze en in beide onderzoeken werd in de ene groep eerst clonidine gegeven en dan de placebo, en in de andere groep eerst de placebo en dan clonidine. In het eerste onderzoek werd geen enkel effect van clonidine gevonden. In het tweede onderzoek werd aanvankelijk eveneens geen effect gevonden, maar na de cross-over wel: de mensen die eerst clonidine kregen en toen placebo gingen veel minder effectiviteit aangeven, en de mensen die eerst een placebo kregen en toen clonidine veel meer effectiviteit. Het verschil was dat in het tweede onderzoek de opzet precies verteld was, inclusief wanneer de cross-over zou plaatsvinden, en in het eerste onderzoek alleen maar globaal gezegd werd dat men 'dummies' zou kunnen krijgen. Hoewel de onderzoeken ook nog wel in andere opzichten van elkaar verschilden, was de meest plausibele verklaring dat door de informatie in het tweede onderzoek de mensen extra goed gingen opletten wat voor uitwerking de tabletten hadden en een duidelijk verschil opmerkten na de cross-over; met als gevolg onblinding en een interpretatie van effectiviteit.

4. Aanmerkelijke percentages patiënten in de placeboconditie staken hun deelname aan PDRT-AD's vroegtijdig vanwege *lack of efficacy*, dus bij (nog) hoge depressiviteitsscores. Die beslissing kan mede berusten op de waarneming 'geen bijwerkingen', terwijl de waarneming: 'wel bijwerkingen' (= 'effectief middel') aanleiding kan zijn staking te ontraden, gezien de ervaring dat antidepressiva vaak pas na weken hun effectiviteit laten zien. Als dat het geval is en eindpunt- of *intention to treat*-analyses worden uitgevoerd (gebaseerd op de laatst beschikbare scores), is overschatting van de effectiviteit van het antidepressivum het gevolg; de tijd heeft zijn heilzame werking minder in de placeboconditie kunnen uitoefenen dan in de antidepressivumconditie. Waarschijnlijk heeft onblinding op die manier een rol gespeeld in twee grote PDRT-AD's die door Stassen, Angst en Delini-Stula (1996) worden besproken. De drop-out in de placebogroepen (voornamelijk vanwege gebrek aan effect) vond veel eerder plaats dan in de antidepressivumgroepen (voornamelijk vanwege bijwerkingen) – het grote merendeel in de placeboconditie zelfs binnen twee weken, terwijl men gewoonlijk vier weken neemt om de effectiviteit van een antidepressivum af te wachten. Gegevens van Niklson, Reimitz en Sennel (1997), medewerkers van Organon, doen een soortgelijk effect van onblinding vermoeden. De gegevens betroffen twee grote, drie-armige PDRT-AD's naar de effectiviteit van een nieuw antidepressivum, met in de derde arm imipramine, waaraan 22 onderzoekscentra deelnamen. Zowel imipramine als het nieuwe antidepressivum bleken volgens een eindpuntanalyse minimaal en niet-significant effectiever dan de placebo: ongeveer één punt op de Hamiltonschaal, N = 386. (Om niet-verklaarde redenen werd slechts een selectie van alle 527 patiënten in de berekeningen betrokken.) Niklson et al. maakten vervolgens een onderscheid tussen onderzoekscentra die wel en niet een gemiddeld verschil tussen placebo en imipramine opmerkten van minimaal twee Hamiltonpunten ten gunste van imipramine. Volgens deze indeling werd in acht *discriminative centers* vastgesteld dat zowel

imipramine als het nieuwe antidepressivum gemiddeld vier Hamiltonpunten effectiever was dan de placebo, terwijl volgens de veertien non-discriminative centers de placebo anderhalve punt effectiever was dan beide andere middelen. Ook werd gerapporteerd dat in de discriminative centers minder patiënten in de antidepressivum condities afhaakten (20 procent) dan in de non-discriminative centers (36 procent), terwijl de trend in de placeboconditie andersom was: 38 procent versus 30 procent. 'Thus the remarkable differences in drop-out rates between both types of centers is an important determinant of their differential therapeutic results' (p. 48). De interpretatie van de auteurs was ongeveer dat centra die de antidepressieve effectiviteit van een erkend antidepressivum als imipramine konden onderkennen, dat ook met betrekking tot het nieuwe antidepressivum konden. Een voor de hand liggende mogelijkheid is natuurlijk ook dat onblinding tot bias leidde in de discriminative centers. Een moeilijkheid met de bias-interpretatie is evenwel dat, als op deze manier de effectiviteit van antidepressiva vaak overschat wordt, men zou kunnen verwachten dat 'completer'-analyses, waarin alleen patiënten betrokken worden die de trial afmaakten, gemiddeld een geringere effectiviteit zouden laten zien dan eindpuntanalyses. Dit lijkt echter zeker niet het geval te zijn ( Bollini et al., 1999).

Ten slotte moet nog worden opgemerkt dat een overschatting van de effectiviteit via de bijwerkingen nog vergroot kan worden door een slordige of inadequate uitvoering van het onderzoek. Als bijvoorbeeld een onderzoeker weet dat een patiënt die forse bijwerkingen ondervond, in 'groep A' zat, weet hij meteen (praktisch zeker) dat in groep A het verum gegeven wordt. Als hij dan ook van volgende patiënten weet aan welke groep ze toegewezen worden, is al sprake van onblinding vóór de patiënt nog maar zijn entree gemaakt heeft. Als dan verder het toewijzingschema bekend is, kan bovendien nog een selectie-bias optreden, doordat de onderzoeker bijvoorbeeld even wacht met het includeren van een patiënt tot de groep aan de beurt is waar hij hem het liefste aan toewijst. Dergelijke en andere onvolkomenheden in het onderzoek liggen waarschijnlijk ten grondslag aan de algemene bevinding dat medische onderzoeken een grotere effectiviteit van een verum (of meer algemeen: een specifieke behandeling) laten zien als de rapportering van de concrete werkwijze te wensen overlaat (cf. Schultz, 2000). (Vroeger werd ook nogal eens gesignaleerd dat placebo en verum qua uiterlijk soms duidelijk onderscheidbaar zijn. Misschien is de mogelijkheid dat onblinding ook hierdoor kan ontstaan tegenwoordig geen probleem meer.) De precieze werkwijze wordt gewoonlijk nauwelijks beschreven in rapporten van PDRT-AD's ( Even et al., 2000; Ogundie, Boardman, & Masterton, 1999) en ook trouwens niet in die van PDRT's naar de effectiviteit van antipsychotica ( Thornley & Adams, 1998); vaak wordt niet veel meer gerapporteerd dan: 'Het onderzoek werd dubbelblind uitgevoerd.' Het valt te vrezen dat een gebrekkige uitvoering van PDRT-AD's soms mede tot een overschatting van de effectiviteit van antidepressiva heeft geleid.

## Beschouwing

Volgens de naïeve logica van een placebo-gecontroleerde, dubbelblinde, gerandomizeerde proefopzet hoeft men deze slechts toe te passen om precies en objectief de specifieke effectiviteit van een farmacon vast te stellen. Soms krijgt men de indruk dat deze 'dream of the mechanization of judgment' ( Gigerenzer, Swijtink, Porter, Daston, & Krüger, 1989) sterk leeft bij sommige voorvechters van evidence based medicine, waaronder schrijver dezes zich in beginsel overigens ook rekent, en in het bijzonder

bij fervente ontwikkelaars van algemene richtlijnen voor behandeling. Wendt en Cameron (1961) merken op dat een placebo-effect mede afhankelijk is van

‘the attitude of the subject towards the experiment, the subject’s perception of the experiment, the subject’s expectations as affected by the total procedures and his environment both in and out of the laboratory, and the kinds of drugs being used.’

Hetzelfde geldt voor de oordelen van een onderzoeker en de resultaten van een PDRT. Daardoor is een PDRT slecht vergelijkbaar met een onderzoek naar, zeg, de houdbaarheid van twee soorten gloeilampen. Een realistischer denkkader is een PDRT op te vatten als een soort toneelstuk dat niet op de planken maar in het volle leven opgevoerd wordt, met acteurs van wie een bepaalde, ‘objectief-oordelende’ rol gevraagd wordt – maar die in de concrete vormgeving van die rol altijd eigen percepties van de situatie, verwachtingen en belangen mee kunnen betrekken.

Juist vanuit die optiek is een PDRT duidelijk anders van aard dan de situatie waarin een farmacon gegeven wordt in de klinische praktijk (zie ook Kleijnen, Goetzsche, Kunz, Oxman, & Chalmers, 1997). Zoals besproken kan de informed consent-procedure bij PDRT’s naar de effectiviteit van psychofarmaca, naast of in samenhang met andere factoren zoals de opstelling of bias van de clinicus/beoordelaar, medebepalend zijn voor zowel een onderschatting als een overschatting van de effectiviteit daarvan. De mogelijkheden van vertekening zijn hier vooral een punt van zorg, omdat deze PDRT’s doorgaans nogal variabele resultaten opleveren, een gemiddeld gering verschil laten zien tussen verum en placebo, en noodzakelijkerwijs geheel berusten op subjectieve oordelen (Van den Burg, 1998a, b). Het is mogelijk dat overschattingen en onderschattingen (soms) min of meer tegen elkaar wegvallen in bijvoorbeeld gemiddelde effectgroottes die uit meta-analyses naar voren komen. Voor PDRT-AD’s geldt echter, zoals besproken, dat er weinig evidentie is voor onderschatting en dat er een aantal aanwijzingen zijn voor overschatting.

Ook als men overschatting verre van bewezen acht, is achterdocht op zijn plaats, gezien de pure mogelijkheid daarvan door onblindering via de bijwerkingen en het feit dat men eerder een positieve dan een negatieve bias kan verwachten bij alle betrokkenen. In principe kan men ervan uitgaan dat zowel patiënt, onderzoeker als sponsor hopen op een positief resultaat. Maar in het bijzonder zijn de hoop en de belangen van de sponsor groot. Bijna alle PDRT’s worden betaald door de industrie en vaak worden ze uitgevoerd om een nieuw middel geregistreerd te krijgen. Daarbij is de mate waarin veel onderzoekers in de wereld van de (psycho)farmacologie, financieel en anderszins, betrekkingen onderhouden met de industrie langzamerhand buitengewoon zorgwekkend geworden (Angell, 2000; Fava, 1998; Van den Burg, 2000).

---

### **Abstract**

*Contrasted are two possible effects of the informed consent-procedure in placebo-controlled, double-blind, randomised trials for appraising the efficacy of psychotropic drugs: 1 underestimation of the efficacy of the drug because of doubts about the nature of the pill and 2 overestimation of the efficacy due to intentness on the question what has been administered and, consequently, unblinding by the side-effects of the drug. Regarding the studies performed*

into the efficacy of antidepressant drugs, the data do not appear to point at underestimation, while they strongly suggest overestimation. The issues show that the trials yield less 'objective' results than usually is assumed.

## Referenties

- Angell, M. (2000). Is academic medicine for sale? *New England Journal of Medicine*, 342, 1516-1518.
- Anseau, M. (1992). The Atlantic gap: clinical trials in Europe and the United States. *Biological Psychiatry*, 31, 109-111.
- Bergmann, J-F., Chassany, O., Gandiol, J., Deblois, P., Kanis, J.A., Segrestaa, J-M., Caulin, C., & Dahan, R. (1994). A randomised clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain. *Clinical Trials and Meta-Analysis*, 29, 41-47.
- Bollini, P., Pampallona, S., Tibaldi, G., Kupelnick, B., & Munizza, C. (1999). Effectiveness of antidepressants; meta-analysis of dose-effect relationships in randomised clinical trials. *British Journal of Psychiatry*, 174, 297-303.
- Burg, W. van den (1998a). De (in)effectiviteit van psychofarmaca. *Directieve therapie (Dth)*, 18, 5-68.
- Burg, W. van den (1998b). De (in)effectiviteit van psychofarmaca; antwoord aan Kahn en Cath. *Directieve therapie (Dth)*, 18, 196-211.
- Burg, W. van den (2000). Over psychofarmaca, serotonine en de tucht van de markt. *Maandblad voor Geestelijke volksgezondheid (MGv)*, 55, 303-305.
- Cooney, J.M. (1998). Active placebos in antidepressant trials. *British Journal of Psychiatry*, 172, 541-542.
- Edlund, M.J., Craig, T.J., & Richardson, M.A. (1985). Informed consent as a form of volunteer bias. *American Journal of Psychiatry*, 142, 624-627.
- Even, C., Siobud-Dorocant, E., & Dardennes, R.M. (2000). Critical approach to antidepressant trials: blindness protection is necessary, feasible and measurable. *British Journal of Psychiatry*, 177, 47-51.
- Faden, R.R., & Beauchamp, T.L. (1986). *A History and Theory of Informed Consent*. New York: Oxford University Press.
- Fava, G.A. (1998). All our dreams are sold. *Psychotherapy and Psychosomatics*, 67, 191-193.
- Fisher, S., & Greenberg, R.P. (1993). How sound is the double-blind design for evaluating psychotropic drugs? *Journal of Nervous and Mental Disease*, 181, 345-350.
- Fisher, S., & Greenberg, R.P. (1997). The curse of the placebo: fanciful pursuit of a pure biological therapy. In S. Fisher & R.P. Greenberg (Eds.), *From placebo to panacea: putting psychiatric drugs to the test*. New York: Wiley, 1997.
- Gigerenzer, G., Swijtink, Z., Porter, Th., Daston, L. & Krüger, L. (1989). *The Empire of Chance*. Cambridge: Cambridge University Press.
- Greenberg, R.P., Bornstein, R.F., Zborowski, M.J., Fisher, S., & Greenberg, M.D. (1994). A meta-analysis of fluoxetine outcome in the treatment of depression. *Journal of Nervous and Mental Disease*, 182, 547-551.
- Healy, D. (1998). Commentary: meta-analysis of trials comparing antidepressants with active placebos. *British Journal of Psychiatry*, 172, 232-234.
- Heimberg, R.G. (1997) The use of pill placebo as a control strategy in the evaluation of psychotherapy: reply to Klein. *Treatment*, 1, comment 2. ([www.journals.apa.org/treatment](http://www.journals.apa.org/treatment)).
- Hill, A.B. (1990). Memories of the British streptomycin trial in tuberculosis. *Controlled Clinical Trials*, 11, 77-79.
- Hotopf, M., Churchill, R., & Lewis, G. (1999). Pragmatic randomized controlled trials in psychiatry. *British Journal of Psychiatry*, 175, 217-223.
- Hughes, J.R., Gulliver, S.B., Amori, G., Mireault, G.C., & Fenwick, J.F. (1989). Effect of instructions and nicotine on smoking cessation, withdrawal symptoms and self-administration of nicotine gum. *Psychopharmacology*, 99, 486-491.
- Joyce, C.R.B. (1982). Placebos and other comparative treatments. *British Journal of Clinical Pharmacology*, 13, 313-318.
- Khan, A., Warner, H.A., & Brown, W.A. (2000). Symptom reduction and suicide risk in patients treated with placebo in antidepressant clinical trials. *Archives of General Psychiatry*, 57, 311-317.

- Kienle, H. (1996). A critique of the double-blind clinical trial. Part 1. *Alternative Therapies in Health and Medicine*, 2, 74-80.
- Kirsch, I., & Rosadino, M.J. (1993). Do double-blind studies with informed consent yield externally valid results? *Psychopharmacology*, 110, 437-442.
- Kleijnen, J., Goetzsche, P., Kunz, R.A., Oxman, A.D., & Chalmers, I. (1997). So what's so special about randomization? In A. Maynard & I. Chalmers (Eds.), *Non-random Reflections on Health Services Research*. London: BMJ Publishing Group.
- Kramer, M.S., & Shapiro, S.H. (1984). Scientific challenges in the application of randomized trials. *Journal of the American Medical Association*, 252, 2739-2745.
- Licht, R.W., Gouliaev, G., Vestergaard, P., & Frydenberg, M. (1997). Generalisability of results from randomised drug trials. *British Journal of Psychiatry*, 170, 264-267.
- Moncrieff, J., Wessely, S., & Hardy, R. (1998). Meta-analysis of trials comparing antidepressants with active placebos. *British Journal of Psychiatry*, 172, 227-231.
- Mulrow, C.D., Williams, J.W., Trivedi, M., et al. (1999). *Treatment of Depression: Newer Pharmacotherapies*. Evidence Report/Technology Assessment No. 7, AHCPR Publication No. 99-E014. Rockville MD: Agency for Health Care Policy and Research. www.ahcpr.gov.
- Ney, P.G., Collins, C., & Spensor, C. (1986). Double blind: double talk or are there ways to do better research. *Medical Hypotheses*, 21, 119-126.
- Niklson, I.A., Reimitz, P.-E., & Sennef, C. (1997). Factors that influence the outcome of placebo-controlled antidepressant clinical trials. *Psychopharmacology Bulletin*, 33, 41-51.
- Ogundie, L.O., Boardman, A.P., & Masterton, A. (1999). Randomisation in clinical trials. *British Journal of Psychiatry*, 175, 581-584.
- Orne, M.T. (1969). Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R.L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 143-179). New York: Academic Press.
- Oxtoby, A., Jones, A., & Robinson, M. (1989). Is your 'double-blind' design truly double-blind? *British Journal of Psychiatry*, 155, 700-701.
- Paul, G.L., Tobias, L.L., & Holly, B.L. (1972). Maintenance psychotropic drugs in the presence of active treatment programs. *Archives of General Psychiatry*, 27, 106-115.
- Rickels, K. (1986). Use of placebo in clinical trials. *Psychopharmacology Bulletin*, 22, 19-24.
- Schultz, K.F. (2000). Assessing allocation concealment and blinding in randomized controlled trials: why bother? *Evidence Based Medicine*, 5, 36-37.
- Schweizer, E., & Rickels, K. (1997). Placebo response in generalized anxiety: its effect on the outcome of clinical trials. *Journal of Clinical Psychiatry*, 58 (suppl.11), 30-38.
- Silverman, W.A. (1995). Effectiveness, efficiency... and subjective choice. *Perspectives in Biology and Medicine*, 38, 480-495.
- Silverman, W.A., & Altman, D.G. (1996). Patients' preferences and randomised trials. *British Medical Journal*, 347, 171-174. Commentaar, 1118-1119.
- Skovlund, E. (1991). Should we tell trial patients that they might receive placebo? *Lancet*, 337, 1041.
- Spilker, B., & Cramer, J.A. (1992). *Patient recruitment in clinical trials*. New York: Raven Press.
- Stassen, H.H., Angst, J., & Delini-Stusa, A. (1996). Delayed onset of action of antidepressant drugs? Survey of results of Zürich meta-analyses. *Pharmacopsychiatry*, 29, 87-96.
- Swartzman, L.C., & Burkell, J. (1998). Expectations and the placebo effect in clinical drug trials: why we should not turn a blind eye to unblinding, and other cautionary notes. *Clinical Pharmacology & Therapeutics*, 64, 1-7.
- Thase, M.E. (1999). How should efficacy be evaluated in randomized clinical trials of treatments for depression? *Journal of Clinical Psychiatry*, 60 (suppl.4), 23-32.
- Thornley, B., & Adams, C. (1998). Content and quality of 2000 controlled trials in schizophrenia over 50 years. *British Medical Journal*, 317, 1181-1184.
- Trivedi, M.H., & Rush, A.J. (1994). Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*, 11, 33-43.
- Veatch, R.M. (1987). *The patient as partner*. Indiana University Press.
- Wechsler, H., Grosser, G.H., & Greenblatt, M. (1965). Research evaluating antidepressant medications on hospitalized mental patients: a survey of published reports during a five-year period. *Journal of Nervous and Mental Disease*, 141, 231-239.
- Wendt, G.R., & Cameron, J.S. (1961). Chemical studies of behavior: V. Procedures in drug experimentation with college students. *Journal of Psychology*, 51, 173-211.