

Streken en kuren va

W e kennen allemaal de enthousiaste verhalen van therapeuten, medicijnmannen, piskijkers, astrologen, homeopaten, iriscopisten, Jomanda en andere kwakzalvers over hun eigen kunsten, het succes van hun 'therapie' of de precisie van hun voorspellingen. Erger nog: velen van ons geloven hen. Zo bezien is het niet zo gek dat onderzoekers hebben geprobeerd therapieën en medicatie aan de strengst mogelijke onderzoekseisen te onderwerpen. De gedachte is: laat eerst maar eens zien wat je waard bent in strikt gecontroleerd onderzoek (die vorm van onderzoek wordt aangeduid met de naam *randomized controlled trials*, kortweg: RCT's). Dat is in een notendop de filosofie van *evidence-based medicine*, of ruimer *evidence-based practice*: alleen medicijnen, medische ingrepen en therapieën die hebben bewezen dat ze echt effect sorteren, komen als behandelingsoptie in aanmerking en worden vergoed. We geloven de arts of therapeut niet meer op zijn mooie praatjes, we willen eerst harde bewijzen zien dat iets werkt. Ook mensen zoals ik, die niet weglopen met de onbuigzame en misschien wel 'onmenselijke' – d.w.z. niet voor mensen ontworpen – natuurwetenschappelijke onderzoeksmethoden, moeten toegeven dat enig wantrouwen tegenover al te menselijke goedgelovigheid op zijn plaats is.



Beeld: Mark Quinn, *Emotional deox* (1995)

So far so good. Maar vervolgens is er iets misgegaan. In de wetenschap waar die strenge wetenschappelijke methoden het eerst zijn toegepast, de farmacolo-

gie, is het meteen grondig fout gelopen. Dat wil ik eerst laten zien. Vervolgens wil ik het hebben over de psychotherapie, de sector waarin ik werkzaam ben, zij het niet als therapeut. Ook daar mankeert nogal wat aan de schone schijn die de dompteurs van het *evidence*beest ons voorspiegelen.

Wie na het boek 'The constant gardener' van John Le Carré (of de film) wil weten hoe het nu écht zit met de weldaden en vooral de misdaden van de farmaceutische industrie, moet David Healy lezen. Zeker zo spannend. Hij onthult in zijn geschiedschrijving van de psychofarmacologie schokkende feiten, ook al zijn het geen moorden-met-voorbedachten-rade, zoals bij Le Carré. Healy was in eerdere geschriften nog betrekkelijk mild in zijn oordeel, maar in zijn magnum opus (Healy, 2002) heeft hij geen goed woord meer over voor de 'wetenschappelijke' methoden van de farmaceutische industrie. Kort samengevat: ooit werd de research gedaan door onafhankelijke onderzoekers, maar deze werden al spoedig vervangen door werknemers van het farmaceutische bedrijf zelf. Tegenwoordig worden vooraanstaande klinici ingehuurd om als schaamlap te dienen: zij tekenen voor onderzoeksverslagen die ze nooit gelezen, laat staan geschreven hebben. Als lezer ben je geneigd dat relaas met een flinke schep zout te nemen, totdat je de noot opzoekt die bij deze alinea hoort: 'Alle feiten in deze passage zijn gebaseerd op praktijken die ik niet alleen heb gezien, maar waaraan ik zelf heb deelgenomen' (Healy, 2002, noot 100, p. 438). Deze klokkenkluiders heeft eerst zelf vuile handen gemaakt en is vervolgens met slaande deuren vertrokken om de wereld rond te doen van alles wat er vies en voos is aan de pillenindustrie. Zijn boeken, vooral 'The creation of psychopharmacology', hebben hun uitwerking niet gemist: hij heeft velen wakker geschud en de farmaceutische industrie staat in een kwaaiere reuk dan ooit tevoren.

Volgens Healy bestaat er bijvoor-



beeld onderzoek 'dat laat zien dat persoonlijkheidstreken van patiënten voor 50 procent kunnen voorspellen hoe groot de kans is dat iemand reageert op antidepressiva (...) en ook op antipsychotica' (Healy, 2002, p. 314). Maar voor dat onderzoek heeft de farmaceutische industrie geen belangstelling: zij wil het toepassingsdomein van een middel niet inperken maar liefst zo omvangrijk mogelijk maken. Dat veel te veel patiënten psychofarmaca krijgen toegediend, is al een ramp, maar dat wetenschappelijk onderzoek naar het verband tussen stoornissen, geneesmiddelen en persoonlijkheidsstructuren niet van de grond komt, is zo mogelijk nog erger. RCT's kunnen wel degelijk zo worden uitgevoerd dat eruit blijkt welk percentage patiënten naar alle waarschijnlijkheid een bepaalde farmacotherapie 'heilzaam en aanvaardbaar' zal vinden. Maar zulk onderzoek wordt niet gedaan, stelt Healy vast en hij concludeert: 'Als wetenschap wil zeggen dat we *alle* mogelijke gegevens over een verschijnsel verzamelen, dan is de moderne psychiatrie ondanks haar retoriek allesbehalve wetenschappelijk' (Healy, 2002, p. 315).

Andere wanpraktijken die Healy constateert, zijn het *meer dan één keer* publiceren van positieve onderzoeksresultaten en het *niet* publiceren van negatieve resultaten. Wanneer die niet-gedeepte resultaten worden meegeteld, blijkt de identiteit van de sponsor die het onderzoek subsidieert de beste voorspeller van het onderzoeksresultaat te zijn. Als we Healy mogen gelo-

n het evidencebeest

ven staan de zaken er in de psychiatrie beroerd voor: hij komt tot een vernietigend oordeel over de psychofarmacologie en de biologische psychiatrie. 'Voeg bij de niet-publicatie van onderzoeksgegevens het feit dat het distributienetwerk van deze door *ghostwriters* geschreven artikelen veel beter gelied is dan de *reguliere* wetenschappelijke kanalen voor de verspreiding van *onafhankelijke* onderzoeksresultaten en het wordt duidelijk in welke mate de informatiebeheersing een industriële aangelegenheid is geworden. Waar de psychofarmacologische literatuur ooit het gezag uitstraalde van clinici die uit eigen waarneming wisten waarover zij het hadden, is aan het eind van de 20^{ste} eeuw een steeds groter deel van de onderzoeksliteratuur verworven tot het psychiatrische equivalent van een *Big Mac*' (Healy, 2002, pp. 310/11). Kortom: geen wetenschap, maar reclameslogans, commercie en propaganda, onder meer om de 'consument' rijp te maken voor de acceptatie van een *biologische* psychiatrie. Inmiddels zijn psychiatrische vakbladen dan ook verplicht om psychiaters en psychofarmacologen te vragen naar mogelijke belangenverstrengeling. 'Werkt u voor Eli Lilly? Dan raden we de lezer aan uw betoog niet voor zoete koek te slikken!' is de impliciete boodschap.



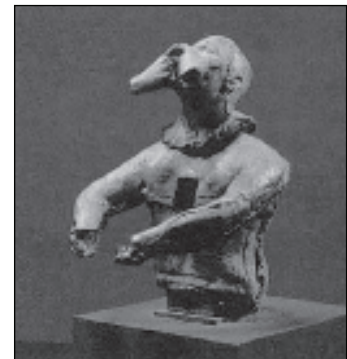
'Maar is dit alleen een probleem bij farmacotherapie?' vraagt Healy zich af. Het is een retorische vraag en het antwoord luidt: Nee, het geldt ook voor de psychotherapie. Ook daar staan grote financiële belangen op het spel, zij het

nog niet van een 'industrie' als de farmaceutische, maar wel van specialistische verenigingen en beroepsgroepen. Healy waarschuwt dan ook voor een al te simpele conclusie die de psychotherapie zou ontzien: 'Oppervlakkige lezing van deze geschiedenis (van de psychofarmacologie, ml.) suggereert wellicht dat dit het probleem is: een wetenschap die is gecorrumpeerd door het bedrijfsleven (de farmaceutische industrie, ml.). Maar het feit dat een soortgelijke dynamiek het domein van de psychotherapie heeft aangetast, waarschuwt tegen de al te gemakzuchtige aanvaarding van een dergelijke conclusie.' (Healy, 2002, p. 332).

In de psychotherapie is de eis van *evidence-based practice* pas veel later gesteld dan in de psychofarmacologie. Het zijn het vooral de gedragstherapeuten geweest die voorop hebben gelopen bij effect(iviteits)onderzoek van hun therapie. In 1953 joeg de overigens nogal omstreden psycholoog Hans Jürgen Eysenck in zijn boek 'Uses and abuses of psychology' de knuppel in het hoenderhok van psychologen, psychotherapeuten en vooral psychoanalytici (die toentertijd het psychotherapeutische landschap beheersten) door hen uit te dagen: het effect van psychotherapie zou volgens hem niet uitstijgen boven het 'natuurlijke' herstel dat patiënten laten zien die op een wachtlijst voor therapie staan. De gedragstherapeuten waren de eersten die de knuppel oppakten. Zij wilden zich niet bij dit vonnis neerleggen, temeer omdat zij altijd al hadden gepretendeerd dat hun therapie wetenschappelijk het best was onderbouwd.

In de jaren '70 toonden meta-analyses van Smith en Glass de gunstige effecten van psychotherapie aan en zij weerlegden daarmee de stelling van Eysenck. Toch bleef men in grote delen van de gezondheidszorg van mening dat psychologische therapieën voor bepaalde stoornissen ofwel domweg niet werkten of inferieur waren

aan farmacotherapie. In 1995 probeerde de *Division 12 (Clinical psychology) task force on promotion and dissemination of psychological procedures* van de *American psychological association* (APA) psychologische behandelingen te promoten door criteria te publiceren waarmee 'empirisch ondersteunde' therapieën voor bepaalde stoornissen kunnen worden vastgesteld. Deze werkgroep kwam met achttien behandelingen waarvoor de empirische steun in RCT's goed genoeg werd geacht. De meeste daarvan waren vormen van cognitieve gedragstherapie (CGT). Het rapport leidde niet alleen tot herwaar-



dering van en geestdrift voor psychologische therapieën, maar ook en vooral tot erg veel gekissebis. Enerzijds vergrootte het de erkenning van aantoonbaar effectieve psychotherapieën onder het grote publiek, bij beleidsmakers en in opleidingsprogramma's. Anderzijds wonden veel psychotherapeuten zich op over het feit dat de werkgroep zijn aandacht exclusief had gericht op korte, geprotocolleerde therapieën en op *specifieke* therapie-effecten in plaats van op de *gemeenschappelijke* factoren. Bovendien was lang niet iedereen onder de indruk van het bewijsmateriaal dat de 'taskforce' aandroeg voor de superioriteit van de empirisch ondersteunde therapieën.

Een vooraanstaand criticus van bovengenoemde werkgroep en van de geselecteerde en gepropageerde 'empirisch ondersteunde' psychotherapieën is Drew Westen, duizendpoot, veel-schrijver en psychoanalytisch georiën-



teerd therapeut en onderzoeker. Hij schetst de situatie aan het begin van de 21^e eeuw als volgt: 'De therapieën die wij meta-analytisch hebben onderzocht, zijn bewezen even effectief als of effectiever dan farmacotherapieën, die het voordeel hadden dat de farmaceutische industrie miljarden dollars kon steken in onderzoek en marketing ervan.' Maar dat is nog pas de helft van het verhaal: 'Met uitzondering van CGT voor paniek herstelde de meerderheid van de patiënten in de therapieën voor alle stoornissen die we hebben onderzocht *niet*' (Westen *et al.*, 2004, p. 650). Met andere woorden: kortdurende psychotherapieën doen de symptomen weliswaar snel afnemen, maar genezing of volledig herstel wordt niet bereikt.

De belangrijkste kritiek zal ik kort samenvatten. Het voornaamste kritiekpunt is dat veel onderzoek een *self-fulfilling prophecy* is. We worden, al dan niet opzettelijk, *bedrogen*. In de eerste plaats bepaalt de therapievoorkeur van de onderzoekers *welke* behandeling(en) worden onderzocht. Omdat de meeste onderzoekers CGT-*-minded* zijn, is CGT de meest onderzochte therapie. Hierboven heb ik de CGT-stroming als geheel nog om haar onderzoeksijver geprezen, maar, in de tweede plaats, worden ook de onderzoeksresultaten in hoge mate bepaald door de therapievoorkeur van de onderzoekers. Het gaat hier om de veelbesproken, maar zelden verdisconteerde 'loyaliteitseffecten' (*allegiance effects*). Lester Luborsky stelde in 1999 vast dat hij louter op grond van de therapievoorkeur van de onderzoeker(s) de uitkomsten van effectonderzoek waarin therapieën onderling op hun werkzaamheid werden vergeleken met vrij grote zekerheid kon voorspellen. Onlangs deed Luborsky dat onderzoek nog eens over (Luborsky & Barrett, 2005) en het resultaat was identiek: als je weet wie het

(effect)onderzoek doet, weet je ook wat eruit komt. Zoiets zet al die harde-feitenpretenties (niet alleen van *effect*-onderzoek trouwens) wel in een heel ander daglicht!

Hoe zetten onderzoekers de resultaten naar hun hand? Bijvoorbeeld door de succescijfers van hun therapie uit te drukken als percentage van het aantal patiënten dat de behandeling voltooit. Afvallers (*drop-outs*, hier kom ik nog op terug) en potentiële deelnemers van het onderzoek die om allerlei redenen als niet-buikbaar werden afgewezen, blijven buiten beschouwing. Dit flatteert de behandelresultaten van RCT's enorm. Zelf als de precieze aard van de succescijfers al in de publicatie wordt vermeld, dan nog kraait geen haan naar het hoe en waarom van zo veel afvallers.

Kwalijk is ook dat onderzoek dat geen significante uitkomsten oplevert, nauwelijks wordt gepubliceerd. Dat komt niet alleen omdat onderzoekers liever met positieve en spectaculaire resultaten op de proppen komen, maar ook omdat tijdschriftredacties vaak niet bereid zijn negatieve resultaten, die net zo informatief zijn als significante positieve uitkomsten, te publiceren. Daarom schetsen de uitkomsten van meta-analyses, voor zover het ongepubliceerde onderzoek buiten beschouwing blijft, systematisch een te rooskleurig beeld van de werkelijke situatie.



Er wordt in de psychotherapie enorm veel poeha gemaakt over wat de ene therapie van de andere onderscheidt, maar naar alle waarschijnlijkheid zijn de hoofdvormen van psychotherapie (cognitieve/gedragstherapeutische, clientgerichte/experientiele en psychoanalytische/psychodynamische) ongeveer even effectief, *mits* je het vooroordeel van de onderzoeker(s) neutraliseert. Toch zijn het vooral de ultrakorte cognitief-gedragstherapeutische behande-

lingen die de richtlijnen voor verschillende stoornissen domineren. Maar het is zacht gezegd lang niet zeker dat die korte, geprotocolleerde therapieën echt de beste therapieën zijn. Ik som kort wat redenen op waarom dat zelfs onwaarschijnlijk is:

1 Een protocol ontstaat als een min of meer toevallig tot stand gekomen, gestandaardiseerde versie van een behandeling en 'veel van de procedurele details van een therapiepakket worden tijdens het schrijven van het behandelprotocol noodgedwongen beslist op niet-empirische gronden. Moeten sessies een of twee uur duren? Wekelijks of tweewekelijks? Individueel of groep? Totaal 20 sessies, of 12 of 16?' (Haaga, 2004, p. 674). Die protocollering is nodig opdat de onderzoeker weet dat de therapeut doet wat hij wordt verondersteld te doen en dat alle aan het onderzoek deelnemende therapeuten *hetzelfde* doen. Wanneer een protocol 'empirisch ondersteund' is, betekent dat geenszins dat de specificaties ervan optimaal zijn (Haaga, 2004). Maar eenmaal effectief betoond, gaat het protocol een eigen leven leiden. Niemand twijfelt meer aan de (effectiviteit van de) ingrediënten van de therapie. Andere ingrediënten, die wellicht (nog) beter zijn, worden niet meer onderzocht.

2 RCT's verhullen meer dan ze onthullen. Bij een gebleken effect weet je slechts dat de behandeling voor de *proefgroep-als-geheel* werkt. Je weet niet *wat* er precies genezend werkt en evenmin bij *wie* het effect optreedt en bij wie niet. Het is dus geenszins zeker dat de richtlijn de beste therapie voorschrijft voor een *individueel* met al zijn persoonlijke kenmerken en eigenaardigheden.

3 De steekproef in een RCT is nooit klinisch representatief. Onderzocht worden stevast patiënten met één welomschreven stoornis-volgens-het-boekje (dat boekje is de DSM, de bijbel van de psychiatrische diagnoses). De selectie is erop gericht de homogeniteit van de diagnose te maximaliseren. Veel potentiële deelnemers aan het onderzoek vallen af, want de meeste patiënten hebben volgens de criteria van de DSM *meer* dan één stoornis. Als deze patiënten voor of

tijdens het onderzoek 'gedropt' worden – men spreekt van *drop-outs*, vandaar –, dan vermindert dat de generaliseerbaarheid naar de doelgroep, want die is 'zelden symptoomzuiver'. Bij dat drop-pen gaat het niet kinderachtig toe: 'In de gemiddelde RCT (...) wordt tussen eenderde en tweederde deel van de patiënten die zich aanmelden, uitgesloten van behandeling, en de categorieën patiënten die worden uitgesloten, lijken vaak niet alleen representatiever maar zijn (...) ook therapieresistenter gebleken' (Westen *et al.*, 2004, p. 658). De onderzochte en 'empirisch ondersteunde' therapieën lijken dus vooral te zijn toegesneden op de niet of nauwelijks bestaande *ideale* patiënt.



4 De meeste richtlijntherapieën zijn domweg te kort voor een goed resultaat, want de meeste stoornissen laten zich niet in drie maanden genezen. RCT's omvatten zelden *follow-up* onderzoek over een voldoende lange periode. Cognitieve gedragstherapie is tamelijk effectief als het gaat om onmiddellijke symptoomvermindering, maar het is *niet* zo dat de gemiddelde patiënt herstelt en hersteld *blijft* over langere tijd. De gerapporteerde resultaten betreffen de korte termijn en het is allesbehalve zeker dat ze gedurende langere perioden na het einde van de therapie overeind blijven.

5 In RCT's wordt veel te weinig gebruikgemaakt van relevante controlegroepen. Meestal bestaat de controlegroep uit een wachtlijst of geen behandeling. Dat is de vergelijking die Eysenck in 1953 eiste voor fatsoenlijk effectonderzoek. Maar het is niet genoeg, want inmiddels is ten overvloede aangetoond dat eigenlijk *alles* wat onderzoekers gedurende tien of twintig sessies met patiënten doen en waarvan ze geloven dat het zal werken tot betere resultaten leidt dan experimentele con-

ditie die zijn bedoeld om *niet* te werken. 'Het enige wat kan worden geconcludeerd uit zulk onderzoek is dat iets dat bedoeld is om effectief te zijn superieur is aan iets dat bedoeld is *niet* effectief te zijn' (Westen *et al.*, 2005, p. 428).

Als je de effectiviteit van psychotherapie *in het algemeen* wilt bewijzen, dan kies je voor een zwakke controle: een wachtlijst of geen behandeling. Maar als je de werkzaamheid en de superioriteit van een *specifieke* psychotherapeutische aanpak wilt bewijzen, dan is een wachtlijst of 'geen behandeling' of zelfs *treatment as usual* een zwakke bod. Westen adviseert onderzoekers te stoppen met iedere vorm van placebocontrole die bedoeld is *niet* te werken. 'Als het doel is *evidence-based practice* de weg te wijzen, zouden onderzoekers zich nog slechts moeten verlaten op twee typen controlecondities: (a) therapieën die worden uitgevoerd door *ervaren* klinici die proberen een *zo goed mogelijk* behandelresultaat te bereiken en die niet worden beperkt in de duur van de therapie door ofwel de onderzoekers ofwel subsidiënten, en (b) verschillende duren van dezelfde experimentele behandeling, waardoor niet alleen gecontroleerd wordt voor 'therapievoorkeur' (loyaliteit, *allegiance*) van de therapeut(en) en de onderzoeker(s), maar waardoor tevens empirische gegevens beschikbaar komen over de optimale therapieduur' (Westen *et al.*, 2005, p. 430).

De conclusie moet *niet* luiden: weg met alle *randomized controlled trials*. Maar RCT's moeten niet langer opgevat worden als de *enige* vorm van onderzoek en misschien evenmin als 'gouden standaard'. Daarvoor zijn ze te kwetsbaar gebleken voor misinterpretatie en misbruik. Er is in veel RCT's gesjoemeld en de invloed van *allegiance effects* op onderzoeksresultaten is verontrustend groot gebleken. Westen stelt daarom voor in ieder onderzoeksteam ten minste één potentiële dwarsligger op te nemen: een aanhanger van een andere 'therapieschool' of, in het geval van psychofarmaca, een onderzoeker die niet werkt voor de desbetreffende pillenfabrikant. Maar het zou natuurlijk nog mooier zijn als de onderzoekssubsidiënten dit als voorwaarde voor financiële steun zouden stellen. De redacties van wetenschappelijke tijdschriften moeten ook veeleisender worden voor

zij een onderzoeksverslag accepteren. Hetzelfde geldt voor instanties als de Amerikaanse *Food and drug administration* (FDA) of het Nederlandse College ter Beoordeling van Geneesmiddelen, die beslissen over de consequenties van het door de farmaceutische industrie uitgevoerde en gepresenteerde onderzoek, en daarmee over de vraag of een medicijn op de markt komt.

Wetenschap is een beetje de ideologie van onze tijd. Wat uit onderzoek blijkt, wordt als waarheid geaccepteerd. Ten onrechte (vaak). Vereist is dus: een kritische houding, en misschien zelfs regelrecht wantrouwen, tegenover de conclusies en verregaande pretenties van 'wetenschappelijke' theorieën en, vooral, *onderzoek*.

✍ Max Lauteslager

Max Lauteslager is psycholoog en werkt voor de Stichting Tijdschrift voor Psychotherapie. Correspondentieadres: maxlout@xs4all.nl

De hier afgedrukte tekst is een sterk ingekorte versie van mijn artikel 'Het evidence-beest heeft kuren. Gebruik en misbruik van EBP, RCT- en EST-methodologie' in het Tijdschrift voor Psychotherapie, 32 (2006), 347-366. Alle cursiveringen in citaten zijn van de auteur.

Literatuur

- Haaga, D.A.F. (2004). A healthy dose of criticism for randomized trials: Comment on Westen, Novotny, and Thompson-Brenner (2004). *Psychological Bulletin*, 130, 674-676.
- Healy, D. (2002). *The creation of psychopharmacology*. Cambridge: Harvard University Press.
- Luborsky, L.B., & Barrett, M.S. (2005). Theoretical allegiance. In J.C. Norcross, L.E. Beutler & R.F. Levant (Eds.), *Evidence-based practices in mental health. Debate and dialogue on the fundamental questions* (pp. 257-267). Washington DC: American Psychological Association.
- Vandereycken, W. (2006). 'Over bittere en vergulde pillen: psychiatrie in het licht (of de schaduw) van de farmaceutische industrie'. *Tijdschrift voor Psychiatrie*, 48, 119-129.
- Westen, D., Novotny, C.M., & Thompson-Brenner, H. (2004). The empirical status of empirical supported psychotherapies: Assumptions, findings and reporting in controlled clinical trials. *Psychological Bulletin*, 130, 631-663.
- Westen, D., Novotny, C.M., & Thompson-Brenner, H. (2005). EBP ≠ EST: Reply to Crits-Christoph et al. (2005) and Weisz et al. (2005). *Psychological Bulletin*, 131, 427-433.

Voetnoten bij dit stuk kunnen per e-mail worden aangevraagd bij de Deviant-redactie.