

De effectiviteit van SSRI's en andere antidepressiva

Randomized controlled trials bij depressieve patiënten

Sinds de ontdekking van de eerste antidepressiva, vijftig jaar geleden, zijn een paar duizend onderzoeken uitgevoerd naar de effectiviteit van deze middelen, soms in vergelijking met elkaar, soms ten opzichte van een placebo, een andersoortige behandeling of helemaal niets. Naar aanleiding daarvan zijn honderden overzichtsartikelen gepubliceerd. Van een betrekkelijk nieuwe vorm daarvan, de meta-analyse (waarbij de uitkomsten van vele individuele onderzoeken cijfermatig gecombineerd worden) vond Anderson er anno 2001 in de literatuur 108. Hij selecteerde er achttien, teneinde zo precies mogelijk te beoordelen in hoeverre de diverse oude en nieuwe antidepressiva, waaronder de SSRI's, in effectiviteit verschillen. Daarmee beoefende hij een nieuw genre studie: een soort meta-analyse van meta-analyses.

Ondanks de bijkans obsceen grote hoeveelheid onderzoek bestaan er nog vele onduidelijkheden over de effectiviteit van antidepressiva, en meta-analyses of meta-analyses van meta-analyses verhelpen die slechts zeer ten dele. Het belangrijkste probleem is dat wat in de literatuur verschijnt geen getrouwe afspiegeling vormt van wat het onderzoek heeft laten zien of wat fatsoenlijk onderzoek zou kunnen laten zien. De reden is dat de farmaceutische industrie grotendeels het onderzoek in handen heeft en grotendeels bepaalt wat en hoe er over haar producten geschreven wordt. Het gevolg is dat, zoals Richard Smith (2004) onlangs schreef bij zijn afscheid als hoofdredacteur van het *British Medical Journal*: '... the medical journals are more an extension of the marketing arm of pharmaceutical companies than independent scientific forums.' Massage, vervorming en achterhouding van de eigenlijke onderzoeksgegevens komen dan ook op grote schaal voor (bijvoorbeeld Melander e.a., 2003). Nogal wat literatuur is waarschijnlijk *ghostwritten*, dat wil zeggen geschreven door *medical*

VAN DEN BURG

communication agencies in opdracht van de farmaceutische bedrijven; daartoe uitgenodigde 'auteurs-onderzoekers', die gewoonlijk niet de eigenlijke ruwe onderzoeksgegevens gezien hebben, hoeven daaraan dan alleen maar hun naam te geven, meestal tegen fikse betaling (Healy, 2004; Healy & Cattell, 2003). Dehue gaat in dit nummer verder in op een aantal onsmakelijkheden van 'de marketing'.

Op de publicatie-*bias* kom ik later in dit artikel terug, maar centraal staan hier een aantal andere, vooral methodologische problemen die de interpretatie van 'de cijfers' betreffen. Het artikel is geschreven met een lezer voor ogen die weinig vertrouwd is met het gangbare onderzoek of met statistiek. De cijfers die ik geef, moeten als globale indicaties gezien worden van wat de gepubliceerde literatuur laat zien; het betreffen meestal resultaten van meta-analyses. Het uitgangspunt is dat alle gangbare antidepressiva, inclusief de SSRI's, gemiddeld ongeveer even effectief zijn – iets wat in grote lijnen algemeen aanvaard wordt. Voorzover dat mogelijk niet het geval is, bijvoorbeeld bij bepaalde subgroepen van depressies, zal dit in beperkte mate aan de orde komen.

Ik zal me beperken tot de effectiviteit van SSRI's zoals die bij depressies naar voren komt in *randomised controlled trials* (RCT's). Het gaat hier om de effectiviteit van wat wel de acute behandeling genoemd wordt. *At random* krijgen depressieve patiënten gedurende een aantal weken een antidepressivum of een placebo en vervolgens worden de resultaten vergeleken. Deze proefopzet geldt als de gouden standaard voor de vaststelling van de effectiviteit van een geneesmiddel.

Placebo-gecontroleerde RCT's

In de simpelste vorm van deze RCT zijn er twee groepen, de antidepressivum-groep en de placebogroep, maar soms hebben RCT's meer 'armen', bijvoorbeeld een placebo-arm en twee antidepressivum-armen. In tegenstelling tot vroeger worden de patiënten precies op de hoogte gebracht van de opzet van het onderzoek en weten ze dus dat ze een placebo kunnen krijgen; ze moeten dan voor hun deelname expliciet *informed consent* geven. Noch aan de deelnemers noch aan hun beoordelaars wordt gezegd wie wat gekregen heeft, zodat de RCT in opzet dubbelblind is. Meestal moeten deelnemers een 'major depression' hebben volgens het DSM-systeem en moeten ze een bepaalde ernst van depressiviteit bereikt hebben. Dit wordt meestal vastgesteld met de Hamilton-schaal voor depressies. Deze schaal speelt in meerdere opzichten een centrale rol in de RCT's (zie hieronder). Een veelgebruikte, korte versie met zeventien items staat in figuur 1 afgebeeld (in de vertaling van Dijkstra uit 1974); langere

DE EFFECTIVITEIT VAN SSRI'S EN ANDERE ANTIDEPRESSIVA

versies beslaan meestal 21 of 24 items. Op grond van een intake-interview beoordeelt een clinicus in hoeverre er sprake is van een gedeprimeerde stemming, schuldgevoelens et cetera, soms op een vijfpuntsschaal, soms op een driepuntsschaal. Is volgens deze beoordelaar de stemming uiterst gedeprimeerd, dan kan hij het meest rechtse alternatief aankruisen en wordt de patiënt voor dit item 4 punten toegerekend. Per item worden nadere omschrijvingen gegeven voor wanneer een 0, 1, 2, 3 of 4, dan wel 0, 1 of 2 te geven, maar het zal duidelijk zijn dat beoordelaars nog wel eens verschillend zullen scoren. Om deze verschillen zo klein mogelijk te houden worden beoordelaars vaak getraind vóór het begin van een RCT. De som van alle scores, de 'Hamilton-score', is de meest gebruikte maat voor de ernst van de depressie. Vaak dienen RCT deelnemers een Hamilton-score van minimaal 18 te hebben, maar soms is het criterium 20 (mede afhankelijk van de schaalversie) of 14. Behalve een te lage Hamilton-score zijn andere redenen voor uitsluiting meestal: zeer ernstige depressiviteit en suicidaliteit, psychotische symptomen, het gelijktijdig voorkomen van andere lichamelijke of psychische aandoeningen (comorbiditeit), een voorgeschiedenis van manische of hypomane episodes, de leeftijd (meestal worden alleen niet-oude volwassenen opgenomen) en verslaving aan alcohol of drugs.

237

Figuur 1 Depressie-beoordelingsschaal van Hamilton

1	Gedeprimeerde stemming	0 1 2 3 4
2	Schuldgevoelens	0 1 2 3 4
3	Suicide	0 1 2 3 4
4	Inslaapmoeilijkheden	0 1 2
5	Slapeloosheid	0 1 2
6	Vroeg ontwaken (doorslaapstoornis)	0 1 2
7	Werk en belangstelling	0 1 2 3 4
8	Geremdheid	0 1 2 3 4
9	Agitatie	0 1 2
10	Angst en spanning	0 1 2 3 4
11	Angstequivalenten	0 1 2 3 4
12	Somatische symptomen (gastrointestinaal)	0 1 2
13	Somatische symptomen (algemeen)	0 1 2
14	Seksualiteit	0 1 2
15	Hypochondrie	0 1 2 3 4
16	Gewichtsverlies	0 1 2
17	Inzicht	0 1 2

VAN DEN BURG

In toenemende mate worden, vooral in de Verenigde Staten, de RCT's uitgevoerd in commerciële onderzoeksinstellingen ('research mills'), worden de patiënten via advertenties of TV-spots geworven en wordt hun een beloning in het vooruitzicht gesteld, bijvoorbeeld geld of extra gezondheidszorg. Een RCT kan per patiënt vier weken duren en een enkele keer zestien weken, maar bedraagt meestal zes of zeven weken.

Drie uitkomstmaten

Percentages responders

Verreweg de meest gebruikte uitkomstmaat is het percentage *responders*. Het gaat dan om het aantal mensen van wie de Hamilton-score aan het eind van de RCT minstens gehalveerd is. De grootste en meest recente meta-analyse waarbij deze maat werd gebruikt, betrof 81 RCT's, bijna alle met een duur van zes tot acht weken, waarin SSRI's en andere moderne antidepressiva vergeleken werden met een placebo (Mulrow e.a., 1999). Meer dan tienduizend personen namen deel, met een gemiddelde Hamilton-score van 24. Het percentage *responders* na gebruik van het antidepressivum bedroeg gemiddeld 50% en na gebruik van een placebo 32%. Ook een andersoortige meta-analyse komt uit op 18% verschil (53% versus 35%; Bollini e.a., 1999). Aldus kan het verschil bijkijns de gepubliceerde literatuur begroot worden op een kleine 20%.

Een probleem is dat dit percentageverschil, klein als het is, als maat een bedrieglijke maat is en een te rooskleurig beeld kan geven van de effectiviteit van antidepressiva. De maat lijkt aan te geven dat men goed twee groepen patiënten kan onderscheiden, namelijk zij die wel en zij die niet goed gereageerd hebben op de behandeling. Met de bijmaak: zij die wel en niet van hun depressie afgeholpen zijn. Het criterium van halvering kan echter een zeer onbeduidend verschil in Hamilton-punten doen verkeren in forse percentageverschillen tussen de medicatie- en placebogroep, afhankelijk van het aantal personen dat eindscores om en nabij het halveringspunt heeft (Moncrieff, 2001). De verkeerde suggestie wordt nog sterker gewekt als het percentageverschil, zoals soms gebeurt, wordt omgezet in een andere maat: *number needed to treat* (NNT). Deze maat geeft aan hoeveel patiënten men moet behandelen om één extra responder te 'scoren'; alsof dit responderschap op zich een kostbare winst representeert.

Los van dit punt is het percentage responders meettheoretisch een merkwaardige maat. Impliciet wordt ervan uitgegaan dat een verbetering van 30 naar 15 punten even groot is als de verbetering van 20 naar 10 punten. Waarom dat het geval zou zijn, is bij mijn weten nooit duidelijk gemaakt. Overigens maakt dit idee deel uit van het moeilijke probleem in hoeverre de

DE EFFECTIVITEIT VAN SSRI'S EN ANDERE ANTIDEPRESSIVA

puntenaantallen goed corresponderen met zoiets als ernst van de depressie (zie ook hieronder). De Hamilton-schaal werd 45 jaar geleden door Max Hamilton min of meer met een timmermansoog geconstrueerd.

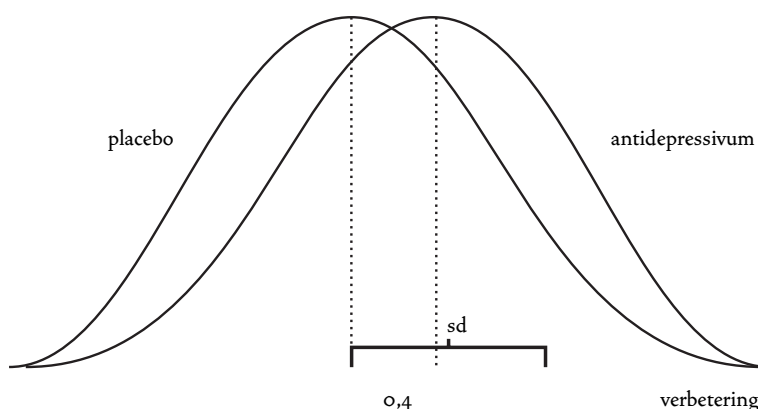
Effectgrootte

Een tweede maat die nogal eens gebruikt wordt voor meta-analytische berekeningen, is de effectgrootte. Hoewel de precieze berekeningswijze kan verschillen, gaat het hier naar de geest om de gemiddelde verbetering in de medicatiegroep minus de gemiddelde verbetering in de placebogroep, gedeeld door de standaarddeviatie van de verbeteringsscores (als maat voor de variabiliteit van deze scores). Verschillende meta-analyses komen op verschillende effectgroottes uit (Kirsch & Sapirstein, 1998), maar gemiddeld lijkt de effectgrootte ongeveer 0,4 te bedragen, de waarde die Anderson en Tomenson (1994) berekenden. Hoe groot dat is, is in figuur 2 zeer schematisch weergegeven: er is net gedaan alsof de verbeteringsscores een standaard-normale toevalsverdeling hebben, wat vaak niet zo'n goede weergave zal zijn van de werkelijkheid.

Een voordeel van deze maat is dat de uitkomsten ook uitgerekend kunnen worden voor andere schalen dan de Hamilton. Daardoor wordt bijvoorbeeld een vergelijking mogelijk tussen de effectiviteit zoals door de klinici-beoordelaars vastgelegd (bijvoorbeeld op de Hamiltonschaal) en de effectiviteit volgens de patiënten zelf op een zelfbeoordelingsschaal; daarvan zijn er diverse ontworpen. De resultaten zullen verderop aan de orde komen.

239

Figuur 2 Een effectgrootte van 0,4 (sd = standaarddeviatie)



VAN DEN BURG

Hamilton-punten

Omdat de Hamilton-schaal zo'n centrale plaats inneemt in het onderzoek en overal bekend is, is dit een zeer directe en eenvoudige maat. Het gaat hier gewoon om de vooruitgang in termen van het aantal Hamilton-punten in de antidepressivumgroep minus dat aantal punten in de placebogroep. Deze maat heeft weinig aandacht gekregen in RCT-rapportages. De tabel bevat gegevens die de producent van venlafaxine (Efexor) verstrekke (Entsuah e.a., 1995). Het verschil bedroeg 2,1 punten. Soortgelijke gegevens werden door de producent van fluoxetine (Prozac) gepubliceerd (Tollefson e.a., 1994; zie ook Van den Burg, 1998). Deze lieten een verschil van 2,4 punten zien, maar dan op een Hamilton-schaal met 21 items. Volgens beide artikelen waren de gegevens afkomstig van een aantal niet nader genoemde RCT's.

Tabel Afname van depressiviteit gedurende een zesweekse behandeling met venlafaxine (n=862) of placebo (n=491)

tijdstip	venlafaxine	placebo
aanvang	25.0	24.9
week 1	21.3	21.4
week 2	18.5	18.9
week 3	16.5	17.8
week 4	15.2	16.9
week 6	14.4	16.4

Gemiddelde scores op de Hamilton-schaal (Entsuah, Rudolph & Chitra, 1995)

Een recente studie betrof de resultaten van RCT's die de Amerikaanse Food and Drug Administration (FDA) had ontvangen ten behoeve van de registratie als antidepressivum van fluoxetine, paroxetine, sertraline, venlafaxine, nefazodon en citalopram (Kirsch e.a., 2002). Dankzij de recente Freedom of Information Act konden deze gegevens, die voordien geheim waren, opgevraagd worden. Men kan aannemen dat het probleem van de publicatie-*bias* hier een veel minder belangrijke rol speelt dan bij andere analyses. Het verschil tussen antidepressivum en placebo bedroeg hier gemiddeld 1,8 punten. Ter vergelijking: blijkens soortgelijke gegevens is de effectiviteit van de klassieke antidepressiva niet beduidend hoger. Storosum e.a. (2001) berekenden daarvoor een gemiddeld placebo-antidepressivumverschil van 2,7 punten uit 32 (drie-armige, zie boven) RCT's die aan het Nederlandse College ter Beoordeling

DE EFFECTIVITEIT VAN SSRI'S EN ANDERE ANTIDEPRESSIVA

van Geneesmiddelen (CBG) in de periode 1979-1991 waren overgelegd in het kader van aanvragen voor de registratie van nieuwe antidepressiva. (Waarschijnlijk mag het CBG, anders dan in de Verenigde Staten, nog geen mededelingen doen over de resultaten van meer recente RCT's.) In termen van het percentage responders bedroeg het verschil 14% en 11% voor de subgroep van 29 RCT's met een minimale duur van zes weken (Storosum e.a., 2004).

241

Verdere problemen

Het is duidelijk dat het placebo/SSRI-verschil zeer klein is. Uitgaande van de gepubliceerde literatuur, zou het verschil in percentage responders (zoals gezegd een effectiviteit majorerende maat) een kleine 20% bedragen. Zouden we dan de vuistregel van Thase (2002) toepassen ter verdiscontering van de publicatie-*bias* en dit verschil halveren, dan komen we net als Thase uit op 10%. Op grond van de gepubliceerde literatuur werd de effectgrootte op zo'n 0,4 begroot, maar is die volgens dezelfde vuistregel wellicht niet veel meer dan 0,2. Ervan uitgaande dat een dergelijke *bias* geen rol speelt bij de bij de FDA ingeleverde onderzoeken (dus dat alle RCT's daar adequaat zijn gerapporteerd), dan bedraagt het SSRI-placeboverschil minder dan 2 Hamilton-punten. Scherp gesteld betekent dit dat als een clinicus-beoordelaar voor een 'placebo-patiënt' bijvoorbeeld een 1 in plaats van een 2 zou invullen op het item 'agitatie' en bijvoorbeeld een 2 in plaats van een 3 op het item 'angst en spanning' (zie figuur 1), de gemiddelde winst van zes weken antidepressivum boven een placebo al bereikt is.

Hoewel het misschien overdreven lijkt bij zulke kleine verschillen, is het dienstig om nog nader in te gaan op een paar punten die voor een deel tot overschatting van de effectiviteit kunnen leiden:

Dubbelblindheid. In theorie weten noch de patiënten noch de clinicus-beoordelaars wie een placebo of het antidepressivum gekregen heeft. In de praktijk zijn er echter bijwerkingen en kunnen de patiënten ongewone lichamelijke sensaties ervaren, waardoor de beoordelaar vaak redelijk goed kan gokken wie in welke groep zit (Moncrieff, 2001). Als die beoordelaar dan groot vertrouwen heeft in het medicijn of graag het succes ervan bevestigd wil zien, kan dit makkelijk leiden tot overschatting van de effectiviteit. Het probleem is telkens en telkens weer gesignaleerd, maar erg veel aandacht bij de uitvoering van de RCT's heeft het niet gekregen. Zelden wordt gepoogd om door navraag, bijvoorbeeld na de eerste innames van de tabletten, er achter te komen in hoeverre de 'dubbelblindheid' nog intact is. Ook is zelden geprobeerd het probleem op te lossen door het gebruik van 'actieve placebo's': placebo's die bijwerkingen vertonen die vergelijkbaar zijn met die van het antidepressivum,

VAN DEN BURG

maar niet de veronderstelde, op de depressie werkende stof bevatten. Moncrieff, Wessely en Hardy (1998) vonden in de literatuur enkele zeer oude RCT's waarbij als placebo atropine gebruikt werd, een anticholinerge stof waardoor men net als door de klassieke antidepressiva een droge mond krijgt. Volgens een berekening was de effectgrootte in deze RCT's slechts 0,21, dus de helft van de in figuur 2 gehanteerde waarde van 0,4. (Deze berekening was echter voor discussie vatbaar.) Een andere bevinding die wijst op de mogelijke invloed van 'onblinding' kwam naar voren uit een meta-analyse van de RCT's waarin fluoxetine werd vergeleken met een placebo (Greenberg e.a., 1994). Er bleek een duidelijke samenhang te bestaan tussen de effectgrootte en de gerapporteerde hoeveelheid bijwerkingen. Zie (voor een nadere discussie, Van den Burg 2001). Het is heel goed mogelijk dat de grote variatie in de placebo - antidepressivumverschillen die de RCT's plegen te laten zien, voor een belangrijk deel zijn toe te schrijven aan de variabele effecten van onblinding.

De stakers. Een probleem met onduidelijke consequenties is hoe de gegevens in de effectiviteitsberekeningen meegenomen moeten worden van de ruim 30% van de patiënten die hun deelname aan een RCT vroegtijdig staken – blijkens 281 RCT's gemiddeld 11% vanwege bijwerkingen, 9% vanwege 'geen effectiviteit' en 11% om andere, meestal niet gespecificeerde redenen (Mulrow e.a., 1999). Eén mogelijkheid is dat men die helemaal niet meerekent; de berekeningen worden dan gebaseerd op de *completers only*. Een belangrijk methodologisch bezwaar is dan dat de basis van het experimentele onderzoek ondergraven wordt, namelijk de randomisatie waardoor generalisatie mogelijk is (althans voorzover de onderzoeksgroep een goede afspiegeling is van de groep waarvoor men de middelen in de praktijk wil gebruiken). Verder: wat zegt het überhaupt als duidelijke effectiviteit gesuggereerd wordt wanneer minder dan 50% van de patiënten de RCT afmaakt (zoals weleens voorkomt)? Meestal wordt daarom de voorkeur gegeven aan een LOCF-oplossing (*last observation carried forward*). Dan wordt de laatste score van een staker, mits die een minimumperiode aan het onderzoek heeft deelgenomen, als eindscore beschouwd. Onduidelijk is in hoeverre, door de bank genomen, de twee berekeningswijzen tot verschillende conclusies leiden. Soms wordt gevonden dat LOCF-analyses een kleinere effectiviteit laten zien dan *completers only* analyses, maar het omgekeerde kwam naar voren uit het FDA-bestand dat Kirsch e.a. (2002) analyseerden. (Voor een voorbeeld hoe de LOCF-procedure door feitelijke niet-blindheid kan leiden tot overschatting van de effectiviteit, zie Van den Burg, 2001.)

Zelfbeoordelingen. De mening van de patiënten zelf wordt meestal niet gevraagd, althans niet gepubliceerd. En de enige meta-analyse waarin de effectgrootte volgens zelfbeoordelingen vergeleken werd met de effectgrootte volgens de clinicus/beoordelaar, leverde op dat deze veel kleiner was, in feite

nihil (Greenberg e.a., 1992). Er is wat voor te zeggen om niet al te veel te varen op zelfbeoordelingen, omdat sommige minder geletterde personen misschien niet goed overweg kunnen met de zelfbeoordelingsschalen en omdat iemand anders wellicht beter kan beoordelen hoe het met je gaat dan jijzelf, maar deze uitkomst versterkt het vertrouwen in de effectiviteit natuurlijk niet. Overigens is een gangbaar commentaar op zo'n bevinding, namelijk dat de invloed van de placebopil zó groot is dat de effectiviteit van het antidepressivum niet meer te zien is, weinig plausibel (Van den Burg, 2003). Door de *informed consent*-procedure weet men dat men een placebo kan krijgen en als men dan door lichamelijke gewaarwordingen zeker weet dat men *the real thing* gekregen heeft, kan juist dit tot optimisme leiden. Het placebo-effect werkt dan dus niet via de placebo maar via het antidepressivum.

243

Hamilton-scores. Een fundamenteel maar moeilijk probleem betreft de vergelijkbaarheid van verschillen van Hamilton-scores. Hierboven wees ik al op de eigenaardige impliciete aanname, bij de definitie van responders, dat die van 30 punten naar 15 punten even groot is als de afname van 20 naar 10. Maar er zijn ook andersoortige interpretatieproblemen. Duidt het door Storosum e.a. (2001) gevonden verschil van 2,7 punten erop dat klassieke antidepressiva toch iets effectiever zijn dan SSRI's, die slechts 1,8 punten verschil te zien gaven ten opzichte van een placebo? Afgezien van de futiliteit van de vraag en afgezien van kwesties of het misschien om verschillende typen patiënten ging en of die 0,9 punt verschil statistisch significant is, kan mijns inziens die vraag niet goed beantwoord worden. De reden is de toch tamelijk willekeurige wijze waarop de Hamilton-schaal is samengesteld. Weliswaar dekken de items redelijk goed wat als typerend voor depressiviteit beschouwd wordt, maar het onderlinge gewicht dat aan de items gegeven zou moeten worden, staat geenszins vast, en een belangrijk symptoom als concentratiestoornissen ontbreekt. Een relatief groot gewicht wordt gegeven aan slaapproblemen. Als dan de klassieke middelen door de bank genomen wat meer de slaap bevorderen dan SSRI's, wat naar alle waarschijnlijkheid het geval is, zou dit bijvoorbeeld een reden kunnen zijn waarom deze middelen er op een schaal als de Hamilton beter uitkomen. Dit punt roept overigens de meer algemene vraag op, of niet elk middel dat een beetje de slaap bevordert of sdeert als 'antidepressivum' op de markt gebracht kan worden, na de benodigde RCTs met schalen als de Hamilton als criterium.

Generaliseerbaarheid. Zoals gezegd, moet men gewoonlijk, om in aanmerking te komen voor deelname aan een RCT, op zijn minst tamelijk depressief zijn, bijvoorbeeld een Hamilton-score van 18 of meer hebben, en gelden verder allerlei andere uitsluitingscriteria. Maar aan die voorwaarden voldoet slechts een klein percentage van de depressieve patiënten die men in de praktijk tegenkomt; in een onderzoek van Zimmerman, Mattia en Posternak

VAN DEN BURG

(2002) slechts 8%. Geldt enige effectiviteit (voor zover daar sprake van is) ook voor andere groepen? Dit is in het bijzonder relevant gezien de overtuiging, ook wel gesteund door onderzoek (Khan e.a. 2002), dat antidepressiva bij de lichtere gevallen beduidend minder effectief zijn. De gemiddelde Hamilton-score van de patiënten waarop de meta-analyse van Mulrow e.a. (1999) betrekking had, bedroeg 24, maar de honderdduizenden die tegenwoordig SSRI's voorgeschreven krijgen (op geleide van de marketingboodschap dat depressies te weinig onderkend en behandeld worden), zullen in merendeel bij de lichtere gevallen horen. Des te kleiner zal het verschil met een placebo in deze groepen zijn.

Discussie en conclusie

Wie de stelling wil verdedigen dat volgens de gouden standaard van de RCT, de effectiviteit van SSRI's en andere antidepressiva bij depressieve patiënten gemiddeld praktisch nihil is, heeft een sterke zaak – zeker voor wat betreft de lichtere gevallen. Betreft men daarbij dat steeds duidelijker wordt dat veel gebruikers grote moeite hebben om weer van deze middelen af te komen als zij dat willen en dat de ongewenste effecten vaak hinderlijk en soms zeer ernstig zijn (zie ook Bos & Vriens, in dit nummer), dan zijn aan deze gegevens weinig argumenten te ontleen voor het voorschrijven ervan. De artikelen in dit nummer van Treffers en Rinne voor wat betreft kinderen, en van Bijl en Toenders voor wat betreft ouderen, hebben dezelfde teneur.

Niettemin geloof ik, maar uitsluitend op de wankelende basis van ervaringen van mensen die ik ken, dat de middelen soms behulpzaam zijn. De vraag is dan welk gewenst effect ze kunnen hebben en waarom dat niet of nauwelijks op de standaard-beoordelingsschalen tot uiting komt. Zelf heb ik de indruk dat één gewenst effect precies in het verlengde ligt van een belangrijke psychologische reden om het gebruik weer te staken (Bolling & Kohlenberg, 2004): minder makkelijk overweldigd worden door emotionaliteit in het contact met anderen, met als negatief pendant: verlies van 'affectabiliteit'. Helaas zijn deze vragen nauwelijks onderwerp van onderzoek. De onderzoeksagenda wordt bijna geheel bepaald door de commerciële belangen van de industrie, die er in geslaagd is, met steun van daaraan gelieerde kopstukken uit de wereld van de psychofarmacologie, honderdduizenden Nederlanders langdurig antidepressiva te laten gebruiken. Er is weinig reden om te veronderstellen dat dit de volksgezondheid ten goede is gekomen.

☞

DE EFFECTIVITEIT VAN SSRI'S EN ANDERE ANTIDEPRESSIVA

Literatuur

- Anderson, I.M. (2001). Meta-analytical studies on new antidepressants. *British Medical Bulletin*, 57, 161-178.
- Anderson, I.M., & Tomenson, B.M. (1994). The efficacy of selective serotonin reuptake inhibitors in depression: a meta-analysis of studies against tricyclic antidepressants. *Journal of Psychopharmacology*, 8, 238-249.
- Bijl, D., & Toenders, W.G.M. (2005) Antidepressiva bij ouderen. *Maandblad Geestelijke volksgezondheid*, 60, 311-324.
- Bolling, M.Y., & Kohlenberg, R.J. (2004). Reasons for quitting serotonin reuptake inhibitor therapy: paradoxical psychological side effects and patient satisfaction. *Psychotherapy and Psychosomatics*, 73, 380-385.
- Bollini, P., Pampallona, S., Tibaldi, G., Kupelnick, B. & Munizza, C. (1999). Effectiveness of antidepressants. *British Journal of Psychiatry*, 174, 297-303.
- Burg, W. van den (1998). De (in)effectiviteit van psychofarmaca. *Directieve Therapie*, 18, 5-68.
- Burg, W. van den (2001). De rol van 'informed consent' in placebo-gecontroleerde, dubbelblinde, gerandomiseerde onderzoeken naar de effectiviteit van psychofarmaca. *Directieve Therapie*, 21, 368-383.
- Burg, W. van den (2003). Placebo-effecten: het belang van de definitie en de 'informed consent' procedure. *De Psycholoog*, 38, 649-654.
- Dehue, T. (2005). Verdienen aan neerslachtigheid. *Maandblad Geestelijke volksgezondheid*, 60, 223-234.
- Entsuah, A.R., Rudolph, R.L. & Chitra, R. (1995). Effectiveness of venlafaxine treatment in a broad spectrum of depressed patients: a meta-analysis. *Psychopharmacology Bulletin*, 31, 759-766.
- Greenberg, R.P., Bornstein, R.F., Greenberg, M.D. & Fisher, S (1992). A meta-analysis of antidepressant outcome under 'blinder' conditions. *Journal of Consulting and Clinical Psychology*, 60, 664-669.
- 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. (2004) Shaping the intimate: influences on the experience of everyday nerves. *Social Studies of Science*, 34, 219-245.
- Healy, D, Cattell D (2003) Interface between authorship, industry and science in the domain of therapeutics. *British Journal of Psychiatry*, 183, 22-27.
- Khan, A, Leventhal, R.M, Khan, S.R. & Brown, W.A. (2002) Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration Database. *Journal of Clinical Psychopharmacology*, 22, 40-45.
- Kirsch, I., & Sapirstein, G. (1998) Listening to Prozac but hearing placebo: a meta-analysis of antidepressant medication. *Prevention & Treatment*, 1, article 0002a. www.journals.apa.org/prevention.
- Kirsch, I., Moore, T.J., Scoboria, A. & Nicholls, S.S. (2002). The emperor's new drugs: an analysis of antidepressant medication data submitted to the U.S. Food and Drug administration. *Prevention & Treatment*, 5, article 23. www.journals.apa.org/prevention.
- Menander, H, Ahlqvist-Rastad, J., Meijer, G. & Beermann, B. (2003). Evidence b(i)ased medicine - selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *British Medical Journal*, 326, 1171-1175.
- Moncrieff, J. (2001). Are antidepressants overrated? A review of methodological problems in antidepressant trials. *Journal of Nervous and Mental Disease*, 189, 288-298.
- 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., e.a. (1999). *Treatment of depression: Newer pharmacotherapies*. Rockville, MD: Agency for Health Care Policy and Research. (www.ahrq.gov/clinic/epcsums/deprsumm.htm).
- Smith, R. (2004). Travelling but never arriving: reflections of a retiring editor. *British Medical Journal*, 323, 588.
- Storosum, J.G., Elferink, A.J.A., Van Zwieten, B.J., Van den Brink, W, Gersons, B.P.R. Van Strik, R. & Broekmans A.W. (2001). Short-term efficacy of tricyclic antidepressants revisited: a meta-analytic study. *European Neuropsychopharmacology*, 11, 173-180.

VAN DEN BURG

- Storosum, J.G., Elferink, A.J.A., Van Zwieten, B.J., Van den Brink, W. & Huyser, J. (2004). Natural course and placebo response in short-term, placebo-controlled studies in major depression: a meta-analysis of published and non-published studies. *Pharmacopsychiatry*, 37, 32-36.
- Thase, M.E. (2002). Antidepressant effects: the suit may be small, but the fabric is real. *Prevention & Treatment*, 5, article 32. www.journals.apa.org/prevention.
- Tollefson, G.D., Holman, S.L., Saylor, M.E. & Potvin, J.H. (1994). Fluoxetine, placebo, and tricyclic antidepressants in major depression with and without anxious features. *Journal of Clinical Psychiatry*, 55, 55-59.
- Treffers, P.D.A., & Rinne, M. (2005). SSRI's bij kinderen en adolescenten (2). Werkzaamheid. *Maandblad Geestelijke volksgezondheid*, 60, 260-275.
- Zimmerman, M., Mattia, J.I. & Posternak, M.A. (2002) Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *American Journal of Psychiatry*, 159, 469-473.

Summary

Burg, W. van den
 'Efficacy of SSRI's and other antidepressants. Randomised controlled trials on depressed patients'
 Randomised studies of depressed patients have found SSRI's and other modern antidepressants to be little more effective than placebo. The article examines three outcome measures used in such trials and discusses a number of problems inherent in the research.

Personalia

Dr W. van den Burg (1944), psycholoog, is als docent/onderzoeker verbonden aan de Afdeling Neuropsychologie van de vakgroep Psychologie van de Rijksuniversiteit Groningen. In 1993 promoveerde hij aan dezelfde universiteit op het proefschrift *In search of biological determinants of depressions*. Eerder verschenen van hem in dit tijdschrift onder meer de artikelen 'De werkzaamheid van antidepressiva: blikvernuwing in de biologische psychiatrie' (MGv 94-11) en 'Biologische psychiatrie, de mode en de pretenties. Over R. Kahn's Gids pillen psychiatrie' (MGv 98-7/8). Hij is sinds 1997 redacteur van dit tijdschrift.