

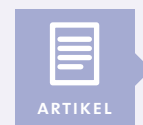
# Evidence-b(i)ased psychiatrie: onderzoek naar bewijs voor effectiviteit en veiligheid van antidepressiva bij depressie en angststoornissen<sup>1</sup>

Y.A. DE VRIES, A.M. ROEST, P. DE JONGE

- ACHTERGROND** Antidepressiva blijven controversieel, onder andere vanwege beschuldigingen dat teleurstellende resultaten zijn achtergehouden en vanwege de bescheiden gemiddelde effectiviteit.
- DOEL** Onderzoeken van bias in de onderzoeksliteratuur over antidepressiva en de mogelijkheden om te voorspellen welke patiënten met depressie of angststoornissen wél veel baat hebben bij antidepressiva.
- METHODE** We hebben bias onderzocht door informatie van de Amerikaanse geneesmiddelenautoriteit (FDA) te vergelijken met de gepubliceerde literatuur. Om respons te voorspellen werden patiëntengegevens uit gerandomiseerde studies gebruikt.
- RESULTATEN** Van alle studies naar antidepressiva voor depressie of angststoornissen was respectievelijk 50% en 72% positief, tegenover 95% en 96% van alle gepubliceerde studies. Veiligheidsuitkomsten werden slecht gerapporteerd in gepubliceerde artikelen en ongepubliceerde studies werden vaak 'gebundeld' in pooled-trialspublicaties met positieve conclusies. Wat respons op antidepressiva betreft vonden wij een associatie tussen symptoomernst en effectiviteit bij sommige, maar niet alle angststoornissen; er is eerder inconsistent bewijs gevonden voor deze associatie bij depressie. Verder hadden patiënten met een depressie die al snel iets herstelden vaker een goede respons, ongeacht welke symptomen verminderden.
- CONCLUSIE** Deze bevindingen tonen aan dat de onderzoeksliteratuur over antidepressiva sterk vertekend is door bias. Ernst en vroeg herstel voorspelden een goede respons, maar meer informatie is nodig om betere voorspellingen te doen. De toegenomen beschikbaarheid van patiëntengegevens maakt dit hopelijk binnenkort mogelijk.

TIJDSCHRIFT VOOR PSYCHIATRIE 61(2019)9, 635-643

**TREFWOORDEN** angststoornissen, antidepressiva, bias, depressie.



ARTIKEL



Depressies en angststoornissen zijn veelvoorkomende psychische stoornissen, met jaarlijkse prevalenties van respectievelijk 5,2% en 10,1% (de Graaf e.a. 2012). Hoewel er duidelijke verschillen bestaan tussen deze stoornissen, bijvoorbeeld wat betreft symptomen en ontstaansleeftijd (Kessler e.a. 2005), komen ze ook veel samen voor, wat wellicht een gedeelde kwetsbaarheid reflecteert (Krueger

1999). Bovendien is de behandeling vergelijkbaar, aangezien de belangrijkste behandelopties voor zowel depressies als angststoornissen bestaan uit antidepressiva en/of psychotherapie (Spijker e.a. 2013; van Balkom e.a. 2013). Het gebruik van antidepressiva is de afgelopen decennia toegenomen (Verweij & Houben-van Hertzen 2013), maar er is ook de nodige controverse geweest over deze medicij-

nen en over de vraag of het toegenomen gebruik wel gepast is, onder andere vanwege beschuldigingen dat farmaceutische bedrijven tegenvallende resultaten verborgen hebben gehouden. In 2008 is een invloedrijke meta-analyse van studies naar antidepressiva voor de behandeling van depressie gepubliceerd (Turner e.a. 2008). Hieruit bleek dat maar 51% van *alle* studies positief was, dat wil zeggen dat het antidepressivum een statistisch significant beter resultaat gaf dan een placebo. Werd daarentegen alleen naar de *gepubliceerde* literatuur gekeken, dan leek het alsof 94% van alle studies positief was. Het effect van antidepressiva was hierdoor met 32% overschat.

Het niet publiceren van volledige studies is slechts één van de mechanismen waardoor de onderzoeksliteratuur vertekend kan raken. Meer recent is er ook aandacht gekomen voor andere vormen van *bias*.

- *Outcome reporting bias* treedt op als uitkomsten die wel in het protocol stonden niet in het artikel worden vermeld (of andersom), of als de status van primaire en secundaire uitkomsten wordt omgedraaid, zodat een negatieve studie opeens positief wordt.

- We spreken van *spin* als de uitkomsten wél correct worden vermeld in de resultatensectie van een artikel, maar de auteurs desondanks concluderen dat de interventie effectief was, bijvoorbeeld door secundaire uitkomsten te benadrukken.

- Bias kan ook optreden doordat negatieve studies minder opvallen dan positieve studies, bijvoorbeeld doordat ze minder geciteerd worden (*citatiebias*).

Er is nog veel onbekend over de precieze effecten van bias op de antidepressivaliteratuur. Zo weten we veel minder over bias in de literatuur over angststoornissen, de invloed van bias op de schijnbare veiligheid, en de invloed van subtielere vormen van bias zoals spin of citatiebias.

Daarnaast vormen de honderden gerandomiseerde studies die in de afgelopen decennia gedaan zijn een schat van informatie die nog onvolledig benut is. We weten dat de gemiddelde effectiviteit van antidepressiva bescheiden is, maar sommige patiënten hebben waarschijnlijk wél veel baat bij antidepressiva, terwijl anderen er helemaal niet door herstellen. De gegevens uit deze gerandomiseerde studies kunnen gebruikt worden om voorspellers voor een geslaagde behandeling te zoeken.

In dit artikel bespreken wij een aantal studies uit het proefschrift van de eerste auteur waarin wij een antwoord proberen te geven op twee vragen. Ten eerste, wat is het effect van bias geweest op de antidepressivaliteratuur bij depressie en bij angststoornissen? Ten tweede, kunnen we aan de hand van symptomen patiënten met een depressie of een angststoornis identificeren die baat zullen hebben bij antidepressiva?

## AUTEURS

**YMKJE ANNA DE VRIES**, postdoctoraal onderzoeker, Ontwikkelingspsychologie, afd. Psychologie, Rijksuniversiteit Groningen, en Interdisciplinair Centrum Psychopathologie en Emotieregulatie, Universitair Medisch Centrum Groningen.

**ANNELIEKE M. ROEST**, universitair docent, Ontwikkelingspsychologie, afd. Psychologie, Rijksuniversiteit Groningen, en Interdisciplinair Centrum Psychopathologie en Emotieregulatie, Universitair Medisch Centrum Groningen.

**PETER DE JONGE**, hoogleraar Ontwikkelingspsychologie, afd. Psychologie, Rijksuniversiteit Groningen, en Interdisciplinair Centrum Psychopathologie en Emotieregulatie, Universitair Medisch Centrum Groningen.

## CORRESPONDENTIEADRES

Dr. Ymkje Anna de Vries, Grote Kruisstraat 2/1, 9712 TS Groningen.  
E-mail: ya.de.vries@rug.nl

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 16-5-2019.

## METHODE

Voor ons onderzoek naar bias hebben wij gebruikgemaakt van reviews van de Amerikaanse Food and Drug Administration (FDA) (Turner 2004). Wanneer farmaceutische bedrijven een medicijn voor een bepaalde ziekte op de Amerikaanse markt willen brengen, moeten zij de effectiviteit en veiligheid aantonen met gerandomiseerde studies. Het protocol voor deze studies moeten zij vóór het starten van de studie delen met de FDA en de resultaten worden later door FDA-clinici en -statistici geanalyseerd. Daardoor kunnen deze reviews gebruikt worden als een registratiesysteem en resultatendatabase.

Wij hebben reviews aangevraagd voor alle antidepressiva die sinds 1987 zijn goedgekeurd voor de behandeling van depressie of angststoornissen. Voor depressie ging het om selectieve serotonineheropnameremmers (SSRI's, bijv. paroxetine), serotonine-noradrenalineheropnameremmers (SNRI's, bijv. venlafaxine) en een aantal andere antidepressiva (bijv. mirtazapine), voor angststoornissen alleen om SSRI's en SNRI's.

Uit deze reviews hebben we informatie geëxtraheerd over alle gerandomiseerde, placebogecontroleerde studies die voor de goedkeuring van een medicijn waren uitgevoerd.

Studies worden gezien als positief als er volgens de FDA een statistisch significant ( $p < 0,05$ ) resultaat was op de primaire uitkomst(en), volgens de analyse zoals die gespecificeerd was in het protocol van de studie. De primaire uitkomst was meestal de score op een stoornisspecifieke vragenlijst (bijvoorbeeld de *Hamilton Anxiety Rating Scale* voor gegeneraliseerde angststoornis); voor paniekstoornis was de primaire uitkomst gewoonlijk het aantal paniekaanvallen.

Vervolgens hebben we in de literatuur gezocht naar deze trials, die bijvoorbeeld aan de studieduur en de steekproefgrootte te herkennen zijn. Door deze twee bronnen van informatie (FDA en gepubliceerde literatuur) te vergelijken, kunnen we bias detecteren. We hebben daarnaast voor beide bronnen een *standardized mean difference* (Hedges'  $g$ ) berekend als effectgrootte.

Om onderzoek te doen naar het voorspellen van respons hebben we individuele patiëntgegevens aangevraagd via Clinical Study Data Request, een *data-sharing platform*. Deze gegevens hebben we gebruikt om statistische modellen te bouwen die aan de hand van informatie over symptomen een goede respons op antidepressiva voorspellen. Als statistische modellen hebben we *mixed-effects (generalized) lineaire modellen* en *least absolute shrinkage and selection operator (LASSO)*-regressiemodellen gebruikt.

## METHODE

### Effect van bias op de antidepressivaliteratuur

#### EFFECTIVITEIT VAN ANTI-DEPRESSIVA

In navolging van het eerdere onderzoek bij depressie (Turner e.a. 2008), besproken in de introductie, hebben we allereerst gekeken naar het effect van bias op de schijnbare effectiviteit van antidepressiva voor angststoornissen (Roest e.a. 2015). De FDA-reviews bespraken 57 antidepressivastudies voor gegeneraliseerde angststoornis, sociale fobie, dwangstoornis, posttraumatische stressstoornis (PTSS) en paniekstoornis, waarvan er 41 (72%) een statistisch significant resultaat hadden op de primaire uitkomst(en) en 16 niet. Slechts 9 van de 16 negatieve studies waren gepubliceerd (tegenover 40 van de 41 positieve studies). Daarvan bleken er bovendien drie met outcome reporting bias gepubliceerd te zijn en nog eens drie andere met spin. Daardoor had uiteindelijk maar liefst 96% van de gepubliceerde artikelen een positieve conclusie. De effectgrootte volgens de gepubliceerde literatuur was echter niet statistisch significant groter dan die volgens de FDA (0,38 vs. 0,33).

Daarna zijn we teruggekeerd naar het onderzoek over antidepressiva bij depressie. Veel onderzoek naar bias tot nu toe bestudeert namelijk slechts één of enkele vormen

van bias, en zo ook het oorspronkelijke onderzoek van Turner e.a., waarin bijvoorbeeld alleen naar studiepublicatiebias en outcome reporting bias is gekeken. Wij hebben daarom gekeken naar het cumulatieve effect van zowel deze vormen van bias als spin en citatiebias binnen de literatuur over depressie (de Vries e.a. 2018a). Daarnaast hebben we ook vier nieuwe antidepressiva geïncludeerd die recent door de FDA zijn goedgekeurd voor depressie (vilazodon, desvenlafaxine, levomilnacipran, en vortioxetine; alleen vortioxetine is ook in Nederland te verkrijgen), zodat we konden zien of de prevalentie van bias aan het veranderen is.

Hierdoor hadden we uiteindelijk een cohort van 105 antidepressivastudies, waarvan er 52 negatief waren volgens de FDA (FIGUUR 1A). Bijna alle positieve studies (52 van de 53) waren gepubliceerd, tegenover maar 25 van de 52 negatieve studies. Slechts 77 studies waren dus gepubliceerd, waarvan 25 negatief waren volgens de FDA (FIGUUR 1B).

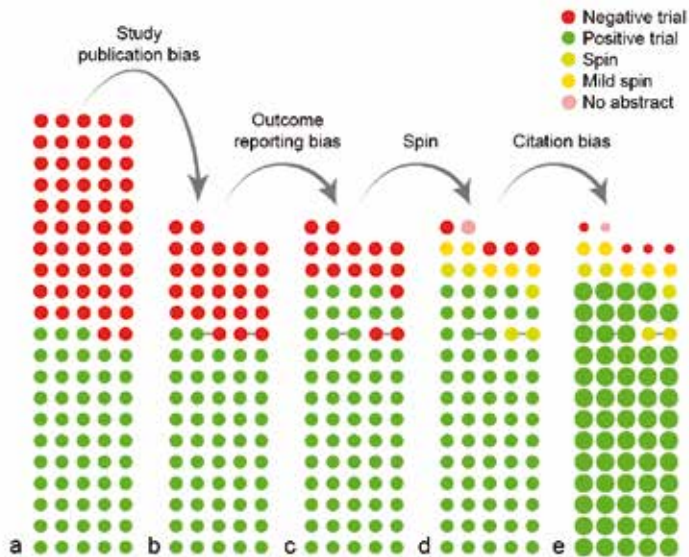
10 van deze 25 negatieve studies werden in de gepubliceerde literatuur echter opeens positieve studies, dankzij outcome reporting bias (FIGUUR 1C). Van de 15 resterende negatieve trials bevatten de meeste duidelijke spin of mildere spin (bijvoorbeeld door te concluderen dat het antidepressivum in ieder geval wel numeriek beter was dan placebo); slechts 4 van de 77 gepubliceerde studies rapporteerden duidelijk dat het antidepressivum niet effectief bleken was in die studie (FIGUUR 1D).

Studies met positieve resultaten bleken ook nog eens drie keer zo vaak geciteerd te worden als studies met negatieve resultaten (FIGUUR 1E). Wanneer we alle vormen van bias optellen, zijn de meeste negatieve resultaten dus überhaupt niet terug te vinden in de gepubliceerde literatuur, en de weinige negatieve resultaten die wel gepubliceerd worden, zijn dankzij spin en citatiebias moeilijk te vinden. Er waren echter ook positieve bevindingen: alle studies die sinds 2004 gedaan zijn (voor nieuwe antidepressiva) bleken gepubliceerd te zijn, ook al waren 9 van deze recente studies negatief. Bij de studies voor de twee meest recent goedgekeurde antidepressiva (levomilnacipran en vortioxetine) was daarnaast ook geen sprake van outcome reporting bias of spin.

#### VEILIGHEID VAN ANTI-DEPRESSIVA

We hebben ook het effect van bias op de schijnbare veiligheid van antidepressiva voor zowel depressie als angststoornissen bestudeerd (de Vries e.a. 2016). Hierbij hebben we specifiek gekeken naar staken van de behandeling en *serious adverse events (SAE's)*, dat wil zeggen gebeurtenissen met ernstige gevolgen zoals overlijden of ziekenhuisopname (wegens lichamelijke of psychische problemen). Omdat we wat bias in het rapporteren van veiligheidsuitkomsten betreft geen grote verschillen verwachtten tussen

**FIGUUR 1** Het oorspronkelijke, complete cohort van studies (a), en het cumulatieve effect van bias (b-e): elke cirkel staat voor een studie; cirkels die met een grijze streep aan elkaar verbonden zijn, staan voor studies die gebundeld gepubliceerd zijn. In figuur 1e staat de diameter van de cirkel voor het (relatieve) gemiddelde aantal citaties voor die categorie studies (overgenomen uit De Vries e.a. (2018a))



studies voor depressie en studies voor angststoornissen, hebben we deze studies samengenomen.

We konden in totaal 133 studies includeren, waarvan 102 gepubliceerd waren (in 97 artikelen). We vonden dat de kans op uitval correct werd weergegeven in de gepubliceerde literatuur. Volgens de FDA staakte zowel in de placebo- als in de antidepressivagroep gemiddeld ongeveer 30% van de participanten voortijdig de medicatie. In de placebogroep hield gemiddeld ongeveer 5% op vanwege bijwerkingen, tegenover 12% in de antidepressivagroep. De gemiddelde cijfers op basis van de 97 gepubliceerde artikelen kwamen hiermee overeen.

SAE's werden in de meeste gepubliceerde artikelen (63%) überhaupt niet vermeld. Wanneer er wel informatie gegeven werd, was die informatie vaak niet compleet of kwamen de gegevens niet overeen met die van de FDA. Er werd bijvoorbeeld niet gezegd hoe vaak zulke gebeurtenissen voorkwamen in de placebogroep, waardoor niet te beoordelen is of er zorgelijk veel SAE's in de antidepressivumgroep voorkwamen. Soms werden ook alleen SAE's vermeld waarvan de onderzoekers dachten dat ze mogelijk aan het medicijn gerelateerd waren, maar dit oordeel is subjectief en mag geen reden zijn om zorgelijke gebeurtenissen niet te vermelden. Deze problemen leiden er dus toe dat de gepubliceerde literatuur geen goede bron is voor informatie over potentieel ernstige bijwerkingen van antidepressiva.

#### IS TE VEEL PUBLICEREN OOK EEN PROBLEEM?

Wat betreft bias gaat het meestal om het *niet* publiceren van resultaten. Farmaceutische bedrijven publiceren echter ook vaak gebundelde publicaties (*pooled-trials publications*), waarin meerdere studies samen geanalyseerd worden. Deze publicaties worden gewoonlijk niet, of alleen onder strenge voorwaarden, meegeteld als het gaat over de vraag of studies gepubliceerd zijn, uit zorg dat deze gebundelde publicaties vatbaar zijn voor bias. Het is bijvoorbeeld vaak onduidelijk waarom de geïncludeerde studies geselecteerd zijn. Gebundelde publicaties kunnen echter ook waardevolle informatie bevatten, bijvoorbeeld over zeldzame bijwerkingen. Wij hebben daarom onderzocht of de ongepubliceerde antidepressivastudies voor depressie wél in gebundelde publicaties opgenomen zijn, en zo ja, wat de kenmerken van deze publicaties zijn (de Vries e.a. 2018b). Van de 105 studies met antidepressiva voor depressie die wij hadden geïncludeerd in ons eerdere onderzoek naar cumulatieve bias waren er 32 niet als zelfstandig artikel gepubliceerd (4 van deze studies hadden we in ons eerdere onderzoek naar cumulatieve bias gezien als wél gepubliceerd, omdat wij voor dat onderzoek onder bepaalde voorwaarden ook gebundelde publicaties als publicatie accepteerden. In dit onderzoek beschouwden we deze studies als ongepubliceerd).

23 van deze 32 studies waren echter wél opgenomen in minstens één van de in totaal 107 gebundelde publicaties die wij gevonden hebben. Slechts 13 van deze publicaties

gingen over de primaire onderzoeksvraag van de geïncludeerde studies (d.w.z. effectiviteit ten opzichte van placebo) en maar vier presenteerden ook de resultaten voor de individuele studies. Hoewel bijna al deze gebundelde publicaties deels of zelfs volledig bestonden uit negatieve studies, hadden maar 5 publicaties een ontmoedigende conclusie; in alle gevallen kwam dat doordat het antidepressivum meer bijwerkingen gaf dan placebo. Deze gebundelde publicaties leken soms ook overbodig: zo waren er maar liefst 8 publicaties die de effectiviteit van venlafaxine vergeleken met die van SSR's.

Aangezien de resultaten van de individuele studies zelden gepresenteerd werden, lijken gebundelde publicaties vooral een manier te zijn om de literatuur te overspoelen met positief gekleurde artikelen. We kunnen dus concluderen dat niet alleen te weinig, maar ook te veel publiceren tot een vertekende literatuur kan leiden.

## Voorspellen effect

### VOORSPELLEN VÓÓR DE START VAN DE BEHANDELING (BIJ ANGSTSTOORNISSEN)

Kunnen we aan de hand van symptomen identificeren wie er baat zullen hebben bij antidepressiva? Eerder onderzoek suggereerde dat patiënten met een lichte depressie ongeveer net zoveel baat hebben bij een placebo als bij een antidepressivum (Kirsch e.a. 2008; Fournier e.a. 2010). Meer recent onderzoek vond daarentegen geen relatie tussen depressie-ernst en de effectiviteit van antidepressiva (Gibbons e.a. 2012; Rabinowitz e.a. 2016; Furukawa e.a. 2018). Wat angststoornissen betreft, is nog nauwelijks onderzoek gedaan naar de rol van de ernst van de stoornis. Wij hebben daarom gegevens aangevraagd van 8979 patiënten uit 29 studies met antidepressiva voor gegeneraliseerde angststoornis, sociale angst, dwangstoornis, PTSS of paniekstoornis (de Vries e.a. 2018c). Voor gegeneraliseerde angststoornis vonden we een relatie tussen ernst en effectiviteit: voor patiënten met relatief lichte klachten was de effectgrootte van antidepressiva ongeveer 0,25, terwijl dit voor mensen met ernstige klachten bijna verdubbelde naar ongeveer 0,45. Ook voor paniekstoornis leek er een verband te bestaan tussen ernst en effectiviteit. Het antidepressivum-placeboverschil was slechts 0,2 à 0,3 paniekaanvallen per 2 weken voor patiënten die voor de behandeling  $\leq 6$  paniekaanvallen per 2 weken hadden. Voor patiënten die bij aanvang 30 paniekaanvallen per 2 weken hadden, was het antidepressivum-placeboverschil ruim 2 paniekaanvallen per 2 weken, en voor patiënten met nog ernstigere klachten was het verschil nog groter. Voor sociale angst, dwangstoornis en PTSS vonden we echter geen relatie tussen ernst en de effectiviteit van antidepressiva.

Deze resultaten suggereren dat de balans tussen effectiviteit en bijwerkingen nadeliger is voor patiënten met relatief lichte klachten van gegeneraliseerde angststoornis of paniekstoornis, maar dat deze balans voor sociale angst, dwangstoornis en PTSS gelijk blijft. Er kunnen echter andere overwegingen zijn om ook voor deze angststoornissen terughoudend te zijn met antidepressiva bij minder ernstige klachten, bijvoorbeeld het verwachte beloop van de stoornis of de effectiviteit van andere behandelingen.

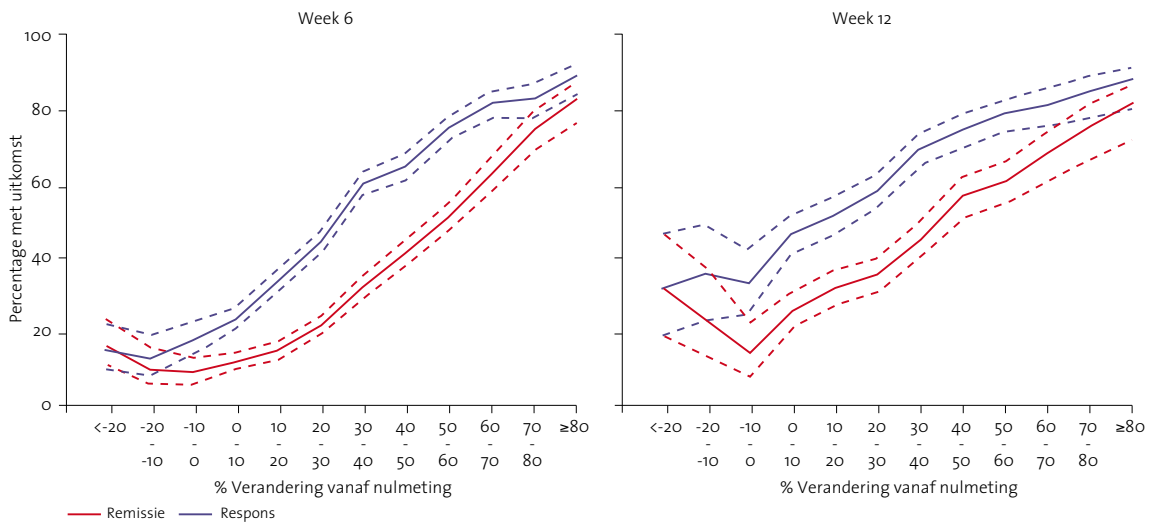
### VOORSPELLEN VROEG NÁ DE START VAN DE BEHANDELING (BIJ DEPRESSIE)

Omdat het nog moeilijk is om respons te voorspellen vóór de start van de behandeling, wilden wij ook kijken of het mogelijk is om deze vroeg in de behandeling te voorspellen. We hebben ons hierbij gericht op depressie. De huidige depressierichtlijnen bevelen aan om het effect na 4 tot 6 weken te beoordelen, maar op groepsniveau is al binnen een week een effect van antidepressiva op depressie zichtbaar (Taylor e.a. 2006). Een lichte verbetering ( $\geq 20\%$ ) binnen de eerste twee weken voorspelt een goede respons, en sommige studies suggereren dat de kans op een goede respons voor patiënten die geen vroege verbetering laten zien zo klein is dat de behandeling aangepast moet worden, terwijl andere studies dat tegenspreken (Uher e.a. 2011; Gorwood e.a. 2013).

De depressie-totaalscore kan echter betekenisvolle variatie in individuele symptomen verhullen (Fried & Nesse 2015) en daarom hebben wij onderzocht of we beter kunnen voorspellen door ook naar individuele symptomen te kijken (de Vries e.a. 2019). We hebben gebruikgemaakt van de *Hamilton Depression Rating Scale* (HAM-D)-scores van 8.242 patiënten uit 30 depressiestudies om respons ( $\geq 50\%$  verbetering) en remissie na 6 en 12 weken behandeling te voorspellen. Daarvoor hebben we drie modellen vergeleken: een *baseline model*, met alleen de totaalscore en de individuele HAM-D-items bij de nulmeting; een *total improvement model*, met daarnaast ook vroege verbetering op de totaalscore ( $\geq 20\%$  na 2 weken); en een *item improvement model*, met ook nog eens vroege verbetering in individuele HAM-D-items.

We vonden dat toevoeging van individuele symptomen inderdaad leidde tot iets betere voorspellingen, maar het verschil tussen de modellen was niet groot: de *area under the curve* (AUC) was bijvoorbeeld 0,73 voor het *total improvement model* en 0,77 voor het *item improvement model* voor respons op week 6, waarbij een AUC van 0,5 aangeeft dat het model niet beter voorspelt dan kans en een AUC van 1 dat het model perfect voorspelt. Door een variabele te dichotomiseren raak je bovendien veel informatie kwijt, dus we hebben ook gekeken naar vroege verbetering als continue variabele (verschil ten opzichte van de nulmeting). Toen

**FIGUUR 2** Kans op respons of remissie na 6 (links) of 12 (rechts) weken behandeling aan de hand van de procentuele verbetering ten opzichte van de nulmeting na 2 weken behandeling. De onderbroken lijnen geven het 95%-betrouwbaarheidsinterval aan (overgenomen uit De Vries e.a. (2019))



hadden individuele symptomen geen toegevoegde waarde meer.

Het lijkt er dus op dat vroeg herstel een goede uitkomst voorspelt, ongeacht welke specifieke symptomen verminderen. We vonden echter ook dat patiënten zonder vroege verbetering nog redelijk vaak goede uitkomsten hadden, zeker na 12 weken behandeling. **FIGUUR 2** laat de kans op respons en remissie na 6 en 12 weken behandeling zien aan de hand van de procentuele verbetering binnen de eerste 2 weken. Zelfs patiënten die eerst verslechteren, hebben nog bijna 40% kans op respons na 12 weken.

Het lijkt dus toch wat voorbarig om al na 2 weken de farmacotherapeutische behandelstrategie voor depressie aan te passen, des te meer omdat er weinig bewijs is voor de effectiviteit van dosisverhoging, zeker van SSRIs (Dold e.a. 2017), of voor switchen naar een ander antidepressivum (Bschor e.a. 2018), ook niet bij zeer vroege non-respons (Tadic e.a. 2016). Mensen zonder vroege verbetering vormen echter zeker een risicogroep waarbij intensiever contact en monitoring, en wellicht toevoeging van psychotherapie, aangewezen zijn.

## DISCUSSIE

### Bias in de antidepressivaliteratuur

Het besproken onderzoek heeft aangetoond dat er verschillende vormen van bias voorkomen in de antidepressivaliteratuur, die het moeilijk maken om baten en risico's goed te evalueren. Hoewel studiepublicatiebias al ruim een halve eeuw bekend is, worden er pas in de laatste decennia pogingen gedaan deze tegen te gaan. Het belangrijkste middel hierbij is verplichte trialregistratie. In 2005 heeft

het International Committee of Medical Journal Editors (ICMJE) prospectieve registratie van klinische trials als voorwaarde gesteld om in bij het ICMJE aangesloten tijdschriften te publiceren. In 2007 is in Amerika de FDA Amendments Act (FDAAA) aangenomen, die het wettelijk verplicht stelt om klinische trials die onder FDA-jurisdicte vallen te registreren.

Onze bevinding dat er aanzienlijk minder bias voorkomt bij recente studies voor nieuwe antidepressiva suggereert dat deze registratieverplichting, of wellicht de bredere aandacht voor bias, effect heeft op publicatiepraktijken. De wetenschappelijke literatuur zal dus in de toekomst hopelijk betrouwbaarder zijn dan voorheen, maar de bestaande literatuur wordt zelden gecorrigeerd. Omdat de 'oude' antidepressiva nog veel gebruikt worden, blijft dit dus een probleem, al kan dit hopelijk enigszins gecompenseerd worden doordat het ware verhaal nu ook naar buiten is gebracht.

Trialregistratie kan echter niet alle vormen van bias tegengaan. Zo zullen citatiebias en gebundelde-publicatiebias waarschijnlijk blijven bestaan, en blijft het bijvoorbeeld mogelijk om negatieve studies in minder bekende tijdschriften geplaatst te krijgen. Een mogelijke andere optie, die verder gaat dan trialregistratie en die zou kunnen voorkomen dat negatieve studies in minder bekende tijdschriften belanden, is het invoeren van *registered reports*, waarbij een tijdschrift een studie accepteert voor publicatie op basis van de introductie en methode, vóórdat de studie uitgevoerd is en de resultaten bekend zijn.

Ook meta-analyses, die niet beïnvloed (zouden moeten) worden door bijvoorbeeld spin of citatiebias, zijn een belangrijk hulpmiddel voor het tegengaan van dit soort

bias, maar de interpretatie van meta-analyses is niet 100% objectief. Bovendien zit er een vertraging van enkele jaren tussen publicatie van de oorspronkelijke trials en publicatie van (onafhankelijke) meta-analyses, juist in die cruciale periode wanneer een nieuw medicijn marktaandeel moet veroveren. De betrouwbaarheid van meta-analyses van secundaire uitkomsten of veiligheidsuitkomsten wordt ook nog altijd sterk bedreigd door selectieve publicatie van bemoedigende resultaten.

## Beperkingen

Een beperking van ons onderzoek is dat wij alleen hebben kunnen kijken naar studies die gesponsord zijn door de farmaceutische industrie; het is mogelijk dat er minder bias voorkomt in academische studies, waar financiële motieven een kleinere rol spelen. Hoewel recente academische studies meestal prospectief geregistreerd zijn, geldt dit niet voor oudere studies. Het is daarom erg moeilijk om onderzoek te doen naar bias in deze studies.

Een andere beperking is dat wij geen onderzoek hebben kunnen doen naar bias bij oudere antidepressiva, zoals de tricyclische antidepressiva, omdat deze te lang geleden zijn goedgekeurd.

## Voorspellen van respons

Het is duidelijk geworden dat er meer informatie nodig is dan alleen over symptomen om echt goed te voorspellen wie er zal reageren op antidepressiva. Denk aan ziektegeschiedenis, traumatische ervaringen, of psychosociale omstandigheden, maar ook aan invasievere, biologische informatie zoals biomarkers of hersenscans. Daarnaast is meer onderzoek nodig naar het voorspellen van respons op de ene specifieke behandeling versus de andere, bijvoorbeeld antidepressiva versus cognitieve gedragstherapie.

Dit soort onderzoek is nog relatief zeldzaam en wordt ook vaak in kleine steekproeven uitgevoerd, waardoor de betrouwbaarheid en generaliseerbaarheid twijfelachtig zijn.

De hoop is dat we uiteindelijk zulke predictiemodellen kunnen gebruiken om voor iedere patiënt een behandeling te kunnen kiezen die bijna zeker zal werken. Maar waarschijnlijk zijn er ook patiënten die op geen enkele beschikbare behandeling goed reageren. Wanneer predictiemodellen ver genoeg zijn ontwikkeld, kunnen we ze daarom wellicht ook gebruiken om specifieke patiëntensubgroepen te identificeren waarvoor nieuwe behandelingen noodzakelijk zijn.

## CONCLUSIES

Uit de studies die we in dit artikel hebben besproken, blijkt dat de antidepressivaliteratuur ernstig vertekend is door bias. Er is enige reden tot optimisme, omdat negatieve studies tegenwoordig veel vaker wél netjes gepubliceerd worden, maar het blijft belangrijk om als lezer kritisch te zijn en je bij elk artikel af te vragen of auteurs aan de resultaten konden 'sleutelen' of daarin begrensd werden door bijvoorbeeld een prospectieve registratie. Het voorspellen van respons op antidepressiva blijft voorlopig moeilijk, maar door de betere beschikbaarheid van individuele patiëntengegevens uit gerandomiseerde studies kan hierin hopelijk in de komende jaren vooruitgang worden geboekt.

## NOOT

1 Dit artikel is gebaseerd op het proefschrift van de eerste auteur, *'Evidence-b(i)ased psychiatry'* (Groningen: Rijksuniversiteit Groningen; 2018), te downloaden via: <http://hdl.handle.net/11370/bb50a2eb-15f1-4676-b3fo-e3a4726fcd6>

## LITERATUUR

- Balkom AJLM van, van Vliet IM, Emmelkamp PMG, Bockting CLH, Spijker J, Hermens MLM, e.a. Multidisciplinaire richtlijn angststoornissen (3e revisie). Richtlijn voor de diagnostiek, behandeling en begeleiding van volwassen patiënten met een angststoornis. Utrecht: Trimbos-instituut; 2013.
- Bschor T, Kern H, Henssler J, Baethge C. Switching the antidepressant after nonresponse in adults with major depression. *J Clin Psychiatry* 2018; 79: 11-8.
- Dold M, Bartova L, Rupprecht R, Kasper S. Dose escalation of antidepressants in unipolar depression: a meta-analysis of double-blind, randomized controlled trials. *Psychother Psychosom* 2017; 86: 283-91.
- Fournier JC, Derubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC, e.a. Antidepressant drug effects and depression severity. *JAMA* 2010; 303: 47-53.
- Fried EI, Nesse RM. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Med* 2015; 13: 72.
- Furukawa TA, Maruo K, Noma H, Tanaka S, Imai H, Shinohara K, e.a. Initial severity of major depression and efficacy of new generation antidepressants: individual participant data meta-analysis. *Acta Psychiatr Scand* 2018; 137: 450-8.
- Gibbons RD, Hur K, Brown CH, Davis JM, Mann JJ. Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. *Arch Gen Psychiatry* 2012; 69: 572-9.

- Gorwood P, Bayle F, Vaiva G, Courtet P, Corruble E, Llorca P-M. Is it worth assessing progress as early as week 2 to adapt antidepressive treatment strategy? Results from a study on agomelatine and a global meta-analysis. *Eur Psychiatry* 2013; 28: 362-71.
- Graaf R de, ten Have M, van Gool C, van Dorsselaer S. Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. *Soc Psychiatry Psychiatr Epidemiol* 2012; 47: 203-13.
- Kessler RC, Berglund PA, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005; 62: 593-603.
- Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLOS Med* 2008; 5: e45.
- Krueger RF. The structure of common mental disorders. *Arch Gen Psychiatry* 1999; 56: 921-6.
- Rabinowitz J, Werbeloff N, Mandel FS, Menard F, Marangell L, Kapur S. Initial depression severity and response to antidepressants v. placebo: patient-level data analysis from 34 randomised controlled trials. *Br J Psychiatry* 2016; 209: 427-8.
- Roest AM, de Jonge P, Williams CD, de Vries YA, Schoevers RA, Turner EH. Reporting bias in clinical trials investigating the efficacy of second-generation antidepressants in the treatment of anxiety disorders: a report of 2 meta-analyses. *JAMA Psychiatry* 2015; 72: 500-10.
- Spijker J, Bockting CLH, Meeuwissen JAC, van Vliet IM, Emmelkamp PMG, Hermens MLM, e.a. Multidisciplinaire richtlijn depressie (3e revisie). Utrecht: Trimbos-instituut; 2013.
- Tadic A, Wachtlin D, Berger M, Braus DF, van Calker D, Dahmen N, e.a. Randomized controlled study of early medication change for non-improvers to antidepressant therapy in major depression-The EMC trial. *Eur Neuropsychopharmacol* 2016; 26: 705-16.
- Taylor MJ, Freemantle N, Geddes JR, Bhagwagar Z. Early onset of selective serotonin reuptake inhibitor antidepressant action: a systematic review and meta-analysis. *Arch Gen Psychiatry* 2006; 63: 1217-23.
- Turner EH. A taxpayer-funded clinical trials registry and results database. *PLOS Med* 2004; 1: e60.
- Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008; 358: 252-60.
- Uher R, Mors O, Rietschel M, Rajewska-Rager A, Petrovic A, Zobel A, e.a. Early and delayed onset of response to antidepressants in individual trajectories of change during treatment of major depression: A secondary analysis of data from the genome-based therapeutic drugs for depression (GENDEP) study. *J Clin Psychiatry* 2011; 72: 1478-84.
- Verweij G, Houben-van Hertem M. Bevolkingstrends 2013: depressiviteit en antidepressiva in Nederland. Den Haag/Heerlen: CBS; 2013.
- Vries YA de, Roest AM, Beijers L, Turner EH, de Jonge P. Bias in the reporting of harms in clinical trials of second-generation antidepressants for depression and anxiety: A meta-analysis. *Eur Neuropsychopharmacol* 2016; 26: 1752-9.
- Vries YA de, Roest AM, de Jonge P, Cuijpers P, Munafò MR, Bastiaansen JA. The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: the case of depression. *Psychol Med* 2018a; 48: 2453-5.
- Vries YA de, Roest AM, Turner EH, de Jonge P. Hiding negative trials by pooling them: a secondary analysis of pooled-trials publication bias in FDA-registered antidepressant trials. *Psychol Med* 2018b: 1-7.
- Vries YA de, Roest AM, Burgerhof JGM, de Jonge P. Initial severity and antidepressant efficacy for anxiety disorders, obsessive-compulsive disorder, and posttraumatic stress disorder: An individual patient data meta-analysis. *Depress Anxiety* 2018c; 35: 515-22.
- Vries YA de, Roest AM, Bos EH, Burgerhof JGM, van Loo HM, de Jonge P. Predicting antidepressant response by monitoring early improvement of individual symptoms of depression: individual patient data meta-analysis. *Br J Psychiatry* 2019; 214: 4-10.



## SUMMARY

# Evidence-b(i)ased psychiatry: research into the evidence for effectiveness and safety of antidepressants in depression and anxiety disorders

Y.A. DE VRIES, A.M. ROEST, P. DE JONGE

- BACKGROUND** Antidepressants remain controversial, partly due to allegations that disappointing results were buried and because of their modest average efficacy.
- AIM** To investigate bias in the antidepressant literature and the possibilities for predicting which patients with depression or anxiety do receive significant benefits from antidepressants.
- METHOD** We investigated bias by comparing information from the US Food and Drug Administration with the published literature. To predict response, we used patient data from randomized trials.
- RESULTS** Of all studies on depression or anxiety, 50% and 72% were positive, compared to 95% and 96% of all published studies. Safety outcomes were poorly reported in published articles and unpublished studies were often 'bundled' into pooled-trials publications with positive conclusions. We found an association between severity and antidepressant efficacy for some, but not all, anxiety disorders; previous research has found inconsistent evidence for this association for depression. Furthermore, patients with depression that showed early improvement were more likely to attain a good response, irrespective of which symptoms improved.
- CONCLUSION** These results demonstrate the severe impact of bias on the antidepressant literature. Severity and early improvement predicted a good response, but more information is needed to improve predictions. The increased accessibility of individual patient data will hopefully soon enable further progress in this area.

TIJDSCHRIFT VOOR PSYCHIATRIE 61(2019)9, 635-643

**KEY WORDS** antidepressants, anxiety disorders, bias, depression