

Evaluatie van een behandeling en betrouwbaarheid van de meetinstrumenten

of hoe een depressie behandelen
zonder te weten hoe deze te meten...

Drempelwaarden voor de effectgrootte van de behandeling maar niet voor de werkzaamheid

Met hun meta-analyse over de werkzaamheid van antidepressiva gooiden Kirsch et al. in 2008 de knuppel in het hoenderhok van de antidepressiva¹. Minerva besprak deze meta-analyse en wees op het probleem van de drempelwaarde om een klinisch relevant effect van antidepressiva aan te tonen². De drempel voor klinisch relevante winst van antidepressiva versus placebo die Kirsch et al. in hun meta-analyse kozen, was slechts bereikt bij patiënten met een zeer ernstige depressie. De auteurs namen dit arbitraire afkappunt over van NICE. Cohen et al. introduceerden deze drempelwaarden (een gestandaardiseerd gemiddeld verschil van 0,20 voor weinig effect, 0,50 voor een matig effect en 0,80 voor een groot effect) echter niet om een verschil in klinische werkzaamheid aan te tonen, maar wel om de effectgrootte van behandelingen onderling te kunnen vergelijken³. Kirsch et al. volgden de auteurs van de NICE-aanbevelingen in hun keuze voor een afkappunt van 0,50 om te bewijzen dat een behandeling werkzaam is. De keuze voor een drempelwaarde waarvan de klinische relevantie voor discussie vatbaar is, zorgde ervoor dat de betrouwbaarheid van hun conclusies in vraag werd gesteld.

Betrouwbaarheid van het meetinstrument

Bij de kritische evaluatie van de werkzaamheid van antidepressiva komt de betrouwbaarheid van de gebruikte meetinstrumenten als nieuw discussiepunt naar voor.

Minerva verwees in een bespreking over de behandeling van mineure depressie naar de meta-analyse van Fournier et al.^{4,5}. Ook deze meta-analyse toonde aan dat bij patiënten met milde of matige depressie de te verwachten winst van antidepressiva ten opzichte van placebo kleiner is dan Cohen's drempelwaarde van 0,20.

In functie van de initiële scores op de Hamilton Depression Rating Scale (HDRS) zien de drempelwaarden voor werkzaamheid er als volgt uit: een matig effect (Cohen waarde 0,50) vanaf een initiële HDRS-score van 25, en een groter effect (Cohen waarde 0,81) vanaf een initiële HDRS-score van 27, wat overeenkomt met een zeer ernstige depressie (HDRS-score ≥ 25).

Eind 2011 stelden Scandinavische onderzoekers zich vragen bij het gebruik van de HDRS als meetinstrument in verschillende studies (o.a. in de meta-analyse van Fournier et al.)⁶. Ze kregen toegang tot de initiële scores en de evolutie van de HDRS-vragenlijst van 597 patiënten (vijf van de zes RCT's van de meta-analyse van Fournier et al.) en stelden vast dat de precisie van de HDRS afneemt naarmate de ernst van de depressie vermindert, dus naarmate de patiënt verbetert. Het lijkt dus weinig betrouwbaar om op die manier de klinische

evolutie te vergelijken van patiënten die starten op verschillende niveaus van depressie. De auteurs merken verder nog op dat de patiënten in de verschillende RCT's de HDRS-items op een andere manier interpreteerden. Dat zorgt voor een bijkomend risico van bias. Ze besluiten dat de geringe werkzaamheid kan te wijten zijn aan de beperkte precisie van het meetinstrument en aan de lage sensitiviteit om verandering te meten, vooral bij milde tot matige depressie.

Op de consensusconferentie van het RIZIV over antidepressiva in 2006, wees Van Praag op de onzekerheden in de diagnostiek van depressie^{7,8}. De diagnostische beelden zijn volgens Van Praag qua symptomen heterogeen, ze overlappen elkaar en ze zijn onvoldoende afgebakend, ook van het normale. Volgens hem wordt de factor 'ernst' te weinig in rekening gebracht, worden de symptomen niet 'gewogen', dat wil zeggen niet beoordeeld in functie van hun diagnostische waarde, krijgen de oorzakelijke factoren weinig aandacht en heeft het diagnostische beeld weinig voorspellende waarde. Van Praag beklemtoonde dat het nodig is om bij de evaluatie van de oorzaken rekening te houden met het stressniveau en de ernst van de stressfactoren, met het persoonlijke weerstandsvermogen en met de mogelijke invloed van persoonlijkheidsfactoren op de depressieve stoornis. Hij besloot dat het hoog tijd is om de diagnostische aanpak van depressie te veranderen.

Tot op de dag van vandaag heeft men daar nog maar weinig gehoor aan gegeven...

Referenties

1. Kirsch I, Deacon BJ, Huedo-Medina TB, et al. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* 2008;5:260-7.
2. De Meyere M. Verband tussen ernst van depressie en werkzaamheid van nieuwere antidepressiva? *Minerva* 2008;7(8):118-9.
3. Turner EH, Rosenthal R. Efficacy of antidepressants. [Editorial] *BMJ* 2008;336:516-7.
4. Declercq T, Poelman T. Zijn antidepressiva en benzodiazepines effectief voor de behandeling van mineure depressie? *Minerva* 2011;10(7):84-5.
5. Fournier JC, DeRubeis RJ, Hollon SD, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010;303:47-53.
6. Isacsson C, Adler M. Randomized clinical trials underestimate the efficacy of antidepressants in less severe depression. *Acta Psychiatr Scand* 2012;25:453-9.
7. Het doelmatig gebruik van antidepressiva voor de behandeling van depressie. Consensusconferentie RIZIV, Brussel 11 mei 2006.
8. Van Praag HM. Onvrede met de diagnostiek van depressie. Implicaties voor het onderzoek van antidepressiva. Consensusconferentie RIZIV 11 mei 2006. Het doelmatig gebruik van antidepressiva voor de behandeling van depressie.