

Het rapporteren van statistische maten in het Tijdschrift voor Psychiatrie: zijn statistische richtlijnen nodig?

R.P. BRUFFAERTS

Wierdsma en Van Os (2013) breken een lans voor een streng(er) methodologisch beleid bij de beschrijving van onderzoeksresultaten in het Tijdschrift voor Psychiatrie en onderbouwen dit met voorbeelden uit gepubliceerde artikelen. Ze volgen hiermee een aantal recente internationale trends (onder andere Tressoldi e.a. 2013; Cumming 2012; Fritz e.a. 2012). De auteurs bekritiseren terecht de eenzijdige aandacht voor de 'p ≤ 0,05-regel' en geven niet enkel aan waarom deze aandacht te eenzijdig en dus te beperkt is, maar geven ook een aantal richtlijnen en suggesties voor auteurs (waarvoor ik hen dank).

Bruikbare technieken Toch heb ik een aantal commentaren bij hun stuk. Een eerste betreft over hun aanname dat de meeste complexe analysetechnieken voor iedereen beschikbaar zijn. Hier lijken de auteurs ervan uit te gaan dat beschikbaarheid synoniem is voor bruikbaarheid. Het is wél zo dat er anno 2013 inderdaad méér en betere complexe statistische modellen 'op de markt' zijn, maar of deze ook daadwerkelijk worden gebruikt, betwijfelen Nietert e.a. (2012). Complexe technieken zijn immers nog niet (ver) doorgedrongen in de opleidingscurricula van de reguliere psycholoog of psychiater. Het dissemineren van nieuwe statistische technieken loopt traditioneel stroef en traag (Van Nierop 2009). In literatuur wordt de behoefte beschreven aan methoden die nieuwe statistische kennis op een snellere wijze dissemineren. Daarbij is het verhogen van het aantal methodologische vakken in opleidingscurricula één voorbeeld (Picat e.a. 2013)

en het publiceren van (de bruikbaarheid van) nieuwe statistische methoden in niet-statistische tijdschriften een ander (Dearing 2008).

Kansverdeling Een tweede commentaar heeft betrekking op de vaak fout gekozen toetsingsgrootheden of kansverdelingen. Om een boude uitspraak te doen: binnen de menswetenschappen gaan we vaak uit van een normaalverdeling, terwijl daar redelijk weinig evidentie voor is (Grayson 2004). Vanuit een psychometrisch en statistisch gezichtspunt is dit een radicaal probleem omdat het structureel van aard is; in strikte methodologische zin kunnen dus weinig tot geen statistische modellen worden gebruikt die ontworpen zijn voor normaalverdeelde variabelen. Bovendien is er de grote minderheid van gepubliceerde onderzoeken dat daadwerkelijk rapporteert over het al dan niet normaal verdeeld zijn van de variabelen (Jackson e.a. 2009). De suggesties in het artikel van Wierdsma en Van Os dekken slechts ten dele de mogelijke fouten die worden gegenereerd door het schenden van de normaalverdeling.

Koppeling methodologie en onderzoeksvragen Ten derde hebben de auteurs het ook over het gebruik van effectmaten en betrouwbaarheidsintervallen. Ze vonden dat er slechts weinig artikelen in het Tijdschrift voor Psychiatrie ook daadwerkelijk effectmaten en/of betrouwbaarheidsintervallen rapporteren. Bovendien verdient het enige aanbeveling om ook nog de verklaarde variantie als maat toe te voegen als te

rapporteren parameter. Door systematisch rapporteren van de combinatie van p-waarden, effectmaten, betrouwbaarheidsintervallen en de hoeveelheid verklaarde variantie kan de lezer een betere inschatting maken van de (kwaliteit van de) gebruikte statistische modellen en de data die deze modellen genereren.

Tegelijk impliceert dit dat de individuele auteur die een artikel wil publiceren in het Tijdschrift zich voldoende moet inwerken in een aantal abstracte statistische noties die hem of haar in staat stellen statistische uitkomsten klinisch en onderzoeksmatig te interpreteren.

Hierin ligt wellicht de grootste uitdaging: hoewel goede rapportage nodig is, mag de nadruk níét liggen op het louter rapporteren van deze maten omdat deze of gene tijdschriftredactie een specifiek beleid aanhangt. Het rapporteren impliceert immers niet dat deze maten ook effectief worden gebruikt in de interpretatie van de data (Fidler e.a. 2004). De wetenschappelijkheid wordt slechts verhoogd wanneer de statistische parameters worden gebruikt in de interpretatie van de data. Met andere woorden: wanneer er een koppeling is tussen de (formele) methodologie en de (inhoudelijke) onderzoeksvragen en -resultaten.

Hoewel deze uitspraak evident is en op weinig weerstand zal stuiten, blijft de vraag op welke wijze dit kan gebeuren. Counsell e.a. (2011) geven een goede voorzet in de internationale literatuur, maar veel zal afhangen van het beleid dat tijdschriftredacties hanteren inzake methodologische aspecten van de artikelen die ze publiceren. Het is in dezen wellicht belangrijk de vraag te stellen of, en zo ja, in welke mate de redactie van het Tijdschrift voor Psychiatrie het wenselijk acht een aantal statistische richtlijnen op te nemen in de richtlijnen voor auteurs (www.tijdschriftvoorpsychiatrie.nl).

LITERATUUR

- Counsell N, Cortina-Borja M, Lehtonen A, Stein A. Modelling psychiatric measures using skew-normal distributions. *Eur Psychiatry* 2011; 26: 112-4.
- Dearing JW Evolution of diffusion and dissemination theory. *J Public*

Health Manag Pract 2008; 14: 99-108.

- Fidler F, Thomasen N, Cummin G, Finch S, Leeman J. Editors can lead researchers to confidence intervals but can't make them think. *Psychol Science* 2004; 15: 119-26.
- Fritz A, Scherndl T, Kuhlberger A. A comprehensive review of reporting practices in psychological journals: are effect sizes really enough? *Theory & Psychology* 2012; 23: 98-122.
- Grayson D. Some myths and legends in quantitative psychology. *Understanding Statistics* 2004; 3: 101-134.
- Jackson DL, Gillaspay J, Purc-Stephenson R. Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychol Meth* 2009; 14: 6-23.
- Nierop E van. Why do statistical journals have low impact factors? *Statistica Neerlandica* 2009; 63: 52-62.
- Nieter PJ, Wahlquist AE, Herbert TL. Characteristics of recent biostatistical methods adopted by researchers publishing in general/internal medicine journals. *Statist Med* 2012; 32: 1-10.
- Picat MQ, Savès M, Asselineau J, Dumoulin M, Coureau G, Salmi LR, e.a. Méthodes biostatistiques et épidémiologiques employées pour la recherche biomédicale: implications pour la formations médicale initiale. *Rev Epidemiol Sante Publique* 2013; 61: 261-8.
- Tressoldi PE, Giofré D, Sella F, Cumming G. High impact = high statistical standards? Not necessarily so. *PLoS ONE* 2013; 8(2): e56180. DOI: 10.1371/journal.pone.0056180.
- Wierdsma AI, van Os J. Effecten op schaal en betrouwbaarheidsintervallen als alternatief voor $p < 0,05$. *Tijdschr Psychiatr* 2013; 55: 471-80.

AUTEUR

RONNY BRUFFAERTS, licentiaat in de psychologie, doctor in de medische wetenschappen, psychoanalytisch psychotherapeut en hoofddocent Psychiatrie, KU Leuven.

Correspondentieadres: prof. dr. Ronny Bruffaerts, Universitair Psychiatrisch Centrum – KU Leuven (UPC-KUL), Herestraat 49, B-3000 Leuven.

E-mail: ronny.bruffaerts@med.kuleuven.be.

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 30-5-2013.

TITLE IN ENGLISH Reporting statistical measures in the Tijdschrift voor Psychiatrie: is there a need for adopting statistical guidelines?