

## Methodologische kwaliteit en bias in RCT's

P. Chevalier

Om de methodologische kwaliteit van RCT's in een systematisch literatuuroverzicht te evalueren verwijzen auteurs vaak naar de Jadad-score. Minerva benadrukte reeds dat de Jadad-score minstens drie punten zou moeten bedragen, maar dat daarnaast ook een intention to treat analyse aanwezig moet zijn. In de rubriek 'methodologische beschouwingen' vermelden we ook of bias in de besproken studie mogelijk is, waardoor men de interne validiteit van de studie kan bespreken. Betekenen de termen validiteit, kwaliteit en bias alle hetzelfde? Wat is hierbij de plaats van de Jadad-score?

### • Validiteit, kwaliteit en bias

De termen validiteit, kwaliteit en bias worden vaak ten onrechte door elkaar gebruikt. Het is nochtans nuttig de verschillen in betekenis te kennen. De validiteit van een studie geldt zowel 'extern' (toepasbaarheid en extrapoleerbaarheid) als 'intern' (methodologische kwaliteit van de studie). De evaluatie van de methodologische kwaliteit van een studie richt zich niet alleen op het zoeken naar bias (zijn de resultaten betrouwbaar), maar ook op het beoordelen van andere elementen (ethische goedkeuring, correcte steekproefberekening, correcte rapportering). Wanneer bias of vertekening optreedt wijken de resultaten of de interpretatie van een onderzoek af van de werkelijkheid door een systematische fout. Deze vertekening kan het werkelijke resultaat zowel onder- als overschatten, ofwel niet beïnvloeden. Het is dus correcter om te spreken over het 'risico van bias'.

### • Jadad-score en alternatieven

De Jadad-score evalueert het risico van sommige vormen van bias in een studie. Meer en meer beschouwt men deze score echter als vrij onvolledig<sup>1</sup>. Deze score heeft als voordeel dat ze een gemakkelijk reproduceerbaar scoregetal weergeeft. De Cochrane Collaboration stelt een andere aanpak voor. Zij gebruikt geen score- of controlelijst, maar zoekt op zes domeinen naar het risico van bias. Is het risico van bias aanwezig, afwezig of onzeker voor de toewijzingsmethode en concealment of allocation bij randomisatie, dubbelblinde uitvoering, verwerking van onvolledige gegevens, selectieve rapportering van gegevens en andere mogelijke vormen van bias? Hier volgt wat uitleg over deze begrippen en de evaluatie ervan.

De *toewijzing bij randomisatie* is correct als deze gebaseerd is op een gestandaardiseerde 'random numbers'-tabel, op een kop- of munt- of een gelijkaardige methode. De toewijzing gebeurt niet correct wanneer deze gebaseerd is op niet-toevallige gegevens (bv. geboortedatum, dag van opname in de studie, voorkeur van de patiënt of van de arts). *Concealment of allocation* wijst op een centrale toewijzing, met sequentiële genummerde toewijzing van omslagen of van de behandeling. Het is niet correct wanneer onderzoekers en deelnemers op de hoogte zijn van de toewijzing of de toewijzing kunnen wijzigen zoals bij open random getallen, een alternerend of een rotatiesysteem. *Blinding* wordt als 'correct' aanzien als zowel de deelnemers als de onderzoekers deze respecteren, als men oordeelt dat blinding niet nodig is (eindpunten en resultaten worden niet beïnvloed door het niet-blinden) of als de evaluatie blind gebeurt terwijl de niet-blinding van deelnemers en onderzoekers geen mogelijke bias veroorzaakt. Blinding is 'niet correct' als er geen blinding is terwijl niet-blinden de resultaten wel kan beïnvloeden. De *gegevens* kunnen we beschouwen als *compleet* als ze alle aanwezig zijn of als ze gedeeltelijk ontbreken. In het laatste geval moeten ze in de verschillende studie-armen evenveel ontbreken en omwille van dezelfde redenen. Gegevens mogen ook gedeeltelijk ontbreken indien ze de effectgrootte niet beïnvloeden of indien men ze op een correcte manier verrekent. In alle andere gevallen worden de gegevens aanzien als incompleet. De *vermelding van de resultaten* is *niet-selectief* als men de resultaten weergeeft voor alle eindpunten die beschreven zijn in het oorspronkelijke protocol. De vermelding is wel selectief als de resultaten voor sommige vooropgestelde eindpunten ontbreken, als de resultaten werden bekomen met meetinstrumenten of met analysetechnieken die vooraf niet waren vastgelegd of als het resultaten zijn van een niet vooraf vastgelegd primair eindpunt. Er zijn nog talrijke andere *risico's van bias* die moeten uitgesloten worden, onder meer de studie-opzet zelf, vroegtijdige stopzetting, sterk afwijkende basiskarakteristieken tussen de studie-armen, fraude of een ander belangrijk probleem.

Deze evaluatiemethode laat toe om de meest klassieke vormen van bias op te sporen: selectiebias (basiskarakteristieken van de verschillende studiegroepen), performance bias (toediening van complementaire behandelingen of complementaire blootstelling aan risico's), migration bias (stopzetten van de behandeling), detectiebias (evaluatie van de resultaten), notification bias (genoteerde en niet-genoteerde waarnemingen).

### • Scoren of domeinen exploreren?

De BMJ publiceerde in 2009 een studie waarbij de auteurs vaststelden dat de correlatie tussen de verschillende methodes om het risico van bias op te sporen gering is<sup>2</sup>. Ze vergeleken de methode van de Cochrane Collaboration (op basis van domeinen), de Jadad-score en de methode van Schultz (concealment of allocation, intention to treat analyse en dubbelblinding)<sup>3</sup>. Ze stelden vast dat de interobservervariatie (vijf onderzoekers van hetzelfde universitair centrum) sterk varieerde naargelang de onderzochte domeinen van de Cochrane Collaboration. De variatie was over het algemeen groter voor domeinen waar men op een meer subjectieve manier de mogelijkheid van bias moest beoordelen. De techniek van de Cochrane Collaboration is duidelijk minder goed reproduceerbaar, tenminste voor sommige domeinen. Bij afwezigheid van een score met een valide drempelwaarde voor de evaluatie van een studie, stelt zich de vraag op welke basis (welke waarde) men zich finaal kan baseren om te besluiten of de resultaten van een studie al of niet in de meta-analyse kunnen geïncludeerd worden. Is dat een echt probleem? Leiden de verschillen tussen de observatoren tot andere keuzes die de betrouwbaarheid van de gepoolde resultaten in het gedrang brengen? Op deze vraag is er nog geen antwoord.

### Referenties

1. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 [updated September 2009]. The Cochrane Collaboration, 2009. Available from <http://www.cochrane.org/resources/handbook/>
2. Hartling L, Ospina M, Liang Y, et al. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. BMJ 2009;339:b4012.
3. Schultz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408-12.