

## Equivalentiestudies tegenover inferioriteits- en superioriteitsstudies

P. Chevalier

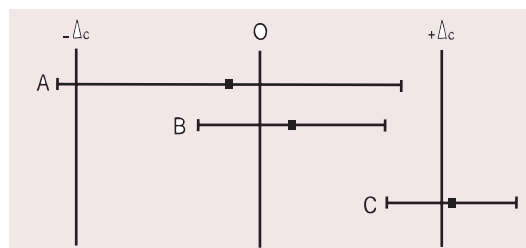
Met een statistische toets besluiten dat twee behandelingen volledig equivalent zijn, is onmogelijk: de kans dat het verschil tussen beide behandelingen nul is, ... is nul. Indien echter het klinische verschil verwaarloosbaar klein is, kunnen we wel aannemen dat er equivalentie bestaat, zonder dat daarom het verschil tussen beide behandelingen gelijk is aan nul<sup>1</sup>. Equivalentiestudies worden veel minder uitgevoerd dan non-inferioriteitsstudies, behalve in het domein van de bio-equivalentiestudies<sup>2</sup>.

Een equivalentiestudie onderzoekt of het gevonden verschil, met haar 95% betrouwbaarheidsinterval, volledig tussen de vooraf bepaalde equivalentiedrempels ligt. De marge voor equivalentie ( $\Delta_c$ ) is het resultaat van een consensus tussen experts en is gebaseerd op literatuuronderzoek, bij voorkeur op - indien beschikbaar - een meta-analyse. Deze marge moet kleiner zijn dan het afkappunt voor klinisch relevante superioriteit. Over het algemeen kiest men een waarde die de helft bedraagt van het afkappunt voor superioriteit<sup>1</sup>. Soms kiest men voor twee derde en naargelang de context voor een nog kleinere waarde (een vierde bijvoorbeeld)<sup>3</sup>.

### Illustratie

Nemen we verschillende equivalentiestudies, met een resultaat en betrouwbaarheidsinterval voor eenzelfde primair eindpunt op het vlak van werkzaamheid. Het resultaat (het verschil tussen twee behandelingen) wordt weergegeven door een blokje en het 95% betrouwbaarheidsinterval van dit verschil door een horizontale lijn aan beide kanten van het blokje. Vergelijken we de grenzen van deze betrouwbaarheidsintervallen met de vooraf afgesproken equivalentiemarges ( $\Delta_c$ ). Om te kunnen spreken over equivalentie, moeten de waarden van de betrouwbaarheidsintervallen binnen de equivalentiemarges  $-\Delta_c$  en  $+\Delta_c$  liggen.

In het voorbeeld A kunnen we niet besluiten dat beide behandelingen equivalent zijn. In het voorbeeld B is er wel equivalentie aangetoond. In het voorbeeld C kunnen we evenmin besluiten dat beide behandelingen equivalent zijn.



De studie over acute jichtartritis, besproken in dit nummer, is een voorbeeld van een equivalentiestudie<sup>4</sup>. De auteurs besluiten dat prednisolon en naproxen equivalent zijn. Dit besluit is gebaseerd op het resultaat voor pijnreductie (bij een jichtaanval), gemeten op een VAS van 100 mm. Het verschil tussen prednisolon en naproxen is 1,3 mm (95% BI van -9,8 tot 7,1) en de waarden van het betrouwbaarheidsinterval liggen dus binnen de vooropgestelde marge van 10 mm. Deze marge is echter louter arbitrair gekozen en is gebaseerd op een andere studie over jicht waarin de superioriteitsmarge was vastgelegd op 13 mm. Deze marge was op haar beurt gebaseerd op een andere studie over posttraumatische pijn waarin een verschil van 13 mm met een BI van 10 tot 17 mm aanzien werd als een klinisch relevante verandering. De keuze voor 10 mm als equivalentiemarge in een studie over de behandeling van een jichtaanval is dus erg vatbaar voor discussie. Een equivalentiedrempel van minstens de helft van de superioriteitsmarge, in dit geval 13/2 (of beter nog 10/2), zou correcter geweest zijn en zou aangetoond hebben dat er geen equivalentie is (voorbeeld A in de illustratie).

### Verskil tussen equivalentie- en non-inferioriteitsstudie

Om te kunnen besluiten dat twee behandelingen equivalent zijn, moet het vastgestelde verschil met zijn betrouwbaarheidsinterval binnen de equivalentiemarges (positief en negatief) liggen. Om te kunnen spreken over non-inferioriteit moet de ondergrens van het betrouwbaarheidsinterval hoger zijn dan de vastgelegde non-inferioriteitsmarge ( $-\Delta_c$ ). In dat geval kan behandeling X equivalent of superieur zijn aan behandeling Y.

Bij equivalentiestudies kan het klinisch relevant zijn om een asymmetrie in de equivalentiemarges te aanvaarden tussen de onderste waarde (-) en de bovenste waarde (+), waarbij de laatste het dubbele kan bedragen van de eerste.

Om de steekproefgrootte te berekenen van een equivalentiestudie zijn er geen preciese, expliciete formules beschikbaar zoals wel het geval is voor non-inferioriteitsstudies, maar experts stellen wel bepaalde formules voor<sup>3</sup>.

### Superioriteit, non-inferioriteit en equivalentie

Indien er in een superioriteitsstudie geen verschil is vastgesteld tussen twee behandelingen, kunnen we dan spreken over equivalentie of over non-inferioriteit tussen beide behandelingen? Het antwoord hierop is duidelijk 'neen'<sup>5</sup>. Vanuit methodologisch standpunt wil 'geen bewijs van verschil' niet zeggen dat er 'bewijs is van geen verschil' (non-inferioriteit) of van equivalentie (klinisch niet relevant, verwaarloosbaar verschil). Elke vorm van onderzoek vereist een specifieke methodologie. Voor non-inferioriteits- en equivalentiestudies is dit vooral een per protocol analyse en voor superioriteitsstudies een intention to treat analyse.

Het is niet mogelijk om in het kort alle elementen aan te geven waarmee we rekening moeten houden bij het interpreteren van de methodologische waarde en de resultaten van een equivalentiestudie. Minerva zal bij de analyse van dergelijke studies steeds ingaan op de aspecten die nuttig kunnen zijn voor de clinicus.

### Referenties

- Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous methods. *BMJ* 1996;313:36-9.
- Bogaert M, Chevalier P. De klinische equivalentie van generieken. [Editoriaal] *Minerva* 2009;8(4):41.
- Grouin JM, Coste M. Essais de non-infériorité et d'équivalence: méthodologie et analyse. In *Essais cliniques: théorie, pratique et critique*. Médecine-Sciences Editions Flammarion 2006.
- Poelman T. Behandeling van acute jichtartritis met prednisolon? *Minerva* 2009;8(7):100-1.
- Conzalez CD, Bolaños R, de Sereday M. Editorial on hypothesis and objectives in clinical trials: superiority, equivalence and non-inferiority. *Thromb J* 2009;7:3.