



### Vergelijking van twee behandelingen in non-inferioriteitsstudies: beperkingen van de courant gebruikte hazard ratio (HR)

Pierre Chevalier, médecin généraliste

Minerva publiceerde in 2009 een algemene methodologische bijdrage over het nut, de beperkingen en de valkuilen van non-inferioriteitsstudies (1). Later gingen we meer specifiek in op de problemen met de marges voor non-inferioriteit (2) en met de drempel voor een klinisch relevante verbetering (3).

In non-inferioriteitsstudies gebruikt men vaak de hazard ratio (HR) om het verschil te berekenen tussen 2 behandelingen. Uno et al. publiceerden in 2015 een onderzoek waarin ze de beperkingen van deze werkwijze illustreren aan de hand van de gegevens uit 2 studies (4). Ze stellen ook alternatieven voor die een meer precieze inschatting van het verschil en een vanuit klinisch oogpunt beter interpreteerbaar verschil geven.

Eén van de 2 voorbeelden die Uno et al. gebruiken in hun onderzoek, is een publicatie van 2 studies die het effect evalueren op cardiovasculaire gebeurtenissen en bloeddruk van celecoxib voor de preventie van colorectaal carcinoom (5). Minerva gaf in 2007 commentaar op één van beide studies (6). De uitkomstmaat voor cardiovasculaire veiligheid is de tijd tot een samengestelde uitkomstmaat van cardiovasculaire sterfte, niet-fataal myocardinfarct, niet-fataal CVA en hartfalen.

*Voor de patiënten die gedurende 36 maanden behandeld worden met celecoxib 400 mg, bedraagt de HR voor cardiovasculaire veiligheid versus placebo 3,35 (95% van 1,44 tot 7,81 en  $p=0,005$ ). Dat komt overeen met een absolute risicotoename van 2,0% (95% BI van 0,8 tot 3,2%) over 36 maanden.*

De precisie van de schatting van de resultaten op basis van de HR hangt hoofdzakelijk af van het aantal vastgestelde gebeurtenissen en niet onmiddellijk van de blootstellingstijd of van de steekproefgrootte van de onderzochte populatie. Als de incidentie van de gebeurtenissen laag is, zal in de studie een vrijwel onmogelijk hoog aantal gebeurtenissen moeten optreden om het vooropgestelde criterium van non-inferioriteit voor de HR te kunnen bereiken.

Als het specifiek gaat over de evaluatie van de veiligheid van een behandeling, is de blootstellingstijd klinisch belangrijker dan het aantal vastgestelde gebeurtenissen in een non-inferioriteitsstudie. Als de blootstellingstijd van de patiënt voldoende lang is, zijn er methodologische alternatieven voor de HR die klinisch beter interpreteerbaar zijn, zelfs bij een laag aantal gebeurtenissen.

*In de celecoxibstudies verandert het (fictieve) toevoegen van 1 000 periodes supplementaire blootstellingstijd met hetzelfde aantal gebeurtenissen weinig aan de HR: 3,29 (95% BI van 1,41 tot 7,67). De absolute risicotoename vermindert echter naar 0,9% (95% BI van 0,3 tot 1,6%) en is een meer precieze schatting (minder ruim betrouwbaarheidsinterval). Het absolute risicoverschil (toename of afname) is in dit geval preciezer en instructiever dan de HR.*

Een ander probleem bij de interpretatie van de HR is de variabiliteit ervan tijdens de studieduur. Als de HR niet constant is in de tijd (wat betekent dat er niet voldaan is aan de proportional hazards

aanname, of eenvoudiger uitgedrukt dat de cumulatieve incidentiecurves niet parallel verlopen), is de klinische betekenis van de HR niet duidelijk. Er bestaat geen test met voldoende power om uit te maken of het proportioneel hazards model geschikt is.

*In de celecoxibstudies is er tijdens de eerste 10 maanden van de behandeling geen duidelijke scheiding van de cumulatieve incidentiecurves voor celecoxib en placebo en op een bepaald moment kruisen de curves elkaar zelfs. Dat betekent dat er mogelijk niet voldaan is aan de hypothese van een proportionele evolutie van de hazards.*

Om de onzekerheden bij de klinische interpretatie van de HR in non-inferioriteitsstudies op te vangen, stellen Uno et al. voor om de resultaten uit te drukken in absoluut risicoverschil of in ‘**restricted mean survival time**’ (RMST).

*Bij 36 maanden behandeling met celecoxib is er, zoals hoger vermeld, een absolute risicotoename van 2,0% (95% BI van 0,8 tot 3,2%). Over dezelfde periode bedraagt de RMST 35,33 maanden in de celecoxibgroep en 35,76 maanden in de placebogroep. Het verschil bedraagt dus 0,43 maanden (95% BI van 0,08 tot 0,78) wat overeenkomt met 13 dagen (95% BI van 2 tot 24). Anders uitgedrukt toont de analyse op basis van de RMST aan dat de periode waarin patiënten behandeld met celecoxib vrij zijn van cardiovasculaire gebeurtenissen gemiddeld hoogstens 24 dagen korter is dan de periode waarin de patiënten uit de placebogroep vrij zijn van cardiovasculaire gebeurtenissen.*

## Besluit

Deze publicatie vestigt de aandacht op een ander probleem dan de drempel voor non-inferioriteit bij de interpretatie van de resultaten van non-inferioriteitsstudies, namelijk het gebruik van de HR voor de evaluatie van de veiligheid (ongewenste effecten) van een behandeling. In dit geval is de blootstellingstijd aan de te evalueren behandeling de belangrijkste parameter. Als de incidentie van ongewenste effecten laag is, zijn het absolute risicoverschil en de ‘restricted mean survival time’ (RMST) tot een specifiek punt in de tijd vanuit klinisch oogpunt gemakkelijker te interpreteren dan de HR.

## Referenties

1. Chevalier P. Non-inferioriteitsstudies: het nut, de beperkingen en de valkuilen. *Minerva* 2009;8(6):88.
2. Chevalier P. Drempelwaarden voor een klinisch relevant verschil. *Minerva* 2011;10(8):103.
3. Chevalier P. Het minimaal klinisch relevante verschil. *Minerva* 2015;14(3):37.
4. Uno H, Wittes J, Fu H, et al. Alternatives to hazard ratios for comparing the efficacy or safety of therapies in noninferiority studies. *Ann Intern Med* 2015;163:127-34.
5. Solomon SD, Pfeffer MA, McMurray JJ, et al; APC and PreSAP Trial Investigators. Effect of celecoxib on cardiovascular events and blood pressure in two trials for the prevention of colorectal adenomas. *Circulation* 2006;114:1028-35.
6. Chevalier P, van Driel M. Coxibs, andere NSAID's en cardiovasculair risico. *Minerva* 2006;5(10):161-3.