



Résultats dans les études de non infériorité : limites des rapports de hasards

Pierre Chevalier, médecin généraliste

Nous avons déjà abordé dans la revue Minerva plusieurs aspects généraux des études de non infériorité (1) mais aussi le concept particulier de la borne de non infériorité (2) et de sa pertinence clinique (3).

Uno et al. (4) analysent et illustrent par des exemples un autre aspect limitant de ce type d'étude, la mention des résultats en rapport de hasards (RH, hazard ratio – HR), en proposant de meilleures alternatives, plus précises et plus faciles à interpréter d'un point de vue clinique.

Un des exemples cités par Uno et al. est celui d'une publication évaluant la sécurité du célécoxib au point de vue cardiovasculaire et pression artérielle dans deux études sur la prévention des adénomes colorectaux (5), études évoquées précédemment dans la revue Minerva (6). Pour le critère de jugement composite de la sécurité cardiovasculaire (décès cardiovasculaire, infarctus du myocarde non fatal, AVC non fatal ou insuffisance cardiaque), le RH versus placebo est de 3,35 (avec IC à 95% de 1,44 à 7,81 ; $p = 0,005$) pour les patients sous 400 mg de célécoxib, sur 36 mois de traitement. Ceci correspond à une augmentation absolue de risque de 2,0% (avec IC à 95% de 0,8 à 3,2%), sur 36 mois.

La précision de l'estimation des résultats donnés par un RH dépend principalement du nombre d'événements observés et non directement du temps d'exposition au traitement ou de la taille de l'échantillon de population étudié. Si l'incidence d'événements concernés est faible, la recherche nécessite un nombre d'événements démesurément important pour espérer atteindre le critère de non infériorité dans le RH.

Dans le domaine particulier de l'évaluation de la sécurité d'un traitement, la durée d'exposition au traitement est cliniquement plus importante que le nombre d'événements observés en étude de non infériorité. Si la durée d'observation est suffisamment longue, d'autres procédures qu'un RH permettent d'apporter des preuves de sécurité cliniquement mieux interprétables, même si le nombre d'événements est faible.

Dans l'étude précitée avec le célécoxib (5), l'addition (fictive) de 1000 périodes de temps d'observation supplémentaires, avec le même nombre d'événements, modifie peu le RH : 3,29 avec IC à 95% de 1,41 à 7,67. Par contre, l'augmentation absolue de risque passe à 0,9% (avec IC à 95% de 0,3 à 1,6%), résultat plus précis (IC moins large). La modification absolue de risque (augmentation ou réduction) est un outil plus précis et instructif que le RH dans ce cas.

Un autre problème pour l'interprétation du RH est sa variation dans le temps de l'étude. S'il n'est pas constant dans le temps (= violation de l'hypothèse de hasards proportionnels, plus simplement dit, évolution non « symétrique » des courbes d'incidence), la signification clinique du RH est difficile à situer. Il n'existe pas de test performant pour déterminer l'adéquation du modèle des hasards proportionnels.

Dans l'étude précitée avec le célécoxib (5), il n'y a pas de séparation nette (et même un croisement momentané) des courbes d'incidence pendant les dix premiers mois, ce qui indique une possible violation de l'hypothèse d'évolution proportionnelle des hasards.

Pour pallier aux incertitudes dans l'interprétation clinique d'un RH dans une étude de non infériorité, Uno et al. proposent de recourir soit à la modification absolue de risque soit au **temps moyen de survie limité (TMSL)** à un délai spécifique.

Dans l'étude précitée avec le célécoxib (5), comme mentionné plus haut, l'augmentation absolue de risque avec le célécoxib est de 2,0% (avec IC à 95% de 0,8 à 3,2%) sur 36 mois. Sur la même période, les TMSL sont de 35,33 mois sous célécoxib et de 35,76 mois sous placebo, soit une différence de 0,43 mois (avec IC à 95% de 0,08 à 0,78) soit environ 13 jours (avec IC à 95% de 2 à 24). Autrement dit, les patients traités par célécoxib n'ont pas présenté d'évènement cardiovasculaire sur une période en moyenne au maximum 24 jours plus courte que celle pour les sujets sous placebo...

Conclusion

Cette publication attire notre attention sur une autre caractéristique (que la borne de non infériorité) plus difficile à interpréter dans une étude de non infériorité : le rapport de hasards pour la sécurité (effets indésirables). Le temps d'exposition au traitement évalué est dans ce cas le paramètre le plus important et en cas d'incidence faible de survenue d'effets indésirables, la modification absolue de risque (MAR) et le temps moyen de survie limité (TMSL) à un délai spécifique sont plus faciles à interpréter au point de vue clinique.

Références

1. Chevalier P. Etude de non infériorité : intérêt, limites et pièges. *MinervaF* 2009;8(7):100.
2. Chevalier P. Pertinence clinique d'un seuil de différence. *MinervaF* 2011;10(8):103.
3. Chevalier P. La différence minimale cliniquement pertinente. *MinervaF* 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, autres AINS et risque athérombotique. *MinervaF* 2007;6(1):8-10.