

## ● LOCF ou pas LOCF ? Quand les données manquent...

P. Chevalier

Dans leur RCT évaluant l'efficacité préventive de l'ajout d'un inhibiteur de la pompe à protons à un COXIB, Chan et coll. enregistrent 6% d'arrêts d'étude dans le groupe célecoxib plus ésomeprazole et 7% dans le groupe célecoxib plus placebo. Il y a donc des données « manquantes » au terme prévu pour l'étude (de 13 mois). Soucieux d'analyser au mieux les résultats, les auteurs pratiquent une analyse en intention de traiter et tentent de remplacer les données manquantes par la technique du Last Observation Carried Forward<sup>1</sup> qui consiste à considérer la dernière observation de chaque participant comme son résultat « définitif » même si le moment de cette dernière observation ne coïncide pas avec le moment prévu pour la mesure. Les données manquantes peuvent-elles influencer les résultats d'une étude ? Cette technique du LOCF visant à les remplacer est-elle valide et fiable ?

### Données manquantes

L'importance de l'absence de données en fonction de la puissance estimée pour l'étude est claire : un échantillon d'un nombre de personnes est calculé, et dans leur calcul, les auteurs peuvent ajouter une marge en fonction d'un nombre prévu de sorties d'étude qui ne ruinerait pas la puissance de l'étude ; il faut donc qu'ils estiment ce risque correctement et que le nombre de sorties d'étude ne soit pas dépassé, sinon la puissance devient insuffisante. Une autre conséquence d'un manque de données est plus complexe : l'absence de données pour certaines variables (différentes) pour certains sujets (différents) peut fausser une analyse (en régression multiple) qui tente de déterminer quelle variable influence le plus les résultats. Un échantillon fort important minimise aussi ce problème<sup>2</sup>.

Par contre, un autre problème ne peut être résolu par une augmentation de la population incluse, problème lié au motif de la sortie d'étude, qui peut être variable. Une personne incluse dans une étude peut s'estimer non améliorée et chercher un autre traitement, parfois en raison d'une moindre observance<sup>3</sup> ; si la moitié des personnes qui quittent l'étude le font en raison d'une inefficacité du traitement et que le traitement 1 est plus efficace que le traitement 2, l'ampleur de la différence pour les résultats sera moindre. Les personnes peuvent quitter l'étude en raison d'effets indésirables insupportables, parfois en fonction de la dose du médicament<sup>3</sup>. Elles peuvent également abandonner parce qu'elles se sentent mieux et ne voient pas l'intérêt de se rendre à une visite de contrôle ; si la moitié des personnes quittent pour ce motif, l'ampleur de la différence pour le résultat est également diminuée<sup>2</sup>. Si les motifs d'arrêts d'étude sont différents selon les groupes et, surtout, si la proportion de sorties d'étude est différente selon les groupes, les résultats peuvent être faussés dans un sens ou dans un autre. Il est donc intéressant d'analyser les motifs de sortie d'études et de voir si les pourcentages sont identiques dans les différents bras d'étude, toute asymétrie rendant l'interprétation des résultats plus difficile.

### Remplacer les données manquantes ?

La méthode la plus aisée pour remplacer des données manquantes est celle du LOCF. Elle est cependant vivement critiquée<sup>3-5</sup>. En cas de proportion de sorties d'étude semblable dans le groupe principe actif et dans le groupe placebo, une analyse LOCF sous-estime généralement l'ampleur de l'effet. En cas de proportion non identique, les résultats peuvent être biaisés dans un sens ou dans un autre. Les statisticiens ont développé différentes méthodes, basées sur l'hypothèse que les données sont manquantes par hasard (MAR Missing At Random : manque en lien avec l'une ou l'autre variable mais non avec le résultat) : approches basées sur le rapport de vraisemblance, estimations pondérées ou, plus utilisée, imputation multiple. La fragilité de telles analyses est liée à une incapacité dans la majorité des cas de vérifier que les données sont manquantes par hasard et non pour une autre raison (NMAR Not Missing At Random : manque lié à la pathologie ou à l'intervention (par exemple manque d'efficacité, effet indésirable, observance)). Certains experts proposent donc de considérer ces analyses comme des analyses de sensibilité<sup>4</sup>. Dans cette optique, Sackett proposait<sup>6</sup> un « worst-case scenario » à comparer avec un modèle

Evaluer la proportion de données manquantes dans une étude est important pour en interpréter les résultats. Toute asymétrie entre les bras de traitement est particulièrement préoccupante. Pour corriger l'effet de ces absences, un recours à des techniques statistiques plus élaborées que le LOCF est possible mais l'intérêt de celles-ci se situe au niveau d'une analyse de sensibilité.

centrant les patients avec données manquantes. Ce worst-case scenario fait l'hypothèse que dans le groupe intervention tous les patients avec données manquantes présentent le critère de jugement primaire (avec un effet péjoratif pour le résultat de ce groupe) alors que dans le groupe placebo aucun des patients avec données manquantes ne présente un événement correspondant au critère (ce qui améliore globalement le résultat moyen de ce groupe). Il reconnaît que la plupart de ses collègues estiment que ce modèle est extrême et que ce type d'hypothèse peut aussi conduire à une interprétation erronée. La méthode, plus complexe, d'imputation multiple recueille davantage d'assentiment parmi les experts.

### Références

1. Chevalier P. Prévention de l'ulcère récidivant : COXIBs + IPP ? *MinervaF* 2007;6(9):142-3.
2. Streiner DL. The case of the missing data: methods of dealing with dropouts and other research vagaries. *Can J Psychiatry* 2002;47:68-75.
3. Streiner DL. Missing data and the trouble with LOCF. *Evid Based Ment Health* 2008;1:3-5.
4. Vandembroucke JP, von Elm E, Altman DG, et al; STROBE initiative. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Ann Intern Med* 2007;147:W163-94.
5. Lane P. Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. *Pharm Stat* 2008;7:93-106.
6. Sackett DL. Decide how to handle missing data in the analysis. In Haynes RB, Sackett DL, Guyatt GH, Tugwell P. *Clinical epidemiology. How to do clinical practice research*. Lippincott Williams & Wilkins edition 2006; 496 pp.