

## Appariement sur les scores de propension

La méthode la plus précise pour mesurer l'effet d'un traitement est de mener une étude expérimentale où un certain nombre de patients sont randomisés en deux groupes, l'un qui reçoit le traitement (groupe intervention), et l'autre qui ne reçoit pas de traitement, un autre traitement ou un placebo (groupe témoin). Du moment que la randomisation est effectuée correctement et en aveugle, il n'y a pratiquement aucun risque de différences entre les deux groupes quant aux caractéristiques de base. Si, après un certain temps, on veut analyser la différence entre le traitement et le témoin pour un critère de jugement, il est toujours possible d'appliquer une correction afin de prendre en compte les facteurs de confusion fortuits.

Nombreux sont les traitements pour lesquels il n'existe pas ou que peu d'études contrôlées, pour toute une série de raisons. Mais, si on ne dispose pas de recherche expérimentale, on peut tout de même souvent trouver des études d'observation. Dans ce type d'étude, les chercheurs enregistrent, de manière prospective ou rétrospective, un certain nombre de paramètres dans un groupe de patients atteints d'une pathologie déterminée qui, de leur propre initiative ou après avoir consulté un médecin, suivent ou non un traitement donné. Ces données peuvent ensuite être utilisées pour rechercher un lien entre le traitement ou l'absence de traitement et la disparition/prévention de la pathologie. Mais comme ces études sont menées sans randomisation, il se peut que le lien observé soit dû non pas au traitement mais à des différences dans les variables secondaires des participants. Pour tenir compte de ces facteurs de confusion, on a le choix entre plusieurs techniques de méthodologie. On peut par exemple corriger les résultats au moment de l'analyse statistique pour tenir compte des facteurs de confusion. Il est également possible de procéder, avant l'analyse, à l'appariement du groupe traité et du groupe non traité. Plus le nombre de facteurs de confusion pour lesquels on procède à un appariement est élevé, plus le groupe recevant le traitement et le groupe non traité seront comparables et moins le lien observé dépendra des facteurs de confusion, autrement dit, plus le lien s'approchera du résultat d'une étude contrôlée randomisée. Il y a cependant un revers à la médaille : plus le nombre de variables pour lesquelles on procède à un appariement est élevé, plus il sera difficile de trouver pour chaque patient recevant le traitement un patient non traité qui lui corresponde, et plus le nombre d'observations exclues de l'analyse sera élevé. Du fait de cette perte de puissance, l'existence d'un lien sera plus difficile à démontrer. Que faire pour éviter ce problème ?

Partons de la constatation suivante : un facteur de confusion qui est lié au résultat d'un traitement joue souvent un rôle déterminant dans le fait de recevoir ou non le traitement. La probabilité de devoir suivre un traitement déterminé (la « propension » à se le voir prescrire) dépendra donc d'un certain nombre de facteurs de confusion (zie voorbeeld in kader).

**Comment procéder à l'appariement des cas et des témoins de la population d'étude sur ces facteurs de confusion sans trop perdre au niveau des observations ?**

Après l'identification des facteurs de confusion, il est possible de les utiliser comme variables prédictives dans une analyse à variables multiples par régression logistique, le traitement étant utilisé comme variable indépendante. Il découle de cette analyse un score de 0 à 1. Le score de « propension » (terme qui signifie « tendance ») désigne la probabilité d'être exposé à un traitement

Exemple issu de l'analyse « Prévention primaire et secondaire de la coronaropathie : bêtabloquant ? » (p. 99)<sup>1</sup>

Cette d'observation prospective<sup>1</sup> visait à explorer, après quatre ans, le lien entre la prise d'un bêtabloquant, par comparaison à l'absence de ce traitement, et l'incidence d'événements cardiovasculaires dans trois cohortes, à savoir : des patients ayant des antécédents d'infarctus du myocarde, des patients ayant des antécédents de coronaropathie sans infarctus du myocarde et des patients ayant des facteurs de risque de coronaropathie. On sait que l'hypertension est un facteur de risque de développement d'événements cardiovasculaires. Il s'agit donc d'un facteur de confusion qui va biaiser le résultat du bêtabloquant. Mais, le fait de recevoir un bêtabloquant ou de ne pas en recevoir dépend aussi de la présence ou de l'absence d'une hypertension. Cela vaut aussi pour d'autres facteurs de confusion, comme l'âge, l'hyperlipidémie...

donné en fonction de variables présentes au moment de l'instauration du traitement. À l'aide d'un programme informatique, on peut ensuite apparier le groupe sous traitement à celui qui n'y est pas exposé en veillant à ce que la différence absolue entre les scores de propension soit la plus petite possible. Avec une perte minimale d'observations, on obtient ainsi deux groupes pour lesquels la probabilité d'être exposé à un traitement est pratiquement égale et qui diffèrent donc moins quant à un certain nombre de facteurs de confusion.

Dans le même exemple : les chercheurs ont préalablement prévu que 27 facteurs de base sont susceptibles d'influencer le lien entre la prise d'un bêtabloquant et le critère de jugement (événements cardiovasculaires). Les deux groupes ont ensuite été appariés, la différence absolue quant au score de propension ne pouvant excéder 0,6 DS. Après cet appariement, on a observé que les 27 facteurs de confusion étaient répartis de manière égale entre les deux groupes.

Mais on a beau tenter de procéder à un appariement sur les scores de propension pour un grand nombre de facteurs de confusion, il restera toujours des facteurs de confusion que l'on ne connaît pas et pour lesquels il n'y aura donc pas d'appariement. C'est la raison pour laquelle, lorsque l'on veut rechercher de manière rigoureuse l'effet d'un traitement, la conception d'étude la plus valide demeure celle des études randomisées.

### Conclusion

Le score de propension est le résultat d'une analyse à variables multiples par régression logistique portant sur un ensemble de facteurs de confusion ; il exprime la probabilité, pour une personne, d'être exposée ou non à un traitement sur base de ces facteurs de confusion. Dans une étude d'observation, l'appariement de deux groupes sur les scores de propension permet d'éliminer un grand nombre de facteurs de confusion sans perte importante d'observations.

### Références

1. De Cort P. Prévention primaire et secondaire de la coronaropathie : bêtabloquant ? *MinervaF* 2013;12(8):99-100.
2. Bangalore S, Steg G, Deedwania P, et al; REACH Registry Investigators.  $\beta$ -blocker use and clinical outcomes in stable outpatients with and without coronary artery disease. *JAMA* 2012;308:1340-9.