

Facteur de confusion

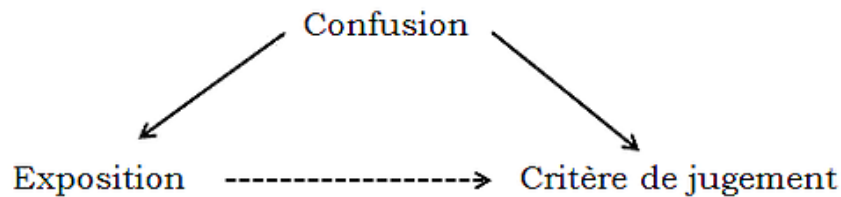
Reda Arab

1 Concept

Lorsque l'on cherche à établir un lien de causalité (on parle en français d'une exposition qui entraîne un critère de jugement), il faut faire attention aux facteurs de confusion.

On parle de facteur de confusion en Inférence Causale lorsqu'une variable a une influence à la fois sur le traitement/l'exposition (treatment/exposure) et sur le critère de jugement (outcome). On peut aussi le voir comme le fait d'avoir une influence sur la façon dont le traitement est assigné et le critère de jugement.

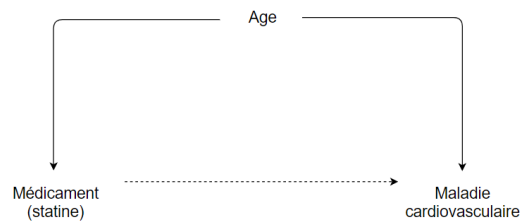
Pourquoi ? : Parce que cela peut mener à des conclusions erronées comme nous allons le voir à travers deux exemples.



2 Exemple 1 : Effet d'un médicament sur les maladies cardiovasculaires

Supposons que nous cherchions à établir l'effet d'un médicament (comme la statine) sur les maladies cardiovasculaires. L'exposition est donc le fait de prendre un médicament ou non et le critère de jugement est le fait de développer une maladie cardiovasculaire ou non (par exemple l'année qui suit).

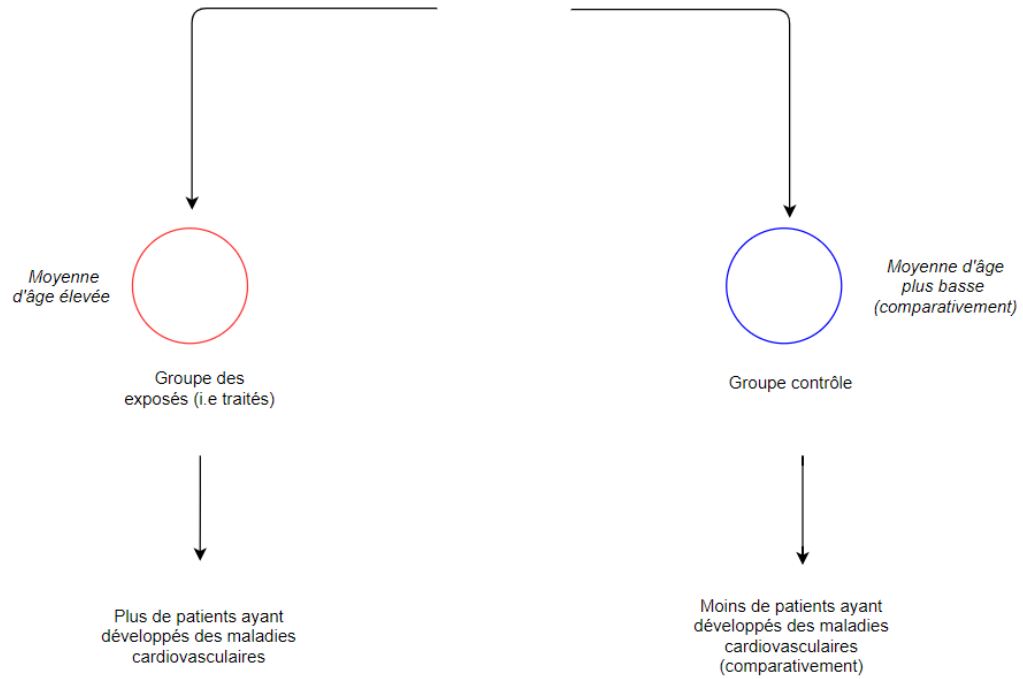
Mais supposons que les personnes âgées sont plus enclines à prendre ce médicament que les personnes plus jeunes. Et par la même occasion les personnes âgées ont plus de probabilité de développer des maladies cardiovasculaires. Notre facteur de confusion ici est *l'âge*.



Où est se situe le problème ?

Le problème est le suivant : pour savoir s'il y a un effet du médicament, quelqu'un pourrait comparer le nombre de personnes ayant développé des maladies cardiovasculaires dans le groupe des exposés (ceux qui prennent de la statine) et dans le groupe témoin (ceux qui n'en prennent pas). Or, on pourrait se retrouver avec le groupe des exposés ayant une moyenne d'âge significativement plus élevée que le groupe témoin, étant donné que les personnes âgées ont plus de probabilité de prendre ce médicament. Et, de plus, les personnes âgées développent des maladies cardiovasculaires avec une plus grande probabilité.

Exemple de situation pouvant mener
à des conclusions erronées

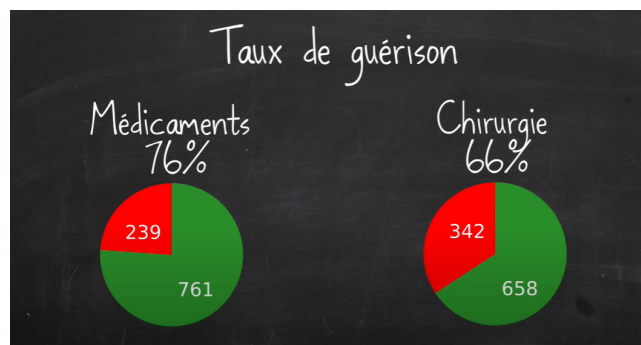


Conclusion : On pourrait alors conclure que la statine ne fonctionne pas (voir que le médicament est néfaste) étant donné que l'on pourrait vraisemblablement trouver plus de personnes ayant développé des maladies cardiovasculaires parmi le groupe des exposés avec pourtant les deux groupes de même taille.

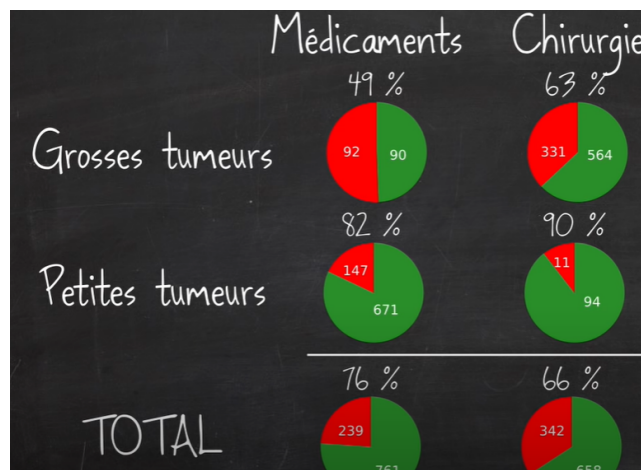
3 Exemple 2 : Effet d'un type d'intervention (médicament ou chirurgie) sur le taux de guérison

Cet exemple et les deux diagrammes ont été tirés de la vidéo de ScienceEtonnante dont vous pouvez avoir accès en cliquant sur le lien suivant : https://www.youtube.com/watch?v=vs_Zzf_vL2I.

Supposons que l'on ait deux traitements pour traiter les tumeurs : médicaments et chirurgie. On souhaite ainsi savoir quelle est le meilleur traitement.



En voyant ce premier graphe, on peut alors conclure que les médicaments fonctionnent significativement mieux que la chirurgie. Essayons d'observer des données plus précises.

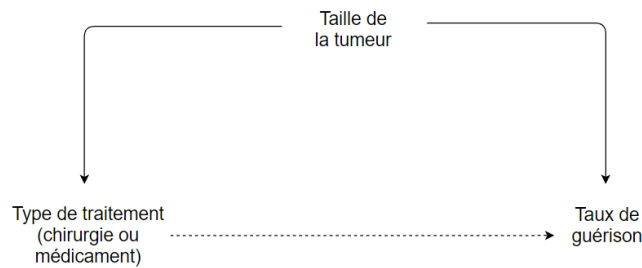


En séparant selon que la taille des tumeurs, on observe que la chirurgie est meilleure dans tous les cas.

Quel est le problème ?

Ici, la chirurgie est beaucoup plus employée dans le cas des grosses tumeurs que dans le cas des petites tumeurs (comme on le voit sur les diagrammes). Or, les grosses tumeurs impliquent un taux de guérison qui est bien plus faible que les petites tumeurs.

On a ainsi un facteur de confusion qui est *la taille de la tumeur* qui influe à la fois sur l'assignation du traitement et sur le taux de guérison.



Conclusion : En analysant les données sans prendre en compte le facteur de confusion, on peut arriver à une conclusion erronée qui décrirait la prise de médicaments comme plus efficace par rapport à la chirurgie. Le facteur de confusion *taille de la tumeur* qui agit à la fois sur quel traitement utiliser et le taux de guérison est ainsi important à prendre en compte dans une analyse.