

Y a-t-il relation de cause à effet ?

La science statistique de l'inférence causale

Dans le langage courant, on appelle « cause » la raison ou l'origine d'un phénomène, mais il est plus difficile que l'on pense de définir le concept de causalité précisément. Dès lors, comment peut-on identifier une cause et en quantifier les effets, surtout si l'on ne peut pas recourir à une expérience contrôlée pour des raisons logistiques, matérielles ou éthiques ?

Christian Genest
Erica E. M. Moodie
Université McGill

Tel ou tel traitement, vaccin ou médicament est-il efficace ou non ? C'est une question centrale en médecine qui a d'ailleurs souvent fait la une des journaux depuis 2020, alors que divers moyens de se protéger contre la COVID-19 étaient mis de l'avant. Mais comment peut-on s'assurer qu'une thérapie est efficace si elle n'a pas toujours l'effet escompté ? En effet même une forte corrélation entre deux phénomènes n'est pas garante d'un lien de cause à effet : le nombre de consultations pour des coups de soleil a beau être plus grand quand les bars laitiers vendent beaucoup de cornets, la crème glacée n'est pas pour autant la cause des insolation !

La notion de causalité n'est pas aussi facile à définir qu'on pourrait le croire. Elle a passionné les philosophes et les théologiens bien avant qu'elle préoccupe les statisticiens. Les premiers écrits connus à ce sujet remontent à Platon, au 4^e siècle avant notre ère. La question a même amené de grands esprits tels Immanuel Kant et Friedrich Nietzsche à s'interroger sur la nature déterministe de l'univers et l'existence du libre arbitre.

Dans une perspective déterministe, tout événement est prévisible (au degré de précision voulue) à condition de disposer d'une connaissance parfaite du passé et des lois de la nature. Le hasard n'est alors qu'une perception attribuable à une connaissance partielle de l'état du monde. La théorie

statistique de l'inférence causale adopte implicitement ce point de vue en supposant qu'une fois tous les facteurs pertinents pris en compte, un traitement a toujours le même effet.

Ceci étant on peut espérer pouvoir mesurer l'effet d'un traitement si l'on connaît les conditions précises dans lesquelles il a été administré. Néanmoins, cet objectif est souvent illusoire parce que les conditions peuvent être difficiles à déterminer, mesurer ou contrôler. Il est alors possible que l'effet observé soit dû à des facteurs inconnus plutôt qu'au traitement. Pour parvenir à isoler un éventuel effet de traitement, on doit donc se résoudre à formuler des hypothèses et à faire appel à des approches d'analyse statistique dont nous allons décrire le fondement.



Comment faire des comparaisons équitables?

Pour mesurer l'efficacité d'un traitement, il suffit en principe d'administrer celui-ci à un certain nombre de sujets et de comparer leur évolution à celle des membres d'un groupe témoin. Pour que la comparaison soit équitable, il faut toutefois s'assurer que la manière dont les sujets sont assignés au groupe traitement ou au groupe contrôle ne fausse pas les résultats en s'appuyant, par exemple, sur leur état de santé initial ou sur leurs chances de bien répondre à la thérapie.

Cette idée n'est pas nouvelle. Dès 1364, le savant florentin Pétrarque estimait que pour pouvoir comparer deux traitements de façon adéquate, il fallait que les groupes soient constitués d'un nombre égal de sujets malades qui soient tous du même âge, de mêmes mœurs et de même tempérament en plus d'évoluer dans le même environnement. C'est dans ces conditions seulement, pensait-il, qu'une comparaison pouvait être équitable.

De nos jours, les essais cliniques cherchent le plus souvent à assurer cette équité en assignant les traitements au hasard, c'est-à-dire essentiellement par un jet de dé (bien qu'en pratique, la randomisation soit plutôt faite par ordinateur). Cette façon de procéder garantit que les profils des groupes sont alors comparables en terme d'âge, de

sexe ou de toute autre caractéristique d'intérêt. Toute variation entre les groupes (hormis le traitement) n'est alors que le fruit du hasard.

Par exemple, supposons que l'on veuille étudier l'effet d'un régime à faible teneur en sel sur la prévention de l'hypertension artérielle (HTA) chez des sujets qui y sont prédisposés. Certains facteurs de risque déjà documentés doivent alors être pris en compte tels que l'âge et le sexe, puisque les hommes et les individus de plus de 60 ans sont davantage atteints. En déterminant au hasard qui suivra ou pas le régime faible en sel, on s'assure que les différents facteurs de risque tels l'âge, le sexe, les antécédents familiaux et autres sont présents en proportions à peu près égales au sein des deux groupes. Si une variation significative de la HTA est observée entre les groupes à la fin de l'étude, on pourra alors conclure que le traitement est en cause.

En somme, le recours à la randomisation permet de déduire qu'une corrélation est bel et bien le résultat d'une relation de cause à effet entre un traitement et un résultat observé. C'est le principe sur lequel reposent tous les essais cliniques et c'est ainsi, entre autres, que Santé Canada procède pour approuver un nouveau traitement ou un médicament.

Affaire classée?

S'il suffit de randomiser les traitements pour pouvoir établir un lien de cause à effet, pourquoi ne le fait-on pas toujours? Parce qu'il existe de nombreuses situations dans lesquelles cela s'avère impossible pour des raisons éthiques, matérielles ou logistiques. Si on soupçonne par exemple qu'un produit chimique est cancérigène, il serait contraire à l'éthique d'exposer volontairement des sujets à ses effets. Dans bien des cas, il s'avère aussi que l'exposition doit être prolongée pour qu'une augmentation de l'incidence du cancer soit observée. Avant de pouvoir conclure, il faudrait donc suivre les participants pendant des décennies, ce qui est coûteux et complexe.





Figure 1

Au début de la pandémie, divers traitements ont été proposés par des médecins qui cherchaient des moyens d'atténuer les effets de la COVID-19. Pour juger de l'efficacité de ces approches, on ne disposait alors que de données disparates et anecdotiques, par opposition à celles émanant d'un essai clinique contrôlé. Bien que l'on ait pu observer que tel ou tel traitement semblait mieux fonctionner qu'un autre pour certains groupes de personnes (par exemple les fumeurs ou les gens obèses), il est possible que les caractéristiques de ces sujets aient été associées à des facteurs cachés qui augmentent les chances de succès a priori. À l'inverse, certains traitements non homologués sont utilisés en dernier recours chez des patients particulièrement mal-en-point ; pas étonnant alors que leur efficacité soit faible !

Le problème des essais cliniques non randomisés, c'est précisément que les traitements peuvent avoir été choisis, sciemment ou pas, en fonction de facteurs connus ou cachés qui sont susceptibles d'influencer le résultat.

Pour mieux comprendre en quoi cela pose problème, revenons à l'exemple portant sur l'évaluation d'un régime à faible teneur en sel comme moyen de réduction du risque de HTA chez les hommes de plus de 60 ans et examinons en quoi les antécédents familiaux peuvent influencer les résultats d'une étude non randomisée.

La population cible est représentée à la figure 1, où les sujets en bleu ont des antécédents familiaux de HTA et les autres pas. Supposons que les sujets du sous-ensemble en jaune aient décidé par eux-mêmes de suivre un régime faible en sel. Comme la proportion d'hommes à risque élevé de HTA est plus grande dans ce groupe (5/10) que dans le reste de la population-cible (5/35), il serait trompeur de juger de l'efficacité du traitement en comparant (disons l'année suivante) la proportion de sujets hypertendus dans les deux groupes.

En effet, s'il existe une prédisposition familiale à la HTA, on s'attendrait à ce que de nombreux participants à l'étude finissent par être hypertendus, même en l'absence de toute forme de traitement ou si celui-ci s'avérait inefficace. Comment tenir compte

d'un tel handicap, qui peut jouer contre le traitement ou dans certains cas, qui sait, en sa faveur ?

Avant de pouvoir répondre à cette question, il est nécessaire d'introduire un peu de formalisme.

Formulation mathématique

Pour un sujet donné, appelons Y la valeur future de la variable d'intérêt (par exemple la tension artérielle moyenne du sujet dans un an). Présumément, la valeur de Y ne sera pas la même selon que ce sujet suive un régime à faible teneur en sel ou non. Si on dénote $Z = 1$ le fait que le sujet suive le traitement et $Z = 0$ le fait qu'il ne le suive pas, alors

$$Y = Z \times Y_1 + (1 - Z) \times Y_0,$$

où Y_z est la valeur future de la variable d'intérêt si $Z = z$. La distribution de Y_z (et en particulier sa moyenne) varie selon que $z = 0$ ou $z = 1$.

En terme d'espérance mathématique, l'effet moyen du traitement est alors égal à

$$E(Y_1) - E(Y_0).$$

Cette différence serait facile à estimer si, au lieu de comparer des groupes déséquilibrés comme dans la figure 2, on pouvait traiter tous les sujets et en même temps n'en traiter aucun de façon à comparer deux groupes identiques, comme dans la figure 3.

C'est toutefois impossible et



Figure 2

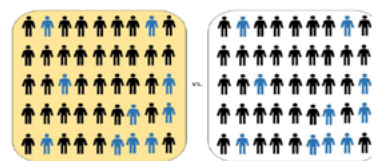


Figure 3

pour chaque sujet, on observera donc plutôt Y_0 ou Y_1 , mais pas les deux. Ce que les données permettent d'estimer, c'est alors une différence d'espérances conditionnelles, à savoir

$$E(Y_1 | Z = 1) - E(Y_0 | Z = 0).$$

Ceci ne pose pas problème dans le cadre d'une étude randomisée, car le fait d'avoir été assigné au groupe $z = 1$ ou au groupe $z = 0$ n'influence pas la distribution des variables Y_0 et Y_1 . On a donc

$$E(Y_0 | Z = 0) = E(Y_0), \quad (1)$$

$$E(Y_1 | Z = 1) = E(Y_1) \quad (2)$$

et l'effet de traitement peut être estimé par la différence entre les moyennes des deux groupes.

Dans une étude non randomisée, en revanche, la probabilité d'avoir $Z = 1$ peut dépendre de certaines caractéristiques du sujet, par exemple le fait qu'il ait des antécédents familiaux de HTA (dénote $X = 1$) ou qu'il n'en ait pas ($X = 0$). Les identités (1) et (2) ne sont alors plus valables et

$$E(Y_1 - Y_0) \neq E(Y_1 | Z = 1) - E(Y_0 | Z = 0).$$

Dans l'exemple portant sur la HTA, on peut résumer la situation au moyen des diagrammes des figures 4 et 5, dans lesquelles X représente la variable « antécédents familiaux » du sujet, Y dénote sa tension artérielle dans un an et Z ses chances de suivre le régime à basse teneur en sel. La direction de la flèche entre deux variables indique laquelle influence l'autre. Ainsi, dans la figure 4, X influence à la fois Y et Z , comme dans une étude non randomisée, alors que dans la figure 5 correspondant à une étude randomisée, X et Z ont un effet sur Y , mais X n'affecte pas Z .

Retour sur la randomisation

La clef de voûte d'une étude randomisée, ce sont les identités (1) et (2), c'est-à-dire le fait que l'assignation aux groupes traitement ($Z = 1$) et contrôle ($Z = 0$) ne dépend pas de facteurs influents ou « covariables » tels que X . Il suffit pour cela que l'assignation se fasse au hasard, mais il n'est pas nécessaire pour autant que la probabilité d'avoir $Z = 1$ soit égale à $1/2$.

Si par exemple on décidait que $\Pr(Z = 1) = 1/3$ (en ne faisant suivre le régime à un sujet que si, disons, on obtient 5 ou 6 lors du jet d'un dé équilibré), le groupe contrôle serait alors à peu près deux fois plus grand que le groupe traitement, mais on s'attendrait quand même à ce que les caractéristiques des deux groupes soient les mêmes en moyenne, en raison de la randomisation.

Or, on peut pousser le raisonnement un peu plus loin et faire en sorte que la probabilité de $Z = 1$ dépende de la valeur de X . Supposons qu'en présence d'antécédents familiaux, il y ait une chance sur deux que le sujet adhère au traitement, dénoté $\Pr(Z = 1 | X = 1) = 1/2$, mais qu'en l'absence de tels antécédents, $\Pr(Z = 1 | X = 0) = 1/7$.

Dans pareil cas, les sujets ayant des antécédents familiaux seraient surreprésentés dans le groupe traitement, ce qui empêcherait toute comparaison équitable avec le groupe contrôle. En revanche, et c'est là le nœud de l'affaire, il serait encore possible de comparer de façon équitable les sujets traités et non traités parmi ceux pour lesquels $X = 1$, puisqu'il y aurait eu randomisation au sein de cette strate. De même, on pourrait encore comparer de façon équitable les sujets traités et non traités parmi ceux pour lesquels $X = 0$.

Comment s'y prendre pour imiter un essai randomisé stratifié ?

Si l'étude est non randomisée, on n'a généralement aucune idée de la valeur de $\Pr(Z = 1 | X = x)$ pour $x = 0$ ou 1 . On peut toutefois essayer de la déduire des données. Pour la population représentée à la figure 1, par exemple, on a

$$\Pr(Z = 1 | X = 1) = 5/10 = 1/2,$$

$$\Pr(Z = 1 | X = 0) = 5/35 = 1/7.$$

En supposant que l'assignation des sujets aux différents groupes ait été faite au hasard selon ces probabilités, on peut alors comparer strate par strate le groupe traitement au groupe contrôle et calculer un effet de traitement pour les sujets ayant des antécédents familiaux, soit

$$\mu_{oui} = E(Y | Z = 1, X = 1) - E(Y | Z = 0, X = 1)$$

ainsi que pour les sujets qui n'en ont pas, soit

$$\mu_{non} = E(Y | Z = 1, X = 0) - E(Y | Z = 0, X = 0).$$

On peut ensuite estimer l'effet de traitement global en faisant une moyenne pondérée par strates, à savoir

$$\mu = 10/45 \times \mu_{oui} + 35/45 \times \mu_{non},$$

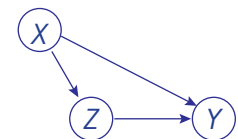


Figure 4

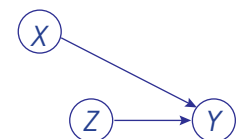


Figure 5

puisque 10 des 45 sujets représentés à la figure 1 avaient des antécédents familiaux et que 35 n'en avaient pas. On peut alors tirer des conclusions fiables quant à la valeur du traitement, sous réserve que l'assignation aux deux groupes ait été faite au hasard au sein de chacune des strates.

Évidemment, les choses se compliquent lorsque la variable X peut prendre plusieurs valeurs différentes, voire même une infinité, comme dans le cas de l'indice de masse corporelle, par exemple. Dans la pratique, il est aussi fréquent que le résultat (Y) et la probabilité d'être traité ($Z = 1$) soient influencées par un vecteur de covariables. Même si toutes les composantes de X sont discrètes, certaines combinaisons de catégories peuvent être rares ou n'avoir jamais été observées, ce qui rend encore plus périlleux le calcul du « score de pension » $\Pr(Z = 1 | X = x)$.

Approche par régression

La régression offre une autre façon de faire des comparaisons équitables. Cette méthode cherche à établir un lien entre la valeur moyenne de la variable d'intérêt, disons la tension artérielle Y mesurée dans un an, et plusieurs variables explicatives telles que le traitement reçu, l'âge, le sexe ou l'indice de masse corporelle.

Dans le cas le plus simple où il n'y a qu'une seule variable explicative X , disons l'âge, on peut supposer par exemple que l'espérance de Y est une fonction linéaire de X , c'est-à-dire

$$E(Y | X = x) = \beta_0 + \beta_1 x, \quad (3)$$

où β_0 est l'ordonnée à l'origine et β_1 est la pente de la droite. Pour estimer les valeurs de ces deux paramètres, on doit disposer d'observations $(x_1, y_1), \dots, (x_n, y_n)$ formant un échantillon de taille n .

Si toutes ces paires de points sont alignées, le calcul de β_0 et de β_1 est un jeu d'enfant. Dans la pratique, on obtient plutôt un graphique semblable à celui de la figure 6, parce que même si la relation (3) est un juste reflet de la réalité, la valeur Y_i observée chez le sujet i âgé de x_i années ne sera pas égale à $\beta_0 + \beta_1 x_i$. En effet, la tension artérielle fluctue

constamment et sera donc vraisemblablement différente de sa valeur moyenne au moment de la prise de mesure.

Pour tenir compte de cette variation individuelle, on suppose que la valeur future Y_i de la tension artérielle du sujet i d'âge $X = x_i$ s'exprime sous la forme

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

où ε_i représente l'écart à la moyenne spécifique à ce sujet.

Dans ces circonstances, estimer la pente et l'ordonnée à l'origine revient à trouver la droite qui s'ajuste le mieux aux données. Dans la figure 6, trois solutions possibles sont proposées : une droite en tiret (---), l'autre en pointillé (...), et la troisième en trait continu. Chacune des droites exprime une relation possible entre l'âge du sujet (abscisse, en années) et sa tension artérielle diastolique moyenne (ordonnée, en mmHg). La première est clairement un mauvais choix car elle passe en-dessous de la plupart des points. Cependant, il est plus difficile de choisir entre les deux autres.

Pour objectiver la recherche d'une solution, on a souvent recours au principe des moindres carrés, qui consiste à trouver les valeurs de β_0 et de β_1 pour lesquelles la fonction

$$L(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$

est minimisée. Il se trouve ici que c'est la ligne pleine qui correspond à la solution.

L'approche par régression peut être étendue et adaptée au cas où l'espérance de Y dépend de plusieurs covariables. En particulier, si l'on soupçonne que la tension artérielle des sujets dépend à la fois de leur âge X et de leur adhérence à un régime à faible teneur en sel ($Z = 1$) ou non ($Z = 0$), on peut supposer que

$$E(Y | X = x, Z = z) = \beta_0 + \beta_1 x + \beta_2 z, \quad (4)$$

voire même

$$E(Y | X = x, Z = z) = \beta_0 + \beta_1 x + \beta_2 z + \beta_3 xz. \quad (5)$$

Après avoir estimé les paramètres par la méthode des moindres carrés, on peut alors tracer les droites correspondant aux groupes traitement ($z = 1$) et contrôle ($z = 0$). C'est ce qui a été fait pour le modèle (5) dans la

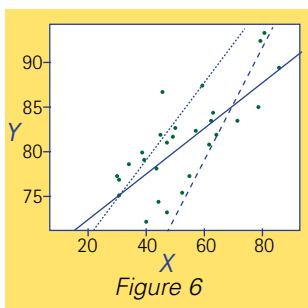


Figure 6

figure 7, où deux droites ont été tracées, une en bleu pour le groupe traitement ($z = 1$) et l'autre en rouge pour le groupe contrôle ($z = 0$).

Dans le modèle (4), les deux groupes ont la même pente, soit β_1 , mais leurs ordonnées à l'origine sont différentes, soit β_0 pour le groupe $z = 0$ et $\beta_0 + \beta_2$ pour le groupe $z = 1$. Dans le modèle (5), les deux groupes ont aussi des ordonnées à l'origine différentes, les mêmes que dans le modèle (4), mais en plus leurs pentes diffèrent : la pente est encore β_1 pour le groupe $z = 0$ mais elle vaut maintenant $\beta_1 + \beta_3$ pour le groupe $z = 1$.

Une fois que les paramètres ont été estimés pour chaque valeur $X = x$, on peut aisément calculer l'effet du traitement. Pour le modèle (5), on trouve

$$\begin{aligned} \mu_x &= E(Y|X=x, Z=1) - E(Y|X=x, Z=0) \\ &= \beta_2 + \beta_3 x. \end{aligned}$$

Si la variable X est le seul facteur susceptible d'influencer la valeur de Y , hormis le traitement lui-même, la relation peut alors être considérée comme causale. Si la variable X est vectorielle, on doit alors faire appel à une version plus complexe du modèle de régression et à la notion de score de propension.

Un exemple concret

Pour illustrer les méthodes présentées plus haut, considérons la greffe de cellules-souches hématopoïétiques allogéniques (AHCT), qui offre des perspectives de guérison aux hémopathies malignes. Cette procédure consiste à transplanter chez un patient des cellules-souches récoltées de la moelle osseuse, du sang périphérique ou du cordon ombilical du donneur. Ces cellules-souches engendrent la production de plaquettes et de globules blancs et rouges, ce qui permet à terme de restaurer le système immunitaire du receveur, dont l'intégrité est compromise. Malheureusement, la maladie aiguë du greffon contre l'hôte (MAGH) est observée chez certains receveurs de cellules-souches allogéniques, entraînant diverses complications pouvant provoquer la mort dans 20 à 40% des cas. Si cette maladie se développe et résiste aux tentatives de traitement précoces,

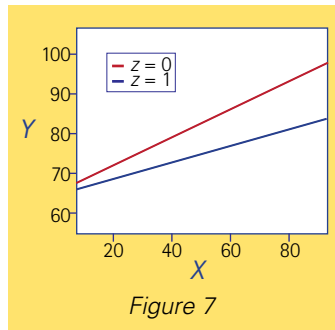


Figure 7

elle peut être combattue au moyen d'immunosuppresseurs standard ou de thérapies non spécifiques, dites NHTL, de déplétion massive des lymphocytes T (la lettre T est l'abréviation de thymus, nom de l'organe dans lequel leur développement s'achève). Des données non expérimentales suggèrent toutefois que les thérapies NHTL sont nocives.

La figure 8 indique la proportion de patients exempts de MAGH en fonction du temps, exprimé en mois depuis la greffe. La courbe rouge correspond aux patients ayant reçu la thérapie NHTL ; la courbe bleue correspond au traitement standard. La courbe rouge descend beaucoup plus rapidement que la bleue, de sorte qu'après 20 mois, par exemple, à peine 15% des patients sous thérapie NHTL restent exempts de MAGH, alors que près de 30% des patients sous traitement standard ne sont pas encore affectés.

À première vue, il semble donc que la thérapie NHTL n'ait aucun avenir. Toutefois, une analyse plus poussée des données révèle que le cancer des patients ayant reçu la thérapie NHTL était généralement plus avancé au moment de la greffe et que ces personnes avaient des risques de MAGH plus grands en raison d'un mauvais appariement donneur-receveur. Ces deux facteurs sont susceptibles d'influencer les résultats. En effet, compte tenu des différences entre les profils des patients assignés à l'un ou l'autre des deux traitements, on s'attendrait a priori à obtenir de moins bons résultats pour ceux qui sont sous thérapie NHTL, peu importe sa nature.

La régression linéaire s'avère ici insuffisante pour analyser adéquatement les données de cette étude, mais en faisant appel à des techniques plus avancées du même acabit, on peut montrer que si la thérapie NHTL conduit généralement à de moins bons résultats, elle s'avère néanmoins bénéfique pour certains types de sujets. Il appert aussi qu'en négligeant les facteurs mentionnés plus haut, on peut être amené à recommander le mauvais traitement chez environ 5% des patients.

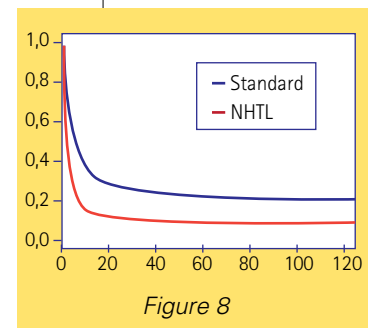


Figure 8