

Prebunking : les vaccinés hospitalisés

En août 2021, pour encourager sa population à se vacciner, le Ministère des Solidarités et de la Santé en France a créé un spot TV qui conclut comme suit :

[Pour un vaccin efficace à 89%, la surprise de l'hospitalisation en cas de vaccination est 11% de la surprise en cas de non-vaccination, ce qui correspond à un rapport de surprise de 1 neuvième. Bon, y a des subtilités derrière cette notion d'efficacité vaccinale, et on y reviendra d'ailleurs dans la prochaine vidéo !]

[à savoir ne pas être vacciné]

[On peut débattre de tout sauf des chiffres. Aujourd'hui, en France, 8 personnes sur 10 hospitalisées à cause du covid ne sont pas vaccinées.]

<https://www.youtube.com/watch?v=iLi5kkTGfRo>

De la même façon, en Suisse, le Conseiller d'État Valaisan Mathias Reynard signale le fait que 90% des hospitalisés ne sont pas vaccinés. Cette statistique est ainsi utilisée pour insister sur l'importance et l'efficacité de la vaccination de masse, et j'imagine que beaucoup de pro-vax parmi vous l'auraient retweeté si vous étiez sur Twitter et si vous aviez vu ce chiffre.

<https://twitter.com/MathiasReynard/status/1425810217694945281>

Cependant, aujourd'hui, je vais essayer de vous convaincre que ce genre de statistique est en fait très problématique à partager. Non pas parce qu'elle est erronée. Mais parce qu'elle est en fait trompeuse. En particulier, elle ne permet en fait pas du tout de conclure à l'efficacité du vaccin, et donc en l'importance de se faire vacciner. En tout cas pas seule.

Pire encore, dans les mois à venir, surtout en France, la mise en avant de telles statistiques va même probablement conduire à un chaos informationnel. Et à l'approche des élections, avec de telles habitudes, on risque fort d'assister à la popularité de mouvements très mal informés, au triomphe de campagnes de désinformations et à des révoltes violentes dangereuses.

$P[A|B] \neq P[B|A]$

Si vous avez suivi cette série, vous m'avez entendu répéter l'expression suivante : la vraisemblance des données n'est pas la crédence de la théorie. Ou dit autrement, la probabilité $P[D|T]$ de données D sachant une théorie T n'est en générale pas égale à la probabilité inverse $P[T|D]$ de la théorie T sachant les données D.

De façon plus générale, pour toutes variables A et B, la probabilité $P[A|B]$ de A sachant B n'est pas égale à la probabilité inverse $P[B|A]$ de B sachant A. Et dans le cas des vaccinés hospitalisés, cette distinction est absolument critique à avoir en tête.

En effet, notamment dans un pays où le taux de vaccination est très élevé, comme c'est désormais le cas en France, il faut en fait s'attendre à ce que la probabilité $P[V|H]$ d'être vacciné sachant qu'on est hospitalisé soit élevée, alors que la probabilité inverse $P[H|V]$ d'être hospitalisé sachant qu'on est vacciné est faible.

Ou dit autrement, ce qu'il va probablement finir par arriver, c'est que la plupart des vaccinés ne seront jamais hospitalisés, quand bien même la plupart des hospitalisés seront vaccinés.

<https://covidtracker.fr/vaccintracker/>

Et le jour où ça va arriver... aïe, aïe, aïe, aïe, aïe... ça va être la cacophonie avec des arguments moisis de partout, le tout appuyé sur des statistiques parfaitement factuelles, mais trompeuses... Surtout avec les élections présidentielles, et des candidats qui vont constamment chercher à plaire à leur électorat...

D'ailleurs, de tels chiffres ont d'ailleurs déjà été constatés dans le cas d'Israël, où le taux de vaccination a atteint 90% chez les personnes âgées, et où la majorité des hospitalisés sont des personnes vaccinées. Clairement, ce taux de vaccinés parmi les hospitalisés est un très mauvais chiffre à communiquer, car il est si facile de mal l'interpréter.

https://www.liberation.fr/checknews/covid-19-est-il-vrai-que-la-majorite-des-personnes-hospitalises-en-israel-sont-vaccinees-20210809_DNY3EN73LBFFJB7CUTWTYCFUI/

Et malheureusement, les sceptiques de l'efficacité vaccinale, ou les campagnes de désinformation qui veulent affaiblir un pays ennemi, vont probablement se faire une joie de relayer massivement de telles statistiques. Or, comme on va le voir, prébunker ou débunker ces statistiques de vaccinés parmi les hospitalisés, c'est déjà horriblement difficile, même auprès de gens assez à l'aise en maths et suffisamment en mode explorateur pour réfléchir calmement !

Surtout qu'il n'y aura pas uniquement cette statistique à débunker. Tellement de chiffres circulent, venant de tellement de pays aux contextes différents, qu'en cherry-picking à gauche et à droite, on peut trouver toutes sortes de statistiques qui suggèrent tout et son contraire, en exploitant typiquement la confusion omniprésente entre une probabilité conditionnelle $P[A|B]$ et la probabilité inverse $P[B|A]$.

Or, même des scientifiques, et même des statisticiens, ont un mal fou à distinguer la probabilité de A sachant B de la probabilité de B sachant A. Ce sophisme est même encore au cœur de la méthode scientifique par p-value !

<https://www.youtube.com/watch?v=jy9b1HSqtSk>

Et bien sûr, dans les mois à venir, on risque aussi d'assister à des interviews de gens qui ne voulaient pas être vaccinés, mais qui à cause des circonstances se sont sentis contraints d'être

vaccinés, et qui ont néanmoins dû être hospitalisés, avec possiblement des séquelles à long terme, et qui témoigneront de l'inutilité des vaccins, probablement avec un esprit très contestataire... Et ça risque de monter en flèche et d'influencer ensuite les positionnements politiques... aïe, aïe, aïe, aïe, aïe...

Bref. Comme vous le comprenez, je suis très loin d'être confiant pour le futur. À moins d'efforts massifs pour soutenir l'information de qualité, sa production et surtout sa recommandation, la crise épistémique risque de continuer à avoir des conséquences dramatiques, à la fois sur le plan médical, mais aussi en termes de géopolitique, d'environnement et de justice sociale.

#Tournesol wink wink...

Trois exemples intuitifs

Bon, c'est bien beau de se plaindre de la mésinformation. Mais il ne suffit pas de la noyer pour la combattre. Si on veut améliorer le niveau épistémique, il est critique de parvenir à mieux réfléchir soi-même, et à aider les autres à mieux réfléchir. Et donc, pour finir cette vidéo, je vais essayer de vous aider à mieux sentir la raison pour laquelle on peut avoir $P[A|B] \neq P[B|A]$, pour que vous ayez davantage le réflexe de vous demander si une statistique du genre "beaucoup de B sont A" est vraiment ce qu'il est pertinent de constater.

Commençons par des exemples pour avoir une intuition instinctive. Et pour cela, je vous invite vivement à revoir l'épisode 21, où on avait vu que la plupart des musulmans ne sont en fait pas arabes, quand bien même la plupart des arabes sont musulmans. En effet, certains arabes ont d'autres religions. Certains sont athées, d'autres catholiques. Mais ceux-ci sont rares parmi les arabes. La plupart des arabes sont musulmans.

Néanmoins la plupart des musulmans ne sont pas arabes. Et ça, c'est parce qu'il y a énormément de gens en Asie, et qu'une fraction, certes faible, mais non-négligeable, d'entre eux sont musulmans, notamment en Indonésie, en Chine ou en Inde. Gardez donc bien cet exemple en tête. La plupart des arabes sont musulmans, mais la plupart des musulmans ne sont pas arabes. Et c'est pas mal parce qu'il y a énormément de gens, notamment en Asie, qui sont musulmans mais ne sont pas arabes. Plus encore que de gens qui sont arabes.

Ou pour prendre un autre exemple encore tiré de l'épisode 22, la plupart des prix Nobel aiment le chocolat, mais la plupart de ceux qui aiment le chocolat n'ont pas de prix Nobel. Cet exemple devrait paraître un peu plus convaincant... Tout ça parce que, de façon encore plus massive, il y a énormément de gens qui aiment le chocolat et n'ont pas de prix Nobel. Beaucoup plus que de gens qui ont un prix Nobel.

Enfin un troisième exemple classique, c'est celui du test médical imparfait pour des maladies très rares, comme en parlent 3Blue1Brown, Veritasium ou encore Monsieur Phi. Même si le test

est valide à 99%, si la maladie ne touche qu'une personne sur un million, alors les malades ont presque tous un test positif, mais ceux qui ont un test positif sont quasiment tous non-malades.

Et bien, de la même façon, en Israël, la plupart des hospitalisés sont vaccinés, mais la quasi-totalité des vaccinés ne sont pas hospitalisés. Tout ça parce qu'il y a énormément de gens qui sont vaccinés et ne sont pas hospitalisés. Beaucoup plus que de gens qui sont hospitalisés.

Voilà. Je vous invite à remuer tous ces exemples dans votre tête, et à en chercher d'autres. C'est souvent beaucoup plus efficace de retenir un principe, si on parvient à trouver soi-même un exemple marquant qui l'illustre parfaitement.

Explication avec le curseur de Turing

Pour les gens un peu plus matheux parmi vous, je vous propose une explication plus fine de ce phénomène, qui fait appel au curseur de Turing. Ça va demander plus d'efforts de votre part, mais si vous saisissez très bien le curseur de Turing, alors vous pourrez résoudre beaucoup plus de paradoxes apparents qu'avec une poignée d'exemples représentatifs d'un cas très précis.

Rappelez-vous, le curseur de Turing permet d'appliquer rigoureusement la formule de Bayes dans un cas où il n'y a que deux théories à comparer. Ça tombe bien, c'est notre cas, si on considère les théories "pas vacciné" et "vacciné" pour un individu donné.

Mettons donc pas vacciné et vacciné à chaque bout d'une droite. La position initiale du curseur représente ensuite la probabilité qu'un français tiré au hasard soit vacciné. Au fil des derniers mois, ce curseur n'a cessé de se déplacer vers la droite, au fur et à mesure que le taux de vaccination augmente. En fait, il penche désormais vers la droite puisque le taux de vaccination a dépassé 50%. Et avec, parmi 100 français, 63 personnes ayant reçu 2 doses en France et 28 n'en ayant reçu aucune, on en est même à un peu plus d'un bit du côté de la vaccination.

[$\log_2(63/28) \approx 1,1$]

OK. Maintenant, essayons d'estimer la probabilité qu'une personne soit vaccinée, sachant qu'elle est hospitalisée. Pour cela, il faut estimer comment la donnée "la personne est hospitalisée" déplace le curseur de Turing. Et ça, ça dépend en fait du rapport de surprise. À quel point devrions-nous être davantage surpris par l'hospitalisation d'un individu, sachant qu'elle est vaccinée, par opposition à sachant qu'elle n'est pas vaccinée ?

Alors, je vous épargne les détails de calcul, mais cette question est en fait directement reliée à l'efficacité du vaccin. Pour un vaccin efficace à 90%, on devrait être 9 fois plus surpris par

l'hospitalisation, si la personne est en fait vaccinée. C'est même ça, la définition de l'efficacité — à des subtilités importantes près, dont on parlera dans la prochaine vidéo.

[$\log_2(9) \approx 3,2$]

Or un facteur de 9, ça correspond à un peu plus de 3 bits d'information. En déplaçant le curseur de 3 bits, dans la direction de la théorie qui explique mieux l'hospitalisation, à savoir être non-vacciné, on constate que, a posteriori, le curseur de Turing penche 2 bits en faveur de la non-vaccination. Ou dit autrement, on s'attend alors à ce qu'il soit 4 fois plus probable qu'un hospitalisé soit vacciné que non-vacciné. Ou dit autrement, sur 5 hospitalisés, 4 ne devraient pas être vaccinés.

OK. Mais là où tout cela devient piégeux, c'est si la vaccination augmente. Imaginons que désormais 90% des français sont doublement vaccinés, contre 10% de non-vaccinés. Désormais, a priori, avant de savoir si un individu est hospitalisé, le curseur de Turing penche de 3,2 bits en faveur de la vaccination. Mais alors, a posteriori, après le déplacement de 3,2 bits dû à la donnée "l'individu est hospitalisé", on se retrouve à 0. Autrement dit, si 90% des français sont doublement vaccinés et si le vaccin est efficace à 90%, alors, parmi les hospitalisés, il y aura autant de vaccinés que de non-vaccinés.

Et en fait, mon analyse ici est en fait trop simplifiée, car elle omet plein de facteurs de confusion, comme l'âge de la personne, ses antécédents médicaux et son application des gestes barrières. Si les personnes âgées sont plus vaccinées, si les personnes déjà infectées se vaccinent moins et si les vaccinés prennent moins soin de leurs gestes barrières, alors il faut s'attendre à ce que la plupart des hospitalisés soient vaccinés.

<https://www.blick.ch/fr/news/suisse/lofsp-publie-des-chiffres-concrets-seules-des-personnes-de-plus-de-80-ans-sont-mortes-du-coronavirus-parmi-les-vaccines-id16729639.html>

Et ce, même si le vaccin demeure efficace à 90%...