Société de Calcul Mathématique, S. A.
Algorithmes et Optimisation

Méthodes probabilistes en Epidémiologie :

Analyse critique de l'Etude [Draper]

Société de Calcul Mathématique SA

rédaction : Bernard Beauzamy et Manon Baradat

juillet 2009

Nous réalisons une analyse critique de l'étude :

Cancer infantile en lien avec la distance aux lignes hautes tensions de distribution de l'électricité en Angleterre et au Pays de Galles : une étude cas – témoins.

Gerald Draper, *Directeur de recherche* (*honorary senior research fellow)*1, Tim Vincent, *Chargé de recherche (research officer)*1, Mary E Kroll, *Statisticien (statistician)*, John Swanson, *Conseiller scientifique* (*scientific adviser).*

I. Présentation générale

Le résultat principal de l'étude est : comparé aux enfants qui vivent à plus de 600 m d'une ligne à la naissance, les enfants vivant à moins de 200 m ont un risque relatif de leucémie de 1.69 et ceux entre 200 et 600 un risque relatif de 1.23.

(Le "risque relatif" est défini par rapport à la population de référence : il est mis à 1 pour celle-ci, et un risque de 1.6 signifie qu'il y a aura 1.6 fois plus de leucémies dans la population test que dans la population de référence, pour 1 000 habitants.)

Mais le résultat de l'étude est en contradiction avec les chiffres que celle-ci fournit : comme dit Poincaré, le calcul des probabilités ne devrait pas empêcher d'avoir du bon sens.

L'auteur dit en effet :

"L'incidence de la leucémie chez l'enfant, en Angleterre et au Pays de Galles, est de 42 par million et par an. Nous estimons qu'il y a 9.7 millions d'enfants, dont 80 000 à moins de 200 m d'une ligne et 320 000 entre 200 et 600 m. "

Dans ces conditions, le nombre normal de leucémies chez l'enfant, à proximité immédiate des lignes, devrait être, par an :

$$n_1 = 42 \times 10^{-6} \times 80\,000 \approx 3.36$$

et dans la bande 200 – 600 m :

$$n_2 = 42 \times 10^{-6} \times 320\,000 \approx 13.44$$

soit en tout, pour la population à risque : $n_1 + n_2 \approx 16.80$ leucémies par an.

L'étude concerne 33 années (1962 à 1995). On devrait donc s'attendre à un nombre total de leucémies, au voisinage des lignes, de l'ordre de $33 \times (n_1 + n_2) \approx 554$. Or les auteurs ne dénombrent que 322 cas de leucémies au voisinage des lignes (5+19+40+44+61+78+75).

Il y a donc un défaut de bon sens, que l'on peut constater dès le départ. Les auteurs mettent ensuite en œuvre tout un arsenal statistique, que nous allons maintenant analyser, pour obtenir le résultat qu'ils souhaitent ; quand on veut démontrer quelque chose, on y arrive toujours.

Avant d'entrer dans l'analyse statistique, arrêtons-nous un moment sur une remarque de bon sens. Les calculs ci-dessus peuvent avoir logiquement trois conclusions :

- ou bien les lignes HT protègent de la leucémie ;
- ou bien il y a moins d'enfants dans les zones considérées (pyramide des âges différente) ;
- ou bien le recensement n'est pas correct (on n'a pas enregistré tous les cas de leucémies).

De manière générale, dans cette étude, les données ne paraissent pas très fiables. On ne sait pas si les populations étaient les mêmes il y a 10, 20, 30 ans, si les lignes HT étaient en fonction, etc.

II. Analyse statistique de l'étude Draper

L'erreur méthodologique commise par l'étude Draper se trouve dans la définition de la population de référence, obtenue par "appariements" ; nous avons expliqué plus haut (voir Règle 1, chapitre I) en quoi cette approche était erronée. Les "cas-témoin" ne sont pas statistiquement représentatifs de l'ensemble de la population de référence (population générale) ; les auteurs s'en seraient aperçus s'ils avaient fait les remarques de bon sens développées plus haut.

III. Autres erreurs méthodologiques de l'étude Draper

Le choix du domicile à la naissance ne semble pas réellement pertinent pour juger si oui ou non les gens ont habité (et pendant combien de temps ?) au voisinage d'une ligne HT. Il suppose que les sujets sont restés à leur domicile 24 h sur 24 pendant 14 ans. Or, la plupart des enfants passent leur journée à la crèche ou à l'école. Le calcul est donc correct pour l'exposition nocturne mais pas pour la journée. De plus, l'adresse prise en compte est l'adresse de naissance, mais rien ne prouve que les sujets vivent leurs 14 premières années à leur domicile de naissance. L'étude [Huss], au contraire, prend soin d'évaluer le temps de résidence.

Pourquoi les enfants adoptés sont-ils éliminés ? Ils sont exposés comme les autres !

Les informations de distance, reposant sur un code postal, sont douteuses. Les auteurs disent "la moitié des enfants atteints de leucémie dans cette étude ont la même adresse de résidence à la naissance et au moment du diagnostic", ce qui signifie que la moitié a changé d'adresse ! Mais n'oublions pas que le taux de risque est seulement de 42 par million : c'est très faible, les nombres seront très petits (quelques unités, comme on l'a vu), et une variation de $\pm 50\,\%$ est considérable.

Revenons aux chiffres donnés par [Draper] ; nous en extrayons le tableau :

distance à la ligne HT(m)	nb de cas de leucémie
0 – 100	24
100 – 200	40
200 – 300	44
300 – 400	61
400 – 500	78
500 – 600	75

Tableau 1 : chiffres de leucémies en fonction de la distance à la ligne HT

Il s'agit du nombre de cas enregistrés sur 33 ans. Les populations concernées ne sont pas mentionnées exactement : ceci est inacceptable, s'agissant d'une publication, car ces chiffres sont évidemment essentiels. Il peut y avoir plus ou moins de leucémies parce qu'il y a plus ou moins de gens !

Pour les deux premières tranches, 0 – 100 m, et 100 – 200 m, les auteurs donnent une estimation totale de 80 000 personnes (pour les deux tranches) et une estimation totale de 320 000 personnes pour la somme des quatre dernières tranches.

Prenons les deux premières tranches : nous avons 64 cas pour 80 000 personnes en 33 ans, soit une incidence par an de 24 pour un million : nous sommes très au-dessous de la moyenne nationale (42 cas par million et par an).

Prenons maintenant la troisième tranche, 200 – 300 m, et attribuons-lui le quart de la population des quatre dernières tranches, soit à nouveau 80 000 personnes. Nous avons 44 cas pour 80 000 personnes en 33 ans, soit une incidence de 17 par million et par an : nous sommes encore très au-dessous de la moyenne nationale !

Et on constate, contrairement aux affirmations de l'étude, que ce sont les zones les plus voisines des lignes qui ont le plus faible taux de leucémies.

Prenons la dernière tranche, 500 – 600 m ; le champ généré par les lignes HT (500 A à 500 m) est le même que celui généré par une installation de cuisine (1 A à 1 m). Le taux d'incidence des leucémies de l'enfant est cependant de 28 par million et par an : il reste un effet protecteur des lignes HT (puisque la moyenne est 42 par million et par an), bien que nous soyons très loin !

Concernant les cancers autres que la leucémie, le rapport conclut : "aucun excès de risque en lien avec la proximité des lignes n'a été trouvé pour les autres cancers infantiles". Pourtant, aux abords des lignes, le risque relatif est 0.44 et il augmente significativement lorsqu'on s'éloigne. On peut donc conclure plus vigoureusement, si l'on en croit l'étude : il y a une diminution du risque de tumeurs du système nerveux central et du cerveau avec la proximité à la ligne, ou encore : les lignes THT protègent des tumeurs du système nerveux central et des tumeurs du cerveau.

IV. Que retenir de cette étude ?

Si les chiffres rappelés dans le tableau ci-dessus étaient fiables, on pourrait en déduire que, contrairement à ce qu'annoncent les auteurs, les lignes HT protègent contre la leucémie de l'enfant. Cette hypothèse n'est en rien plus absurde que l'hypothèse contraire, selon laquelle les lignes HT favorisent cette leucémie. Les champs magnétiques et électriques sont utilisés de manière thérapeutique, et rien ne dit que, dans certaines conditions, ils ne puissent avoir un effet bénéfique.

Malheureusement, notre conclusion est ici que rien ne permet de l'affirmer : ces données ne sont pas pertinentes. Nous recensons 24 cas de leucémies, dont le domicile à la naissance est à moins de 100 m des lignes, mais combien de temps ces enfants sont-ils restés exposés ? Le domicile à la naissance n'a été conservé que pour la moitié des sujets, disent les auteurs. A l'inverse, bien des enfants, nés n'importe où, ont pu s'établir au voisinage des lignes ; ceux parmi eux qui ont eu une leucémie ne sont pas recensés. Ne disposant d'aucune évaluation de la population soumise à l'influence des lignes HT, nous ne retenons de cette étude aucune conclusion scientifiquement fondée.

Annexe

La physique du problème : lignes HT et champ magnétique

Les auteurs des études statistiques que nous analysons ici semblent faire reposer leurs raisonnements uniquement sur les statistiques (qu'ils maîtrisent mal), et non sur la physique, qu'ils ne maîtrisent pas du tout. La question de savoir si les lignes HT sont dangereuses est tout à fait légitime, encore faut-il qu'elle soit abordée avec les connaissances appropriées.

Une ligne électrique crée un champ électrique et un champ magnétique. Les auteurs semblent vouloir se limiter aux effets du champ magnétique. Pour celui-ci, les faits physiques sont les suivants :

L'intensité du champ magnétique (mesurée en Tesla) est proportionnelle à l'intensité du courant dans la ligne (et non à la tension !!) ; pour une ligne rectiligne de grandes dimensions (cas usuel), l'intensité du champ magnétique s'exprime par la formule :

$$B = c\frac{I}{d} \qquad (1)$$

où c est une constante, I l'intensité du courant dans la ligne (en Ampères) et d la distance à la ligne (en mètres).

La formule est démontrée plus bas.

Par conséquent, si vous êtes à 100 m d'une ligne transportant 500 Ampères, vous recevez le même champ magnétique que si vous êtes à 1 m du câble alimentant une cuisinière (5 Ampères). La fréquence des deux champs est la même (50 Hz).

Dans l'étude [Draper], diverses hypothèses sont faites en ce qui concerne la dépendance du champ par rapport à la distance : on y rencontre $1/d$, $1/d^2$, $1/d^3$; cette ignorance du phénomène physique en cause est tout de même étonnante, car enfin si l'effet était en $1/d^3$, à 100 m il n'en reste plus que le millionième !

Démontrons la formule (1). Elle est claire intuitivement, car dans le cas d'un conducteur rectiligne infini (ou simplement de grande longueur par rapport à la distance à l'observateur) le champ a nécessairement une symétrie cylindrique. La quantité reçue dans une portion de l'espace à distance r et d'épaisseur dr est proportionnelle au périmètre du cercle, soit $2\pi r dr$.

Donnons aussi une démonstration complète, issue de la formule de Biot et Savart :

$$B(M)=\frac{\mu_0}{4\pi}\int_C I\,\frac{\overrightarrow{dl}\wedge\overrightarrow{SM}}{\left\|SM\right\|^3}$$

où $B(M)$ est le champ magnétique en M, μ_0 la perméabilité magnétique du vide, C le conducteur (ici l'axe des x), $\overrightarrow{dl}$ l'élément de longueur sur l'axe Ox, et S le point courant.

Mettant le point M sur l'axe Oy avec l'ordonnée d, nous obtenons :

$$B(M)=\frac{\mu_0 I}{4\pi}\int_{-\infty}^{+\infty}\frac{(x^2+d^2)^{1/2}\sin(\vartheta)dx}{(x^2+d^2)^{3/2}}$$

où ϑ désigne l'angle (Ox,SM), et donc $\sin(\vartheta)=\frac{|x|}{\sqrt{x^2+d^2}}$. Reportant dans l'expression précédente, nous obtenons :

$$B(M)=\frac{\mu_0 I}{2\pi}\int_0^{+\infty}\frac{xdx}{(x^2+d^2)^{3/2}}$$

Le changement de variable $x=d\cdot y$ donne :

$$B(M)=\frac{\mu_0 I}{2\pi d}\int_0^{+\infty}\frac{ydy}{(1^2+y^2)^{3/2}}=\frac{\mu_0 I}{2\pi d},$$

comme annoncé.