

Les Statistiques portent-elles atteinte à la liberté?

Pour un regard critique sur les informations chiffrées

François Sauvageot

Maître de conférences à l'Université Paris 7

Animateur à l'IREM-Paris 7

Comment réagit-on face aux statistiques ?

- Ceux qui possèdent une culture scientifique, peut-être portés par le sentiment de pouvoir appréhender de façon plus rationnelle une problématique, peuvent se laisser piéger par ce qu'ils espèrent être un reflet approximatif de la réalité et qu'ils ne cherchent pas à réfuter.
- Ceux qui au contraire sont réticents face à ces quantifications de « l'humain » sont parfois conduits à se résoudre à accepter ce que les médias appellent grossièrement des « chiffres » et peuvent alors passer du refus à la soumission totale, incapables de contester la validité de ces « chiffres », de les contrôler et d'appréhender ce qu'ils mesurent réellement.

Quelles en sont les conséquences ?

- On peut en toute sincérité convaincre grâce à un argument qui se révèle en fait faux et entraîner des décisions désastreuses, ou plus simplement donner de fausses idées sur une réalité complexe.
- On peut se décharger de toute responsabilité, se dédouaner des décisions prises, en s'abritant derrière ces « chiffres » obscurs mais inattaquables.

Quelles solutions ?

- Il ne s'agit pas d'interdire les statistiques, ou plus généralement les informations quantifiées !
- Je plaide pour une meilleure compréhension des phénomènes, une éducation scientifique **en profondeur** tout particulièrement des professionnels qui « produisent des chiffres », mais peut-être plus encore de ceux qui les utilisent (politiques par exemple) et de ceux qui les diffusent et les vulgarisent (médias par exemple).

Plus de maths au lycée ?

- Oui et non. Il ne sert à rien de savoir calculer une moyenne si on ne comprend pas ce qu'elle représente et comment on peut s'en servir pour dire quelque chose qui a réellement un sens.
- Même des notions simples, comme un pourcentage, nécessitent des explications profondes.
- D'un autre côté, des notions plus techniques comme une probabilité conditionnelle ou un estimateur peuvent être décrites et comprises sans entraîner une abondance de techniques.

Des erreurs de pourcentages

- Un pourcentage mesure le rapport d'une partie à un tout. Si on change le « tout », on change le sens du pourcentage : comparer le « score » d'un candidat à deux élections différentes n'a pas de sens.
- Ainsi lors de l'élection présidentielle de 2002, certains ont pu être frappés par la progression du Front National. Il y a pourtant de nombreux facteurs qui font que c'est une erreur de penser cela :
 - La liste des candidats n'était pas la même et, dans une compétition, le score de chacun dépend de ses opposants !
 - Le taux de participation était plus faible en 2002.
 - Le rapport entre le nombre de personnes ayant voté pour le Front National et le nombre d'**inscrits** est resté constant entre 1988, 1995 et 2002 : environ 11.5%.
- Il faut garder à l'esprit que les « chiffres » médiatisés représentent un classement (comme dans une course) et ils ne sont en aucun cas une description profonde.

Des faits mieux décrits

Un pourcentage est une certaine forme de résumé, mais il est parfois essentiel de fournir plusieurs éclairages de la même situation, plusieurs résumés.

Ainsi il est certainement plus parlant de rappeler que Lionel Jospin a obtenu 7 097 786 voix en 1995 contre 4 610 749 en 2002.

Seul compte, pour gagner, le pourcentage de suffrages exprimés en faveur d'un candidat, mais si l'on veut comprendre les transformations de l'électorat, il faut tenir compte de tous les électeurs potentiels, c'est-à-dire de tous ceux qui sont inscrits sur les listes électorales (voire même de tous ceux qui sont en droit de voter, mais c'est une donnée difficile à obtenir).

Ainsi, rapportés aux inscrits, les scores cités précédemment sont de 17.8% et de 11.2% respectivement : ce n'est pas le Front National qui a progressé, c'est le reste qui s'est effondré autour de lui !

Quel enseignement en tirer ?

- Ce que l'on doit apprendre à l'école et que l'on est en droit d'attendre de tout journaliste et de tout homme politique, c'est qu'il faut évaluer l'électorat par rapport aux votants, pas par rapport aux suffrages exprimés.
- Nul besoin d'être scientifique ou « d'aimer les chiffres » pour acquérir ce réflexe !

A propos des sondages

- Arnold Schwarzenegger, pourtant crédité de 26% des intentions de vote, vient d'être élu gouverneur de la Californie avec 49% des voix.
- Tous les sondages donnaient un classement J. Chirac, L. Jospin, J-M. Le Pen au premier tour en 2002.
- F. Bayrou aurait gagné 1 point dans les sondages grâce à une gifle ...

Première idée forte : les intervalles

- En statistiques, on travaille avec des intervalles, pas avec des nombres. On dit par exemple : entre 14% et 20%.
- Attention, ça ne veut certainement pas dire 17% ! Il se peut que le « pic » de probabilité soit à 16%, autrement dit que les chances **ne soient pas uniformes** dans l'intervalle donné.
- **Aucun** sondage donnant des nombres bruts n'est crédible. Aucun ! En particulier F. Bayrou n'a certainement rien gagné avec sa gifle : statistiquement, il ne s'est **rien** passé.

Seconde idée forte : des groupes, pas des classements !

- La conséquence de ce qui précède est qu'on peut prédire des tendances et des groupements, mais pas prédire un classement. Pas même avec de grosses chances : c'est impossible !
- Par exemple, lors des élections présidentielles de 2002, les sondages ont en fait **toujours** donné trois groupes : un trio de tête, cinq outsiders, un groupe de queue.
- Il n'a **jamais** été possible d'écarter la possibilité que J-M. Le Pen arrive au second tour, ni que J. Chirac n'y parvienne pas. Alors que tous les médias ont parlé du troisième homme (en fait, un des cinq outsiders), ils ont omis de voir qu'il y avait déjà trois prétendants !

Quel enseignement en tirer ?

- Vouloir classer est sans objet. En particulier le classement des lycées donné par les médias est **ridicule**.
- On peut au mieux donner, disons cinq groupes de lycées et dans chaque groupement les lycées sont **indiscernables** statistiquement parlant.
- Il en va de même du classement des hôpitaux.

Du vrai, du faux et de la manipulation

- C'est un phénomène bien connu des physiciens, toute mesure perturbe le phénomène observé. Autrement dit tout sondage d'opinion est faux au moment même où il est rendu public !
- Par ailleurs une fausse élection (sondage) ne remplace pas la vraie et le résultat des fausses élections ne doit pas être un objectif : ce qu'exige la démocratie, c'est un vrai débat d'idées suivi d'une vraie élection.
- Vouloir informer à tout prix, c'est très souvent, dans ce domaine, transmettre des mensonges à très grande échelle. Car une information erronée ou partielle n'est pas un début de vérité, mais un mensonge.
- Certains le savent et l'utilisent. Sous des dehors de sincérité et de « redresseurs de torts » des temps modernes, se cachent des manipulateurs.

Un exemple : Google

- Google est devenu en moins de trois ans le moteur de recherche sur Internet le plus employé au monde (53% des requêtes mondiales, soit 200 millions par jour).
- Il est fondé sur l'algorithme suivant : une page est d'autant plus pertinente qu'elle est fréquemment citée par d'autres.
- En conséquence, pour devenir LA référence du web sur un sujet il suffit d'être cité par de nombreux sites.

Exemples de nouvelles références créées par Google

- Ainsi en cherchant sur google « massive destruction weapons » on tombait sur une page humoristique ironisant « nous n'en avons pas trouvé dans votre pays, il doit y avoir une erreur, veuillez recommencer la recherche » ...
- Un slogan anti-guerre a été reformulé et effacé d'Internet par un expert en stratégie économique de Harvard en 42 jours. Il a juste proposé une autre définition que celle utilisée par Kofi Annan (expliquant qu'il y a deux superpuissances sur la planète : les États-Unis et l'opinion publique) et est devenu LA référence d'Internet, au point que la déclaration de Kofi Annan a quasiment disparu. (<http://www.uzine.net/article1960.html>)

Premier bilan

- Qu'est-ce qu'une information chiffrée ?
 - C'est un choix de paramètres
 - C'est un point de vue sur un problème
 - C'est une façon plus ou moins opportune de décrire un problème
 - C'est une indication en réponse à une question précise et quantifiée.

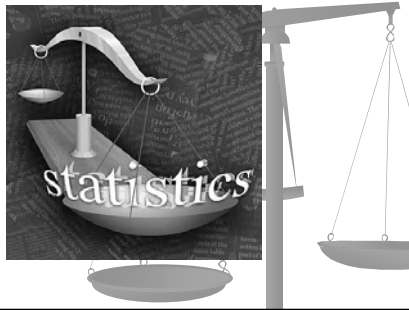
Premier bilan

- Qu'est ce que n'est pas une information chiffrée ?!
 - La description complète des phénomènes intervenant dans un problème
 - L'unique point de vue permettant de comprendre un problème
 - Une certitude absolue et incontestable
 - Une information permettant de légitimer une décision

Nouvelles questions

- Quelle autorité peuvent avoir des informations de nature statistique?
- Peut-on se servir d'une probabilité comme élément de preuve lors d'un procès?
- Des indicateurs statistiques peuvent-ils servir de fondement à une décision politique?

Statistiques et justice



L'affaire Sally Clark

- Le 9 novembre 1999, au Royaume-Uni, Sally Clark est accusée d'avoir tué ses deux enfants :
 - Christopher, âgé de 11 semaines, le 13 décembre 1996 à 21h35.
 - Harry, âgé de 8 semaines, le 26 janvier 1998 à 21h27
- Le 2 octobre 2000, son appel est rejeté.
- Le 29 janvier 2003, elle est libérée en second appel.

Christopher : les faits médicaux

- Christopher est d'abord reconnu comme étant mort de mort naturelle, due à une infection des voies respiratoires inférieures.
- Lors du procès le médecin ayant pratiqué l'autopsie revient sur son diagnostic et assure qu'il n'y a jamais eu aucun signe d'infection et que le bébé a été étouffé.
- Il se base sur des tests sanguins qu'il a menés à l'époque de la mort de Christopher mais dont il n'a pas fait référence dans son rapport post-mortem.

Christopher : les conclusions des experts du procureur

- Dr Williams : étouffement
- Pr Meadow : ni une infection des voies respiratoires inférieures, ni une « Mort Subite du Nourrisson » mais pas non plus une mort naturelle.
- Pr Greene & Dr Keeling : raison inconnue, mais cela pourrait être non naturel.

Christopher : les conclusions des experts de la défense

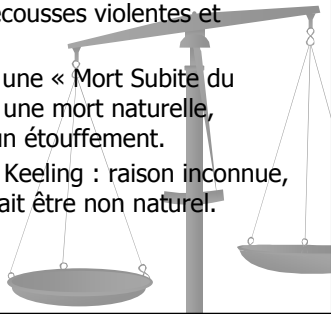
- Pr Berry & Dr Rushton : raison inconnue, c'est-à-dire une maladie non détectée, une mort non naturelle ou une « Mort Subite du Nourrisson ». Le rapport d'autopsie est trop incomplet pour donner lieu à des certitudes.
- Pr David : hémorragie pulmonaire idiopathique, mais une suffocation est possible.

Harry : les faits médicaux

- Des hémorragies (yeux, paupières et moelle épinière)
- Des lésions dans le cerveau
- Dislocation d'une côte
- Le rapport d'autopsie envisage que la mort a été provoquée par de violentes secousses, répétées plusieurs fois.
- Néanmoins la quasi-totalité des éléments énoncés dans le rapport d'autopsie sont sujet à caution, certains étant démontrés comme des artefacts (de la même façon que des centaines de jugements basés sur une analyse ADN sont en train d'être cassés car ils sont entachés d'erreurs de manipulation d'un laboratoire du FBI qui rend les tests totalement inutilisables).

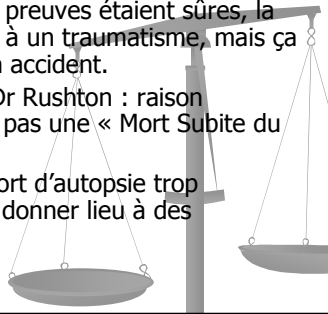
Harry : les conclusions des experts du procureur.

- Dr Williams : secousses violentes et répétées.
- Pr Meadow : ni une « Mort Subite du Nourrisson » ni une mort naturelle, probablement un étouffement.
- Pr Greene & Dr Keeling : raison inconnue, mais cela pourrait être non naturel.



Harry : les conclusions des experts de la défense

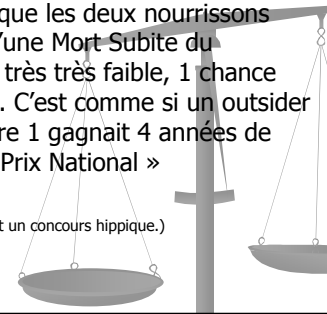
- Pr Berry : si les preuves étaient sûres, la mort serait due à un traumatisme, mais ça pourrait être un accident.
- Dr Whitwell & Dr Rushton : raison inconnue, mais pas une « Mort Subite du Nourrisson ».
- Pr David : rapport d'autopsie trop incomplet pour donner lieu à des certitudes.



L'argument du Pr Meadow

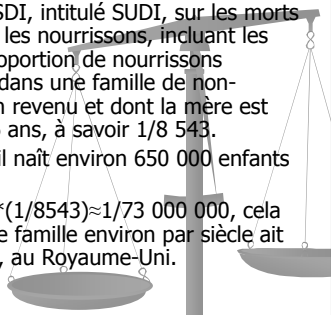
« La probabilité que les deux nourrissons soient morts d'une Mort Subite du Nourrisson est très très faible, 1 chance sur 73 millions. C'est comme si un outsider côté à 80 contre 1 gagnait 4 années de suite le Grand Prix National »

(Le Grand Prix National est un concours hippique.)



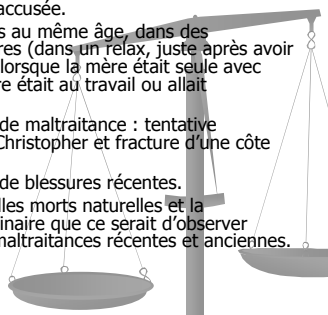
Les bases de l'argument

- Le rapport du CESDI, intitulé SUDI, sur les morts inexplicables chez les nourrissons, incluant les MSN, donne la proportion de nourrissons mourant de MSN dans une famille de non-fumeurs, ayant un revenu et dont la mère est âgée de plus de 26 ans, à savoir 1/8 543.
- Au Royaume-Uni il naît environ 650 000 enfants chaque année.
- Comme $(1/8543) * (1/8543) \approx 1/73\ 000\ 000$, cela fait seulement une famille environ par siècle ait à subir deux MSN, au Royaume-Uni.



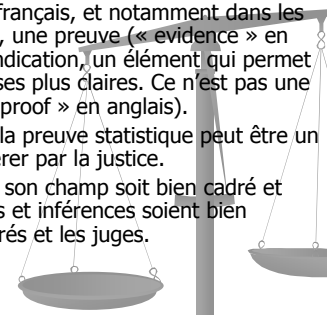
Les conclusions du procès et du premier appel

- Les faits accablent l'accusée.
- Les bébés sont morts au même âge, dans des circonstances similaires (dans un relax, juste après avoir mangé, vers 21h30, lorsque la mère était seule avec l'enfant et que le père était au travail ou allait s'absenter).
- Il y avait des signes de maltraitance : tentative d'étouffement pour Christopher et fracture d'une côte pour Harry.
- Il y avait des signes de blessures récentes.
- La rareté de deux telles morts naturelles et la coïncidence extraordinaire que ce serait d'observer simultanément des maltraitances récentes et anciennes.



La preuve statistique

- Petit rappel : en français, et notamment dans les affaires juridiques, une preuve (« evidence » en anglais) est une indication, un élément qui permet de rendre les choses plus claires. Ce n'est pas une démonstration (« proof » en anglais).
- En tant que telle, la preuve statistique peut être un élément à considérer par la justice.
- Encore faut-il que son champ soit bien cadré et que les arguments et inférences soient bien compris par les jurés et les juges.



La « mystification du procureur »

- C'est un paradoxe très largement débattu et dont il faut se méfier dès que l'on manipule des probabilités.
- L'argument fallacieux consiste à confondre « la probabilité qu'un événement survienne » avec « la probabilité qu'un événement soit survenu dans un cas étudié ».
- Ici c'est confondre « il y a une chance sur 73 millions qu'un double décès naturel survienne » avec « dans le cas d'un double décès, il y a une chance sur 73 millions qu'il soit naturel » ...
- Être innocent d'un crime, c'est se trouver dans la situation où 1° une mort est survenue et 2° un événement extérieur (par exemple une mort naturelle inexplicée) s'est produit.

La « mystification du procureur »

- Par conséquent pour évaluer la probabilité d'innocence de Sally Clark, il faut évaluer la chance qu'il y ait un décès inexplicé (double) **sachant qu'il y a un décès.**
- Autrement dit chercher **le nombre de fois qu'un événement rarissime se produit dans une population très restreinte** (celle de ceux qui ont subi un décès) et non le nombre de fois qu'il se produit au sein de la population totale.

Du rare au très rare

- Quand on analyse des événements rares l'important est de comparer le rare au très rare. Et pareillement avec les événements probables.
- Une chose évidente qui n'a pas été faite était de calculer de la même façon le pourcentage de mères infanticides et récidivistes !
- Ainsi au Royaume-Uni environ 30 enfants sont tués par leur mère chaque année, un nombre comparable au nombre de MSN par an, ce qui fait que les deux hypothèses sont en fait à peu près aussi probables ...
- Un calcul très grossier, utilisant la formule de Bayes, permet de voir que la probabilité que Sally Clark ait tué ses enfants **sachant qu'ils sont morts** est inférieure à 1/3.

A-t-on le droit de transformer la question posée
« Sally Clark est-elle coupable ? »
en une question quantifiée ?

- Assurément, non. L'indication statistique n'a pas d'autre vertu que de donner à réfléchir, stigmatiser les problèmes.
- Elle n'en résout aucun : c'est à l'être humain, par son analyse, de chercher les réponses.
- Idéalement, c'est en cherchant à comprendre la statistique que l'on peut découvrir des clefs que l'on avait omises ...

Quel était l'objet du rapport SUDI dont s'est servi le Pr Meadow ?

- Il s'agit d'un rapport en vue de la prévention du risque de MSN, notamment d'un deuxième tel événement dans une même famille.
- La probabilité 1/8 543 est donnée en comparaison avec la probabilité 1/ 214 obtenue avec des parents fumeurs, sans revenu et avec une mère âgée de moins de 26 ans.
- Notamment, le rapport explicite : « Ainsi, pour les familles comportant plusieurs facteurs de risques pour la MSN, une seconde MSN, bien que très rare, est 1 600 fois plus probable que dans les familles dans ces facteurs. »

Troisième idée forte : mettre en
balance les hypothèses

- **Une telle statistique n'a de pertinence que si elle est utilisée pour comparer deux événements**
- C'est ce qui est fait dans le rapport SUDI.
- Ici il faudrait mettre en balance les deux hypothèses « meurtre » et « mort naturelle ».

Les erreurs

- De façon assez remarquable, personne n'a jamais prétendu que les morts de Christopher et Harry étaient des MSN, puisqu'à chaque fois on a cherché à trouver des explications et qu'on a trouvé des éléments de réponse (par exemple Harry est mort 4 heures après avoir été vacciné). Le calcul même du Pr Meadow est sans objet !
- L'obtention du rapport 1/73 000 000 est également erronée. Inutile d'y revenir, car ce serait une querelle d'experts alors que le problème est ailleurs !

Le bon sens mis en défaut

- Et pourtant le Pr Meadow a marqué l'opinion publique en faisant croire
 - d'une part qu'il y avait 1 chance sur 73 millions pour que Sally Clark soit innocente,
 - d'autre part qu'en prouvant que les morts n'étaient pas des MSN, elles n'étaient pas naturelles.
- Son image avec le concours hippique est également erronée car, bien évidemment, il n'y a pas 650 000 Grand Prix par an, mais un seul et qu'on ne peut pas faire des statistiques sur un seul évènement !

Quatrième idée forte : la taille de la population testée ou questionnée

- Il n'est pas rare qu'en médecine les tests sur la foi desquels sont prises des décisions soient effectués sur des populations aussi petites que 5 individus.
- Par exemple l'étude SUDI analysait les doubles décès à travers l'étude de 24 familles.

Pourquoi Sally Clark a-t-elle été libérée ?

- Contrairement à ce qu'on pourrait penser ou espérer, ce n'est pas grâce aux arguments précédents.
- Il a en fait été prouvé que les autopsies ont été pratiquées en dehors des normes de rigueur, notamment quant à la transmission des faits observés.
- Des analyses sanguines demandées pour Harry lors de l'autopsie avaient des résultats alarmants qui n'ont pas été communiqués. Seule la persévérance du mari et du père de Sally Clark a permis de les découvrir. On peut penser à partir de ces analyses qu'Harry est mort de méningite ou d'un autre type de maladie fulgurante liée au staphylocoque doré.

La couverture médiatique

- Les médias anglo-saxons ont largement couvert le procès et les appels et ont contribué à l'accent mis sur la « preuve statistique ».
- Ils ont également véhiculé l'argument fallacieux dénoncé précédemment : la mystification du procureur.

Polémique

- Faut-il interdire les statistiques dans les affaires de justice ?
- Faut-il demander la présence d'experts en statistiques au barreau ?
- Faut-il que toute personne ait appris à appréhender ou à manipuler des statistiques ?

Tests en médecine



Le dosage HT21 ou « Triple test »

- Le dosage « Human Trisomy 21 » est un dosage sanguin de deux ou trois marqueurs pris dans le sang de la mère (entre 15 et 18 semaines d'aménorrhée).
- A partir de ces dosages est calculé un risque que l'enfant soit affecté d'une anomalie du chromosome 21 (aussi connue sous le nom de mongolisme).

Triple-test et Amniocentèse

- D'une part le triple-test ne détecte pas tous les fœtus atteint de trisomie 21.
- D'autre part, très souvent, il donne un résultat positif alors que le fœtus est normal.
- Lorsque le résultat est positif, une confirmation doit être obtenue par amniocentèse.

Quelques données

860 000 tests
en 1997 et 1998

799 700 enfants normaux (Test HT21 négatif)	59 300 enfants normaux faussement détectés (Test HT21 faussement positif)	700 trisomies 21 détectées par amniocentèse (Test HT21 positif)	300 trisomies 21 non détectées (Test HT21 faussement négatif)
--	---	---	---

Tests positifs 7 %	dont 98.8% faussement dont 1.2% confirmés	soit 6.9% des femmes testées soit 0.08% des femmes testées soit 70% des trisomies
Tests négatifs 93 %	dont 0.04% faussement dont 99.96% vraiment	soit 0.03% des femmes testées soit 30% des trisomies soit 93% des femmes testées

Limites du triple-test

- En fait des études montrent que l'on peut affiner le résultat du test par deux facteurs : l'âge de la mère et la clarté nucale.
- La mesure de la clarté nucale par échographie est un geste difficile et précis.
- L'amniocentèse comporte des risques de fausse couche évalués à 1% environ.

Amniocentèse

- Elle est proposée aux femmes de plus de 38 ans et aux femmes dont le risque calculé est supérieur à 1/250. Elle est alors remboursée.
- Les autres femmes peuvent prendre à leur charge une amniocentèse, avec le coût et le risque de fausse couche que cela représente.

Justifications statistiques ?

- Le test est validé en **comparant** deux phénomènes : la positivité au test et la trisomie 21.
- Que représente le seuil (imposé par la loi) de 1/250 ? Il est destiné à minimiser le risque de faux négatifs, mais par là même il augmente le nombre de faux positifs.
- Il faut en particulier comparer ce seuil 1/250 au nombre de fausses couches dues à l'amniocentèse (1/100) et au nombre de fœtus atteints de trisomie 21 (1/860).

Quelques doutes

- Le triple-test est globalement peu fiable, à la fois quant à la méthode et quant à sa pratique.
- C'est son principe même qui contient ses limites : les modifications biologiques entraînées par une trisomie 21 ne sont ni systématiques, ni exclusives de cette pathologie.

Quelques remarques polémiques

- Si l'on croit au risque dû à l'amniocentèse, il y a donc, sur 860 000 grossesses testées, 700 trisomies 21 détectées, 300 trisomies 21 non détectées et 600 fausses couches ayant entraîné la mort d'un enfant sain.
- On est loin du 70% qui donne confiance et on est un peu effrayé par le 37% de morts « inutiles ».

Trisomie 21 détectée	Trisomie 21 non détectée	Mort d'un enfant sain
44%	19%	37%

Quelques remarques polémiques

- Connait-on l'impact d'une anxiété prolongée (3 à 4 semaines d'attente pour connaître le résultat de l'amniocentèse) durant la grossesse, sur l'enfant ?
- Est-on en droit de pousser des parents à prendre le risque d'une fausse couche pour se prémunir d'une trisomie 21 ?

Quelques évidences jamais dites

- Le risque de fausse couche dû à l'amniocentèse n'est pas uniforme.
- Il dépend du praticien, même si la technique est de mieux en mieux maîtrisée et que ce facteur est, de facto, mineur.
- Mais surtout il y a des populations à risque. Par exemple lorsque l'utérus a développé un fibrome, le risque est plus élevé. En pratique le risque 1/100 ne représente pas grand-chose : il y a des femmes pour qui le risque est bien plus important alors que pour les autres le risque est réellement minime.

Bilan : les réflexes à acquérir

- Pour comparer des pourcentages, il faut les ramener à un « tout » commun.
- Les statistiques donnent des intervalles, pas des nombres.
- On ne peut pas donner un classement individuel statistique, mais un classement par groupes.
- On peut opposer deux hypothèses pour décider laquelle est la plus probable, mais la notion de probabilité « absolue » n'a pas de sens.
- Une statistique n'a de pertinence que si elle s'est opérée sur un grand nombre de cas : dans le cas contraire, l'intervalle obtenu statistiquement est tout simplement tellement gigantesque qu'on ne peut rien en conclure.

